UNIVERSIDADE FEDERAL DE PELOTAS

Programa de Pós-Graduação em Biotecnologia

Tese

Desenvolvimento de ferramentas para a manipulação genética e metodologias para a identificação de fatores de virulência de *Leptospira* spp.

Gustavo Maia de Cerqueira

Pelotas, 2009

GUSTAVO MAIA DE CERQUEIRA

Desenvolvimento de ferramentas para a manipulação genética e metodologias para a identificação de fatores de virulência de *Leptospira* spp.

Tese apresentada ao Programa de Pós-Graduação em Biotecnologia da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Doutor em Ciências (área de conhecimento: Biologia Molecular).

Orientador: Odir Antônio Dellagostin

Pelotas, 2009

Banca examinadora:

Prof. Dr. Alan J. A. McBride, Centro de Pesquisas Gonçalo Moniz

Prof. Dr. Fabiana K. Seixas, Universidade Federal de Pelotas

Prof. Dr. Fábio Leivas Leite, Universidade Federal de Pelotas

Prof. Dr. Odir Antônio Dellagostin, Universidade Federal de Pelotas

1

AGRADECIMENTOS

À Universidade Federal de Pelotas pela oportunidade de realização do Curso de Pós-Graduação em Biotecnologia.

Ao meu orientador, Odir A. Dellagostin, pela valiosa orientação, experiência e confiança dispensada na execução deste trabalho, o qual contribuiu para meu amadurecimento e formação profissional.

Ao meu orientador de doutorado sanduíche, Mathieu Picardeau, pela orientação e confiança dispensadas na execução do meu estágio de doutorando no Institut Pasteur.

A minha esposa e filhos, por todo amor e companheirismo e por sempre me incentivar acreditando em meu potencial.

A minha família, pelo carinho e incentivo.

Aos meus amigos e colegas de laboratório de Biologia Molecular pelo apoio, incentivo e amizade.

Aos colegas, estagiários e amigos do Centro de Biotecnologia, pela amizade e convívio agradável.

Aos colegas e amigos do Institut Pasteur, pela convivência diária, pelos ensinamentos transmitidos e pela amizade desenvolvida.

A todos que direta ou indiretamente contribuíram de alguma forma para a realização deste trabalho.

Muito Obrigado

RESUMO

CERQUEIRA, Gustavo Maia. **Desenvolvimento de ferramentas para a manipulação genética e metodologias para a identificação de fatores de virulência de *Leptospira* spp.** 2009. 145 f. Tese (Doutorado) - Programa de Pós-Graduação em Biotecnologia. Universidade Federal de Pelotas, Pelotas.

Neste estudo, o *Himar1* foi utilizado para a obtenção de um total de 929 mutantes, dos quais, 721 correspondem a seqüências codificadoras (CDS) interrompidas pelo transposon, cobrindo 551 genes diferentes. Alguns mutantes foram avaliados com relação aos efeitos da interrupção gênica sobre a virulência em modelo animal, e dois mutantes atenuados contendo o transposon em genes hipotéticos foram identificados. Outro estudo realizado buscou desenvolver uma nova ferramenta genética para auxiliar no estudo da função de genes específicos. Assim, foi desenvolvido um sistema de expressão induzível para *L. biflexa*. Este sistema de expressão empregou como repórter o gene da proteína verde fluorescente (*gfp*) e o gene da proteína flagelina B (*flaB*). *L. biflexa* fluorescente foi observada, assim como leptospiras que voltaram a ter mobilidade após a adição do agente indutor (IPTG). Finalmente, foi conduzido um estudo para a determinação da presença dos genes da família *lig* nas diversas espécies de leptospiras patogênicas, através da técnica de PCR. O gene *ligB* apareceu amplamente distribuído, enquanto que *ligA* e *ligC* foram detectados apenas em um reduzido número de sorovares. Além disso, a sequência de um fragmento específico de 214 pb do gene *ligB* pode ser usada para a constituição de um novo método para a classificação das espécies patogênicas de *Leptospira*.

ABSTRACT

CERQUEIRA, Gustavo Maia. D**evelopment of tools for the genetic manipulation and methodologies for the identification of virulence factors of *Leptospira* spp.** 2009. 145 p. Thesis (PhD degree) - Programa de Pós-Graduação em Biotecnologia. Universidade Federal de Pelotas, Pelotas.

In this study, *Himar1* was used to obtain a total of 929 transposon mutants, of which, 721 correspond to coding sequences (CDS) disrupted by the transposon in 551 different genes. Some mutants were evaluated regarding the effect of gene disruption to virulence in the hamster model, and two attenuated mutants containing the transposon into hypothetical genes were identified. Another study aimed to develop a new genetic tool to help in the study of specific genes. Thus, an inducible expression system for *L. biflexa* was developed. Such inducible expression system employed as reporter genes the green fluorescence protein (*gfp*) and the gene encoding for the flagelin B protein (*flaB*). Fluorescent *L. biflexa* (GFP) and motile leptospires (FlaB) were observed after induction by IPTG. Finally, another study was conducted to determine, by PCR amplification, the presence of the *lig* genes in different pathogenic species. *ligB* appeared to be ubiquitously distributed, while *ligA* and *ligC* were detected only in a reduced number of serovars. In addition, a 214 bp specific *ligB* fragment was used to constitute a new method for pathogenic *Leptospira* species classification.

RÉSUMÉ

CERQUEIRA, Gustavo Maia. **Développement de nouveaux outils pour la manipulation génétique et methodes pour l'identification de nouveaux éléments liés à la virulence de *Leptospira* spp.** 2009. 145 f. Thèse (Doctorat) - Programa de Pós - Graduação em Biotecnologia. Universidade Federal de Pelotas, Pelotas.

Dans cet étude le transposon Himar1 a été utilisé pour l'obtention d'un total de 929 mutant, parmi lesquels 721 correspondent à des séquences codantes (CDS) interrompues par le transposon dans 551 gènes différent. La virulence de quelques mutants a été évaluée dans le modèle hamster, et deux mutants atténués, lesquelles contiennent une interruption dans des gènes hypothétiques, ont été identifiés. Une autre étude a consisté à développer un nouveau outil génétique pour l'étude de l'expression des gènes. Ainsi, à l'aide du système Lac, il a été développé un système d'expression inductible pour *L. biflexa*. Ce système a utilisé le gène de la protéine fluorescente GFP ou le gène de la flagelline FlaB (dans un mutant flaB immobile) comme gènes rapporteurs. Des bactéries de *L. biflexa* fluorescentes ou mobiles ont été observées après addition d'IPTG. Finalement, on a étudié par PCR la distribuition des gènes *lig* parmi les leptospires pathogènes. Le gène *ligB* a été largement retrouvé, par contre *ligA* et *ligC* ont été détectés uniquement dans un nombre réduit de sérovars. La séquence d'un fragment spécifique de 214 bp du gène *ligB* peut être utilisé pour l'identification des souches pathogènes de *Leptospira*, constituant une nouvelle méthode pour la classification des espèces pathogènes de *Leptospira.*

SUMÁRIO

1. INTRODUÇÃO GERAL

Leptospirose é uma doença infecciosa do tipo antropozoonose, e representa um problema de saúde pública que afeta principalmente países em desenvolvimento [LEVETT, 2001; BHARTI et al., 2003; McBRIDE et al., 2005]. Dentre as espécies patogênicas que compõem o gênero *Leptospira*, *L. interrogans* aparece como o principal agente responsável pela leptospirose. Diversas espécies mamíferas, incluindo roedores, podem servir como reservatórios excretando espiroquetas na urina. Humanos são geralmente infectados através do contato com água e solo contaminados. A maior problemática causada pela leptospirose reside em ambientes rurais pobres e em populações de favelas urbanas nos países em desenvolvimento [LEVETT, 2001; BHARTI et al., 2003; McBRIDE et al., 2005]. Mais de 500.000 casos de leptospirose grave ocorrem a cada ano, com taxa de mortalidade entre 5 e 20% [WHO, 1999]. Pouco se sabe sobre a patogenia de *Leptospira*, o que desfavorece a identificação de novas estratégias de intervenção.

Leptospiras são bactérias altamente móveis e capazes de penetrar as membranas da pele e mucosas e rapidamente se disseminar para outros tecidos, logo após a infecção. Em hospedeiros suscetíveis (humanos e animais de grande e pequeno porte – bovinos, eqüinos, suínos, caninos e felinos entre outros) a infecção sistêmica produz manifestações graves em diversos órgãos, incluindo icterícia, falha renal grave e hemorragia grave nos pulmões e outros órgãos. Entretanto, em reservatórios animais como o rato doméstico, a infecção produz carreamento persistente assintomático nos túbulos renais [LEVETT, 2001; BHARTI et al., 2003; McBRIDE et al., 2005].

Atualmente são conhecidas 8 espécies patogênicas [MOREY et al., 2006; SLACK et al., 2008]. Dentro das diferentes espécies, encontram-se sorovares antigenicamente relacionados, que constituem os sorogrupos. Já foram descritos quase 300 sorovares os quais estão distribuídos entre aproximadamente 28 sorogrupos [LEVETT, 2001].

Há muito se discute sobre a necessidade de uma vacina protetora de amplo espectro contra a leptospirose. Dessa forma, diversas proteínas de membrana externa de *Leptospira* têm sido investigadas para esse fim. Antígenos candidatos incluem as proteínas LipL32 [BRANGER et al., 2001; 2005], OmpL1 e LipL41

[HAAKE et al., 1999], as proteínas LigA e LigB [KOIZUMI e WATANABE, 2004; PALANIAPPAN et al., 2006; SILVA et al., 2007; FAISAL et al., 2008; YAN et al. 2008; FAISAL et al., 2009], entre outros antígenos [CHANG et al., 2007].

Os mecanismos de virulência, e a compreensão fundamental da biologia dos agentes causadores de leptospirose permanencem desconhecidos. Até o momento, apenas poucas proteínas têm sido caracterizadas como fatores ligados ou supostamente associados à virulência [MATSUNAGA et al., 2005; STEVENSON et al., 2007; CHOY et al., 2007; RISTOW et al., 2007; CRODA et al., 2008; MURRAY et al., 2009a,b,c]. Estudos anteriores têm mostrado que leptospiras patogênicas expressam adesinas [BARBOSA et al., 2006; MERIEN et al., 2000], hemolisinas [LEE et al., 2002] e um grande número de lipoproteínas que podem eventualmente atuar na interação patógeno-hospedeiro [CULLEN et al., 2005]. Entretanto, o papel desses supostos fatores de virulência permanece especulativo. A recente publicação da sequência completa dos genomas de sorovares patogênicos e saprófitas de *Leptospira* [REN et al., 2003; NASCIMENTO et al., 2004; BULACH et al., 2006] tem fornecido uma base para auxiliar na compreensão da patogênese de *Leptospira* spp. e permitiram avançar inclusive na tentativa de se identificar novos candidatos ao desenvolvimento de uma vacina. A falta de ferramentas genéticas para manipular leptospiras patogênicas tem dificultado a elucidação do papel destes determinantes na virulência.

A evidência da transferência genética em *L. interrogans*, por meio do uso de um elemento transponível de origem eucariótica [BOURHY et al., 2005], abriu portas para a identificação de proteínas ligadas à virulência de *Leptospira*. Até o momento quatro importantes genes de *Leptospira*, os quais codificam para as proteínas Loa22 [RISTOW et al., 2007], LipL32 [MURRAY et al., 2009b], LigB [CRODA et al., 2008] e Heme oxigenase [MURRAY et al., 2009c] foram interrompidos pela tecnologia de transposon mutagênese, entretanto apenas Loa22 e Heme oxigenase demonstraram uma influência direta sobre a virulência de *L. interrogans*.

Neste estudo, um banco de mutantes de *Leptospira* foi construído onde mais de 900 mutantes foram obtidos pelo método de transposon mutagênese, entretanto, apenas uma pequena parte destes mutantes foi testada quanto à sua virulência em modelo animal [MURRAY et al., 2009a]. Um dos mutantes obtidos, o qual se encontra em análise, perdeu completamente a virulência quando um gene codificador de uma chaperona foi interrompido. Além disso, outras alterações

funcionais foram constatadas em função da perda da expressão desta chaperona. Foi observado que o número de cepas mutantes viáveis com mutação em genes *housekeeping* foi baixo, possivelmente por resultarem em mutações letais. Para possibilitar a obtenção de mutantes letais condicionais, uma nova ferramenta foi desenvolvida. Esta ferramenta é capaz de expressar de forma controlada genes heterólogos ou da própria *Leptospira* sob a regulação de um promotor de *E. coli* ou de *Leptospira*, sendo este último modificado pela inserção de um sítio operador do operon da lactose. Assim, o sistema pode ser facilmente controlado pela utilização de IPTG, um agente indutor não hidrolisável. Por fim, outro estudo foi realizado o qual buscou identificar a presença dos genes *lig* de *Leptospira* entre sorovares patogênicos, intermediários e saprófitas. O único gene que apareceu presente entre todas as cepas patogênicas foi *ligB*, cuja análise da sequência possibilitou a discriminação entre as espécies patogênicas de *Leptospira*.

A tese está apresentada na forma de artigos científicos. Inicialmente é apresentada uma revisão que abrange os campos de taxonomia, filogenia e epidemiologia molecular *Leptospira* spp. Nesta revisão foi apresentado um histórico da utilização de todas as ferramentas já empregadas na classificação de *Leptospira* spp. nos níveis de espécie, sorogrupo, sorovar e isolado. Este artigo “A century of typing of *Leptospira* strains” foi solicitado pelo editor chefe Dr. Michel Tibayrenc para publicação na revista **Infection Genetics and Evolution**. Em seguida, no artigo “Genome-wide transposon mutagenesis in Pathogenic Leptospira species” (artigo 2), foi descrita a construção de um banco de mutantes gerados por transposon mutagênese. Alguns dos mutantes foram usados para infecção em modelo de hamster onde se identificou novos mutantes atenuados. Este artigo foi publicado no periódico **Infection and Immunity**. O artigo 3 trata da construção de um sistema de expressão induzível em *Leptospira* spp. Este sistema pode ser facilmente induzido pelo uso de IPTG e é controlado pelo repressor LacI. Para a validação deste sistema dois repórteres foram testados: o gene que codifica para a proteína GFP e o gene *flaB*. Este último codifica para a proteína flagelina B e o sistema que o inclui foi utilizado para complementar um mutante *flaB*$^{-}$ de *L. biflexa*. Este trabalho será submetido para publicação no periódico **Applied and Environmental Microbiology**.

Na continuidade (artigo 4), a presença e distribuição dos genes *lig* de *Leptospira* foram avaliados por PCR. Neste trabalho foi mostrado que o gene *ligB* é o único presente entre todos os sorovares de *Leptospira* testados. Inclusive, foi

também demonstrado que um fragmento de 214 pb do gene *ligB* é capaz de identificar corretamente as espécies de *Leptospira*. Esse trabalho será submetido para publicação no periódico **Journal of Medical Microbiology.**

2. OBJETIVOS

2.1. OBJETIVOS GERAIS

Os trabalhos aqui gerados tiveram por objetivo desenvolver e utilizar ferramentas e métodos para a manipulação e caracterização genética de *Leptospira* spp.

2.2. OBJETIVOS ESPECÍFICOS

1. Desenvolver um banco de mutantes por transposon mutagênese, utilizando o transposon Himar 1, e demonstrar a aplicabilidade deste método para identificar a participação de proteínas em processos biológicos.
2. Desenvolver um sistema de expressão induzível para *L. biflexa* e avaliar suas potenciais aplicações.
3. Desenvolver um método de detecção das regiões codificadoras das importantes proteínas *lig* de *Leptospira* nas diferentes espécies.

3. ARTIGO 1

A CENTURY OF TYPING OF *Leptospira* STRAINS

(Artigo de revisão solicitado pelo editor chefe do periódico *Infection Genetics and Evolution*)

A century of typing of *Leptospira* strains

Gustavo M. Cerqueira [1, 2] and Mathieu Picardeau [2] *

[1] Institut Pasteur, Unité de Biologie des Spirochètes, Paris, France, [2] Centro de Biotecnologia, Universidade Federal de Pelotas, Pelotas, Brazil

phylogeny/taxonomy/molecular epidemiology

* Corresponding author.

Mailing address: Mathieu Picardeau, Unité de Biologie des Spirochètes, Institut Pasteur, 28 rue du docteur Roux, 75724 Paris Cedex 15, France. Tel: 33 (1) 45 68 83 68. Fax: 33 (1) 40 61 30 01. E-mail: mpicard@pasteur.fr

Introduction

Leptospira and leptospirosis

The spirochetes, which include medically important pathogens such as the causative agents of Lyme disease, syphilis, and leptospirosis, constitute an evolutively unique group of bacteria. First described by Weil in 1886, leptospirosis is a zoonosis of world-wide distribution. Rodents are the main reservoir of the disease, excreting the bacteria into their urine. Humans are usually infected through contaminated water. Leptospirosis has been identified as a re-emerging infectious disease, particularly in tropical and subtropical regions. More than 500,000 cases of severe leptospirosis occur each year, with a mortality rate >10% and severe haemorrhage syndrome >50% (McBride et al., 2005).

Antibiotic therapy is beneficial but must be administered in the early stage of the disease. However, leptospirosis is often lately diagnosed, due to its wide spectrum of symptoms, ranging from a flu-like syndrome to renal failure. The symptoms mimic the clinical presentations of many other diseases, including Dengue fever and Malaria. Laboratory diagnosis of leptospirosis is therefore important in order to provide better patient care. Characterization of leptospires is also essential for understanding the epidemiology of the disease. Serovar is the basic taxon of leptospires and it is defined based on the structural heterogeneity in the carbohydrate component of the lipopolysaccharide (LPS). Historically, the genus *Leptospira* is divided into several hundred serovars of two species: *L. interrogans* and *L. biflexa*, which contained pathogenic and saprophytic strains respectively. Based upon DNA-DNA hybridization data, the genus is now classified into 19 species,(Yasuda *et al*., 1987; Ramadass *et al*., 1992; Perolat *et al*., 1998; Brenner *et al*., 1999; Levett *et al*., 2005; Levett *et al*., 2006; Matthias *et al*., 2008; Slack *et al*., 2008). However, identification is further complicated because serovars of a same serogroup can be distributed between different species.

Taxonomy and classification

Leptospires were first seen in silver-stained tissues from a patient by Stimson in 1907 (Stimson, 1907). At that time, the agent of leptospirosis was identified as *Spirochaeta interrogans* without any more information on the bacterium. The naming of the species was therefore not in conformity with the requirements of the International Code of Nomenclature. The first valid description of saprophytic

Leptospira is given by Wolbach and Binger (Wolbach & Binger, 1914) and that of pathogenic ones by Inada et al. (Inada *et al.*, 1916). The genus *Leptospira* was initially divided into two groups: the pathogenic leptospires referred to as *Leptospira interrogans* sensu lato and the saprophytic leptospires referred to as *L. biflexa* sensu lato. The saprophytes are ubiquitous in the environment and usually found in fresh surface water. Unlike pathogenic leptospires, they have rarely been associated with animal hosts and reportedly do not produce infections in experimental animal models. The two groups also differ in nutritional requirements and other phenotypic properties. For example, growth of pathogenic leptospires is inhibited by the purine analogue 8-azaguanine, whereas saprophytic leptospires are able to grow (Johnson & Rogers, 1964). In addition, pathogenic leptospires have a generation time of about 20 hours and are considered as slow-growing bacteria, in comparison with saprophytes (generation time around 5 hours).

Serological classification

Serotyping is a useful epidemiologic tool because establishing the causative serogroup or serovar is the first step towards identifying reservoirs and generating control strategies. For example, rats are generally maintenance hosts for strains of the Icterohaemorrhagiae serogroup. Agglutination tests for leptospiral antibody were developed soon after the first isolation of leptospires, which occurred a century ago (Martin & Pettit, 1918). The reference method for serological identification, the Microscopic Agglutination Test (MAT), is a complex and fastidious test since it requires live cultures of collection strains that will be used as antigens to detect agglutinating antibodies (Turner, 1968). The results are also difficult to standardize because they depend on the biologist operating the microscope (Levett, 2001). Several studies conducted worldwide are published annually, which employ MAT for the serological identification of leptospires. A previous study evaluated the ability of the MAT to infer the serovar identity of infecting leptospires, but for more than one-half of the patients in that study this was not possible (Levett, 2003). In addition, MAT does not permit early diagnosis because it relies on detection of antibodies to leptospiral antigens and cannot detect infection until one week after the onset of symptoms. In a non-endemic area, if a low titre (>1/100) is obtained in the MAT test with one or a number of antigens, this may indicate leptospirosis. In endemic areas, a high titre (>400-800) against antigens is required to suspect leptospirosis. However,

confirmation of leptospirosis is obtained by examination of a second serum sample about one week later with a 4-fold rise in titre or seroconversion (WHO, 2003). High agglutination of the serum to one of the antigen will then predict the presumptive infecting serogroup (Dikken & Kmety, 1978). The interpretation of the results is complicated by the frequent cross-reactions that occur between serogroups, i.e. cross-reactions against the serogrop Semaranga strain Patoc 1, which are due to the absence of local isolates into leptospiral panels (Ooteman et al., 2006). A broad range of serogroups should be represented in the panel of antigens used in the MAT to maximize the probability of detecting an immune response to one of the 24 reported pathogenic serogroups (Turner, 1968; Faine, 1982). Serogroups have no official taxonomic status but serve the practical purpose of grouping strains that share common antigens. A high degree of cross-reaction occurs between different serogroups, especially when working on acute-phase serum samples. The use of MAT for serovar determination is also complicated by the often observation of similar titers to all serovars of an individual serogroup (Ahmad et al., 2005). To overcome this current serovar determination is done using monoclonal antibodies.

Analysis by cross agglutination absorption test (CAAT) led to the definition of serovar, which is considered as the basic systematic unit of leptospires. Two strains are considered different if, after cross-absorption with adequate amounts of heterologous antigens, at least 10% of the heterologous titre regularly remains in either of the two antisera. For decades CAAT has been used to classify leptospires (Kmety & Dikken, 1993) and now approximately 250 pathogenic serovars have been recognized. About 80 serovars have been identified in *L. interrogans* sensu stricto (Brenner *et al.*, 1999); among them, 60 serovars are validly described (Kmety & Dikken, 1993). Antigenically related serovars constitute serogroups; thus, 24 serogroups have been described in pathogens. The list of serovars is updated periodically and, recently, two new pathogenic serovars have been described (Corney *et al.*, 2008; Valverde M de *et al.*, 2008) (Table 1). However, CAAT is cumbersome and time-consuming for routine typing, mainly because of the time needed for the preparation of immune serum in rabbits. Although serovar identification of isolates is essential to understand the epidemiology of the disease, few laboratories are able to perform CAAT (Terpstra *et al.*, 1985) and therefore most isolates are not identified at the serovar level.

With the emergence of molecular typing methods, it has become increasingly clearer that the serovar concept is no longer fully satisfactory as it may fail to adequately define epidemiologically important strains. For example, molecular typing was found to better discriminate strains from the Grippotyphosa serogroup than serological typing (Hartskeerl *et al.*, 2004; Steinen *et al.*, 1992).

Alternatives have been evaluated to serogroup leptospires with modifications to improve upon data reliability and interpretability. An immunoblotting of whole-cell bacteria was performed and evaluated as an attractive method in comparison with the reference standard MAT, because its simplicity, feasibility and suitability for laboratory diagnosis (Doungchawee et al., 2007). Although immunobloting also produced cross-reaction, it was more easily distinguishable. Immunobloting was also successful in discriminating pathogenic leptospires from the non-pathogenic *L. biflexa*.

Classification based on DNA relatedness

Yasuda *et al.* demonstrated by genomic DNA-DNA hybridization that the different strains of *L. interrogans* constitute not one but at least six distinct species (Yasuda *et al.*, 1987). Based on further DNA hybridization experiments, a total of 19 species, including 13 named species (*L. interrogans*, *L. kirschneri*, *L. borgpetersenii*, *L. santarosai*, *L. noguchii*, *L. weilii*, *L. inadai*, *L. alexanderi*, *L. broomii*, *L. fainei*, *L. biflexa*, *L. meyeri*, and *L. wolbachii*) have been described in the genus *Leptospira* (Brenner *et al.*, 1999; Levett *et al.*, 2005; Levett *et al.*, 2006; Perolat *et al.*, 1998; Slack *et al.*, 2008; Yasuda *et al.*, 1987). More recently, other species, called "*Leptospira licerasiae*" and *"Leptospira wolffii"*, were also described (Matthias *et al.*, 2008; Slack et al., 2008). Based on their pathogenic, saprophytic or doubtful nature, leptospires are clustered in three groups: the pathogens (*L. interrogans*, *L. kirschneri*, *L. borgpetersenii*, *L. santarosai*, *L. noguchii*, *L. weilii*, *L. alexanderi*, *L. wolbachii* and *L.* genomospecies 1), the intermediates (*L. broomii*, *L. inadai*, *L. fainei*, *L. licerasiae* and *L. wolffii*), and the saprophytes (*L. biflexa*, *L. meyeri* and *L.* genomospecies 2-5). Phylogenetic analysis reveals three clades, representing species that contain pathogenic serovars, non-pathogenic serovars and an intermediate group (Matthias *et al.*, 2008; Paster *et al.*, 1991; Schmid *et al.*, 1986). However, there is a poor correlation between the serological and genotypic classification systems (Brenner *et*

al., 1999; Yasuda *et al.*, 1987). A given serogroup is often found in several *Leptospira* species. For instance, the 9 validly described serovars from Bataviae serogroup are distributed among *L. interrogans* sensu stricto species (2 serovars), *L. santarosai* (4 serovars), *L. kirschneri* (1 serovar), *L. noguchii* (1 serovar) and *L. borgpetersenii* (1 serovar) (Levett, 2001). Several studies have thus shown that the system of serogroups was not related to molecular classification. Nowadays, the classification system based on genetic similarities is therefore being used in conjunction with classical antigenic classification.

Phylogeny

Sequencing of the *rrs* gene is a standard approach for differentiating species in all branches of the phylogenetic tree of life. *Leptospira* belongs to the bacterial phylum of spirochetes, which has a deep branching lineage in Bacteria, as indicated by 16S rDNA analysis (Paster *et al.*, 1991). Based on comparative 16S rDNA sequence analysis, Figure 1 depicts the phylogeny of all known *Leptospira* species. The clades contained branches that, with few exceptions, reflected species designations based on the “pathogenic” status (i.e. pathogenic, saprophytic and intermediate strains of unclear pathogenicity), MLST analysis (Ahmed *et al.*, 2006), *S10-spc-α* locus analysis (Victoria *et al.*, 2008), and DNA homology data (Brenner *et al.*, 1999; Yasuda *et al.*, 1987). The “pathogenic” status of the group of intermediates remains controversial. For example, inoculation of the intermediates *L. inadai* and *L. licerasiae*, both recovered from patients, in the hamster model of infection does not produce death or clinical signs of leptospirosis (Schmid *et al.*, 1986; Matthias *et al.*, 2008). The other species *L. fainei* and *L. broomii* have been isolated from humans but have never been tested for reproducibility of the virulence in the hamster model (Perolat et al., 1998; Petersen et al., 2001; Arzouni et al., 2002; Levett et al., 2006). The subgroup of saprophytes, which includes *L. biflexa*, forms the deepest branch within the genus. The emergence of pathogenic leptospires may have arose as the result of changes in the ecology of its mammalian host. However, lateral genetic transfer among leptospires (McBride et al., 2009; Haake *et al.*, 2004; Ralph & McClelland, 1994) prevent the construction of species phylogenetic trees by whole gene sequencing.

Other genes have been examined for the purpose of species discrimination within the genus *Leptospira*, this includes *rpoB* (La Scola et al., 2006), *gyrB* (Slack et

al., 2006) and more recently the partial *ligB* (Cerqueira GM, manuscript in preparation). All proposed taxonomic markers corroborate the findings of the 16S rDNA gene sequences with respect to the clustering of strains into three major groups composed by pathogenic, intermediate and saprophytes.

Molecular typing

In the 1990s, typing methods began to evolve from phenotype-based methods toward genotype-based methods such as pulsed-field gel electrophoresis (PFGE) and PCR-based methods. Identification at the species level is not informative, except in identifying pathogenic species, for epidemiological studies. Since each serovar is usually associated with a particular host, identification of serovars is essential to epidemiological studies and strategies of prevention (Faine *et al.*, 1999). Serovars can be characterized by different molecular methods such as RFLP-based methods (Perolat *et al.*, 1993; Thiermann *et al.*, 1985), AP-PCR (Ralph *et al.*, 1993), and PFGE (Herrmann *et al.*, 1991; Herrmann *et al.*, 1992; Galloway and Levett, 2008).

Ribotyping

Grouping of bacteria by ribotyping (i.e. restriction fragment patterns of digested-chromosomal DNA probed with rRNA) has been frequently used for both taxonomic purposes and subgroup characterization of microorganisms belonging to different genera and species (Grimont & Grimont, 1986). Leptospires possess two sets of 16S and 23S rRNA genes and one or two 5S rRNA gene that are not closely linked to each other, but are dispersed throughout the large chromosome (Baril *et al.*, 1992; Zuerner *et al.*, 1993b). Because of the low number of rRNA genes, this typing method is not very discriminative (Kositanont *et al.*, 2007; Perolat *et al.*, 1993).

Insertion sequences

Bacterial typing methods based on Insertion Sequence (IS) elements have considerable epidemiological value. Two, IS*1500* and IS*1502* (Boursaux-Eude *et al.*, 1995; Zuerner & Huang, 2002), and one, IS*1533* (Zuerner, 1994), were initially identified in the pathogens *L. interrogans* and *L. borgpetersenii*, respectively. The copy number of these IS varies widely between different serovars and among isolates of the same serovar as demonstrated by Southern blot analysis employing as probes the fragments of the IS elements (Boursaux-Eude *et al.*, 1995; Boursaux-

Eude *et al.*, 1998; Zuerner, 1994; Zuerner & Bolin, 1997). It was previously demonstrated that the insertion sequences IS*1500* and IS*1533* can be successfully employed in taxonomy and typing (Zuerner et al., 1995; Zuerner and Bolin, 1990) While IS*1533* is able to differentiate among serovars of *Leptospira* spp. and even into some of them (Zuerner at al., 1993, Zuerner et al., 1995), IS*1500* can discriminate among *L. interrogans* (sensu stricto) isolates (Zuerner and Bolin, 1997). On the other hand, the IS*1500*-based assays developed by Zuerner and Bolin (1997) complement the IS*1533*-based techniques and extend the number of serovars that can be analyzed, although genomic DNA from some serovars tested with the IS*1500* assays failed to yield either PCR products or detectable hybridization patterns (Zuerner and Bolin, 1997). Zuerner and colleagues (2002) identified a new IS*1502* in *L. interrogans* genome that is assumed to have a functional transposase by internal frameshifting. It was undetectable in some strains, suggesting its recent introduction into the genus *Leptospira* (Zuerner et al., 2002). Polymorphisms among *Leptospira* isolates can therefore be revealed by RFLP followed by sequencing (Boursaux-Eude *et al.*, 1995). PCR-based methods were also developed to detect repeated insertion sequences and for the diagnostic identification of *Leptospira* spp. These techniques took advantage of the high-copy number of IS (and consequently the number of priming sites) to increase the sensitivity of detection (Cameron *et al.*, 2008).

Restriction Endonuclease Analysis and Pulsed-Field Gel Electrophoresis

Non-sequence-based methods rely on the comparison of restriction profiles in agarose or acrylamide gels. Restriction Endonuclease Analysis (REA) of total genomic DNA proved to be a reliable method for typing some *Leptospira* strains (Ellis *et al.*, 1988; Ellis *et al.*, 1991; Venkatesha & Ramadass, 2001). However, this technique is labor-intensive and requires significant volumes of culture. In addition, the presence of a high number of bands makes the interpretation and interlaboratory data comparisons difficult. REA can also be performed on PCR products. Identification of leptospires can then be based on the analysis of fragments obtained after restriction of the PCR product amplified from 16S and 23S rDNA (Heinemann *et al.*, 2000; Ralph *et al.*, 1993; Woo *et al.*, 1997), a repetitive sequence (Savio *et al.*, 1994), and *flaB* loci (Kawabata *et al.*, 2001; Woodward & Redstone, 1993). However, these methods exhibit a low discriminatory power.

Genomic macrorestriction using rare cutting endonucleases such as *Not*I followed by pulsed-field gel electrophoresis (PFGE) is considered a powerful typing method for classifying *Leptospira* strains. The use of computer-assisted gel analyses allows the relationship between strains to be measured, allowing dendrograms to be constructed. Interlaboratory data comparisons is also possible due to a recent improvement in the execution of this technique (Galloway & Levett, 2008). Variations in the macrorestriction profiles in both size and number may be due to sequence rearrangements, insertion or deletion of DNA, or base substitution within the restriction sites (Tenover *et al.*, 1995). Hermann et al. found good concordance between results of PFGE and serotyping (Herrmann *et al.*, 1992; Zuerner *et al.*, 1993a). Serovars produced PFGE patterns that were unique to each serovar. For example, classification of strain Dadas I as a new serovar Dadas of serogroup Grippotyphosa was strongly supported by a unique pulsed-field gel electrophoresis pattern (Herrmann *et al.*, 1994). However, discrepancies between PFGE and serological methods have also been described. For example, PFGE was not able to discriminate between *L. interrogans* serovars Icterohaemorrhagiae and Copenhageni. Despite this, almost 90% of the serovars can be identified by unique PFGE patterns, including that belonging to the pathogenic species (Galloway and Levett, 2008).

For all these reasons PFGE is considered as the gold standard for molecular typing of *Leptospira* serovars to which all other techniques are compared to. However, this method is labor intensive and not accessible to most of the laboratories in tropical and sub-tropical countries where the incidence of the disease is the highest.

PCR-Based typing

Randomly amplified polymorphic DNA (RAPD) and arbitrarily primed PCR (AP-PCR)

Randomly amplified polymorphic DNA (RAPD) fingerprinting (Williams *et al.*, 1990) or arbitrarily primed PCR (AP-PCR) (Welsh & McClelland, 1990) use low-stringency PCR amplification involving primers with an arbitrary sequence to generate strain-specific fingerprints. Ralph *et al.,* (1993) used AP-PCR to classify a group of 48 *Leptospira* reference strains into *L. interrogans*, *L. kirschneri*, *L.*

borgpetersenii, and *L. santarosai* and demonstrated that this method was in agreement with 16S rRNA gene sequencing and DNA-DNA homology. The ability of species discrimination was later confirmed by several studies (Brown & Levett, 1997; Ciceroni *et al.*, 2002; Collares-Pereira *et al.*, 2000; Letocart *et al.*, 1997; Ramadass *et al.*, 2002). Random amplified polymorphic DNA (RAPD) fingerprinting has proved useful for epidemiologic investigations of leptospirosis in regions of high endemicity (India) (Natarajaseenivasan *et al.*, 2005; Roy *et al.*, 2004; Roy *et al.*, 2005). This technique provides a simple and rapid identification of leptospires species (Ramadass et al., 1997) and serovar comparison (Corney et al., 1993; Gerritsen et al., 1995), which could be useful in molecular epidemiological studies of leptospirosis. However, these techniques do not allow large scale studies and the reproducibility and interlaboratory comparisons of the data is difficult.

Amplified fragment length polymorphism

AFLP (amplified fragment length polymorphism) is a three-step procedure in which genomic DNA is restricted, ligated with adapters and then fragments are amplified for generation of fingerprints. FAFLP (fluorescent amplified fragment length polymorphism) have been used (Corney *et al.*, 2008; Slack *et al.*, 2006a; Vijayachari *et al.*, 2004). The use of computer-assisted gel analyses makes it appropriate for clustering analysis. However, it requires large quantities of purified DNA in comparison to other PCR-based methods.

Species-specific PCR/DNA sequence analysis

PCR amplification with species-specific primers, in which the conserved 16S rRNA gene or species-specific gene loci are targeted, can be used directly for species identification of *Leptospira* spp. Use of two primer sets is able to discriminate between *L. interrogans* and *L. kirschneri* (Cameron *et al.*, 2008). Quantitative Real-Time PCR with primers specific for Lfb1 is able to distinguish between pathogenic species by analyzing melting curves (Merien *et al.*, 2005). DNA sequence analysis of some highly conserved gene loci can be used as an identification method at the species or subspecies level. DNA sequence analysis of the DNA region amplified by primers G1 and G2 allows the detection of pathogenic species (Gravekamp *et al.*, 1993). Similarly, *rrs*, *gyrB*, *rpoB*, have been used for this purpose with *Leptospira* spp. (Morey et al., 2006; La Scola *et al.*, 2006 ; Slack *et al.*, 2006b). More recently,

(Victoria *et al.*, 2008) demonstrated that one gene belonging to *S10-spc-α* locus, which codes for SecY preprotein translocase, could be satisfactory employed to identify species. However, disadvantages of sequencing a single gene include the occurrence of horizontal transfer among strains from different species and the low extent of polymorphism into some of leptospiral genes. Mosaicism was found to occur among several *Leptospira* spp. genes (Haake et al., 2004, McBride et al., 2009) including the 16S rRNA genes of *L. broomii*, where *L. fainei*- and *L. inadai*-like sequences were identified (Levett et al., 2005). The use of mosaic genes for species classification represents a problem due to misidentification. Otherwise, *Leptospira* spp. have been traditionally classified by the sequencing of the nearly full-length 16S rRNA gene (Postic et al., 2000; Morey et al., 2006), demonstrating agreement with DNA-DNA hybridization, the golden standard for *Leptospira* species classification (Brenner et al., 1999; Yasuda et al., 1987).

The evaluation of partial *rpoB* gene as taxonomic marker revealed a higher number of polymorphic sites, among *Leptospira* strains belonging to different species, than the 16S rRNA gene. Additionally, its shorter length (600 bp vs. 1500 bp) was able to retain the ability to cluster the species in three clear branches containing pathogens, intermediates and saprophyte species (La Scola et al., 2006). A short 504 bp fragment of the *gyrB* gene was also tested in conventional and real-time PCR for the same purpose (Slack et al., 2006), but focused only the identification of pathogenic leptospires. This gene is also more variable than the 16S. One common limitation observed among all taxonomic markers was their failure to distinguish between leptospiral serovars.

Molecular typing at the genomic era

In the past few years, the genome sequences of six strains from the pathogens *L. interrogans* and *L. borgpetersenii* and the saprophyte *L. biflexa* were released (Bulach *et al.*, 2006; Nascimento *et al.*, 2004; Picardeau *et al.*, 2008; Ren *et al.*, 2003). The genomes, with a G + C content ranging between 35% and 41%, possess two circular chromosomes: one of approximately 4 Mb in length, and a smaller of 300 kb in size. The presence of a 74 kb replicon has also been reported in *L. biflexa* (Picardeau *et al.*, 2008). Numerous repeated sequences have been found in *Leptospira* genomes. For example, several insertion sequences (IS) have been

identified (Data not shown). The genomes also contain an abundant number of small repetitive DNA sequences. Among these DNA repeats, structure of short sequence repeats is typical of tandem repeats. Tandem repeats consist of head-to-tail repetitions of short sequence motifs of about 10 to 100 base pairs. Polymorphic tandem repeats also called Variable Number of Tandem Repeats (VNTR) have been extensively used for fingerprinting in higher eukaryotes, including humans.

A database of tandem repeats (Grissa *et al.*, 2008) reveals that *L. interrogans* genomes contain a high number of tandem repeats with sequence motifs of less than 100 bp, which is convenient to observe polymorphism by analyzing PCR products of polymorphic loci on agarose gel. However, these repetitive DNA elements are not found in high number in the pathogen *L. borgpetersenii*, thus MLVA (multi Locus VNTR Analysis) requires primers specific for each species (except for the closely related *L. interrogans* and *L. kirschneri*). MLVA is specific for pathogenic strains and can distinguish between the serovars belonging to the most frequently reported pathogenic species of the genus *Leptospira* (Majed *et al.*, 2005; Pavan *et al.*, 2008; Salaün *et al.*, 2006; Slack *et al.*, 2006a; Slack *et al.*, 2005). Further studies should also evaluate the stability of these minisatellites over time and geographical distribution. VNTR typing could also provide an accessible mean of testing for research and public health laboratories, particularly in developing countries. Further improvements to this method are required, so that MLVA can be applied directly to biological (serum or urine of patients and samples from animals) and environmental samples without the need to culture the pathogen. Genome analysis also allowed the identification of one Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) in the *L. interrogans* serovar Copenhageni strain Fiocruz (Grissa *et al.*, 2008) that may be useful for molecular typing.

Multilocus sequence typing (MLST), a typing method that is based on the partial sequences of 7 housekeeping genes, has also been applied to *Leptospira* spp. (Ahmed *et al.*, 2006; Thaipadungpanit *et al.*, 2007). In these studies, two different set of genes, *adk*, *icdA*, *lipL32*, *lipL41*, *rrs*, *secY* on the one hand and *pntA*, *sucA*, *pfkB*, *tpiA*, *mreA*, *glmU*, *fadD* on the other hand, have been used (Ahmed *et al.*, 2006; Thaipadungpanit *et al.*, 2007). Applications of MLST allow the identification of clusters of closely related isolates in outbreaks and epidemics (Thaipadungpanit *et al.*, 2007). Another study that aimed to reveal new potential loci for MLST identified the *S10-spc-α* locus, which is a 17.5 kb cluster of 32 genes encoding ribosomal

proteins, as a target for *Leptospira* species classification (Victoria *et al.*, 2008). MLST presents a simple and cheap molecular technique that does not require large quantities of purified DNA, the methodology can be implemented as a routine worldwide, produces reliable, reproducible and easy-to-interpret results and whose are widely exchangeable.

Implications for diagnosis and epidemiology

A better understanding of the epidemiology of leptospirosis requires the isolation and serological characterization of leptospires. The isolation of the organism from clinical specimens (usually blood or urine) is difficult and requires several weeks of incubation, but the isolation of leptospires allows for identification of the infecting serovar. The Microscopic Agglutination Test (MAT) remains the gold standard for the serological diagnosis of leptospirosis and usually gives an indication of the presumptive serovar or serogroup of leptospires involved in an infection (Levett, 2003). However, most of the institutions or hospitals may not have facilities to perform this test. More simple and rapid diagnostic tests detecting antibodies are usually used (ELISA-based tests). However, the sensitivity of these tests is usually low. In addition, these tests do not identify the presumptive infecting serogroup. One of the molecular techniques used for the early diagnosis of leptospirosis in recent years is the amplification of specific fragment of leptospiral genomic DNA in clinical samples using PCR. Again, these techniques do not identify at the serovar/serogroup level and have therefore no epidemiological value.

References

Ahmed, N., Devi, S. M., Valverde Mde, L., Vijayachari, P., Machang'u, R. S., Ellis, W. A. & Hartskeerl, R. A. (2006). Multilocus sequence typing method for identification and genotypic classification of pathogenic Leptospira species. *Ann Clin Microbiol Antimicrob* **5**, 28.

Baril, C., Herrmann, J. L., Richaud, C., Margarita, D. & Girons, I. S. (1992). Scattering of the rRNA genes on the physical map of the circular chromosome of Leptospira interrogans serovar icterohaemorrhagiae. *J Bacteriol* **174**, 7566-7571.

Boursaux-Eude, C., Saint Girons, I. & Zuerner, R. L. (1995). IS1500, an IS3-like element from Leptospira interrogans. *Microbiology* **141**, 2165–2173.

Boursaux-Eude, C., Saint Girons, I. & Zuerner, R. L. (1998). Leptospira genomics. *Electrophoresis* **19**, 589–592.

Brenner, D. J., Kaufmann, A. F., Sulzer, K. R., Steigerwalt, A. G., Rogers, F. C. & Weyant, R. S. (1999). Further determination of DNA relatedness between serogroups and serovars in the family *Leptospiraceae* witha proposal for *Leptospira alexanderi* sp. nov. and four new *Leptospira* genomospecies. *Int J Syst Bacteriol* **49**, 839-858.

Brown, P. D. & Levett, P. N. (1997). Differentiation of Leptospira species and serovars by PCR-restriction endonuclease analysis, arbitrarily primed PCR and low-stringency PCR. *J Med Microbiol* **46**, 173-181.

Bulach, D. M., Zuerner, R. L., Wilson, P. & other authors (2006). Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. *Proc Natl Acad Sci USA* **103**, 14560-14565.

Cameron, C. E., Zuerner, R. L., Raverty, S., Colegrove, K. M., Norman, S. A., Lambourn, D. M., Jeffries, S. J. & Gulland, F. M. (2008). Detection of pathogenic Leptospira bacteria in pinniped populations via PCR and identification of a source of

transmission for zoonotic leptospirosis in the marine environment. *J Clin Microbiol* **46**, 1728-1733.

Ciceroni, L., Ciarrocchi, S., Ciervo, A. & other authors (2002). Differentiation of leptospires of the serogroup Pomona by monoclonal antibodies, pulsed-field gel electrophoresis and arbitrarily primed polymerase chain reaction. *Res Microbiol* **153**, 37-44.

Collares-Pereira, M., Korver, H., Cao Thi, B. V., Santos-Reis, M., Bellenger, E., Baranton, G. & Terpstra, W. J. (2000). Analysis of Leptospira isolates from mainland Portugal and the Azores islands. *FEMS Microbiol Lett* **185**, 181-187.

Corney, B. G., Slack, A. T., Symonds, M. L., Dohnt, M. F., McClintock, C. S., McGowan, M. R. & Smythe, L. D. (2008). Leptospira weilii serovar Topaz, a new member of the Tarassovi serogroup isolated from a bovine source in Queensland, Australia. *Int J Syst Evol Microbiol* **58**, 2249-2252.

Dikken, H. & Kmety, E. (1978). Serological typing methods of leptospires. *Methods in Microbiology, London: Academic Press* **eds. Bergan T, Norris JR,**, 259–307.

Ellis, W. A., Thiermann, A. B., Montgomery, J., Handsaker, A., Winter, P. J. & Marshall, R. B. (1988). Restriction endonuclease analysis of Leptospira interrogans serovar hardjo isolates from cattle. *Res Vet Sci* **44**, 375-379.

Ellis, W. A., Montgomery, J. M. & Thiermann, A. B. (1991). Restriction endonuclease analysis as a taxonomic tool in the study of pig isolates belonging to the Australis serogroup of Leptospira interrogans. *J Clin Microbiol* **29**, 957-961.

Faine, S., Adler, B., Bolin, C. & Perolat, P. (1999). *Leptospira and leptospirosis*. Melbourne, Australia: MedScience.

Galloway, R. L. & Levett, P. N. (2008). Evaluation of a modified pulsed-field gel electrophoresis approach for the identification of Leptospira serovars. *Am J Trop Med Hyg* **78**, 628-632.

Gravekamp, C., Van de Kemp, H., Franzen, M., Carrington, D., Schoone, G. J., Van Eys, G. J., Everard, C. O., Hartskeerl, R. A. & Terpstra, W. J. (1993). Detection of seven species of pathogenic leptospires by PCR using two sets of primers. *J Gen Microbiol* **139**, 1691-1700.

Grimont, F. & Grimont, P. A. D. (1986). Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Ann Inst Pasteur* **137B**, 165–175.

Grissa, I., Bouchon, P., Pourcel, C. & Vergnaud, G. (2008). On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. *Biochimie* **90**, 660-668.

Haake, D. A., Suchard, M. A., Kelley, M. M., Dundoo, M., Alt, D. P. & Zuerner, R. L. (2004). Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer. *J Bacteriol* **186**, 2818-2828.

Hartskeerl, R. A., Goris, M. G., Brem, S., Meyer, P., Kopp, H., Gerhards, H. & Wollanke, B. (2004). Classification of leptospira from the eyes of horses suffering from recurrent uveitis. *J Vet Med B Infect Dis Vet Public Health* **51**, 110-115.

Heinemann, M. B., Garcia, J. F., Nunes, C. M., Gregori, F., Higa, Z. M., Vasconcellos, S. A. & Richtzenhain, L. J. (2000). Detection and differentiation of Leptospira spp. serovars in bovine semen by polymerase chain reaction and restriction fragment length polymorphism. *Vet Microbiol* **73**, 261-267.

Herrmann, J. L., Baril, C., Bellenger, E., Perolat, P., Baranton, G. & Saint Girons, I. (1991). Genome conservation in isolates of *Leptospira interrogans*. *J Bacteriol* **173**, 7582-7588.

Herrmann, J. L., Bellenger, E., Perolat, P., Baranton, G. & Saint Girons, I. (1992). Pulsed-field gel electrophoresis of *Not*I digests of leptospiral DNA: a new rapid method of serovar identification. *J Clin Microbiol* **30**, 1696-1702.

Herrmann, J. L., Bakoss, P., Korver, H., Bulu, A. A., Bellenger, E., Terpstra, W. J., Saint Girons, I. & Baranton, G. (1994). A new serovar in the Grippotyphosa serogroup comprising leptospiral isolates from different regions. *Int J Syst Bacteriol* **44**, 362-364.

Inada, R., Ido, Y., Hoki, R., Kakeno, R. & Ito, H. (1916). The etiology, mode of infection and specific therapy of Weil's disease (Spirochaeta icterohaemorrhagiae). *J Exp Med* **23**, 377-403.

Johnson, R. C. & Rogers, F. C. (1964). Differentiation of pathogenic and saprophytic leptospires with 8-azaguanine. *J Bacteriol* **88**, 1618-1623.

Kawabata, H., Dancel, L. A., Villanueva, S. Y., Yanagihara, Y., Koizumi, N. & Watanabe, H. (2001). flaB-polymerase chain reaction (flaB-PCR) and its restriction fragment length polymorphism (RFLP) analysis are an efficient tool for detection and identification of Leptospira spp. *Microbiol Immunol* **45**, 491-496.

Kmety, E. & Dikken, H. (1993). *Classification of the species Leptospira interrogans and history of its serovars*. Groningen Netherlands: University Press.

Kositanont, U., Chotinantakul, K., Phulsuksombati, D. & Tribuddharat, C. (2007). Assessment of Southern blot ribotyping for differentiation of Leptospira strains isolated from field rats. *J Microbiol Methods* **69**, 288-297.

La Scola, B., Bui, L. T., Baranton, G., Khamis, A. & Raoult, D. (2006). Partial rpoB gene sequencing for identification of Leptospira species. *FEMS Microbiol Lett* **263**, 142-147.

Letocart, M., Baranton, G. & Perolat, P. (1997). Rapid identification of pathogenic *Leptospira* species (*Leptospira interrogans*, *L. borgpetersenii*, and *L. kirschneri*) with species-specific DNA probes produced by arbitrarily primed PCR. *J Clin Microbiol* **35**, 248-253.

Levett, P. N. (2001). Leptospirosis. *Clin Microbiol Rev* **14**, 296-326.

Levett, P. N. (2003). Usefulness of serologic analysis as a predictor of the infecting serovar in patients with severe leptospirosis. *Clin Infect Dis* **36**, 447-452.

Levett, P. N., Morey, R. E., Galloway, R., Steigerwalt, A. G. & Ellis, W. A. (2005). Reclassification of Leptospira parva Hovind-Hougen et al., 1982 as Turneriella parva gen. nov., comb. nov. *Int J Syst Evol Microbiol* **55**, 1497–1499.

Levett, P. N., Morey, R. E., Galloway, R. L. & Steigerwalt, A. G. (2006). Leptospira broomii sp. nov., isolated from humans with leptospirosis. *Int J Syst Evol Microbiol* **56**, 671–673.

Majed, Z., Bellenger, E., Postic, D., Pourcel, C., Baranton, G. & M., P. (2005). Identification of variable-number tandem-repeat loci in Leptospira interrogans sensu stricto. *J Clin Microbiol* **43**, 539-545.

Martin, L. & Pettit, A. (1918). Sero-diagnostic de la spirochaetose icterohaemorrhagique. *Bull Mem Soc Med Hop Paris* **42**, 672-675.

Matthias, M. A., Ricaldi, J. N., Cespedes, M. & other authors (2008). Human leptospirosis caused by a new, antigenically unique leptospira associated with a rattus species reservoir in the peruvian Amazon. *PLoS Negl Trop Dis* **2**, e213.

McBride, A. J., Athanazio, D. A., Reis, M. G., Ko, A. I (2005). Leptospirosis. *Curr Opin Infect Dis* **18: 376-386**.

Merien, F., Portnoi, D., Bourhy, P., Charavay, F., Berlioz-Arthaud, A. & Baranton, G. (2005). A rapid and quantitative method for the detection of Leptospira species in human leptospirosis. *FEMS Microbiol Lett* **249: 139-147**.

Naigowit, P., Charoenchai, S., Biaklang, M., Seena, U., Wangroongsarb, P., Sawanpanyalert, P. & Warachit, P. (2007). Identification of clinical isolates of Leptospira spp by pulsed field gel-electrophoresis and microscopic agglutination test. *Southeast Asian J Trop Med Public Health* **38**, 97-103.

Nascimento, A. L., Ko, A. I., Martins, E. A. & other authors (2004). Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *J Bacteriol* **186**, 2164-2172.

Natarajaseenivasan, K., Vijayachari, P., Sharma, S., Roy, S., Sugunan, A. P., Biswas, D. & Sehgal, S. C. (2005). Phylogenetic relatedness among leptospiral strains belonging to same serovar recovered from patients with different clinical syndromes. *Infect Genet Evol* **5**, 185-191.

Paster, B. J., Dewhirst, F. E., Weisburg, W. G. & other authors (1991). Phylogenetic analysis of the spirochetes. *J Bacteriol* **173**, 6101–6109.

Pavan, M. E., Cairó, F., Brihuega, B. & Samartino, L. (2008). Multiple-locus variable-number tandem repeat analysis (MLVA) of Leptospira interrogans serovar Pomona from Argentina reveals four new genotypes. *Comp Immunol Microbiol Infect Dis* **31**, 37-45.

Perolat, P., Lecuyer, I., Postic, D. & Baranton, G. (1993). Diversity of ribosomal DNA fingerprints of *Leptospira* serovars provides a database for subtyping and species assignation. *Res Microbiol* **144**, 5-15.

Perolat, P., Chappel, R. J., Adler, B., Baranton, G., Bulach, D. M., Billinghurst, M. L., Letocart, M., Merien, F. & Serrano, M. S. (1998). *Leptospira fainei* sp. nov., isolated from pigs in Australia. *Int J Syst Bacteriol* **48**, 851-858.

Picardeau, M., Bulach, D. M., Bouchier, C. & other authors (2008). Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis. *PLoS ONE* **3**, e1607.

Ralph, D., McClelland, M., Welsh, J., Baranton, G. & Perolat, P. (1993). Leptospira species categorized by arbitrarily primed polymerase chain reaction (PCR) and by mapped restriction polymorphisms in PCR-amplified rRNA genes. *J Bacteriol* **175**, 973-981.

Ralph, D. & McClelland, M. (1994). Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus. *J Bacteriol* **176**, 5982-5987.

Ramadass, P., Jarvis, B. D. W., Corner, R. J., Penny, D. & Marshall, R. B. (1992). Genetic characterization of pathogenic Leptospira species by DNA hybridization. *Int J Syst Bacteriol* **42**, 215–219.

Ramadass, P., Latha, D., Senthilkumar, A., Srinivasan, P. & Saranya, N. (2002). Arbitrarily primed PCR- A rapid and simple method for typing of leptospiral serovars. *Indian J Med Microbiol* **20**, 25-28.

Ren, S., Fu, G., Jiang, X. & other authors (2003). Unique and physiological and pathogenic features of *Leptospira interrogans* revealed by whole genome sequencing. *Nature* **422**, 888-893.

Roy, S., Biswas, D., Vijayachari, P., Sugunan, A. P. & Sehgal, S. C. (2004). A 22-mer primer enhances discriminatory power of AP-PCR fingerprinting technique in characterization of leptospires. *Trop Med Int Health* **9**, 1203-1209.

Roy, S., Biswas, D., Vijayachari, P., Sugunan, A. P. & Sehgal, S. C. (2005). A clone of Leptospira interrogans sensu stricto is the major cause of leptospirosis in the archipelago of Andaman and Nicobar Islands, India. *Lett Appl Microbiol* **41**, 179-185.

Salaün, L., Mérien, F., Gurianova, S., Baranton, G. & Picardeau, M. (2006). Application of multilocus variable-number tandem-repeat analysis for molecular typing of the agent of leptospirosis. *J Clin Microbiol* **44**, 3954-3962.

Savio, M. L., Rossi, C., Fusi, P., Tagliabue, S. & Pacciarini, M. L. (1994). Detection and identification of Leptospira interrogans serovars by PCR coupled with restriction endonuclease analysis of amplified DNA. *J Clin Microbiol* **32**, 935-941.

Schmid, G. P., Steere, A. C., Kornblatt, A. N., Kaufmann, A. F., Moss, C. W., Johnson, R. C., Hovind-Hougen, K. & Brenner, D. J. (1986). Newly recognized Leptospira species ("Leptospira inadai" serovar lyme) isolated from human skin. *J Clin Microbiol* **24**, 484-486.

Slack, A., Symonds, M., Dohnt, M. & Smythe, L. (2006a). An improved multiple-locus variable number of tandem repeats analysis for Leptospira interrogans serovar Australis: a comparison with fluorescent amplified fragment length polymorphism analysis and its use to redefine the molecular epidemiology of this serovar in Queensland, Australia. *J Med Microbiol* **55**, 1549-1557.

Slack, A. T., Dohnt, M. F., Symonds, M. L. & Smythe, L. D. (2005). Development of a Multiple-Locus Variable Number of Tandem Repeat Analysis (MLVA) for Leptospira interrogans and its application to Leptospira interrogans serovar Australis isolates from Far North Queensland, Australia. *Ann Clin Microbiol Antimicrob* **4**, 10.

Slack, A. T., Symonds, M. L., Dohnt, M. F. & Smythe, L. D. (2006b). Identification of pathogenic Leptospira species by conventional or real-time PCR and sequencing of the DNA gyrase subunit B encoding gene. *BMC Microbiol* **6**, 95.

Slack, A. T., Kalambaheti, T., Symonds, M. L. & other authors (2008). Leptospira wolffii sp. nov., isolated from a human with suspected leptospirosis in Thailand. *Int J Syst Evol Microbiol* **58**, 2305-2308.

Steinen, A. C. M., Schuurman, J. L., Gravekamp, C., Korver, H. & Terpstra, W. J. (1992). Muskrats as carriers of pathogenic leptospires in The Netherlands. *Antonie Van Leeuwenhoek* **61**, 43–50.

Stimson, A. M. (1907). Note on an organism found in yellow-fever tissue. . *Public Health Reports (Washington)* **22**, 541.

Tenover, F. C., Arbeit, R. D., Goering, R. V., Mickelsen, P. A., Murray, B. E., Persing, D. H. & Swaminathan, B. (1995). Interpreting chromosomal DNA

restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* **33**, 2233-2239.

Terpstra, W. J., Korver, H., Van Leeuwen, J., Klatser, P. R. & Kolk, A. H. J. (1985). The classification of Sejroe group serovars of Leptospira interrogans with monoclonal antibodies. *Zentralbl Bakteriol Mikrobiol Hygiene* **259**, 498–506.

Thaipadungpanit, J., Wuthiekanun, V., Chierakul, W. & other authors (2007). A dominant clone of Leptospira interrogans associated with an outbreak of human leptospirosis in Thailand. *PLoS Negl Trop Dis*, e56.

Thiermann, A. B., Handsaker, A. L., Moseley, S. L. & Kingscote, B. (1985). New method for classification of leptospiral isolates belonging to serogroup pomona by restriction endonuclease analysis: serovar kennewicki. *J Clin Microbiol* **21**, 585-587.

Valverde Mde, L., Ramírez, J. M., Montes de Oca, L. G., Goris, M. G., Ahmed, N. & Hartskeerl, R. A. (2008). Arenal, a new Leptospira serovar of serogroup Javanica, isolated from a patient in Costa Rica. *Infect Genet Evol* **8**, 529-533.

Venkatesha, M. D. & Ramadass, P. (2001). Identification of leptospiral isolates by bacterial restriction endonuclease analysis (Brenda). *Indian J Med Microbiol* **19**, 26-29.

Victoria, B., Ahmed, A., Zuerner, R. L., Ahmed, N., Bulach, D. M., Quinteiro, J. & Hartskeerl, R. A. (2008). Conservation of the S10-spc-alpha locus within otherwise highly plastic genomes provides phylogenetic insight into the genus Leptospira. *PLoS ONE* **3**, e2752.

Vijayachari, P., Sehgal, S. C., Goris, M. G., Terpstra, W. J. & Hartskeerl, R. A. (2003). Leptospira interrogans serovar valbuzzi: a cause of severe pulmonary haemorrhages in Andaman Islands. *J Med Microbiol* **52**, 913–918.

Vijayachari, P., Ahmed, N., Sugunan, A. P., Ghousunnissa, S., Rao, K. R., Hasnain, S. E. & Sehgal, S. C. (2004). Use of fluorescent amplified fragment length

polymorphism for molecular epidemiology of leptospirosis in India. *J Clin Microbiol* **42**, 3575-3580.

Welsh, J. & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* **18**, 7213–7218.

WHO (1999). Leptospirosis worldwide, 1999. *Weekly Epidemiol Rec* **74**, 237-242.

WHO (2003). *Human leptospirosis: Guidance for diagnosis, surveillance and control*.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* **18**, 6531–6535.

Wolbach, S. B. & Binger, C. A. L. (1914). Notes on a filterable spirochete from fresh water. Spirocheta biflexa (new species). *J Med Research* **30**, 23-25.

Woo, T. H., Patel, B. K., Smythe, L., Symonds, M. L., Norris, M. A. & Dohnt, M. F. (1997). Comparison of two PCR methods for rapid identification of Leptospira genospecies interrogans. *FEMS Microbiol Lett* **155**, 169-177.

Woodward, M. J. & Redstone, J. S. (1993). Differentiation of leptospira serovars by the polymerase chain reaction and restriction fragment length polymorphism. *Vet Rec* **132**, 325-326.

Yasuda, P. H., Steigerwalt, A. G., Sulzer, K. R., Kaufmann, A. F., Rogers, F. C. & Brenner, D. J. (1987). Deoxyribonucleic acid relatedness between serogroups and serovars in the family *Leptospiraceae* with proposals for seven new *Leptospira* species. *Int J Syst Bacteriol* **37**, 407-415.

Zuerner, R., Herrmann, J. L. & Saint Girons, I. (1993a). Construction and comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity in genomic organization. *J Bacteriol* **175**, 5445-5451.

Zuerner, R. L., Herrmann, J. L. & Saint Girons, I. (1993b). Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. *J Bacteriol* **175**, 5445-5451.

Zuerner, R. L. (1994). Nucleotide sequence analysis of IS1533 from Leptospira borgpetersenii: identification and expression of two IS-encoded proteins. *Plasmid* **31**, 1-11.

Zuerner, R. L. & Bolin, C. A. (1997). Differentiation of Leptospira interrogans isolates by IS1500 hybridization and PCR assays. *J Clin Microbiol* **35**, 2612-2617.

Zuerner, R. L. & Huang, W. M. (2002). Analysis of a Leptospira interrogans locus containing DNA replication genes and a new IS, IS1502. *FEMS Microbiol Lett* **215**, 175-182.

Tables

Table 1

Distribution of serogroups among the several *Leptospira* species[†].

Serogroup/Species	*L. interrogans*	*L. kirschneri*	*L. noguchii*	*L. borgpetersenii*	*L. weilii*	*L. santarosai*	*L. alexanderi*	*L. wolffii*[‡]	*L. licerasiae*[*]	*L. inadai*	*L. fainei*	*L. broomii*	*L. wolbachii*	*L. meyeri*	*L. biflexa*
	P								I				S		
Andaman															1
Australis	9	1	5	1											
Autumnalis	10	5	1	1		1									
Ballum				6											
Bataviae	5	1	2	1		5									
Canicola	11	3								1					
Celledoni				2	3										
Codice													1		
Cynopteri		1				3									
Djasiman	4	1	1												
Grippotyphosa	3	4				1									
Hebdomadis	2	2		4	1	7	2								
Hurstbridge									1		1				
Icterohaemorrhagiae	14	5			1					1					
Javanica				9	3	3	1			1				1	
Lyme										1					
Louisiana	1		3												
Manhao	1				2		2			2					
Mini	2			1	1	4	1							1	
Panama			2												
Pyrogenes	8		1	2	1	7									
Pomona	4	3	2			2									
Ranarum	1													1	
Sarmin	1				1	3									
Sejroe	12			9	1	6								1	
Semaranga														1	1
Shermani			1			3				1					
Tarassovi			1	7	3	12				1					

[†] based on Levett, 2001.

[‡]No cross-reaction based on Slack et al., 2008.

*Titre of 1/100. Matthias et al., 2008.

P – Pathogens, I – Intermediates, S – Saprophytes.

Numbers represent serovars per serogroup, distributed into each respective species. Total of 274 serovars through 15 species and 28 serogroups.

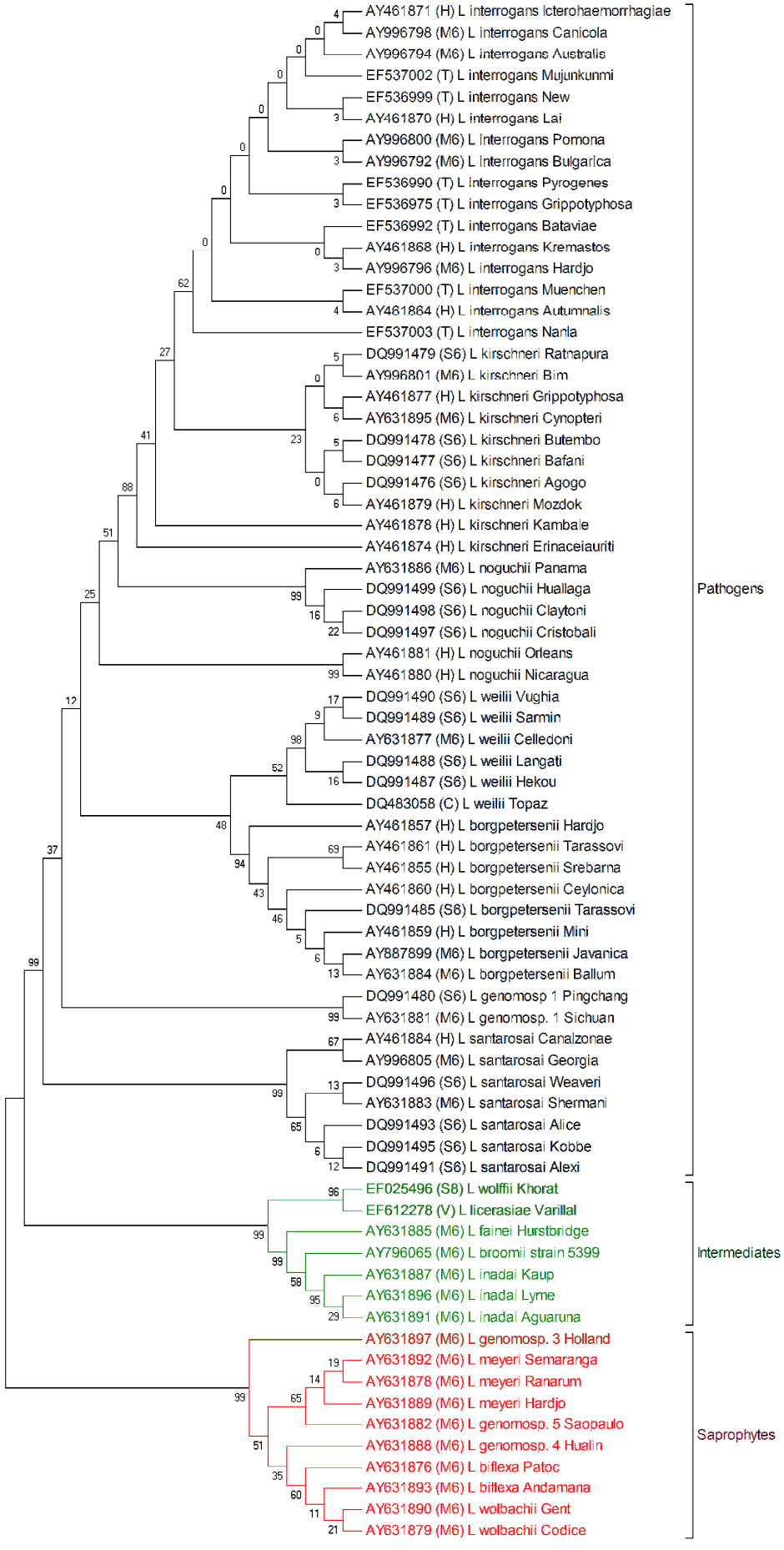

AY461871 (H) L interrogans Icterohaemorrhagiae
AY996798 (M6) L interrogans Canicola
AY996794 (M6) L interrogans Australis
EF537002 (T) L interrogans Mujunkunmi
EF536999 (T) L interrogans New
AY461870 (H) L interrogans Lai
AY996800 (M6) L interrogans Pomona
AY996792 (M6) L interrogans Bulgarica
EF536990 (T) L interrogans Pyrogenes
EF536975 (T) L interrogans Grippotyphosa
EF536992 (T) L interrogans Bataviae
AY461868 (H) L interrogans Kremastos
AY996796 (M6) L interrogans Hardjo
EF537000 (T) L interrogans Muenchen
AY461864 (H) L interrogans Autumnalis
EF537003 (T) L interrogans Nanla
DQ991479 (S6) L kirschneri Ratnapura
AY996801 (M6) L kirschneri Bim
AY461877 (H) L kirschneri Grippotyphosa
AY631895 (M6) L kirschneri Cynopteri
DQ991478 (S6) L kirschneri Butembo
DQ991477 (S6) L kirschneri Bafani
DQ991476 (S6) L kirschneri Agogo
AY461879 (H) L kirschneri Mozdok
AY461878 (H) L kirschneri Kambale
AY461874 (H) L kirschneri Erinaceiauriti
AY631886 (M6) L noguchii Panama
DQ991499 (S6) L noguchii Huallaga
DQ991498 (S6) L noguchii Claytoni
DQ991497 (S6) L noguchii Cristobali
AY461881 (H) L noguchii Orleans
AY461880 (H) L noguchii Nicaragua
DQ991490 (S6) L weilii Vughia
DQ991489 (S6) L weilii Sarmin
AY631877 (M6) L weilii Celledoni
DQ991488 (S6) L weilii Langati
DQ991487 (S6) L weilii Hekou
DQ483058 (C) L weilii Topaz
AY461857 (H) L borgpetersenii Hardjo
AY461861 (H) L borgpetersenii Tarassovi
AY461855 (H) L borgpetersenii Srebarna
AY461860 (H) L borgpetersenii Ceylonica
DQ991485 (S6) L borgpetersenii Tarassovi
AY461859 (H) L borgpetersenii Mini
AY887899 (M6) L borgpetersenii Javanica
AY631884 (M6) L borgpetersenii Ballum
DQ991480 (S6) L genomosp 1 Pingchang
AY631881 (M6) L genomosp. 1 Sichuan
AY461884 (H) L santarosai Canalzonae
AY996805 (M6) L santarosai Georgia
DQ991496 (S6) L santarosai Weaveri
AY631883 (M6) L santarosai Shermani
DQ991493 (S6) L santarosai Alice
DQ991495 (S6) L santarosai Kobbe
DQ991491 (S6) L santarosai Alexi
EF025496 (S8) L wolffii Khorat
EF612278 (V) L licerasiae Varillal
AY631885 (M6) L fainei Hurstbridge
AY796065 (M6) L broomii strain 5399
AY631887 (M6) L inadai Kaup
AY631896 (M6) L inadai Lyme
AY631891 (M6) L inadai Aguaruna
AY631897 (M6) L genomosp. 3 Holland
AY631892 (M6) L meyeri Semaranga
AY631878 (M6) L meyeri Ranarum
AY631889 (M6) L meyeri Hardjo
AY631882 (M6) L genomosp. 5 Saopaulo
AY631888 (M6) L genomosp. 4 Hualin
AY631876 (M6) L biflexa Patoc
AY631893 (M6) L biflexa Andamana
AY631890 (M6) L wolbachii Gent
AY631879 (M6) L wolbachii Codice
Pathogens
Intermediates
Saprophytes

Figure legends

Fig. 1. Phylogenetic tree of the 16S rDNA sequences. The tree was built from a 1155 bp-based alignment of nucleotide sequences of the leptospiral *rrs* genes. This phylogenetic tree was constructed by the Neighbor-joining method, using 1,000 bootstrap replications. Major clusters, containing saprophytes (red), intermediates (green) and pathogens (black), were observed. H, M6, S6, S8, T and V refer to Haake et al., 2004; Morey et al., 2006; Slack et al., 2006; Slack et al., 2008; Thaipadungpanit et al., 2007 and Vinetz et al., 2008, respectively. Accession numbers are presented and they are followed by the species and serovar designations, respectively.

4. ARTIGO 2

GENOME-WIDE TRANSPOSON MUTAGENESIS IN PATHOGENIC *Leptospira* SPP.

(Artigo publicado no periódico Infection and Immunity V. 15 (4): 810-816, 2009)

A minha participação neste artigo foi na transformação de *L. interrogans* Canicola Kito, cujos mutantes gerados constam na tabela suplementar 1 (anexo). Além disso, realizei trabalho de bioinformática para determinar a existência de um sítio consenso de inserção do transposon e participei ativamente na redação do manuscrito.

INFECTION AND IMMUNITY, Feb. 2009, p. 810–816 Vol. 77, No. 2
0019-9567/09/$08.00+0 doi:10.1128/IAI.01293-08

Genome-Wide Transposon Mutagenesis in Pathogenic *Leptospira* Species▽‡

Gerald L. Murray,[1] Viviane Morel,[2] Gustavo M. Cerqueira,[2,3] Julio Croda,[2,4] Amporn Srikram,[5] Rebekah Henry,[1] Albert I. Ko,[4,6] Odir A. Dellagostin,[3] Dieter M. Bulach,[1]† Rasana W. Sermswan,[7] Ben Adler,[1,8]* and Mathieu Picardeau[2]*

Australian Bacterial Pathogenesis Program, Department of Microbiology, Monash University, Clayton, VIC 3800, Australia[1]; Institut Pasteur, Unité de Biologie des Spirochètes, Paris, France[2]; Centro de Biotecnologia, Universidade Federal de Pelotas, P.O. Box 354, 96010-900, Pelotas, RS, Brazil[3]; Gonçalo Moniz Research Center, Oswaldo Cruz Foundation, Brazilian Ministry of Health, Salvador, Brazil[4]; Melioidosis Research Center, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand[5]; Division of International Medicine and Infectious Disease, Weill Medical College of Cornell University, New York, New York[6]; Department of Biochemistry, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand[7]; and Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Department of Microbiology, Monash University, Clayton, VIC 3800, Australia[8]

Received 21 October 2008/Returned for modification 11 November 2008/Accepted 23 November 2008

Leptospira interrogans* is the most common cause of leptospirosis in humans and animals. Genetic analysis of *L. interrogans* has been severely hindered by a lack of tools for genetic manipulation. Recently we developed the *mariner*-based transposon *Himar1* to generate the first defined mutants in *L. interrogans*. In this study, a total of 929 independent transposon mutants were obtained and the location of insertion determined. Of these mutants, 721 were located in the protein coding regions of 551 different genes. While sequence analysis of transposon insertion sites indicated that transposition occurred in an essentially random fashion in the genome, 25 unique transposon mutants were found to exhibit insertions into genes encoding 16S or 23S rRNAs, suggesting these genes are insertional hot spots in the *L. interrogans* genome. In contrast, loci containing notionally essential genes involved in lipopolysaccharide and heme biosynthesis showed few transposon insertions. The effect of gene disruption on the virulence of a selected set of defined mutants was investigated using the hamster model of leptospirosis. Two attenuated mutants with disruptions in hypothetical genes were identified, thus validating the use of transposon mutagenesis for the identification of novel virulence factors in *L. interrogans*. This library provides a valuable resource for the study of gene function in *L. interrogans*. Combined with the genome sequences of *L. interrogans*, this provides an opportunity to investigate genes that contribute to pathogenesis and will provide a better understanding of the biology of *L. interrogans*.

Leptospira interrogans is a spirochete that is the main causative agent of leptospirosis. This zoonosis has emerged as a major public health problem in much of the developing world, with more than 500,000 cases of severe leptospirosis reported each year, for which the mortality rate is more than 10% (17).

The genus *Leptospira* is composed of both saprophytic and pathogenic species. The genome sequences of two epidemic strains of *L. interrogans* serovars Lai and Copenhageni have been determined (20, 25). More recently a human and an animal *L. borgpetersenii* isolate were sequenced (3), and this year, we determined the genome sequence of the saprophyte *L. biflexa* (22). The resulting sequences provide an invaluable source of information for identification of genetic determinants involved in the pathogenicity and environmental biology of the organism. For example, the host-adapted *L. borgpetersenii* genome is 16% smaller and has many more pseudogenes than the *L. interrogans* genome. These findings suggest that genome reduction has resulted in a reduced environmental transmission potential (3). *L. interrogans* has 627 genes that are absent in the *L. biflexa* genome, and more than 500 of these genes have unknown functions, suggesting the presence of novel virulence mechanisms (22). However, the lack of tools for *L. interrogans* genetics has hindered elucidation of the role of these genes in pathogenesis.

Pathogenic leptospires are difficult to propagate under in vitro conditions. *L. interrogans* is a slow-growing organism with a generation time of ~20 h, and colonies take up to 4 weeks to appear on solid medium. Furthermore, unlike saprophytic leptospires, these bacteria are genetically intractable, with no replicating vectors (21, 27), and only one mutant has recently been obtained by homologous recombination (5). The lack of genetic systems has hampered molecular analyses of pathogenic leptospires, with no method to assess directly the role of *L. interrogans* genes in virulence. Recently we demonstrated gene transfer in a pathogenic *Leptospira* strain, involving the transposition of *Himar1*, a transposon of eukaryotic origin (2). We

* Mailing address for Mathieu Picardeau: Unité de Biologie 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. Mailing address for Ben Adler: Department of Microbiology, Monash University, Wellington Road, Victoria 3800, Australia. Phone: 61 3 9905 4815. Fax: 61 3 9905 4811. E-mail: ben.adler@med.monash.edu.au.

† Present address: CSIRO Livestock Industries, Australian Animal Health Laboratory (AAHL), Geelong, VIC 3220, Australia.

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

▽ Published ahead of print on 1 December 2008.

TABLE 1. Bacterial strains used for random transposon mutagenesis

Species	Serovar[a]	Strain[b]	Transformation frequency[c]
L. interrogans	Copenhageni	L1 130 LP	2×10^{-7}
L. interrogans	Copenhageni	L1 130 HP	2×10^{-7}
L. interrogans	Lai	56601 LP	7×10^{-6}
L. interrogans	Lai	56601 HP	9×10^{-6}
L. interrogans	Icterohaemorrhagiae	Verdun HP	1×10^{-6}
L. interrogans	Manilae	L495	8×10^{-6}
L. noguchii	Autumnalis	Bonito	$\leq 1 \times 10^{-8}$
L. interrogans	Canicola	L1 133 LP	9×10^{-6}
L. interrogans	Canicola	Kito LP	9×10^{-6}
L. interrogans	Pomona	PO-06-047	8×10^{-6}
L. weilii	Hebdomadis	EcoChallenge LP	5×10^{-6}

[a] When an entry is underlined, only the serogroup of the studied strain is indicated (the serovar was not identified).
[b] LP, low-passage strain; HP, high-passage strain.
[c] Transformation frequency is defined as the number of transposon mutants divided by the number of cells which survived electroporation (approximately 10%). A transformation frequency of $\leq 1 \times 10^{-8}$ represents the limit of detection in these transformations.

identified genes interrupted by *Himar1* insertion in 35 mutants of *L. interrogans* serovar Lai. Since that study, transposon mutagenesis in *L. interrogans* has allowed the identification of a mutant, lacking expression of Loa22, exhibiting attenuated virulence in animal models (26) and a mutant obtained by insertion of the transposon *Himar1* into a gene encoding heme oxygenase (19).

Low electroporation efficiency and a low growth rate have limited the generation of *L. interrogans* random mutants and mean that the generation of high-coverage libraries, commonplace in most mutagenesis studies, is not feasible for *L. interrogans*. Under these circumstances, each mutant isolated is worth characterizing. Notably, there have been only three studies published on the topic in the last three years. Over this period, we have generated libraries of random mutants for different pathogenic strains. In this study we present a library of approximately 1,000 defined mutants with characterized transposon insertion points. This collection of insertional mutants constitutes an extremely valuable resource for functional studies of pathogenic *Leptospira*. The library will be particularly useful for identifying new genes, validating the functions of predicted proteins, and discovering novel virulence factors.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The strains used for this study are listed in Table 1. All strains were obtained from the collection of the Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil, except the *L. interrogans* serovars Lai and Manilae. *L. interrogans* serovar Manilae was provided by N. Koizumi, National Institute of Infectious Diseases, Tokyo, Japan, while *L. interrogans* serovar Lai was obtained from the National Institute for Communicable Disease Control and Prevention, Beijing, China (25). High-passage strains refer to strains that were subcultured in EMJH liquid medium more than 10 times. All strains were cultured at 30°C in liquid EMJH medium (7, 11) or on EMJH plates containing 1.5% agar. Kanamycin or spectinomycin was added at 40 μg/ml when required.

For UV irradiation, cells were spread at appropriate dilutions on EMJH agar plates and irradiated under UV light (254 nm, 10 $\mu W/cm^2$) for various time periods (from 2 to 10 s). UV sensitivity was evaluated by colony counting, with untreated cells serving as a control. Medium for testing the ability of *Leptospira* strains to use hemin was prepared by supplementing EMJH medium with 50 μM 2,2′-digyridyl (Sigma-Aldrich, St. Louis, MO). Bovine hemin was then added at a final concentration of 10 μM.

Transposon mutagenesis. The plasmid pSC189ColE1 was constructed by amplifying the ColE1 origin of replication from pBluescript II (using primers 5′-AAAATACGTAAGCAAAAGGCCAGGAAC-3′ and 5′-AAAACTGCAGGATCAAAGGATCTTCTTG-3′), and the product was digested with SnaBI and PstI then ligated into similarly digested pSC189, replacing OriR6k. The plasmids pSHT, pKMars (14), and pSC189ColE1 were used to perform random transposon mutagenesis in *L. interrogans* strains as described previously (2). Briefly, *L. interrogans* was grown to exponential phase and then washed and concentrated in water. For electroporation, approximately 10^{10} cells in ~100 μl were mixed with 1 μg of plasmid DNA in 2-mm chilled cuvettes. The electroporator was set to 1.8 kV, 25 μF, and 200 Ω. One milliliter of EMJH medium was immediately added to the cuvette, and the cells were incubated overnight at 30°C. Finally, transformants were plated on EMJH agar plates containing antibiotic. Plates were incubated for 4 weeks at 30°C in sealed plastic bags or wrapped in foil to avoid desiccation. Transformants were then picked and subcultured in 5 ml of EMJH liquid medium. Genomic DNA was extracted, and the *Himar1* insertion site was identified by ligation-mediated PCR (LM-PCR) (15, 24) or direct sequencing (19). Confirmation of genotypes was performed by PCR with primers located in the flanking sequences of the predicted transposon insertion site. We did not observe any kanamycin-resistant colonies that did not contain the transposon.

Hamster model of infection. Four-week-old hamsters were injected intraperitoneally with leptospires at the stated inoculum in 100 μl of EMJH. The 50% lethal dose for *L. interrogans* serovar Manilae was approximately 10 leptospires. Hamsters were monitored for 14 days postinfection and euthanized if moribund in accordance with animal ethics requirements. Lungs were inspected for hemorrhage to confirm infection with *Leptospira*. Culture isolation was performed with kidney tissues from hamsters for approximately half of the strains tested. The genotype of the recovered leptospires was confirmed by PCR amplifying the region across the transposon insertion.

Sequence analysis. The *Himar1* insertion site sequences were compared with the complete genome sequence of *L. interrogans* serovar Lai strain 56601 by using the SpiroScope (http://www.genoscope.cns.fr/agc/mage) (30) and Wasabi (3) databases. Multiple sequence alignments of a conserved ±15-bp region surrounding each insertion point were generated with scripts coded in Perl. Consensus sequences were visualized with Sequence Logo analysis using WebLogo (http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi).

RESULTS AND DISCUSSION

Random transposon mutagenesis in pathogenic *Leptospira* spp. The genetics of the pathogenic *Leptospira* spp. is in its infancy. Transposon mutagenesis is a powerful, broadly applicable tool for the generation of libraries of random mutants. *Himar1*, of the *mariner* family, is one of the most widely used transposons for random mutagenesis in bacteria and other organisms (23). In this article, we describe methods for the use of *Himar1* for transposon mutagenesis in *L. interrogans*.

We have applied the method previously used with the saprophyte *L. biflexa* (14, 15) for use with pathogenic *Leptospira* spp. Initially, the plasmid vector pSC189 (4), containing both the hyperactive transposase C9 and transposon terminal inverted repeats flanking a kanamycin resistance gene, was used

to deliver *Himar1* into the *L. interrogans* genome (2). The only origin of replication present in the plasmid construct was that from the *Escherichia coli* plasmid vectors, which is nonfunctional in *Leptospira* spp. Thus, any resistant colonies arising after electroporation of this plasmid into *L. interrogans* are the result of random insertion into the host genome.

We made a number of modifications of the original vector to potentially improve its use in transforming *L. interrogans*. The ColE1 replication origin was introduced to replace OriR6K from the original pSC189 to simplify preparation of vector DNA. Increased expression of the hyperactive transposase C9 gene by substituting a spirochetal promoter for the native promoter increased the yield of transformants in *L. interrogans* 10-fold (2). In addition, a transposon carrying a spectinomycin resistance gene has been constructed; electroporation of this plasmid construct into *L. interrogans* resulted in spectinomycin-resistant colonies at a frequency similar to that generated by the kanamycin-resistant transposon. Since there is no replicative plasmid vector available for pathogenic *Leptospira*, reintroduction of an intact copy of disrupted genes can be achieved via a transposon with alternative selection (26) or by homologous recombination (5).

Transformation of *L. interrogans* was optimal at 9 kV cm^{-1} for a pulse time of 5 ms. This field strength resulted in approximately 10% viability for all pathogenic strains tested. The *L. interrogans* strains exhibited maximal electrocompetence when harvested in mid- to late exponential growth phase. Use of more than 2 μg of DNA did not significantly improve the yield of transformants, although there was no reduction of transformation efficiency observed when using up to 50 μg of DNA. The inserted transposons remained stable after 100 generations in the absence of antibiotic selection. In addition, all random mutants that were recovered from animals maintained the antibiotic resistance cassette (data not shown), indicating that transposon insertions are extremely stable.

The genus *Leptospira* is composed of more than 16 pathogenic and saprophytic species (12). To identify a strain with improved transformation efficiency, we examined the transformability of laboratory and clinical isolates of pathogenic *Leptospira* spp., including pathogenic strains from *L. noguchii* and *L. weilii* (Table 1), with plasmids delivering *Himar1*. For all the tested strains, transformation of *Himar1* in pathogenic leptospires occurred at a low frequency. There was significant strain-dependent variation in transformation competence, with frequencies varying from 10^{-7} to 9×10^{-6}; some strains were completely resistant to transformation (Table 1). The plating efficiency (the ratio of number of CFU to number of bacteria enumerated in a Petroff-Hauser counting chamber) of pathogenic strains ranged between 70 and 90%, suggesting that the low-transformation efficiency was not due to poor viability of pathogenic strains in solid medium. We did not observe differences in the transformation efficiency between high and low in vitro-passaged variants of the same strain.

The poor transformability of leptospires may reflect the involvement of DNA restriction and modification mechanisms. The genome sequences of *L. interrogans* showed one complete putative type I restriction and modification system (LA3197 to LA3200), which is not found in the saprophyte *L. biflexa*, and a total of 12 putative DNA methyltransferase genes. However, transformation efficiency did not increase in any of the strains when transformation was carried out with plasmid DNA produced from a *dam dcm* double mutant of *E. coli*. In addition, treatment of plasmid DNA with crude protein extracts from *Leptospira* strains (6) prior to electroporation had no effect on the transformation efficiency (data not shown). These results suggest the absence of a strong restriction-modification system in pathogenic leptospires. The transformable character of individual strains could be due to variations in leptospiral cell surface properties, as previously suggested for the poorly transformable mycobacteria and *Borrelia* (8, 29). For example, the low-level-transformable Fiocruz strain was found to aggregate more than did the Lai strain in liquid cultures, reflecting as yet undefined differences in surface properties.

Transposon integration sites were identified by either LM-PCR (24) (304 mutants) or direct genome sequencing of the genomic DNA (19) (624 mutants). LM-PCR is a commonly used technique for amplifying the DNA flanking sequences of transposon insertion sites. However, we have found that this method is laborious and time-consuming. In addition, using this amplification method, we could not amplify insertion sites in 60% of the mutants. Several mutants remained uncharacterized by LM-PCR, despite repeated efforts and modifications to the procedure. Sequencing directly from the chromosome using a primer within the transposon was successful in more than 75% of reactions. Typically, 200 to 1,000 bp of quality sequence was obtained, though only 30 bp or so were required to locate the transposon on the chromosomes. Since signal strength was usually low, reactions were improved with a larger amount of template (up to 2 μg total DNA).

Library of transposon mutants. Sequences were compared with the complete genome sequence of *L. interrogans* serovar Lai strain 56601 to identify the genomic location of the transposon. A total of 929 different genomic sites for transposon insertion were identified in *L. interrogans* strains (see the table in the supplemental material): 617 in *L. interrogans* serovar Manilae strain L495, 250 in *L. interrogans* serovar Lai strain 56601, 32 in *L. interrogans* serogroup Canicola strain Kito, 17 in *L. interrogans* serovar Pomona strain PO-06-047, 9 in *L. interrogans* serovar Copenhageni strain Fiocruz L1-130, and 4 in *L. interrogans* serovar Canicola strain L1-133. The insertion sites of two random mutants were also identified in *L. weilii* serogroup Hebdomadis strain EcoChallenge. All of the sequenced insertion sites could be mapped using the available *L. interrogans* serovar Lai genome sequence. This is consistent with the fact that gene content is highly conserved between *L. interrogans* serovars, with sequences of *L. interrogans* serovars Lai and Copenhageni having 95% identity at the nucleotide level (20). The position of the transposon in every mutant was plotted on a circular map representing the *L. interrogans* serovar Lai strain 56601 chromosomes (Fig. 1).

To evaluate the distribution of *Himar1* in the *L. interrogans* genome and determine any site specificity, we analyzed the insertion site sequences. The two possible orientations of the transposon with respect to the direction of replication or transcription were present in nearly equal proportions, indicating that neither orientation is favored (data not shown). We found that transposon insertion was uniformly distributed across the two chromosomes (4,333 and 358 kb in size). The mapping of 826 insertion sites over a 4,690-kb target genome yields a density of approximately one transposon integration per 5 kb.

1

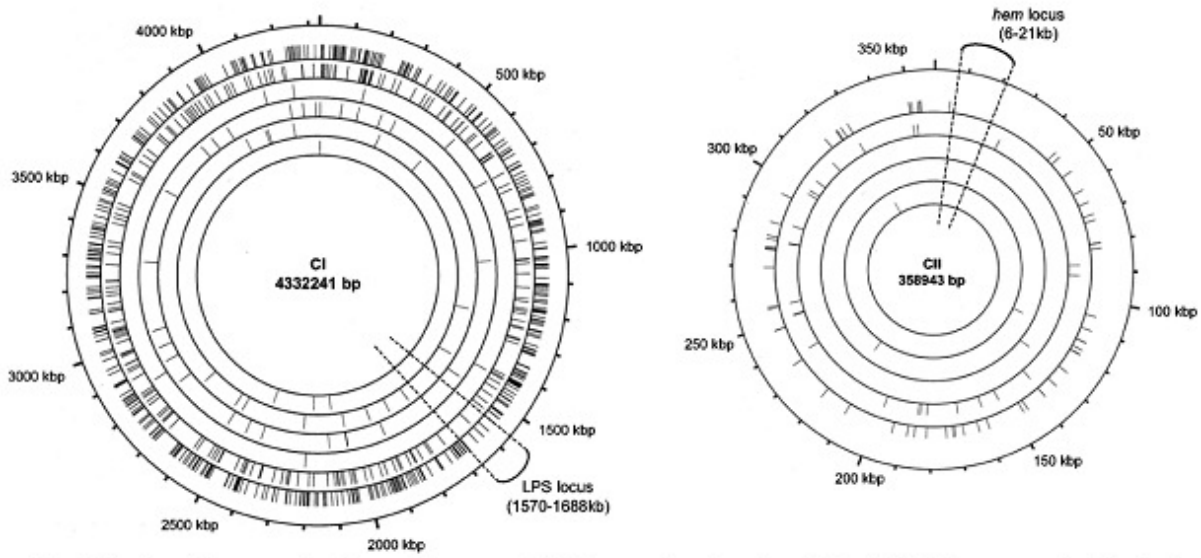


FIG. 1. Mapping of transposon insertions on the genome of *L. interrogans*. Insertion sites of *Himar1* in 826 transposon mutants (excluding insertions into 16S and 23S rRNA and transposases) of *L. interrogans* were mapped onto circular representations. From the outside in: first, coordinates of the circular chromosome; next, insertion sites of the random mutants in (i) *L. interrogans* serovar Manilae strain L495; (ii) *L. interrogans* serovar Lai strain 56601; (iii) *L. interrogans* serovar Copenhageni strain Fiocruz L1-130; (iv) *L. interrogans* serogroup Canicola strain Kito; (v) *L. interrogans* serovar Pomona strain PO-06-047; and (vi) *L. interrogans* serovar Canicola strain L1-133 (no random mutants in the small chromosome for this strain). Positions of the LPS and *hem* loci are indicated on the large (CI) and small (CII) chromosomes, respectively.

Although the profile indicates a random distribution throughout the genome, some regions of the genome showed few insertion sites. These regions generally contained genes that are notionally essential, such as the lipopolysaccharide (LPS) biosynthetic locus in the large chromosome and the heme biosynthetic genes in the small chromosome (Fig. 1). For the LPS locus (position, kilobases 1570 to 1688 of the large chromosome), the few insertion sites (9 insertions, in comparison to 21 predicted, if random insertion was normally distributed) map to an intergenic region or genes encoding hypothetical proteins.

We examined the occurrence of bases in 15-bp sequences upstream and downstream of the target site. Consistent with mariner-based mutagenesis systems used for other bacterial species (23), all *Himar1* insertions in *L. interrogans* occurred at a TA dinucleotide. Statistical target site analyses revealed an absence of any additional target site preference (Fig. 2). The proportion of *Himar1* insertions in coding sequences was 78% (721/929), a frequency that closely approximates the proportion of the genome that is protein coding (75% of the genome). With only one exception (mutants FLaiS270 and AMan990), no two transformants contained a transposon insertion at exactly the same genomic location, further suggesting that *Himar1* inserts randomly into chromosomal DNA. Surprisingly, the transposon insertion sites of several mutants were within the 16S (18 mutants) or 23S (7 mutants) rRNA gene, with each mutant showing a different insertion site. In *Leptospira* spp., rRNA genes are not linked, and *L. interrogans* contains one *rrf* gene, two *rrl* genes, and two *rrs* genes, encoding 5S, 23S, and 16S rRNA molecules, respectively. Whether there is something unusual about the architecture of these highly transcribed regions that favors transposon integration remains to be determined. Excluding insertions in 16S and 23S rRNA genes and transposases, 551 individual genes have been interrupted in *L. interrogans*. Of these, 266 (48%) encode hypothetical proteins. Among the disrupted genes, 437 have orthologs in the pathogen *L. borgpetersenii*, 312 have orthologs in the saprophyte *L. biflexa*, and notably, 139 are unique to pathogenic strains (Table 2) (see the table in the supplemental material).

These observations, together with the high A+T content of the *L. interrogans* genome, suggest that the *mariner* transposition system is suitable for the generation of libraries of random mutants. The *L. interrogans* genome contains approximately

FIG. 2. *Himar1* target site consensus sequence. Sequence logo is drawn from 100 distinct *Himar1* insertion sites in *L. interrogans* serovar Lai strain 56601. The degree of sequence conservation at each position is indicated by the height of letters (maximum of 2 bits for a nucleotide sequence).

1

TABLE 2. *L. interrogans* mutant libraries

Characteristic	Value
No. of mutants with defined insertion locations	929
No. of mutants:	
For *L. interrogans* serovar Manilae strain L495	616
For *L. interrogans* serovar Lai strain 56601	249
For *L. interrogans* serogroup Canicola strain Kito EFS	32
For *L. interrogans* serovar Pomona strain EUA	17
For *L. interrogans* serovar Copenhageni strain Fiocruz L1-130	9
For *L. interrogans* serovar Canicola strain L1-133	4
No. (%) of mutations:	
In chromosome I (92.4% of genome)[a]	826 (90)
In chromosome II (7.6% of genome)[a]	92 (10)
In coding region	721[b]
In transposase	11[c]
In rRNA gene	25[d]
In intergenic region	172
No. of ORFs disrupted:	551
Encoding hypothetical proteins	266
Encoding proteins with predicted *L. biflexa* orthologs	312
Encoding proteins with predicted *L. borgpetersenii* orthologs	437
Encoding pathogen-specific proteins	139

[a] Excluding insertion locations corresponding to multiple locations (transposases).
[b] 473 in *L. interrogans* serovar Manilae strain L495, 200 in *L. interrogans* serovar Lai strain 56601, 24 in *L. interrogans* serogroup Canicola strain Kito, 13 in *L. interrogans* serovar Pomona strain PO-06–047, 7 in *L. interrogans* serovar Copenhageni strain Fiocruz L1–130, and 4 in *L. interrogans* serovar Canicola strain L1–133.
[c] Six in *L. interrogans* serovar Lai strain 56601, four in *L. interrogans* serovar Manilae strain L495, and one in *L. interrogans* serogroup Canicola strain Kito.
[d] Nineteen in *L. interrogans* serovar Manilae strain L495, five in *L. interrogans* serovar Lai strain 56601, and one in *L. interrogans* serogroup Canicola strain Kito.

3,400 predicted protein coding regions (excluding transposases and pseudogenes), of which half have been assigned no biological role whereas the remainder have been assigned roles that await experimental validation. Based on recent whole-genome analyses of essential genes in bacteria (9), it is reasonable to assume that approximately 3,000 out of a total of 3,400 are nonessential and can therefore be mutated. Therefore, at this stage the transposon insertion library for *L. interrogans* is clearly not saturated.

Phenotypic analysis of a subset of mutants. Some mutants were further characterized by comparing their phenotypes to that of the parental strain. *L. interrogans* has periplasmic flagella, essential for motility, that are inserted at each end of the cell and extend toward the middle of the cell body. Approximately 80 genes encode proteins involved in motility (20). Mutants were identified with transposon insertions in putative motility genes, including LA0025 (encoding FliG, one of the four paralogs, associated with the flagellar motor switch in *E. coli*), LA2417 (encoding the flagellar hook protein FlgL-1, one of four paralogs), LA2069 (encoding FliN, a putative flagellar motor switch protein, one of two paralogs), LA2215 (encoding a putative flagellar motor protein, one of three or more paralogs), and LA2592 (encoding FliI, a putative flagellum-specific ATP synthase). Unexpectedly, these mutants were motile in liquid culture and did not exhibit any in vitro growth defects compared to the parental strain (data not shown). This may be due to functional redundancy; as indicated above, these genes of *L. interrogans* have multiple paralogs that may compensate for the motility-associated mutations.

Leptospires have a full nucleotide excision repair system (UvrA, UvrB, UvrC, and UvrD). A mutant with transposon disruption in *uvrB* was assayed for its ability to recover from DNA damage produced by exposure to UV irradiation. In three independent experiments, there were no detectable colonies of the *uvrB* mutant at the lowest UV dose tested, compared to 10% survival for the wild-type strain. This treatment therefore had a significantly greater effect on mortality of the *uvrB* mutant than on that of the wild-type strain. We also identified transposon mutants in a locus containing genes involved in heme acquisition (LB191, encoding a TonB-dependent transporter) and utilization (LB186, encoding a heme oxygenase) (1, 19). The iron chelator dipyridyl was used to produce iron-limited conditions that inhibited the growth of *Leptospira* strains (13). Addition of 10 μM hemin restored the ability of the *L. interrogans* wild-type strain to grow under iron starvation conditions, but not in the mutant strains. These results suggest that disruption of LB186 and LB191, which encode the heme oxygenase and a TonB-dependent receptor (1, 19), resulted in mutants that were impaired in their ability to use hemin as an iron source.

We obtained several mutants exhibiting insertions in the 16S and 23S rRNA genes. The growth rates of all mutants were comparable to that of the parental strain, with no mutants showing altered motility or morphology, consistent with the notion that the mutants are functionally able to overcome inactivation of one of the two copies of the 16S and 23S rRNAs.

To establish a system for the identification of virulence-associated genes, 29 mutants were selected for virulence testing using the hamster model of acute infection (Table 3). Analysis of the *L. interrogans* genome identified few obvious virulence factors, most likely due to the evolutionary distance between *L. interrogans* and prototypic bacterial pathogens. This is consistent with the notion that *Leptospira* has unique virulence mechanisms. Therefore, mutants were selected based on the following criteria for the disrupted gene: the absence of an orthologous gene in *L. biflexa*, a predicted outer membrane location, indicating likelihood of interaction with the host, and a potential role in signaling, motility, or chemotaxis, all of which may be required in the in vivo dissemination of *L. interrogans*. Mutants recovered from host animals were tested for stability of the transposon by PCR. In each mutant tested, the transposon remained in situ, indicating a high degree of stability.

The majority of mutants retained full virulence (Table 3), indicating that the mutagenesis process and the necessary associated in vitro passage do not per se lead to attenuation. Two mutants, with mutations in LA1641 and LA0615, were identified to have lost virulence, with all hamsters surviving infection and exhibiting no lung pathology or signs of disease. Kidneys from these hamsters were also culture negative for *Leptospira*. In both instances, the interrupted gene had no predicted function and showed normal in vitro growth. LA1641 is located in the LPS biosynthesis locus and is found only in *L. interrogans*. The mutant expressed a lower-molecular-weight LPS structure

TABLE 3. Virulence of mutants in hamster model of acute infection

Strain[a]	Location of insertion	Predicted function of mutated gene or description[b]	Location[c]	*L. biflexa*[d]	Dose(s)[e]	Hamster survival[f]
L495		Wild-type control			10^3	0/5
EMJH		Negative control				5/5
M739	LB125	Chemotaxis-related protein	Unknown	Y	10^3, 10^5	0/4
M775*	LB328	OmpA family protein	OM	N	10^3, 10^5	2/8
M776	LA1857	Fur paralogue	C	Y	10^3, 10^5	0/8
M777	intergenic	Intergenic mutant control			10^3, 10^5	0/4
M780	LA2469	CheX, inhibitor of MCP methylation	C	Y	10^3, 10^5	0/4
M789	LA3028	HP, contains leucine-rich repeats	Unknown	N	10^3, 10^5	0/4
M894	LA3258	TonB-dependent receptor	OM	Y	10^3, 10^5	0/4
M880*	LA4161	Metalloprotease	Secreted	N	10^3, 10^5	0/4
M834	LA3881	HP	OM	N	10^3, 10^5	0/4
M895#	LA1641	HP	IM	N	10^3, 10^5	4/4
M1061*	LA0709	HP	IM	N	10^3	1/5
M765*	LA3097	LipL71	OM	Y	10^3	0/5
M1224*	LA1499	HP	OM	N	10^3	0/5
M886*	LA2003	HP	OM	N	10^3	0/5
M1086*	LA4129	HP, ankyrin repeat protein	Unknown	N	10^3	0/5
M874#	LA0615	HP	Unknown	N	10^3	5/5
M421	LA4324	LenE	OM	N	10^3	1/5
M911	LA4209	HP	Unknown	N	10^4	0/5
M977	LA3103	LenB	Unknown	N	10^4	0/5
M983*	LB362	HP	Unknown	N	10^3	0/5
M1065	LA1252	CheB, chemotaxis response regulator	C	Y	10^4	2/5
M1115*	LA0676	MCP	IM	Y	10^5	0/5
M1225*	LA0136	HP, LipL45-related protein	Unknown	Y	10^3	0/5
M1014*	LA2215	Flagellar motor protein/OmpA family protein	Unknown	Y	10^3	2/5
M1073*	LA0137	HP	Unknown	N	10^3	1/5
M968	LA3075	LigC	Unknown	N	10^3	0/5
M1020*	LA0505	HP	Unknown	Y	10^3	0/5
M1059*	LA4122	RNA polymerase sigma subunit	C	Y	10^3	0/5
M1229*	LA3367	HP	Unknown	N	10^3	0/5

[a] Mutants were constructed in *L. interrogans* serovar Manilae L495. *, the genotype of reisolates was confirmed by PCR; #, bacteria could not be recovered from animals.
[b] Predicted function of protein encoded by disrupted gene. HP, hypothetical protein; MCP, methyl-accepting chemotaxis protein.
[c] Predicted subcellular location of disrupted gene product using the psorth software program (www.psort.org/psorth/); C, cytoplasmic; OM, outer membrane; IM, inner membrane.
[d] Indicates the presence of a deduced protein in *L. biflexa* with >50% similarity by BLASTP. Y, yes; N, no.
[e] Mutants were injected intraperitoneally in two doses (10^5 and 10^8 leptospires) into groups of two or four hamsters, or a single dose was injected into groups of five hamsters.
[f] No. surviving/total. Survival data are pooled for both doses if appropriate.

(unpublished data) and was selected for the virulence assay because mutations affecting LPS can lead to attenuation in other bacterial pathogens (10, 18). LA0615 is located downstream of the gene encoding LipL41 and was selected for the virulence assay because the gene is unique to pathogenic species of *Leptospira*. The system outlined here demonstrates the feasibility of using random transposon mutagenesis in conjunction with the hamster animal model to identify novel virulence factors in *L. interrogans*.

A number of mutants of particular interest were examined. These include the *ligC* mutant (LA3075, an intact gene in *L. interrogans* serovar Manilae). Members of the *lig* family of genes in *L. interrogans* encode outer membrane proteins with immunoglobulin-like repeats (16). The lack of attenuation in the *ligC* mutant is consistent with *ligC* being a pseudogene in the pathogenic serovar Copenhageni and the recent observation that mutation of *ligB* does not impair virulence in the hamster model of infection (5). An unexpected finding was that inactivation of a number of chemotaxis-related genes did not result in attenuation. It is possible that chemotaxis is not important in the hamster model of infection, but a more likely explanation is that the mutations may be compensated for by other genes; the *L. interrogans* genome has a high degree of apparent gene duplication and redundancy, with at least 24 chemotaxis genes, including 12 encoding methyl-accepting chemotaxis proteins. Likewise, mutation of the putative OmpA family protein LB328 (with 7 paralogs in the genome), the TonB-dependent receptor LA3258 (with 10 paralogs), or the *fur* gene LA1857 (4 paralogs) may have been compensated for through functional redundancy. Finally, strains carrying mutations in *lenB* and *lenE* (with six paralogs in the genome), which encode proteins binding host extracellular matrix components in vitro (28), did not show an attenuated phenotype (Table 3). Redundancy in the genome may make the identification of virulence factors in *L. interrogans* more difficult; only one attenuated transposon mutant has been described to date, with a mutation in the gene encoding LA0222, an OmpA family protein (26). Although the majority of mutants do not demonstrate an impairment in growth in vivo, further studies may find that these genes play a role under different conditions, such as at the mucosal surface.

This study presents the results of an extensive mutagenesis project generating 929 transposon insertion mutants. Given the low growth rate and genetic intractability of *L. interrogans*,

this work represents a major advance. Clearly, additional work is required to fully understand the phenotypes of randomly constructed mutants. Complementation of the disrupted genes and/or independent generation of further mutants in the same gene will need to be performed to provide confirmation for the phenotypes observed. However, the identification of two apparently attenuated mutants demonstrates the value of this work in identifying novel virulence mechanisms of *L. interrogans*. The use of different routes of inoculation, quantitative PCR, and histopathological analyses may further reveal the role of different genes in spirochete burden and tissue pathology. Further increases in transformation efficiency, through the identification of more transformable strains or the development of new genetic tools, will provide opportunities to generate extensive mutant libraries that may subsequently be used to screen for phenotypes affecting diverse aspects of the physiology of *Leptospira*.

ACKNOWLEDGMENTS

We thank David Haake for kindly providing *Leptospira interrogans* strains EcoChallenge and PO-06-047. We also thank Caroline Boursaux-Eude for graphic support.

This work was supported by the National Health and Medical Research Council, Australia; the Australian Research Council; Institut Pasteur, Paris, France; the French Ministry of Research ANR Jeunes Chercheurs (no. 05-JCJC-0105-01); the Fiocruz-Pasteur Scientific Cooperation Agreement; the Brazilian National Research Council (Instituto Milênio 420067/2005); and the National Institutes of Health (5 R01 AI052473, 2 R01 AI034431, and 2 D43 TW-00919). G.L.M. is supported by a National Health and Medical Research Council (NHMRC) Peter Doherty Fellowship.

REFERENCES

1. **Asuthkar, S., S. Velineni, J. Stadlmann, F. Altmann, and M. Sritharan.** 2007. Expression and characterization of an iron-regulated hemin-binding protein, HbpA, from *Leptospira interrogans* serovar Lai. Infect. Immun. **75:**4582–4591.
2. **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.
3. **Bulach, D. M., R. L. Zuerner, P. Wilson, T. Seemann, A. McGrath, P. A. Cullen, J. Davis, M. Johnson, E. Kuczek, D. P. Alt, B. Peterson-Burch, R. L. Coppel, J. I. Rood, J. K. Davies, and B. Adler.** 2006. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. Proc. Natl. Acad. Sci. USA **103:**14560–14565.
4. **Chiang, S. L., and E. J. Rubin.** 2002. Construction of a *mariner*-based transposon for epitope-tagging and genomic targeting. Gene **296:**179–185.
5. **Croda, J., C. P. Figueira, E. A. J. Wunder, C. S. Santos, M. G. Reis, A. I. Ko, and M. Picardeau.** 2008. Targeted mutagenesis in pathogenic *Leptospira*: disruption of the *ligB* gene does not affect virulence in animal models of leptospirosis. Infect. Immun. **76:**5826–5833.
6. **Donahue, J. P., D. A. Israel, R. M. Peek, M. J. Blaser, and G. G. Miller.** 2000. Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. Mol. Microbiol. **37:**1066–1074.
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. **Etienne, G., F. Laval, C. Villeneuve, P. Dinadayala, A. Abouwarda, D. Zerbib, A. Galamba, and M. Daffé.** 2005. The cell envelope structure and properties of *Mycobacterium smegmatis* mc(2)155: is there a clue for the unique transformability of the strain? Microbiology **151:**2075–2086.
9. **Gil, R., F. J. Silva, J. Peretó, and A. Moya.** 2004. Determination of the core of a minimal bacterial gene set. Microbiol. Mol. Biol. Rev. **68:**518–537.
10. **Harper, M., A. Cox, F. St Michael, I. W. Wilkie, J. D. Boyce, and B. Adler.** 2004. A heptosyltransferase mutant of *Pasteurella multocida* produces a truncated lipopolysaccharide structure and is attenuated in virulence. Infect. Immun. **72:**3436–3443.
11. **Johnson, R. C., and V. G. Harris.** 1967. Differentiation of pathogenic and saprophytic leptospires. J. Bacteriol. **94:**27–31.
12. **Levett, P. N.** 2001. Leptospirosis. Clin. Microbiol. Rev. **14:**296–326.
13. **Louvel, H., S. Bommezzadri, N. Zidane, C. Boursaux-Eude, S. Creno, A. Magnier, Z. Rouy, C. Médigue, I. S. Girons, C. Bouchier, and M. Picardeau.** 2006. Comparative and functional genomic analyses of iron transport and regulation in *Leptospira* spp. J. Bacteriol. **188:**7893–7904.
14. **Louvel, H., and M. Picardeau.** 2007. Genetic manipulation of *Leptospira biflexa*. J. Wiley and Sons, Hoboken, NJ.
15. **Louvel, H., I. Saint Girons, and M. Picardeau.** 2005. Isolation and characterization of FecA- and FecB-mediated iron acquisition systems of the spirochete *Leptospira biflexa* by random insertional mutagenesis. J. Bacteriol. **187:**3249–3254.
16. **Matsunaga, J., M. A. Barocchi, J. Croda, T. A. Young, Y. Sanchez, I. Siqueira, C. A. Bolin, M. G. Reis, L. W. Riley, D. A. Haake, and A. I. Ko.** 2003. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. Mol. Microbiol. **49:**929–945.
17. **McBride, A. J., D. A. Athanazio, M. G. Reis, and A. I. Ko.** 2005. Leptospirosis. Curr. Opin. Infect. Dis. **18:**376–386.
18. **Murray, G. L., S. R. Attridge, and R. Morona.** 2003. Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. Mol. Microbiol. **47:**1395–1406.
19. **Murray, G. L., K. M. Ellis, M. Lo, and B. Adler.** 2008. *Leptospira interrogans* requires a functional heme oxygenase to scavenge iron from hemoglobin. Microbes Infect. **10:**791–797.
20. **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. Degrave, 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.
21. **Picardeau, M., A. Brenot, and I. Saint Girons.** 2001. First evidence for gene replacement in *Leptospira* spp. Inactivation of *L. biflexa flaB* results in nonmotile mutants deficient in endoflagella. Mol. Microbiol. **40:**189–199.
22. **Picardeau, M., D. M. Bulach, C. Bouchier, R. L. Zuerner, N. Zidane, P. J. Wilson, S. Creno, E. S. Kuczek, S. Bommezzadri, J. C. Davis, A. McGrath, M. J. Johnson, C. Boursaux-Eude, T. Seemann, Z. Rouy, R. L. Coppel, J. I. Rood, A. Lajus, J. K. Davies, C. Médigue, and B. Adler.** 2008. Genome sequence of the saprophyte *Leptospira biflexa* provides insights into the evolution of *Leptospira* and the pathogenesis of leptospirosis. PLoS ONE **3:**e1607.
23. **Plasterk, R. H., Z. Izsvák, and Z. Ivics.** 1999. Resident aliens: the *Tc1/mariner* superfamily of transposable elements. Trends Genet. **15:**326–332.
24. **Prod'hom, G., B. Lagier, V. Pelicic, A. J. Hance, B. Gicquel, and C. Guilhot.** 1998. A reliable amplification technique for the characterization of genomic DNA sequences flanking insertion sequences. FEMS Microbiol. Lett. **158:**75–81.
25. **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.
26. **Ristow, P., P. Bourhy, F. W. da Cruz McBride, C. P. Figueira, M. Huerre, P. Ave, I. S. Girons, A. I. Ko, and M. Picardeau.** 2007. The OmpA-like protein Loa22 is essential for leptospiral virulence. PLoS Pathog. **3:**e97.
27. **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.
28. **Stevenson, B., H. A. Choy, M. Pinne, M. L. Rotondi, M. C. Miller, E. Demoll, P. Kraiczy, A. E. Cooley, T. P. Creamer, M. A. Suchard, C. A. Brissette, A. Verma, and D. A. Haake.** 2007. *Leptospira interrogans* endostatin-like outer membrane proteins bind host fibronectin, laminin and regulators of complement. PLoS ONE **2:**e1188.
29. **Tilly, K., A. F. Elias, J. L. Bono, P. Stewart, and P. Rosa.** 2000. DNA exchange and insertional inactivation in spirochetes. J. Mol. Microbiol. Biotechnol. **2:**433–442.
30. **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 result. Nucleic Acids Res. **34:**53–65.

5. ARTIGO 3

DEVELOPMENT OF AN INDUCIBLE EXPRESSION SYSTEM AND CONTROLLED COMPLEMENTATION OF MUTANTS IN *Leptospira biflexa*

(Artigo formatado segundo normas do periódico Applied and Environmental Microbiology)

Development of an Inducible Expression System and Controlled Complementation of Mutants in *Leptospira biflexa*

Running title: INDUCIBLE EXPRESSION SYSTEM FOR *L. biflexa*

Gustavo M. Cerqueira,[1,2] Odir A. Dellagostin,[2] and Mathieu Picardeau[1]*

Unité de Biologie des Spirochètes, Institut Pasteur, 75015, Paris, France[1] and Centro de Biotecnologia, Universidade Federal de Pelotas, P.O. Box 354, 96010-900, Pelotas, RS, Brazil[2]

* Corresponding author. Mailling address: Unité de Biologie des Spirochètes, Institut Pasteur, Rue du Docteur Roux, 25-28, 75015, Paris, France. Phone: 33 (1) 45 68 83 68. 33 (1) 40 61 30 01. E-mail: mpicard@pasteur.fr

ABSTRACT

The sequencing of pathogenic and saprophytic *Leptospira* spp. genomes revealed a number of coding sequences without any function assigned. The development of new genetic tools is an important goal of *Leptospira* research. Although advances have been made in the genetic manipulation of *Leptospira*, there still remains the need for basic molecular systems for assessing protein functions. In this study, we describe for the first time the generation of an inducible expression system for *L. biflexa*. Fluorescence was promoted by GFP-reporter expression. The fluorescence levels were maximum one hour post-induction with 1 mM IPTG and reduced considerably at one week p.i. In addition, the IPTG inducible system was employed in the complementation of a non-motile mutant. A *flaB*$^{-}$ mutant was transformed with the shuttle vector containing an inducible copy of the *flaB* gene and recovered its motility. This system is now available for assessing further the function of the hypothetical genes and for engineering *L. biflexa* for generation of conditional mutants.

INTRODUCTION

The spirochete *Leptospira biflexa* is a saprophytic species which has served as an important model for functional studies and gene characterization in the genus *Leptospira*. Although there is some discrepancy regarding the genome contents of *L. biflexa* and the pathogenic leptospires, 2189 coding sequences were found to be highly conserved or identical to other leptospiral genomes (25, 8, 28). Among these, 1226 hypothetical proteins have their function obscured by the lack of precise information about their role and participation in the biology of *Leptospira* (28). In addition, a number of these are predicted to be essential for *Leptospira*.

Targeted and random mutagenesis studies have enabled the function assignment of several sequences in both saprophytic and pathogenic *Leptospira* (7, 20, 29, 21, 23), however, one limitation remain with respect to the study of essential genes due to the instability of mutants in housekeeping genes. A recent publication from our group has demonstrated the development of a library of mutants generated by random transposon mutagenesis. But, despite the large number of mutants generated, only a small amount of them correspond to leptospiral essential genes (22). Thus, the development of an expression system for *Leptospira*, which may be controlled tightly and in a simple way is of great interest for the clarification of the role of this group of genes, by the potential to generate conditional mutants. At present, similar methods are available for the clinically important spirochete *Borrelia burgdorferi* and it appears to work successfully for the purpose of heterologous protein expression in that species.

Different reporter genes are available which are suitable for assessing the regulation of an inducible system in *L. biflexa*. Among them are the green fluorescent protein (*gfp*) (9, 10, 11, 13, 14), antibiotic resistance genes such as the chloramphenicol acetyltransferase (*cat)* (1, 29, 30) and the luciferase gene (*luc*) (4). Although the instability of the *gfp* and *cat* genes appeared to preclude their use among other spirochetes, new alleles of *gfp* gene in association with different promoters, should be tested for their ability to overcome previous limitations. In addition, alternative reporters may involve the recomplementation of mutant strains in *L. biflexa*. Thus, a *Paramecium caudatum* codon-biased *gfp* gene (34) and a copy of the *L. biflexa flaB* gene were evaluated in the development of an inducible expression system for *Leptospira*.

The major advantages of this system and the rational of this study is that both

reporters have been expressed already in *L. biflexa* or at least in other spirochete species (27, 24). In addition, two different promoters, P*tac* and P*hsp10*, which are responsible for coordinating the reporter's expression, have been previously tested in *L. biflexa* (Unpublished results).

Inducible systems, such as the *lac* operator/repressor and *tet* operator/repressor, were originally derived from *E. coli* (2, 3, 15) and adapted to be utilized in several bacterial species, including spirochetes (12, 16, 17, 18, 26, 32, 35, 4, 5). The development of a leptospiral-inducible expression system, which was adapted to respond to the *E. coli* LacI repressor protein, is expected to improve the genetic study of *Leptospira* as we may be able to coordinate the heterologous expression of any given target. Additionally, this system may serve as a model for the generation of new leptospiral expression systems controlled by promoters induced by a multitude of factors, and thus allow the study of the influence of environmental conditions over reporter's expression. In the present study, we demonstrate the first development and functionality of an IPTG-inducible expression system for *L. biflexa*.

MATERIALS AND METHODS

Bacterial strains and culture conditions. All strains and plasmids used in this study are described in Table 1. *E. coli* strains XL1-Blue (Stratagene, La Jolla, CA) and TOP10 (Invitrogen, Carlsbad, CA) were used as cloning hosts. Culture and transformation of *L. biflexa* were carried out as previously described (19). FlaB inducible expression experiments employed the previously characterized *flaB*$^-$ mutant (Table 1) (27). *L. biflexa* serovar Patoc strain Patoc 1 (Table 1) was utilized for expression experiments involving both *gfp* and spectinomycin resistance reporters. For fluorescence production validation, propagation and analysis, *L. biflexa* strains were cultured in liquid EMJH supplemented with 1% albumin and 50 ng/µl spectinomycin or 50 ng/µl kanamycin when needed. Leptospiral density was assessed by dark-field microscopy and spectrophotometry at 460 nm.

Generation of constructs. Primers used in this study are described as supplemental material in Table 1. The construction containing P*flaB* promoter plus *lacI* repressor gene was initially amplified from the plasmid pJSB104 (Table 1) and cloned into pGSBLe24 and pSLe94 to generate pGL and pSL. The promoters chosen to constitute the inducible system were *Ptac* and *Phsp10* and they were engineered in the reverse primer to contain one *lacO* site and a 6x His-tag. Both promoters were

cloned in pCR2.1 (Invitrogen) and used to transform *E. coli* XL-10 strain by heat-shock. Recombinant colonies were propagated in liquid Luria Bertani (LB) medium supplemented with 50 ng/µl kanamycin and the plasmids were extracted and submitted to restriction analysis and sequencing to confirm the presence of the engineered promoters. To validate the functionality of these promoters we cloned *gfp* downstream of each one and amplified the final construction to clone it into the shuttle vector pSLe94. Both pSTG and pSHG were used to transform *L. biflexa* strain Patoc 1. Recombinant colonies of Patoc 1 strain were propagated in liquid EMJH supplemented with 1% albumin and 50 ng/µl spectinomycin, and used for epifluorescent microscopy analysis. To validate the inducible expression system we first tested *gfp* as the reporter gene. Only *Phsp10* promoter was evaluated in this case. The *gfp* gene was cloned downstream *Phsp10* promoter and the construction was amplified and cloned in both pGL and pSL constructions, and used to transform *E. coli* Top 10 strain (Invitrogen). Recombinant colonies were selected but only those cloned in pSLe94 were selected for further analysis. The *flaB* reporter was also cloned in pCR2.1 downstream *Phsp10* promoter and used to transform *E. coli* Top 10 strain. Then, the construction was amplified and cloned in both pGL and pSL to generate pGLHF and pSLHF, though only the last one was used in the study. *gfp* inducible expression system was introduced by electroporation in Patoc 1 strain, while that containing *flaB* gene was used to transform Patoc 1 *flaB*$^-$ mutant (Table 1). Recombinant colonies were observed in solid EMJH after one week incubation at 30 ºC and then propagated in liquid EMJH as mentioned in the previous section. All constructs were confirmed by restriction analysis and direct sequencing.

Induction assays. Leptospires were cultivated at 30 ºC in the presence of 50 ng/µl spectinomycin or kanamycin when appropriate, up to 10^9 cells/ml and then induced with 1 and 10 mM IPTG, or non-induced. . Culture growth density was determined by optical density using Biomate 3 spectrophotometer (Thermo, USA) at 420 nm.

Epifluorescence microscopy. Slides were prepared using 10^9 spirochetes/ml cultures. One ml of each culture was centrifuged at 4,000 x *g* for 20 min and ressuspended in the same original volume using sterile deionized water. Five microliters (5 x 10^6) of each ressuspension were deposited onto slides and covered with cover slips previously coated with poly-L-Lysine solution (Sigma). Slides were incubated for 4 h in the dark before fluorescence analysis. Conventional epifluorescence microscopy was performed with a Zeiss Axioplan2 fluorescence

microscope mounted with a Hamamatsu charge-coupled device camera, a narrow band GFP filter (480 -/+ 10 nm excitation wavelength; 510 -/+ 10 nm emission wavelength: and exposure time of 2000 ms. Images were acquired using the software program OpenLab. Identical image capture and adjustment settings were used for all images.

Fluorescence assays. For quantification of GFP fluorescence 10^9 spirochetes/ml cultures were harvested (one ml) at 4,000 x *g* for 20 min and ressuspended to 10^7 leptospires/µl (300 µl) in sterile deionized water. The assay was performed in triplicate and 100 µl of each culture were deposited per well in a black microtiter plate Optiplate-96F (Corning Inc., Corning, NY). GFP fluorescence was determined in a plate reader UV-light spectrophotometer, BertholdTech (Mithras, France) using a wavelength range between 485 (excitation) and 535 (emission). Before each measuring the number of cells was normalized. Culture growth was determined by optical density as described in previous section *Induction assays*.

Motility analysis. Leptospires were let to grow up to 10^9 cells/ml and then picked into 1% agar EMJH plates without antibiotics. Plates used for induction were supplemented with 1 mM IPTG. Plates were incubated 1-2 weeks under 30°C.

Immunoblot analyses. Leptospires carrying the GFP inducible expression system was cultivated as previously mentioned in the presence of spectinomycin and when cell density reached 10^7 they were induced with 1 mM IPTG and incubated for one week at 30°C. Approximately 10^9 spirochetes/ml (one ml) were collected and processed for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot. A mouse IgG anti-6x histidine-tag antibody (Clonetech, USA) was diluted 1:2,000 in 1x TBS and incubated with the membrane for 2 h at room temperature. Then a secondary antibody, anti-mouse IgG alkaline phosphatase conjugate, was diluted 1:1,000 and incubated with the membrane for 1 h at room temperature. Membranes were washed for 15 min between each step using 1x TBS supplemented with 0.05% Tween-20 (T-TBS). Blots were developed using the BCIP/NBT alkaline phosphatase chromogenic substrate according to the manufacturer's instructions (Uptima, France).

RESULTS

Functionality of *lac* operator (*lacO*)-containing promoters in *L. biflexa*. The promoters *Ptac* and *Phsp10* were modified and tested for functionality into *L. biflexa* strain Patoc1. *Ptac promoter* derives from pILL2150 plasmid (5). This promoter contains one *lac* operator (*lacO*) site, located between the -35 and -10 regions. *Phsp10* promoter is a leptospiral promoter which was engineered to include one *lacO* site. To demonstrate their functionality into *L. biflexa gfp* was cloned downstream of both promoters into pCR2.1 to constitute pCRTG and pCRHG (Table 1). After selection of green fluorescent *E. coli* colonies both *Ptac-gfp* and *Phsp10-gfp* constructs were amplified and inserted in pSLe94 *E. coli-L. biflexa* shuttle vector to constitute pSTG and pSHG (Table 1 and Fig. 1). This shuttle vector is able to replicate in both *E. coli* and *L. biflexa* due to the presence of the *oriColE1* (*E. coli*) and LE1 (leptospiral phage) origins of replication, respectively. Interestingly, it was noted that *E. coli* colonies transformed by the shuttle vectors took approximately 48 hours for a visible fluorescence, even when transformed by the construction containing *Ptac* promoter. Fluorescent *E. coli* colonies were propagated and the constructs were then used to transform *L. biflexa*. Once fluorescence cannot be directly observed in *Leptospira* colonies, a number of clones either carrying pSTG or pSHG was cultivated and used for epifluorescence microscopy. Low levels of fluorescence were detected among *L. biflexa* transformed with either pSTG or pSHG. The age of the cultures was also evaluated as a parameter affecting fluorescence production, but appeared to have no influence over fluorescence production (data not shown). A non-transformed culture of *L. biflexa* was employed as the control.

Other experiments from our group employed a similar *Phsp10-gfp* construction, but without any *lacO* site, thus expressing constitutively *gfp*, and detected considerable levels of fluorescence among transformed *Leptospira*. Such findings lead us to speculate that the modification of both promoters by the insertion of one *lacO* site may have influenced GFP expression and fluorescence production. However, a system based in low expression levels can be more tightly controlled, what may be of interest for a number of applications.

Development of an IPTG-inducible expression system in *L. biflexa*. It has been previously demonstrated in *B. burgdorferi* the functionality of a LacI repressor-based

inducible system (4). In that study LacI repressor was codon-adapted to better fit the genome content of *B. burgdorferi*. In addition, the *lacI* gene was cloned downstream of a strong promoter (*PflaB* promoter). Considering the similar AT% content between *L. biflexa* and *B. burgdorferi* (approx. 40% GC) we decided to base our system on the pre-engineered *lacI* repressor construction.

The promoters used to constitute this inducible system were *Ptac* and *Phsp10*, as mentioned before. As an inducer, it was chosen IPTG, a nonhydrolyzable inducer of the LacI repressor that is capable of crossing the membrane of numerous prokaryotes (12) and eukaryotes (33), even in the absence of the LacY permease, which is absent from *L. biflexa* genome. To construct the inducible system, the cassette containing the *PflaB* promoter plus the *lacI* gene was amplified from the pJSB104 (Table 1). For further cloning of reporters a *Sma*I site was included in the reverse primer (Fig. 1B). To test the *lac* inducible system, constructs were generated in both pGSBLe24 and pSLe94 *E. coli-L. biflexa* shuttle vectors to express both GFP and FlaB proteins (Fig. 1A and B). For *L. biflexa* transformation purposes only pSLe94-encoded inducible systems were used due to the higher stability of this plasmid, which is motivated by the presence of a higher number of partition genes from LE1 temperate (6) (Fig. 1A). Thus, pSLHG, was electroporated into *L. biflexa* strain Patoc 1, while pSLHF was introduced in *flaB*$^-$ mutant strain (Table 1 and Fig. 1B).

Induction of fluorescence in *L. biflexa.* Inductions were performed by growing Patoc 1 strain and Patoc 1 transformed with the pSLHG shuttle vector up to a cell density between 10^8 and 10^9 spirochetes/ml. Then, cultures used for induction purposes were supplemented with 1 or 10 mM of IPTG. No cell clumps were observed to occur in the cultures that were used for fluorescence assays. Samples to be assayed were collected at 1 week post-induction (p.i.) in triplicate. Treatment of the cultures with IPTG was able to induce fluorescence up to the last time-point evaluated, one week p.i., and the culture treated with 1 mM yielded the highest levels of fluorescence (Fig. 2). Though minimal differences were observed between the fluorescence level induced by 1 and 10 mM IPTG, 1 mM appeared to be an ideal concentration for induction of expression. The non-induced culture also produced some fluorescence signal, when compared to wild-type Patoc 1 strain, demonstrating the occurrence of leakage during repression of protein expression. However, it

appears that expression is lower than when the culture is induced by IPTG. Thus, to evaluate the moment of the highest level of expression we performed kinetics of GFP expression. Fractions of both non-induced and induced cultures were collected at 0, 1, 3, 6, 9, 12 and 24 h, and one week p.i. (Fig. 3). The maximum expression level was achieved immediately after the first hour p.i., and did not change during the first 24 h p.i. However, it was observed a considerable reduction (27%) in the emission of fluorescence at one week p.i. (Fig. 3). Despite this, the fluorescence measured at one week p.i. was still more than 3-fold higher (average 22,579 arbitrary units) than the level observed in the one week non-induced samples (average 7,301 arbitrary units) (Fig. 3). Additionally, for further applications as target mutagenesis of housekeeping genes, it is convenient to employ a system that is able to hold protein expression as long as one week, the minimum time to observe *L. biflexa* colonies in plates.

Western blot analysis was employed to demonstrate that the GFP protein expression is different between induced and non-induced treatments. One mM IPTG was the concentration of choice once it appears as the best concentration between both treatments evaluated for fluorescence induction. For this purpose, the one week time-point was employed and, here, difference in protein expression was also observed between induced and non-induced cultures (Fig. 4).

To evaluate the *in vitro* fluorescence produced by GFP when expressed by the inducible expression system we performed an epifluorescence microscopy. Despite the considerable levels of fluorescence obtained by using this system and the successful detection of GFP by Western blot, leptospires exhibited an *in vivo* fluorescence lower than expected (Data not shown). Explanation for this observation includes the unstable influence of a *lacO* site into the promoter, which may cause reduction in gene transcription. Despite this, clear differences in *in vivo* fluorescence could be noted between the induced culture and controls (Data not shown).

Induction of motility and complementation of a *flaB*$^-$ mutant *L. biflexa.* To demonstrate an immediate application of our system, we decided to use a previously constructed *L. biflexa flaB*$^-$ gene mutant for transformation with shuttle vector containing the *flaB* inducible expression system. Dark-field microscopy analysis showed this mutant lost motility due to the absence of the endoflagella (27). FlaB protein has a central role in leptospiral endoflagella formation. Thus, we cloned this

gene in both pGSBLe24 and pSLe94 (Table 1 and Fig. 1), under the control of *Phsp10* promoter, although only pSLe94 constructs were used for transformation and analysis. Thus, we were able to complement the absence of the native copy of the *flaB* gene and recover the motile phenotype in *flaB*⁻ mutant, under induction with 1 mM IPTG (Fig. 5). Indeed, plating of the strains Patoc 1, *flaB*⁻ mutant, and both induced and non-induced clones was decisive to confirm the functionality of this system into *L. biflexa*. We observed a significant difference in the spreading of the induced clone when compared with both non-induced and non-recomplementated *flaB*⁻ mutant (Fig. 5). The 1 mM IPTG concentration was used to induce motility in the plate, based in previous results obtained with GFP inducible expression system. Interestingly, when the 1 mM induced clone was observed by dark field microscopy the motility appeared to have been recovered, in comparison with the *flaB*⁻ mutant, but not homogeneously as in the wild-type Patoc 1 strain since some of the cells still appeared to be non-motile (data not shown). This is in agreement with the low level of fluorescence presented by the GFP-inducible clones. Also, in this case, we speculate about the low transcription levels of the *flaB* gene or the misfolding of the expressed protein.

DISCUSSION

Inducible expression systems have been developed for a number of bacterial species, including the spirochete *B. burgdorferi*, and have demonstrated a great applicability for genetic studies or as a controllable expression system (12, 18, 4, 5). For *Leptospira* there is no similar tool available, so far. The inducible expression system developed for *B. burgdorferi* by Blevins et al. (4) successfully employed a LacI repressor-based system, which can be regulated by the use of IPTG. In addition, the codon usage of the *lacI* gene used to constitute that system was adapted for borrelial genome content, in order to promote a more strict control of the expression. Considering the simplicity of the functioning of a lac-based system we focused this study in the development of a similar tool which is functional also in *L. biflexa*. The LacI repressor used to engineer our system was based in that described by Blevins et al. (4). Given the codon adaptation of that *lacI* allele for borrelial genome content maximization was achieved in its expression, and this observation was associated with a tighter repression of the expression system. Additionally, in that study, the *lacI* gene was put under the control of the strong *PflaB* promoter.

These features together favored LacI repressor expression in *B. burgdorferi*. Thus, due to the similarity in the AT% content between *Borrelia* and *Leptospira* species genomes (28) and previous evaluation of the functionality of *PflaB* promoter in *L. biflexa* (unpublished results), it seemed rational to base our system in that construction. As an IPTG-inducible promoter we initially choose the *E. coli Ptac* promoter and the *L. biflexa Phsp10* promoter. Both promoters were used to express the *gfp* reporter, which was initially selected to enable us to determine and quantify optimal IPTG treatments. The *gfp*, in this case, was not biased to favor its expression in *L. biflexa*. However, a recent study employing this allele as a reporter obtained high levels of fluorescence in pathogenic *B. burgdorferi* (24). Despite this, *Leptospira* transformed with pSHG plasmid (Table 1) exhibited low *in vitro* fluorescence (data not shown).

The optimal IPTG concentration to induce the pSLHG expression system was determined as 1 mM, which induced a slightly higher level of fluorescence than 10 mM (Fig. 2). The fluorescence levels produced by *L. biflexa* transformed with this construction, and induced by 1 mM IPTG, were highest immediately 1 h postinduction (p.i.), and was kept in similar levels for 24 h p.i. (Fig. 3). The treatment of cultures with 1 mM IPTG has been demonstrate to be optimal and promote the best expression results for other inducible systems (16, 35, 4), however, our findings about the kinetics of the fluorescence levels contrast to that obtained by Blevins and colleagues in *B. burdorferi* (4). In than study, which also employed a spirochete as the model, it was found that the maximum fluorescence levels obtained by Luciferase expression were between 3 and 15 h p.i. A number of reasons may contribute to the differences observed in our study including (*i*) the differences among promoters chosen to induce these reporters, (*ii*) the *in vivo* stability of the plasmid vectors or the protein expressed, or (*iii*) the differences among spirochetal genetic machinery. A fluorescence signal was also detected among non-induced cultures. However, only low signals could be observed among them and which are lower (3-fold) than IPTG-induced cultures (Fig. 2 and 3). Western blot analyses were also performed and confirmed the expression of GFP in IPTG-induced cultures, whereas no protein was detected among that non-induced (Fig. 4). This demonstrates the functionality of the borrelial-adapted LacI as a repressor in *Leptospira.*

In this study, only one *lacO* site was inserted in *Phsp10* promoter, between the -10 and Shine-Delgarno regions. Previous studies mentioned the need to use two

lacO sites to achieve a better control of the expression. However, in our case, we did not observe a high-level of expression. Thus, it was not necessary to reengineer our promoter to obtain a more strict control. In fact, we speculate if the insertion of one *lacO* site into the P*hsp10* promoter was responsible for the reduced expression of GFP. Supporting this hypothesis is the observation of a strong expression of GFP under the control of the native non-modified *Phsp10* promoter, in another study from our group (data not shown). This leads to the conclusion that this system should be modified to make it suitable for maximal expression applications by replacing the *PHsp10* promoter by an *E. coli* Lac-inducible promoter, other than *Ptac*, or a spirochaetal promoter whose expression control is not influenced by the insertion of *lacO* operators.

To validate our inducible system we used the *flaB* gene as a reporter to complement the loss of motility of a *L. biflexa flaB*$^-$ mutant. The *flaB* gene was cloned downstream the inducible *Phsp10* promoter and used to generate pSLHF (Table 1). Plating of the induced culture produced a spreading pattern similar to that found for the wild-type Patoc 1 strain, whereas the non-induced culture behave as the *flaB*$^-$ mutant strain (Fig. 5). Dark-field microscopy analysis of the complemented culture demonstrated a heterogeneous motility among induced leptospires, when compared with the wild-type Patoc 1 strain and Patoc 1 *flaB*$^-$ mutant (Data not shown). Western blot analysis produced negative results concerning the detection of FlaB protein expression (data not shown). Thus, we speculate similarly to GFP inducible system, FlaB is being very poorly expressed.

The successful development of an IPTG inducible expression system opens the opportunity to carry out new studies of molecular genetics in *L. biflexa*. The immediate application of our inducible system will be to generate conditional mutants of essential genes and study their function and participation in *Leptospira* physiology, but a number of applications may involve the use of an inducible expression system.

ACKNOWLEDGEMENTS

We thank Dr. George Chaconas, Dr. Ivo Boneca, Dr. Jon Blevins and Dr. Michael Norgard. These authors have kindly provided the pTM61, the pILL2150 and pJSB104, all of which have been used as templates. GMC was supported by CAPES foundation, Brazilian Government, Ministry of Education. This work was supported by

1 Institute Pasteur, Paris, France; the French Ministry of Research ANR Jeunes
2 Chercheurs (no. 05-JCJC-0105-01).
3

REFERENCES

1. **Alverson, J., S. F. Bundle, C. D. Sohaskey, M. C. Lybecker, and D. S. Samuels.** 2003. Transcriptional regulation of the *ospAB* and *ospC* promoters from *Borrelia burgdorferi*. Mol. Microbiol. **48:**1665–1677.
2. **Beckwith, J. R.** 1978. Lac: the genetic system, p. 11–30. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
3. **Beckwith, J. R.** 1987. The lactose operon, p. 1444–1452. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, DC.
4. **Blevins, J. S., A. T. Revel, A. H. Smith, G. N. Bachlani and M. V. Norgard.** 2007. Adaptation of a luciferase gene reporter and lac expression system to Borrelia burgdorferi. Appl Environ Microbiol. **73**:1501-13.
5. **Boneca, I. G., C. Ecobichon, C. Chaput, A. Mathieu, S. Guadagnini, M. V. Prévost, F. Colland, A. Labigne and H. de Reuse.** 2008. Development of inducible systems to engineer conditional mutants of essential genes of Helicobacter pylori. Appl Environ Microbiol. **74**:2095-102.
6. **Bourhy, P., L. Frangeul, E. Couvé, 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**:3931-40.
7. **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. 187:3255-8.
8. **Bulach, D. M., R. L. Zuerner, P. Wilson, T. Seemann, A. McGrath, P. A. Cullen, J. Davis, M. Johnson, E. Kuczek, D. P. Alt, B. Peterson-Burch, R. L. Coppel, J. I. Rood, J. K. Davies and B. Adler.** 2006. Genome reduction in Leptospira borgpetersenii reflects limited transmission potential. Proc Natl Acad Sci U S A. **103**:14560-5.
9. **Caimano, M. J., C. H. Eggers, K. R. Hazlett, and J. D. Radolf.** 2004. RpoS is not central to the general stress response in *Borrelia burgdorferi* but does control

expression of one or more essential virulence determinants. Infect. Immun. **72:**6433–6445.

10. **Caimano, M. J., C. H. Eggers, C. A. Gonzalez, and J. D. Radolf.** 2005. Alternate sigma factor RpoS is required for the in vivo-specific repression of *Borrelia burgdorferi* plasmid lp54-borne *ospA* and *lp6.6* genes. J. Bacteriol. **187:**7845–7852.
11. **Carroll, J. A., P. E. Stewart, P. Rosa, A. F. Elias, and C. F. Garon.** 2003. An enhanced GFP reporter system to monitor gene expression in *Borrelia burgdorferi*. Microbiology **149:**1819–1828.
12. **Dancz, C. E., A. Haraga, D. A. Portnoy, and D. E. Higgins.** 2002. Inducible control of virulence gene expression in *Listeria monocytogenes*: temporal requirement of listeriolysin O during intracellular infection. J. Bacteriol. **184:**5935–5945.
13. **Eggers, C. H., M. J. Caimano, and J. D. Radolf.** 2004. Analysis of promoter elements involved in the transcriptional initiation of RpoS-dependent *Borrelia burgdorferi* genes. J. Bacteriol. **186:**7390–7402.
14. **Eggers, C. H., M. J. Caimano, and J. D. Radolf.** 2006. Sigma factor selectivity in *Borrelia burgdorferi*: RpoS recognition of the *ospE*/*ospF*/*elp* promoters is dependent on the sequence of the -10 region. Mol. Microbiol. **59:**1859–1875.
15. **Hillen, W., and C. Berens.** 1994. Mechanisms underlying expression of Tn*10* encoded tetracycline resistance. Annu. Rev. Microbiol. **48:**345–369.
16. **Jackson, D. W., K. Suzuki, L. Oakford, J. W. Simecka, M. E. Hart, and T. Romeo.** 2002. Biofilm formation and dispersal under the influence of the global regulator CsrA of *Escherichia coli*. J. Bacteriol. **184:**290–301.
17. **Ji, Y., B. Zhang, S. F. Van Horn, P. Warren, G. Woodnutt, M. K. Burnham, and M. Rosenberg.** 2001. Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. Science **293:**2266–2269.
18. **Kamionka, A., R. Bertram, and W. Hillen.** 2005. Tetracycline-dependent conditional gene knockout in *Bacillus subtilis*. Appl. Environ. Microbiol. **71:**728–733.
19. **Louvel, H. and M., Picardeau.** 2007. Genetic manipulation of Leptospira biflexa. Curr Protoc Microbiol. 2007 Jun;Chapter 12:Unit 12E.4.
20. **Mazouni, K., G. Pehau-Arnaudet, P. England, P. Bourhy, I. Saint Girons and M. Picardeau.** 2006. The scc spirochetal coiled-coil protein forms helix-like

filaments and binds to nucleic acids generating nucleoprotein structures. J Bacteriol. **188**:469-76.

21. **(a) Murray, G.L., A. Srikram, R. Henry, A. Puapairoj, R. Sermswan and Adler B.** 2008. Leptospira interrogans requires heme oxygenase for disease pathogenesis. Microbes Infect. *In press*.
22. **(b) Murray, G. L., V. Morel, G. M. Cerqueira, J. Croda, A. Srikram, R. Henry, A. I. Ko, O. A. Dellagostin, D. M. Bulach, R. Sermswan, B. Adler and M. Picardeau.** 2008. Genome-wide transposon mutagenesis in pathogenic Leptospira spp. Infect Immun. *In press*.
23. **(c) Murray, G. L., A. Srikram, D. E. Hoke, E. A. Jr Wunder, R. Henry, M. Lo, K. Zhang, R. W. Sermswan, A. I. Ko and B. Adler.** 2008. The major surface protein LipL32 is not required for either acute or chronic infection with Leptospira interrogans. Infect Immun. *In press*.
24. **Moriarty, T.J., M. U. Norman, P. Colarusso, T. Bankhead, P. Kubes and G. Chaconas.** 2008. Real-time high resolution 3D imaging of the lyme disease spirochete adhering to and escaping from the vasculature of a living host. PLoS Pathog. **4**:e1000090.
25. **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. Degrave, O. A. Dellagostin, H. El-Dorry, E. S. Ferro, M. I. Ferro, L. R. Furlan, M. Gamberini, E. A. Giglioti, A. Góes-Neto, G. H. Goldman, M. H. Goldman, R. Harakava, S. M. Jerônimo, 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**:2164-72.
26. **Page, A. L., H. Ohayon, P. J. Sansonetti, and C. Parsot.** 1999. The secreted IpaB and IpaC invasins and their cytoplasmic chaperone IpgC are required for intercellular dissemination of *Shigella flexneri*. Cell. Microbiol. **1:**183–193.

27. **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**:189-99.
28. **Picardeau, M., D. M. Bulach, C. Bouchier, R. L. Zuerner, N. Zidane, P. J. Wilson, S. Creno, E. S. Kuczek, S. Bommezzadri, J. C. Davis, A. McGrath, M. J. Johnson, C. Boursaux-Eude, T. Seemann, Z. Rouy, R. L. Coppel, J. I. Rood, A. Lajus, J. K. Davies, C. Médigue and B. Adler.** 2008. Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis. PLoS ONE. **3**:e1607.
29. **Ristow, P., P. Bourhy, F. W. da Cruz McBride, C. P. Figueira, M. Huerre, P. Ave, I. S. Girons, A. I. Ko and M. Picardeau.** 2007. The OmpA-like protein Loa22 is essential for leptospiral virulence. PLoS Pathog. **3**:e97.
30. **Sohaskey, C. D., C. Arnold, and A. G. Barbour.** 1997. Analysis of promoters in *Borrelia burgdorferi* by use of a transiently expressed reporter gene. J. Bacteriol. **179:**6837–6842.
31. **Sohaskey, C. D., W. R. Zuckert, and A. G. Barbour.** 1999. The extended promoters for two outer membrane lipoprotein genes of *Borrelia* spp. uniquely include a T-rich region. Mol. Microbiol. **33:**41–51.
32. **Sun, J., L. Zheng, C. Landwehr, J. Yang, and Y. Ji.** 2005. Identification of a novel essential two-component signal transduction system, YhcSR, in *Staphylococcus aureus*. J. Bacteriol. **187:**7876–7880.
33. **Syroid, D. E., R. I. Tapping, and J. P. Capone.** 1992. Regulated expression of a mammalian nonsense suppressor tRNA gene in vivo and in vitro using the *lac* operator/repressor system. Mol. Cell. Biol. **12:**4271–4278.
34. **Takenaka, Y., N. Haga, T. Harumoto, T. Matsuura and Y. Mitsui.** 2002. Transformation of Paramecium caudatum with a novel expression vector harboring codon-optimized GFP gene. Gene **284**:233-40.
35. **Yansura, D. G., and D. J. Henner.** 1984. Use of the *Escherichia coli lac* repressor and operator to control gene expression in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **81:**439–443.

FIGURE LEGENDS

FIG. 1. Diagram illustrating the two *L. biflexa* shuttle plasmids used in this study (A) and the relevant regions and restriction sites of derivative constructs generated and transformed into *L. biflexa* (B). pSLe94 shuttle vector is more stable than pGSBLe24 due to the presence of a larger number of partition genes. pSTG and pSHG were used to assess the impact of the insertion of one *lacO* operator into *Ptac* and *Phsp10* promoters. pSL was constructed to receive further *lacO*-modified promoters-reporters constructions and generate the pSLHG and pGLHG, which contain a *gfp* inducible system, and pSLHF and pGLHF, which contain a *flab* inducible system under the control of the *lac*-inducible expression system.

FIG. 2. Kinetics of green fluorescence induction from the *Leptospira lac* repressor/operator expression construct (pSLHG). Cultures of Patoc 1 (wt strain) and Patoc 1-*gfp* (containing PSLHG) were induced with various concentrations of IPTG. A culture of Patoc 1 was induced with 10 mM IPTG, while Patoc 1-*gfp* was untreated (0 mM IPTG) or induced with 1 and 10 mM IPTG; samples were collected at one week postinduction. Fluorescence levels (OD) from triplicate samples of each culture were standardized according to a cell density of 1 x 10^8 spirochetes. Four independent induction studies were performed with equivalent results.

FIG. 3. Kinetics of green fluorescence production from the *lac* repressor/operator expression construct (pSLHG). A culture of Patoc 1-*gfp* was untreated (0 mM IPTG) or induced with 1 mM IPTG. Samples were collected at the designated times (h), and fluorescence assays were performed. GFP activities (OD) from triplicate samples of each culture were standardized according to a cell density of 1 x 10^8 spirochetes; results are presented as the mean OD/ 1 x 10^8 bacteria ± standard deviation.

FIG. 4. Western blot analysis of GFP using anti-6x His monoclonal antibodies. Cultures of Patoc 1 and Patoc 1-*gfp* (pSLHG) were untreated or induced with 1 mM IPTG. Cells were collected at one week postinduction. Total protein from 1 x 10^8 spirochetes was loaded in each gel lane for the GFP immunoblot. GFP detection was included to confirm that equivalent concentrations of lysates were loaded per gel

lane. Values at left denote relevant molecular masses (kDa) of Bio-Rad Precision Plus standard (MW).

FIG. 5. Spreading of Patoc 1, Patoc 1 *flaB*$^-$ mutant and Patoc 1-*flaB* (pSLHF). Bacteria were picked into solid agar (1%) and incubated for growth during one week. At One week of growth, the wild-type Patoc 1 strain and induced (1 mM IPTG) Patoc 1-*flaB* presented the same colony morphology and dissemination through the agar. In the absence of IPTG, both the Patoc 1 *flaB*$^-$ mutant and Patoc 1-*flaB* presented as a non-spread colony. Several independent induction studies were performed, and similar results were obtained between Patoc 1 wild-type and induced Patoc 1-*flaB*, and Patoc 1 *flaB*$^-$ mutant and non-induced Patoc 1-*flaB*, either in solid (1%) or soft (0.5%) agar.

TABLES

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Characteristics	Reference
Plasmids		
pCRT	pCR2.1-P*tac*	This study
pCRTS	pCR2.1-P*tac-spc*^{r}	This study
pCRTG	pCR2.1-P*tac-gfp*	This study
pCRH	pCR2.1-P*hsp10*	This study
pCRHS	pCR2.1-P*hsp10-spc*^{r}	This study
pCRHG	pCR2.1-P*hsp10-gfp*	This study
pCRHF	pCR2.1-*Phsp10-flaB*	This study
pJSB104	pJD7::P*pQE30*-Bb*luc*+ and P*flaB*-Bb*lacI* (tandem); Spec/Strep^r	4
pGSBLe24	*ori pGEM-72f ori LE1 lacZ rep spc*^{r}	6
pGL	pGSBLe24-P*flaB*-Bb*lacI*	This study
pGLHG	pGSBLe24-P*flaB*-Bb*lacI*-P*hsp10-gfp*	This study
pGLHF	pGSBLe24-P*flaB*-Bb*lacI*-P*hsp10-flaB*	This study
pSLe94	*ori pGEM-72f ori LE1 lacZ rep parA parB spc*^{r}	6
pSL	pSLe94-P*flaB*-Bb*lacI*	This study
pSTG	pSLe94-P*tac-gfp*	This study
pSHG	PSLe94-P*hsp10-gfp*	This study
pSLHG	pSLe94-P*flaB*-Bb*lacI*-P*hsp10-gfp*	This study
pSLHF	pSLe94-P*flaB*-Bb*lacI*-P*hsp10-flaB*	This study
Strains		
L. biflexa		
Patoc1 wild-type		28
Patoc1 *flaB*$^-$ mutant		27

FIGURES

Figure 1A

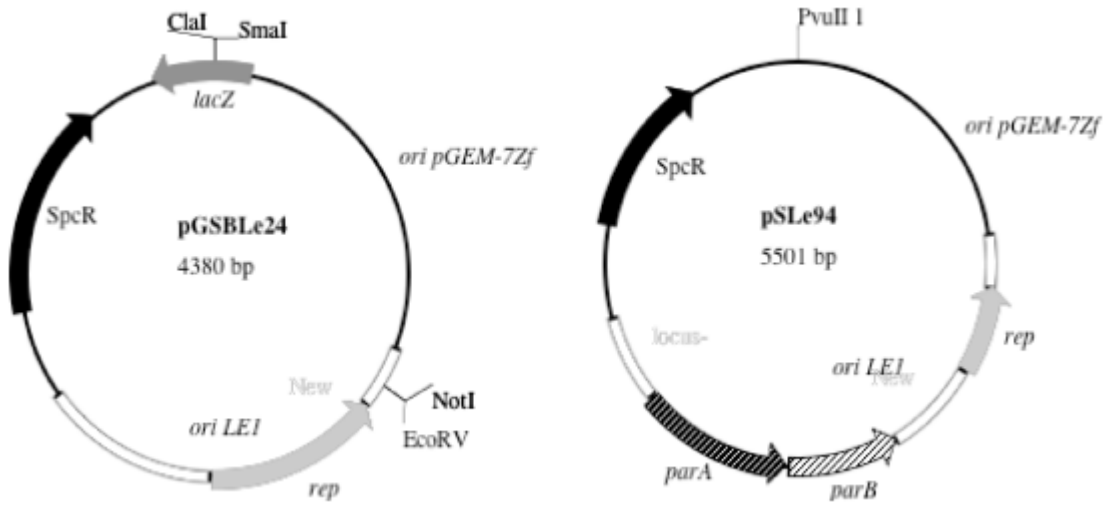


Figure 1B

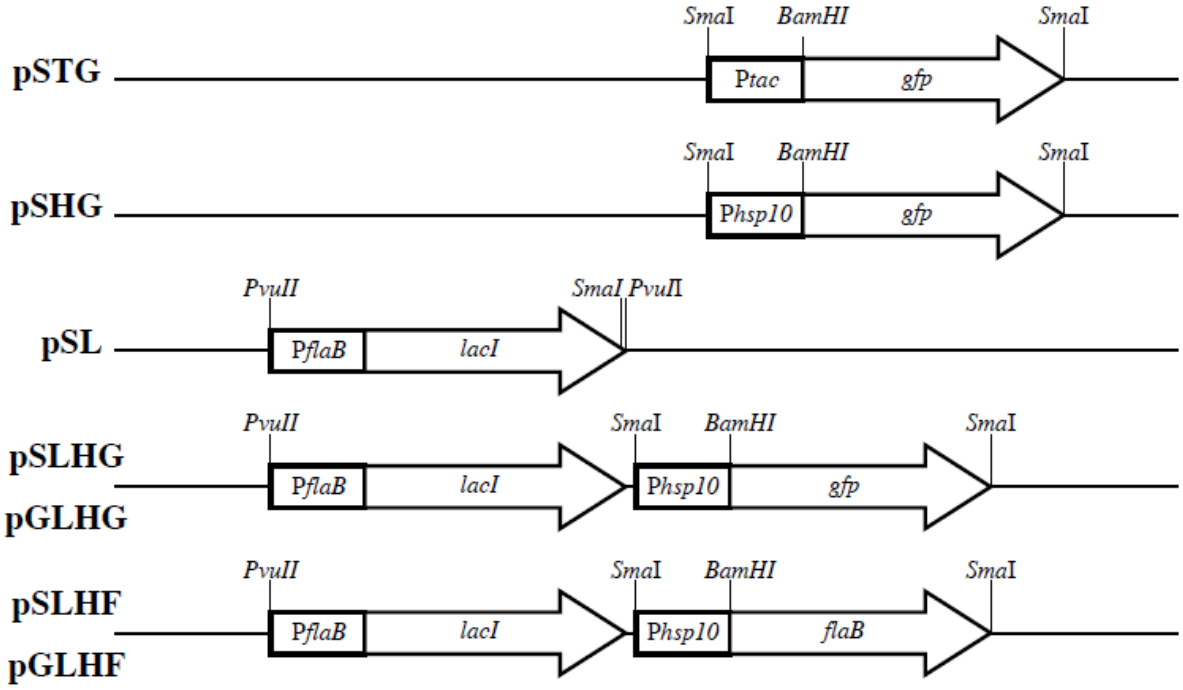

pSTG
SmaI
BamHI
SmaI
Ptac
gfp
pSHG
SmaI
BamHI
SmaI
Phsp10
gfp
pSL
PvuII
SmaI
PvuII
PflaB
lacI
pSLHG
pGLHG
PvuII
SmaI
BamHI
SmaI
PflaB
lacI
Phsp10
gfp
pSLHF
pGLHF
PvuII
SmaI
BamHI
SmaI
PflaB
lacI
Phsp10
flaB

Figure 2

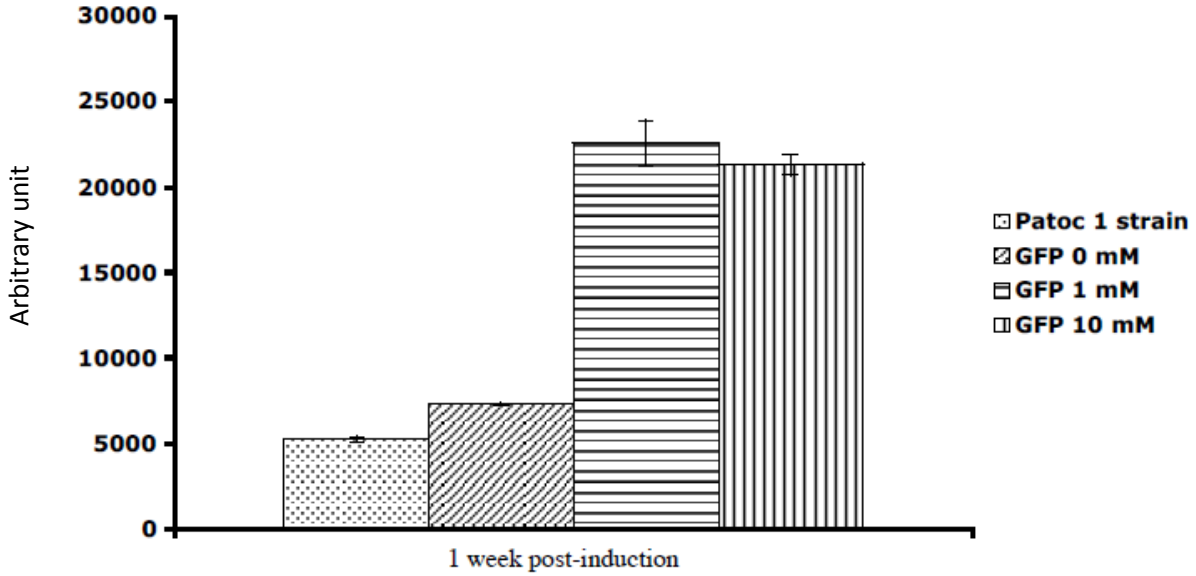

30000
25000
20000
15000
10000
5000
0
Arbitrary unit
1 week post-induction
Patoc 1 strain
GFP 0 mM
GFP 1 mM
GFP 10 mM

Figure 3

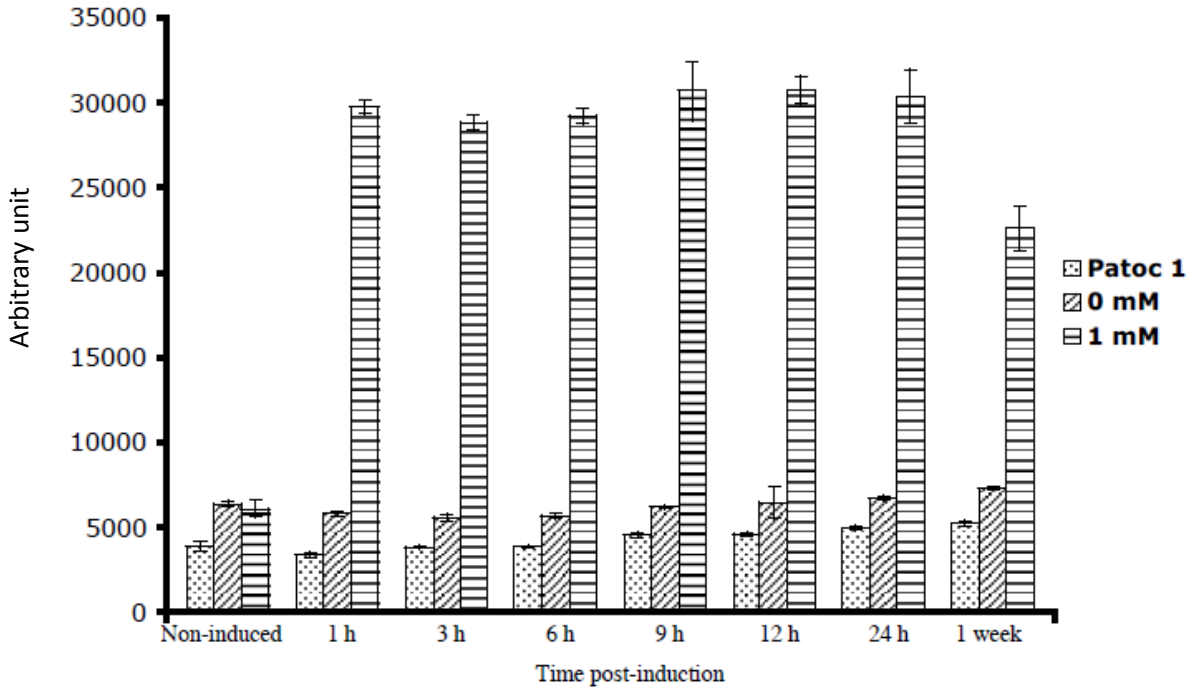

35000
30000
25000
20000
15000
10000
5000
0
Arbitrary unit
Non-induced
1 h
3 h
6 h
9 h
12 h
24 h
1 week
Time post-induction
Patoc 1
0 mM
1 mM

Figure 4

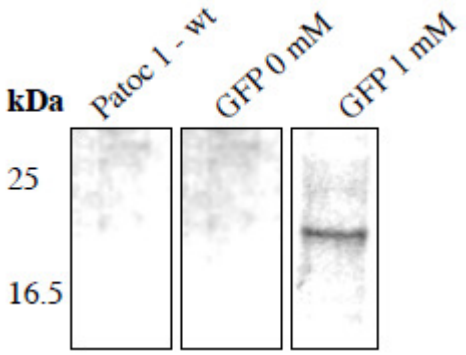
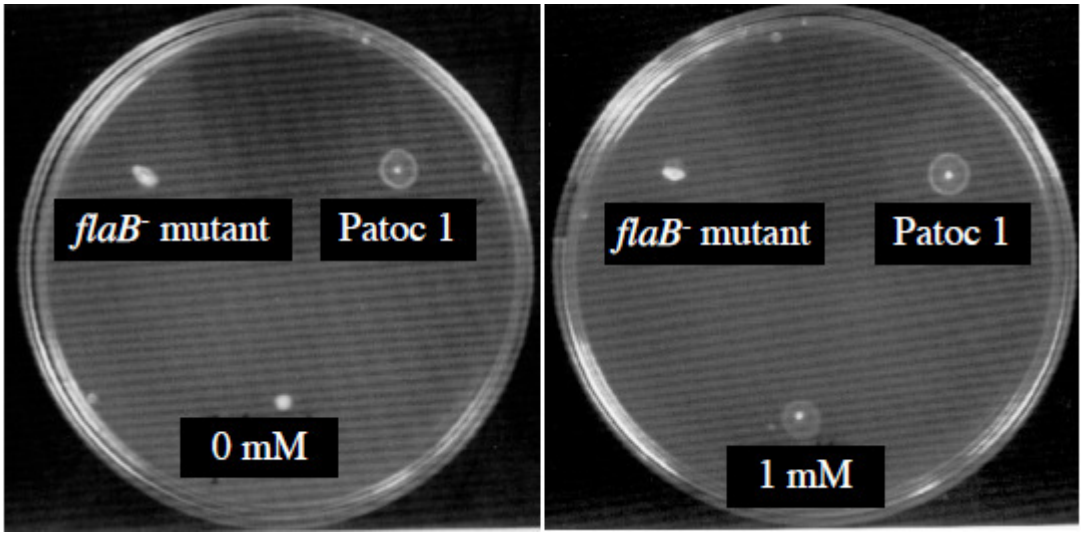


Figure 5



6. ARTIGO 4

DISTRIBUTION OF THE LEPTOSPIRAL IMMUNOGLOBULIN-LIKE (LIG) GENES IN PATHOGENIC *LEPTOSPIRA* SPP. AND APPLICATION OF *LIGB* TO TYPING LEPTOSPIRAL ISOLATES

(Artigo formatado segundo as normas do periódico *Journal of Medical Microbiology*)

Distribution of the Leptospiral immunoglobulin-like (Lig) genes in pathogenic *Leptospira spp.* and application of *ligB* to typing leptospiral isolates

Gustavo M. Cerqueira[1,2], Alan J. A. McBride[3], Mathieu Picardeau[2], Samuel G. Ribeiro,[1] Ângela N. Moreira[1], Viviane Morel[2], Mitermayer G. Reis[3], Albert I. Ko[3,4], Odir A. Dellagostin[1]

[1]Universidade Federal de Pelotas, Centro de Biotecnologia, P.O. Box 354, 96010-900, Pelotas, RS, Brazil;
[2] Institut Pasteur, Unité de Biologie des Spirochètes, 75015, Paris, France;
[3]Centro de Pesquisa Gonçalo Moniz, Fundação Oswaldo Cruz, 40296-710, Salvador, BA, Brazil;
[4]Division of International Medicine and Infectious Disease, Weill Medical College of Cornell University, New York, NY, USA.

Correspondence: Gustavo M. Cerqueira. Tel.: +55 53 3275-7587, Fax: +55 53 3275-7555. E-mail: gcenbiot@ufpel.edu.br

Running title: Distribution and use of *lig* genes

Subject category: Diagnostics, typing and identification

The GenBank accession numbers of the *lig* gene sequences from the *Leptospira* strains described in this study are EU938447 to EU938521.

SUMMARY

The family of Leptospiral immunoglobulin-like (Lig) genes includes *ligA*, *ligB*, and *ligC*. We demonstrate by PCR the presence of the *lig* genes among serovars from a collection of leptospiral strains and clinical isolates. While *ligA* and *ligC* genes appear to be present in a limited number of pathogenic serovars, the *ligB* gene is ubiquitously distributed among all pathogenic strains. None of *lig* genes were detected among intermediate or saprophytic *Leptospira* species. We also show that a short specific PCR fragment of *ligB* can be used to properly identify pathogenic strains of *Leptospira*. These findings demonstrate that *ligB* is widely present among pathogenic *Leptospira* spp. and may be useful for their reliable identification and classification.

Abbreviations: FAFLP - fluorescent amplified fragment length polymorphism; Lig – leptospiral immunoglobulin-like; VNTR – variable number tandem repeat.

INTRODUCTION

Leptospirosis is a re-emerging zoonotic disease caused by *Leptospira* spp. that are transmitted to humans through direct or indirect contact with contaminated urine from a reservoir host, usually rats or other rodents (Faine *et al*., 1999). DNA-DNA hybridization studies have identified 19 *Leptospira* genomospecies to date (Levett 2001; Levett et al 2006; Matthias *et al*., 2008; Slack *et al*., 2008). Among these *L. interrogans*, *L. borgpetersenii*, *L. santarosai*, *L. noguchii*, *L. weilii*, *L. kirschneri* and *L. alexanderi* are considered to be the main agents of leptospirosis (Levett *et al*., 2006). Serologic methods have identified >300 serovars of which more than 200 are considered pathogenic (Faine *et al*., 1999; Levett, 2001; Bharti *et al*., 2003).

The *lig* genes, *ligA*, *ligB* and *ligC*, code for virulence determinants in pathogenic strains (Palaniappan *et al*., 2002; Matsunaga *et al*., 2003; Choy *et al*., 2007; Lin *et al*., 2007). The Lig proteins were identified as markers for the early diagnosis of leptospirosis (Croda *et al*., 2007; Srimanote *et al*., 2008) and as potential vaccine candidates (Koizumi *et al*., 2004; Palaniappan *et al*., 2006; Silva *et al*., 2007; Faisal *et al*., 2008; Yan *et al*., 2008). Previously, we determined that the *lig* genes are highly conserved (70-99%) in virulent pathogenic *Leptospira* isolates (McBride et al 2008). *ligB* was present in all isolates, *ligA* was limited to *L. interrogans* and *L. kirschneri* strains and *ligC* was a pseudogene in several isolates.

Molecular tools employed for the classification of *Leptospira* spp. include PFGE (Herrmann *et al*., 1992; Galloway & Levett, 2008), restriction endonuclease assay (Brown *et al*., 1997), RFLP (Barocchi *et al*., 2001), arbitrarily primed PCR (Perolat *et al*., 1994), FAFLP (Vijayachari *et al*., 2004) and VNTR (Majed *et al*., 2005; Salaün *et al*., 2006; Slack et al 2005). However, these techniques lack reproducibility or have low sensitivity or specificity (Levett, 2006). 16S rDNA sequencing has been used in phylogenetic analyses (Hookey *et al*., 1993) but the *rrs* genes exhibit a low degree of polymorphism, limiting their usefulness in typing. A limitation of the previous investigation of the *lig* genes was the small number of isolates studied (McBride et al 2008). To this end we proposed to determine the presence of the *lig* genes in an expanded collection of strains using a PCR-based assay. Given the lack of a definitive molecular-based method for typing pathogenic leptospires we describe the

use of *ligB* sequencing for the molecular characterization of pathogenic *Leptospira* isolates.

METHODS

Bacterial strains and culture conditions. Reference and clinical strains belonging to nine species and including 39 serovars were obtained from the collections maintained at the Gonçalo Moniz Research Centre, Salvador, Brazil and the National Reference Centre for leptospirosis at the Institut Pasteur, Paris, France. Clinical strains were isolated from both humans and animals and from diverse geographical regions, including Brazil, Russia, Croatia, and Guadeloupe (Majed *et al*., 2005; Silva *et al*., 2008). All strains were cultured at 30 ºC in liquid Ellinghausen-McCullough-Johnson-Harris modified tween 80-bovine albumin medium (Ellinghausen and McCullough, 1965; Johnson and Harris, 1967). The microscopic agglutination test (MAT) was carried out using the standard method for putative serogroup determination (Cole 1973; Levett et al 2003).

Oligonucleotide design. Primers were designed using Vector NTI® 10 software (Invitrogen). *lig* gene sequences deposited in GenBank were aligned, conserved regions were identified and degenerate primers designed. Fragments from each of the *lig* genes were amplified and sequenced with the following primers: *ligA*: PSAF 5'-CKGAWCTTGTRACYTGGARKTCYTC, PSAR 5'-TTGTTAATGTTTTCATRTTAYGGC (*ligA*); *ligB*: PSBF 5'-ACWRVHVHRGYWDCCTGGTCYTCTTC, PSBR 5'-TARRHDGCYBTAATATYCGRWYYTCCTAA; *ligC*: PSCF 5'-GAGAAATAYAATCTCCTTCTTCCGG, PSCR 5'-CCTRTTCGTGTTGGARGAATTCC.

DNA manipulation. Genomic DNA was extracted using the GFX Genomic Blood DNA Purification Kit following the protocol for Gram-negative bacteria recommended by the manufacturer (GE Healthcare). PCR amplification was performed using *Taq* DNA polymerase (Invitrogen) and the following cycling conditions: one denaturing cycle at 94 ºC for 2 min; 35 cycles of denaturing at 94 ºC for 30 s, annealing at 54 ºC

for 30 s and elongation at 72 °C for 45 s; and a final elongation at 72 °C for 10 min. The amplified products were analyzed by 1 % agarose gel electrophoresis.

Sequencing. PCR products were purified by the use of GFX PCR DNA and Gel Band purification kit according to the manufacturer's instructions (GE Healthcare). The sequencing was performed using a MegaBACE 500 DNA sequencer (GE Healthcare) and the Dynamic ET-terminator technology. The assembled sequences were analysed by BLAST alignment (http://www.ncbi.nlm.nih.gov/BLAST) against the available *lig* gene sequences in GenBank. The *lig* sequences were aligned by the use of AlignX® software (Invitrogen).

Phylogenetic analysis. The *ligB* gene sequences from 47 pathogenic strains (Table 1) were used to assemble a phylogenetic tree with the MEGA 4 software (Tamura *et al*., 2007). 16S rRNA gene sequences were obtained from GenBank and aligned as described (Table 1). One thousand bootstrap replications were used to provide confidence in the nodes. The trees were constructed by the Neighbour-joining method using the Jukes-Cantor model (Tamura *et al*., 2007). The synonymous/non-synonymous data were calculated using MEGA 4.1 beta software. *rpoB* sequences used for comparison were obtained from GenBank, and were previously deposited under the accession numbers DQ296129 to DQ296147 (La scola *et al*., 2006).

Southern blotting. A total of 3 µg of genomic DNA was digested with 20 units of *Bam*HI (Invitrogen) and separated by agarose gel electrophoresis. DNA was transferred from the gel to a positively charged Hybond-N nylon membrane (GE Healthcare) with a vacuum blotter (Bio-Rad). Probes to each of the *lig* genes were based on pooled PCR products amplified using the primers described and labelled using the ECL Direct Nucleic Acid Labelling and Detection System (GE Healthcare). Prehybridization was carried out at 42 °C for 1 h in hybridization buffer supplemented with 0.5 M NaCl and 5 % blocking agent. Hybridization was carried out overnight at 42 °C in roller bottles. Following hybridization, the membrane was washed twice for 10 min at 55 °C in wash solution (0.4 % SDS, 0.5× SSC). Finally, the membrane was washed twice in 2× SSC, 5 min per wash at room temperature. After incubation with ECL detection reagents, hybridization products were detected by exposure of the membrane to Hyperfilm ECL X-ray film (GE Healthcare).

RESULTS AND DISCUSSION

Distribution of the *lig* genes in *Leptospira* spp.

In our previous study, pair-wise alignment of the *lig* genes allowed the identification of highly conserved regions within the *lig* genes (interspecies identity ranged from 68 to 99 %) (McBride *et al*., 2008). Based on these regions, primers were designed to successfully amplify *lig* gene fragments from the *Leptospira* strains described in this study (Table 1). For *ligA* the primers spanned nucleotides 3482 – 3693 at the C-terminus, the *ligB* primers spanned nucleotides 2125 – 2504 within the non-identical region and for *ligC* the primers spanned nucleotides 1487 – 1734 (Fig. 1). The expected sizes of the amplicons were 211 bp (*ligA*), 380 bp (*ligB*) and 248 bp (*ligC*). The PCR results indicated that *ligB* was conserved in the genome of 100 % (52/52) of the pathogenic strains tested (Table 1). Notably, *ligA* was limited to *L. interrogans* and *L. kirschneri* strains, and was only found in 26/44 isolates. As well as being present in certain *L. interrogans* and *L. kirschneri* strains, *ligC* was also detected in several *L. noguchii* and *L. weilii* strains (31/44 strains in total).

To confirm the negative PCR results as true negatives a Southern blot analysis was carried out (Table 1). The hybridization results corroborated the PCR assays findings. These results support previous studies that suggested the *lig* genes are only found in pathogenic strains and that of the three *lig* genes only *ligB* was conserved in all pathogenic *Leptospira* strains (Matsunaga *et al*. 2003; McBride *et al*. 2008). The findings presented herein add to the growing body of evidence that suggests the Lig proteins are essential virulence determinants in *Leptospira* spp. (Matsunaga *et al*., 2005; Choy *et al*., 2007, McBride *et al*., 2008). To ensure that the PCR products were not artefacts a selection of amplicons (see Table 1) were sequenced and analyzed with those *lig* gene sequences available in GenBank.

Sequence variability of the *lig* gene fragments

The *ligB* amplicons exhibited considerable DNA sequence polymorphism, particularly at the 5' and 3' ends of the 380 bp fragment. Therefore the *ligB* sequences were trimmed to remove these hypervariable regions and a 214 bp region (nucleotides 2236 – 2449, *L. interrogans* Fiocruz L1-130 strain) was identified that exhibited a high level of conservation. The overall level of pair-wise DNA sequence variability

was determined to be 21.2 ± 3.9% and 20.6 ± 3.8% at amino acid level for the *ligB* amplicon (Fig. 2). This DNA fragment demonstrated some interspecies polymorphism, but it was not significant (Fig. 2). The mean pair-wise DNA sequence variability was 0.8 ± 0.4, 3.7 ± 1.5, 0, 1.2 ± 0.9, 0.9 ± 0.9 and 0% among the *L. interrogans*, *L. kirschneri*, *L. noguchii*, *L. borgpetersenii*, *L. santarosai* and *L. weilii* strains, respectively (0.8 ± 0.4, 3.7 ± 1.5, 0, 1.7 ± 1.2, 0.9 ± 0.9 and 0% at the amino acid level, respectively). Furthermore, 17 *ligB* different orthologs were identified among the 48 *Leptospira* strains that contained one or more base substitutions within the amplified region.

The *ligA* amplicons demonstrated a mean pair-wise variability of 21.5 ± 2.4% among *L. interrogans* strains and 0.8 ± 0.8% among *L. kirschneri* strains (25 ± 3.9 and 0% at the amino acid level, respectively). The overall mean pair-wise DNA sequence variability of the *ligA* amplicons was 22.2 ± 2.7% (26.8 ± 4.4% at the amino acid level) (Fig. 2). The alignment of the *ligA* sequences revealed the presence of indels in some of the *L. interrogans* sequences that corresponded to the loss of an amino acid codon. The *ligC* gene exhibited a mean pair-wise variability of 1.9 ± 1.7 and 0% (1.9 ± 1.8 and 0% at the amino acid level, respectively) among the *L. interrogans and L. kirschneri* strains, respectively. The overall mean pair-wise variability was 4.4 ± 2.8% (4.4 ± 2.8% at the amino acid level) (Fig. 2).

The *lig* genes encode for an important family of outer membrane proteins that are characterized by the presence of immunoglobulin-like domains (Palaniappan *et al*., 2002; Matsunaga *et al*., 2003) and are potential virulence determinants of *Leptospira* spp. (Choy *et al*., 2007; Lin *et al*., 2007). These proteins are surface-exposed and are up-regulated within mammalian hosts (Matsunaga *et al*., 2005; Choy *et al*., 2007). Previous studies have demonstrated their usefulness as markers for diagnosis of leptospirosis (Croda *et al*., 2007; Palaniappan *et al*., 2004; Palaniappan *et al*., 2005; Srimanote *et al*., 2008) and as potential vaccine candidates (Koizumi *et al*., 2004; Palaniappan *et al*., 2006; Silva *et al*., 2007; Faisal *et al*., 2008a,b). More recently, their presence and conservation among virulent pathogenic strains of *Leptospira* spp. was confirmed (McBride *et al*., 2008). Of note, inactivation of *ligB* does not result in attenuation of virulence in animal models (Croda *et al*., 2008). This is probably due to functional redundancy of the Lig proteins as LigA was expressed in the LigB

knockout strain. The findings of this study confirm the ubiquitous nature of LigB in pathogenic *Leptospira* spp. and that LigA and LigC are not present in all strains.

Phylogenetic analysis of *ligB*

The relatedness of the 48 *ligB* 214 bp DNA sequences is presented in Fig. 3A, the *Leptospira* strains resolved into two distinct clusters. Those sequences from *L. interrogans*, *L. kirschneri* and *L. noguchii* grouped together in one cluster, while those from *L. borgpetersenii*, *L. santarosai* and *L. weilii* formed the second cluster. The clustering pattern is similar to the phylogenetic tree based on the full-length *ligB* sequences (McBride *et al.* 2008). The individual *Leptospira* species are easily determined based on the *ligB* internal sequence.

The *ligB* amplicon is situated within a region of the *ligB* gene that was found to be phylogenetically clonal based on a multiple-change-point (MCP) model in the majority of strains (McBride *et al.* 2008). Of the two strains that showed evidence of rearrangements (*L. interrogans* and *L. kirschneri*) the amplicon is located outside these recombination hotspots. The results demonstrate that the internal *ligB* sequence can be used to discriminate *Leptospira* to the species level. Within each major cluster there was evidence of further sub-clustering. For example, 3/5 of the *L. interrogans* serogroup Icterohaemorrhagiae strains clustered together, including serovars Copenhageni and Icterohaemorrhagiae. Within the *L. kirschneri* and *L. borgpetersenii* cluster various sub-clusters were identified but they did not correspond to the serogroups (Fig. 3A). However, there was insufficient discriminatory power to type the serovars beyond the species level. This is a similar situation as reported for the 16S *rrs* gene in *Leptospira* spp. (Morey et al., 2006).

The number of synonymous substitutions (dS) within the *ligB* amplicons was equal in number or higher than those of non-synonymous substitutions (dN) per site. The probability of the existence of recombination among the several *ligB* nucleotide sequences was not confirmed (overall $P = 1.00$) and the ratio between the non-synonymous (dN) and synonymous (dS) substitutions (dN/dS) was 0.34. This supports the hypothesis of sequence stability due to absence of positive selection over this *ligB* locus. Rejection of neutrality hypothesis (positive selection suggestive of recombination) in *ligB* was only seen in the *L. borgpetersenii* Poi and Veldrat

Batavia 46 strains where the dN/dS was 1.72 (P = 0.04). However, this does not preclude the use of *ligB* for species typing, as both belong to the same species. The G + C content of the several *ligB* loci ranged from 37.8 in *L. kirschneri* to 50 in *L. borgpetersenii* (data not shown).

Phylogenetic analysis of the 16S rDNA sequences

The phylogenetic tree based on the available 16S gene sequences (Table 1) is presented in Fig. 3B. The tree describes the relatedness for 36 sequences and the clustering pattern is similar to that described in previous studies (Levett et al. 2006, Haake et al. 2005). The strains clustered according to species: sequences from *L. interrogans*, *L. kirschneri*, *L. noguchii* and *L. santarosai* formed one cluster; while those from *L. borgpetersenii* and *L. weilii* formed a second cluster. The major difference between the predicted relatedness patterns is the clustering of the *L. santarosai* strains. In Fig. 3B these strains clustered with the *L. interrogans*, *L. kirschneri* and *L. noguchii* strains.

Traditionally, 16S rRNA genes sequences have been used for *Leptospira* species classification (Postic *et al*., 2000; Morey *et al*., 2006). However, this gene has few polymorphisms throughout its 1,500 bp length in *Leptospira* spp. (Janda *et al*., 2007). Efforts to identify new markers for species differentiation focused on the evaluation of partial *rpoB* (La Scola *et al*., 2006) and *wzy* (Wangroongsarb *et al*., 2007) polymerases, the gyrase subunit B, *gyrB* (Slack *et al*., 2006), the preprotein translocase *secY* (Victoria *et al*., 2008), and the genes encoding the surface proteins LipL32, LipL41 and OmpL1 (Haake *et al*., 2004; Ahmed *et al*., 2006). The main advantage of selecting housekeeping genes for classification is that the constant selection pressure over these genes in the genome. However, as is the case for the 16S genes, this is associated with a low accumulation of polymorphisms and hence a lower resolution power in terms of strain differentiation. Genes such as *rpoB* and *gyrB* offer the advantage of being shorter and more polymorphic. Recently, La Scola and colleagues (2006) described 3 nucleotides that accounted for the differences between the *L. kirschneri* serovar Cynopteri and *L. interrogans* serovar Canicola *rrs* genes. In addition, Morey and coworkers (2006) reported that the difference between *L. interrogans* and *L. kirschneri* type strains was due to only two nucleotides. This is consistent with descriptions of the high degree of conservation of the 16S rRNA

among other bacterial species (Janda *et al*., 2007). The *rpoB* was found to contain 51 polymorphisms over 600 bp when the Cynopteri and Canicola serovars were compared. In this study, the 214 bp *ligB* sequence contained 23 and 24 polymorphisms between the Cynopteri and Canicola serovars and the *L. interrogans* and *L. kirschneri* type strains, respectively.

The taxonomic analysis performed in this study demonstrated the discriminatory power of the *ligB* gene. We showed that *ligB* is a molecular marker that is able to differentiate the serovars into their respective species (Fig. 3). Recently, we showed that some *ligB* genes contain mosaic sites, but they were located at the carboxy-terminal end of the gene (McBride *et al*., 2008). Furthermore, some of the *ligB* domains were involved in the duplication events that led to the creation of *ligA*. In this study we specifically chose a region outside of the potential mosaic region and that did not include the domains involved in the gene duplication events. In conclusion, the *ligB* molecular typing scheme demonstrates several major advantages (*i*) the ability to differentiate strains to the species level; (*ii*) differentiation between pathogenic and non-pathogenic strains and (*iii*) the potential to be employed in MLST or MVLST analysis for identification of clonal derivation events during the seasonal epidemics and outbreaks associated with urban leptospirosis.

ACKNOWLEDGEMENTS

GMC was supported by CAPES foundation, Brazilian ministry of education. This work was supported by Bio-Manguinhos, Oswaldo Cruz Foundation (09224-7 and PDTIS RVR05), the Brazilian National Research Council (grants 01.06.0298.00 3773/2005, 420067/2005, 554788/2006, 473006/2006-5), the National Institutes of Health (5R01 AI052473, 2D43 TW00919), and the Institut Pasteur.

REFERENCES

Ahmed, N. , Devi, S. M., Valverde, M. A., Vijayachari, P., Machang'u, R. S., Ellis, W. A. & Hartskeerl, R. A. (2006). Multilocus sequence typing method for identification and genotypic classification of pathogenic *Leptospira* species. *Annals Clin. Microbiol. Antimic.* **5**, 28.

Barocchi, M. A., Ko, A. I., Ferrer, S. R., Faria, M. T., Reis, M. G. & Riley, L. W. (2001). Identification of new repetitive element in *Leptospira interrogans* serovar

Copenhageni and its application to PCR-based differentiation of *Leptospira* serogroups. *J. Clin. Microbiol.* **39**, 191-195.

Bulach, D. M., Zuerner, R. L., Wilson, P., Seemann, T., McGrath, A., Cullen, P. A., Davis, J., Johnson, M., Kuczek, E., Alt, D. P., Peterson-Burch, B., Coppel, R. L., Rood, J. L., Davies, J. K. & Adler, B. (2006). Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. *Proc. Nat. Acad. Sci.* **103**, 14560–14565.

Brenner, D. J., Kaufmann, A. F., Sulzer, K. R., Steigerwalt, A. G., Rogers, F. C. & Weyant, R. S. (1999). Further determination of DNA relatedness between serogroups and serovars in the family leptospiraceae with a proposal for *Leptospira alexanderi* sp. nov. and four new *Leptospira* genomospecies. *Int. J. Syst. Bacteriol.* **49**, 839-858.

Brown, P. D. & Levett, P. N. (1997). Differentiation of *Leptospira* species and serovars by PCR-restriction Endonuclease analysis, arbitrarily primed PCR and low-stringency PCR. *J. Med. Microbiol.* **46**, 173-181.

Choy, H. A., Kelley, M. M., Chen, T. L., Moller, A. K., Matsunaga, J. & Haake, D. A. (2007). Physiological osmotic induction of *Leptospira interrogans* adhesion: LigA and LigB bind extracellular matrix proteins and fibrinogen. *Infect. Immun.* **75**, 2441-2450.

Croda, J., Ramos, J. G., Matsunaga, J., Queiroz, A., Homma, A., Riley, L. W., Haake, D. A., Reis, M. G. & Ko, A. I. (2007). *Leptospira* immunoglobulin-like proteins as a serodiagnostic marker for acute leptospirosis. *J. Clin. Microbiol.* **45**, 1528-1534.

Croda J., Figueira C. P., Wunder E. A. Jr, Santos C. S., Reis M. G., Ko A. I. & Picardeau M. Targeted mutagenesis in pathogenic Leptospira species: disruption of the LigB gene does not affect virulence in animal models of leptospirosis. *Infect. Immun.*, **76**(12), 5826-5833.

Faine, S. B., Adler, B., Bolin, C. & Perolat, P. (1999). *Leptospira and leptospirosis.* Melbourne, Australia: MediSci.

(a)Faisal, S. M., Yan, W., Chen, C. S., Palaniappan, R. U., McDonough, S. P. & Chang, Y. F. (2008). Evaluation of protective immunity of *Leptospira* immunoglobulin like protein A (LigA) DNA vaccine against challenge in hamsters. *Vaccine* **26**, 277-287.

(b)Faisal, S. M., Yan, W., McDonough, S. P. & Chang, Y. F. (2008). Leptospira immunoglobulin-like protein A variable region (LigAvar) incorporated in liposomes and PLGA microspheres produces a robust immune response correlating to protective immunity. *Vaccine in press.*

Haake, D. A., Suchard, M. A., Kelley, M. M., Dundoo, M., Alt, D. P. & Zuerner, R. L. (2004). Molecular Evolution and Mosaicism of Leptospiral Outer Membrane Proteins Involves Horizontal DNA Transfer. *J. Bacteriol.* **186**, 2818-2828.

Hamburger, Z. A., Brown, M. S., Isberg, R. R. & Bjorkman, P. J. (1999). Crystal structure of invasin: a bacterial integrin-binding protein. *Science* **286**, 291-295.

Herrmann, J. L., Bellenger, E., Perolat, P., Baranton, G. & Girons, I. S. (1992). Pulsed field gel electrophoresis of NotI digests of leptospiral DNA: a new rapid method of serovars identification. *J. Clin. Microbiol.* **30**, 1696-1702.

Hookey, J. V., Bryden, J. & Gatehouse, L. (1993). The use of 16S rDNA sequence analysis to investigate the phylogeny of Leptospiraceae and related spirochaetes. *J. Gen. Microbiol.* **139**(11), 2585-2590.

Janda, J. M. & Abbott S. L.. (2007). 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. *J. Clin. Microbiol.* **45**, 2761- 2764.

Koizumi, N. & Watanabe, H. (2004). Leptospiral immunoglobulin-like proteins elicit protective immunity. *Vaccine* **22**, 1545-1552.

La Scola, B., Bui, L. T. M., Baranton, G., Khamis, A. & Raoult, D. (2006). Partial *rpo*B gene sequencing for identification of *Leptospira* species. *FEMS Microbiol. Lett.* **263**, 142-147.

Levett, P. N. (2001). Leptospirosis. *Clin. Microbiol. Rev.* **14**, 296–326.

Levett, P. N., Morey, R. E., Galloway, R. L. & Steigerwalt, A. G. (2006). *Leptospira broomii* sp. nov., isolated from humans with Leptospirosis. *Int. J. Syst. Evol. Microbiol.* **56**, 671-673.

Lin, Y. P. & Chang, Y. F. (2007). A domain of the *Leptospira* LigB contributes to high affinity binding of fibronectin. *Biochem. Biophys. Res. Commun.* **362**, 443-448.

Luo, Y., Frey, E. A., Pfuetzner, R. A., Creagh, A. L., Knoechel, D. G., Haynes, C. A., Finlay, B. B. & Strynadka, N. C. (2000). Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* **405**, 1073-1077.

Majed, Z., Bellenger, E., Postic, D., Pourcel, C., Baranton, G. & Picardeau, M. (2005). Identification of variable-number tandem-repeat loci in *Leptospira interrogans* sensu stricto. *J. Clin. Microbiol.* **43**, 539-545.

Matsunaga, J., Barocchi, M. A., Croda, J., Young, T. A., Sanchez, Y., Siqueira, I., Bolin, C. A., Reis, M. G., Riley, L. W., Haake, D. A. & Ko, A. I. (2003). Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. *Mol. Microbiol.* **49**, 929-945.

Matsunaga, J., Sanchez, Y., Xu, X. & Haake, D. A. (2005). Osmolarity, a key environmental signal controlling expression of leptospiral proteins LigA and LigB and the extracellular release of LigA. *Infect. Immun.* **73**, 70-78.

Matthias, M. A., Ricaldi, J. N., Cespedes, M., Diaz, M. M., Galloway, R. L., Saito, M., Steigerwalt A. G., Patra, K. P, Ore, C. V., Gotuzzo, E., Gilman, R. H., Levett, P. N. & Vinetz, J. M. (2008). Human leptospirosis caused by a new, antigenically unique *Leptospira* associated with a *Rattus* species reservoir in the Peruvian Amazon. *PLoS Negl. Trop. Dis.* **2**, e213.

McBride, A. J. A., Cerqueira, G. M., Suchard, M. A., Moreira, A. N., Zuerner, R. L., Reis, M. G., Haake, D. A., Ko, A. I. & Dellagostin, O. A. (2008). Genetic diversity of the Leptospiral immunoglobulin-like (Lig) genes in pathogenic Leptospira spp. *Infect. Genet. Evol. in press*.

Morey, R. E., Galloway, R. L., Bragg, S. L., Steigerwalt, A. G., Mayer, L. W. & Levett, P. N. (2006). Species-specific identification of Leptospiraceae by 16S rRNA gene sequencing. *J. Clin. Microbiol.* **44**, 3510–3516.

Nascimento, A. L., Ko, A. I., Martins, E. A., Monteiro-Vitorello, C. B., Ho, P. L., Haake, D. A., Verjovski-Almeida, S., Hartskeerl, R. A., Marques, M. V., Oliveira, M. C., Menck, C. F., Leite, L. C., Carrer, H., Coutinho, L. L., Degrave, W. M., Dellagostin, O. A., El-Dorry, H., Ferro, E. S., Ferro, M. I., Furlan, L. R., Gamberini, M., Giglioti, E. A., Góes-Neto, A., Goldman, G. H., Goldman, M. H., Harakava, R., Jerônimo, S. M., Junqueira-de-Azevedo, I. L., Kimura, E. T., Kuramae, E. E., Lemos, E. G., Lemos, M. V., Marino, C. L., Nunes, L. R., de Oliveira, R. C., Pereira, G. G., Reis, M. S., Schriefer, A., Siqueira, W. J., Sommer, P., Tsai, S. M., Simpson, A. J., Ferro, J. A., Camargo, L. E., Kitajima, J. P., Setubal, J. C. & Van Sluys, M. A. (2004). Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *J. Bacteriol.* **186**, 2164- 2172.

438 **Nightingale, K. K., Lyles, K., Ayodele, M., Jalan, P., Nielsen, R. & Wiedmann, M.**
439 **(2006).** Novel method to identify source-associated phylogenetic clustering shows
440 that Listeria monocytogenes includes niche-adapted clonal groups with distinct
441 ecological preferences. *J. Clin. Microbiol.* **44**, 3742-3751.
442 **Palaniappan, R. U., Chang, Y. F., Jusuf, S. S., Artiushin, S., Timoney, J. F.,**
443 **McDonough, S. P., Barr, S. C., Divers, T. J., Simpson, K. W., McDonough, P. L.**
444 **& Mohammed, H. O. (2002).** Cloning and molecular characterization of an
445 immunogenic LigA protein of *Leptospira interrogans*. *Infect. Immun.* **70**, 5924-5930.
446 **Palaniappan, R. U., Chang, Y. F., Hassan, F., McDonough, S. P., Pough, M.,**
447 **Barr, S. C., Simpson, K. W., Mohammed, H. O., Shin, S., McDonough, P.,**
448 **Zuerner, R. L., Qu, J. & Roe, B. (2004).** Expression of leptospiral immunoglobulin-
449 like protein by *Leptospira interrogans* and evaluation of its diagnostic potential in a
450 kinetic ELISA. *J Med Microbiol* **53**, 975-984.
451 **Palaniappan, R. U., Chang, Y. F., Chang, C. F., Pan, M. J., Yang, C.W.,**
452 **Harpending, P., McDonough, S. P., Dubovi, E., Divers, T., Qu, J. & Roe, B.**
453 **(2005).** Evaluation of *lig*-based conventional and real time PCR for the detection of
454 pathogenic leptospires. *Mol. Cell Probes* **19**, 111-117.
455 **Palaniappan, R. U., McDonough, S. P., Divers, T. J., Chen, C. S., Pan, M. J.,**
456 **Matsumoto, M. & Chang, Y. F. (2006).** Immunoprotection of recombinant leptospiral
457 immunoglobulin-like protein A against *Leptospira interrogans* serovar Pomona
458 infection. *Infect. Immun.* **74**, 1745-1750.
459 **Perolat, P., Merien, F., Ellis, W. A. & Baranton, G. (1994).** Characterization of
460 *Leptospira* isolates from serovars hardjo by Ribotyping, arbitrarily primed PCR, and
461 mapped restriction site polymorphisms. *J. Clin. Microbiol.* **32**, 1949-1957.
462 **Postic, D., Riquelme-Sertour, N., Merien, F., Perolat, P. & Baranton, G. (2000).**
463 Interest of partial 16S rDNA gene sequences to resolve heterogeneities between
464 *Leptospira* collections: application to *L. meyeri*. *Res. Microbiol.* **151**(5), 333-341.
465 **Qiu, W. G., Bruno, J. F., McCai,g W. D., Xu, Y., Livey, I., Schriefer, M. E. & Luft,**
466 **B. J. (2008).** Wide distribution of a high-virulence Borrelia burgdorferi clone in Europe
467 and North America. *Emerg. Infect. Dis.* **14**(7), 1097-1104.
468 **Ren, S. X., Fu, G., Jiang, X. G., Zeng, R., Miao, Y. G., Xu, H., Zhang, Y. X., Xiong,**
469 **H., Lu, G., Lu, L. F., Jiang, H. Q., Jia, J., Tu, Y. F., Jiang, J. X., Gu, W. Y., Zhang,**
470 **Y. Q., Cai, Z., Sheng, H. H., Yin, H. F., Zhang, Y., Zhu, G. F., Wan, M., Huang, H.**
471 **L., Qian, Z., Wang, S. Y., Ma, W., Yao, Z. J., Shen, Y., Qiang, B. Q., Xia, Q. C.,**

Guo, X. K., Danchin, A., Saint Girons, I., Somerville, R. L., Wen, Y. M., Shi, M. H., Chen, Z., Xu, J. G. & Zhao, G.P. (2003). Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature* **422**, 888-893.

Salaün, L., Mérien, F., Gurianova, S., Baranton, G. & Picardeau, M. (2006). Application of multilocus variable-number tandem-repeat analysis for molecular typing of the agent of leptospirosis. *J. Clin. Microbiol.* **44**, 3954-62.

Silva, E. F., Brod, C. S., Cerqueira, G. M., Bourscheidt, D., Seyffert, N., Queiroz, A., Santos, C. S., Ko, A. I. & Dellagostin, O. A. (2007). Isolation of *Leptospira noguchii* from sheep. *Vet. Microbiol.* **121**(1-2), 144-149.

Silva, E. F., Medeiros, M. A., McBride, A. J., Matsunaga, J., Esteves, G. S., Ramos, J. G., Santos, C. S., Croda, J., Homma, A., Dellagostin, O. A., Haake, D. A., Reis, M. G. & Ko, A. I. (2007). The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. *Vaccine* **25**, 6277-6286.

Silva, E. F., Santos, C. S., Athanazio, D. A., Seyffert, N., Seixas, F. K., Cerqueira, G. M., Fagundes, M. Q., Brod, C. S., Reis, M. G., Dellagostin, O. A. & Ko, A. I. (2008). Characterization of virulence of *Leptospira* isolates in a hamster model. *Vaccine* **26**(31), 3892-6.

Slack, A. T., Symonds, M. L., Dohnt, M. F. & Smythe, L. D. (2006). Identification of pathogenic *Leptospira* species by conventional or real-time PCR and sequencing of the DNA gyrase subunit B encoding gene. *BMC Microbiol.* **6**, 95.

Srimanote, P., Wongdeethai, N., Jieanampunkul, P., Samonkiert, S., Leepiyasakulchai, C., Kalambaheti, T. & Prachayasittikul, V. (2008). Recombinant LigA for leptospirosis diagnosis and *ligA* among the *Leptospira* spp. clinical isolates. *J. Microbiol. Methods* **72**, 73-81.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596-1599.

Victoria, B., Ahmed, A., Zurner, R. L., Ahmed, N., Bulach, D. M., Quinteiro, J. & Hartskeerl, R. A. (2008). Conservation of the S10-spc-alpha locus within otherwise highly plastic genomes provides phylogenetic insight into the genus *Leptospira*. *PLoS ONE* **3**(7), e2752.

Vijayachari, P., Ahmed, N., Sugunan, A.P., Ghousunnisa, S., Rao, K. R., Hasnain, S. E. & Sehgal, S. C. (2004). Use of Fluorescent Amplified Fragment Length Polymorphism for Molecular Epidemiology of Leptospirosis In India. *J. Clin. Microbiol.* **42**, 3575- 3580.

Wangroongsarb, P., Chanket, T., Gunlabun, K., Long, do H., Satheanmethakul, P., Jetanadee, S., Thaipadungpanit, J., Wuthiekanun, V., Peacock, S. J., Blacksell, S. D., Smythe, L. D., Bulach, D. M. & Kalambaheti, T. (2007). Molecular typing of *Leptospira* spp. based on putative O-antigen polymerase gene (wzy), the benefit over 16S rRNA gene sequence. *FEMS Microbiol. Lett.* **271**, 170-179.

Wiedmann, M., Bruce, J. L., Keating, C., Johnson, A. E., McDonough, P. L. & Batt, C. A. Ribotypes and virulence gene polymorphisms suggest three distinct Listeria monocytogenes lineages with differences in pathogenic potential. *Infect. Immun.*, **65**(7), 2707-2716.

Yasuda, B. H., Steigerwalt, A. G., Sulzer, L. R., Kauhnann, A. F., Rogers, F. & Brenner, D. J. (1987). Deoxyribonucleic acid relatedness between serogroups and serovars in the family Leptospiraceae with proposals for seven new *Leptospira* species. *Int. J. Syst. Bacteriol.* **37**, 407-415.

524 **Legends to figures**

525 Fig. 1. Diagram of the main features of (A) *ligA*, (B) *ligB* and (C) *ligC* genes. The solid
526 black regions indicate the exact location of the target PCR amplicon. In (C) the PCR
527 product included the stop codon (X) present in *ligC* from the *L. interrogans* Fiocruz
528 L1-130 genome. The identical domains (green), the non-identical domains of *ligA*
529 (red) and *ligB*, (blue) and the carboxy-terminal domains of *ligB* (hatched) and *ligC*
530 (boxed) are indicated.
531

532 Fig. 2. Comparison of the variability of the DNA sequences from the 16S (*rrs*), *rpoB*,
533 *ligA*, *ligB*, and *ligC* genes from *Leptospira* spp., the error bars indicate standard
534 deviation. The number of individual sequences used for the determination of
535 sequence variability is indicated (n). The presence (+) and absence (-) of each gene
536 in pathogenic (P), intermediate (I), and saprophyte (S) strains is shown. The
537 nucleotide positions used during the alignment analysis were: 75 – 1255 (*rrs*), 1891 –
538 2462 (*rpoB*), 3482 – 3693 (*ligA*), 2236 – 2449 (*ligB*) and 1487 – 1734 (*ligC*).
539

540 Fig. 3. Unrooted phylogenetic trees were constructed from (A) the *ligB* (214 bp) and
541 (B) the 16S gene (1181 bp). The bootstrap consensus values are indicated. Asterisks
542 indicate serogroups instead of serovars.

543 **TABLES**

Table 1. Distribution of *lig* genes.

	Species	Serovar	Strain	16S[†]	PCR		
					ligA	*ligB*	*ligC*
Pathogens	*L. interrogans*	Australis	Ballico	+ (FJ154556)	+	+	+*
		Autumnalis	Akiyami A	+ (FJ154543)	+	+	+
		Bataviae	Van Tienen	+ (FJ154566)	+	+	+
		Bratislava	Jez Bratislava	+ (FJ154547)	+	+	+
		Canicola	Hond Utrech IV	+ (FJ154561)	+	+	+*
		Canicola	Kito	+	+‡	+‡	+‡
		Canicola	Mex 1	+	+‡	+‡	+*‡
		Copenhageni	Fiocruz L1-130	+ (AY461869)	+‡	+‡	+‡
		Copenhageni	M 20	+ (FJ154542)	+	+	+
		Hardjo-prajitno	Hardjoprajitno	+ (FJ154553)	+	+	+
		Hebdomadis	Hebdomadis	+ (FJ154551)	+	+	+
		Icterohaemorrhagiae	RGA[§]	+ (FJ154549)	+	+	+
		Kennewicki	LT 1026	+ (FJ154571)	+	+	+
		Lai	56601	(AY461870)	-	+	+
		Lai	Lai	+	+	+	+
		Manilae	LT 398	+ (FJ154545)	+	+	+
		Muenchen	Munchen C90	+ (FJ154565)	+	+	+
		Pomona	PO-06-047	+	+‡	+‡	+‡
		Pomona	Pomona	+ (FJ154544)	+	+	+
		Wolffi	3705	+ (FJ154558)	+	+	+
	L. kirschneri	Cynopteri	3522C[§]	+ (FJ154546)	+*	+	+*
		Djatzi	HS 26	+	+*	+	+
		Erinaceiauriti	Erinaceus auritus 670	+ (FJ154560)	+*	+	+
		Grippotyphosa	2.002.297[¶]	+	ND	+*	ND
		Grippotyphosa	2.002.306[¶]	+	ND	+*	ND
		Grippotyphosa	2000.11.449[¶]	+	ND	+*	ND
		Grippotyphosa	RM52	+	+‡	+‡	+‡

		Kambale	Kambale	+ (FJ154562)	+	+	+
		Mozdok	5621	+ (FJ154559)	+	+	+
		Ramisi	Musa	+ (FJ154573)	+*	+	+*
	L. borgpetersenii	ND	2E02	+	-*‡	+*‡	-
		Ceylonica	Piyasena	+ (FJ154596)	-	+	-
		Istrica	M 18¶	+	ND	+*	ND
		Javanica	Veldrat Batavia 46§	+ (FJ154600)	-	+	-
		ND	2002.10.110¶	+	ND	+*	ND
		Mini	Sari	+ (FJ154592)	-	+*	-
		Poi	Poi	+ (FJ154597)	-	+	-
		Hardjo	L550	NC008508	-	+	-
		Hardjo	JB197	NC008510	-	+	-
		Tarassovi	Perepelicin	+ (FJ154595)	-	+*	-
	L. noguchii	Bataviae‖	Cascata	+ (EU349495)	-‡	+‡	-‡
		Orleans	LSU 2580	+ (FJ154588)	-	+	-
		Panama	CZ214 K§	+ (FJ154582)	-	+*	+
	L. weilii	Hebdomadis‖	Eco-Challenge	+ (AY034037)	-‡	+‡	+‡
		Celledoni	Celledoni§	+ (FJ154580)	-	+	-
		Coxi	Cox	+	-	+*	+*
		Vughia	LT 89-68	+ (FJ154590)	-	+	+
		ND	2007.025.92¶	+	ND	+	ND
	L. santarosai	Alexi	HS 616	+ (FJ154585)	-	+	-
		Shermani	LT 821§	+(AY631883)	ND	+	ND
		Trinidad	TRVL 34056	+ (FJ154598)	-	+*	-
		ND	2008.010.55¶	+	ND	+	ND
Intermediates	*L. fainei*	Hurstbridge	But 6§	+ (FJ154578)	-	-	-
	L. inadai	Lyme	10§	+	ND	-	ND
Saprophytes	*L. meyeri*	Semaranga	Veldrat Semarang 173§	+	-	-	-
	L. biflexa	Semaranga	Patoc 1§	+	-‡	-‡	-‡

* PCR products not sequenced.

‡ Confirmed by Southern blot analysis.

ND - Not determined.
§ Type strain.
¶ Clinical isolate.
† Internal PCR control (Postic *et al*., 2000).
ǂ Serogroup.
in brackets: accession numbers

FIGURES

Figure 1

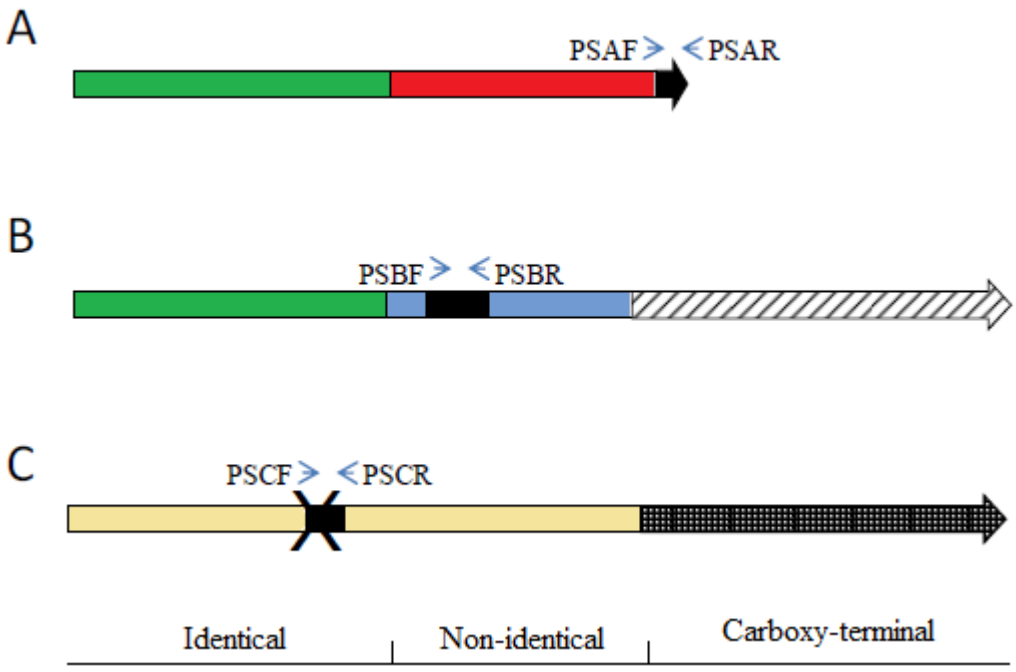


Figure 2

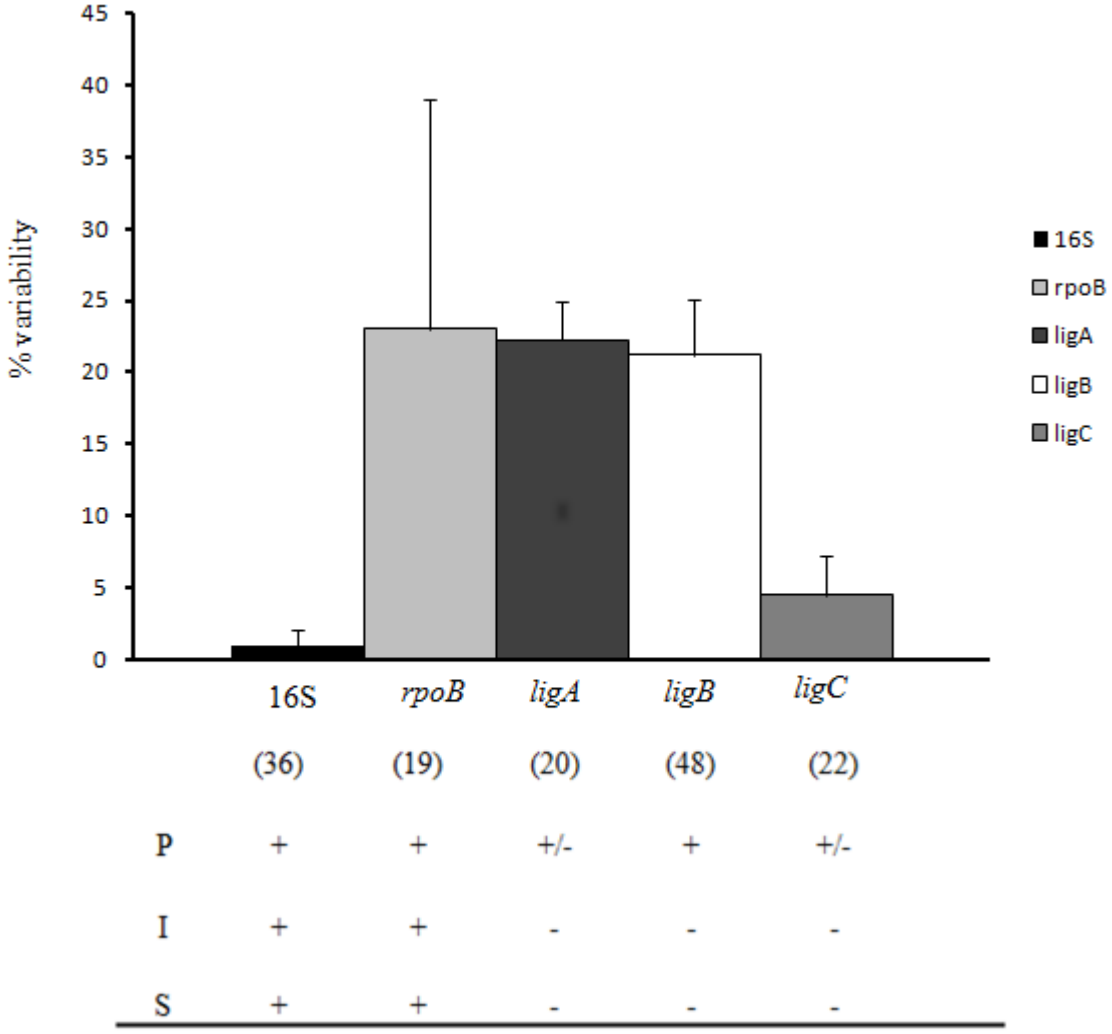
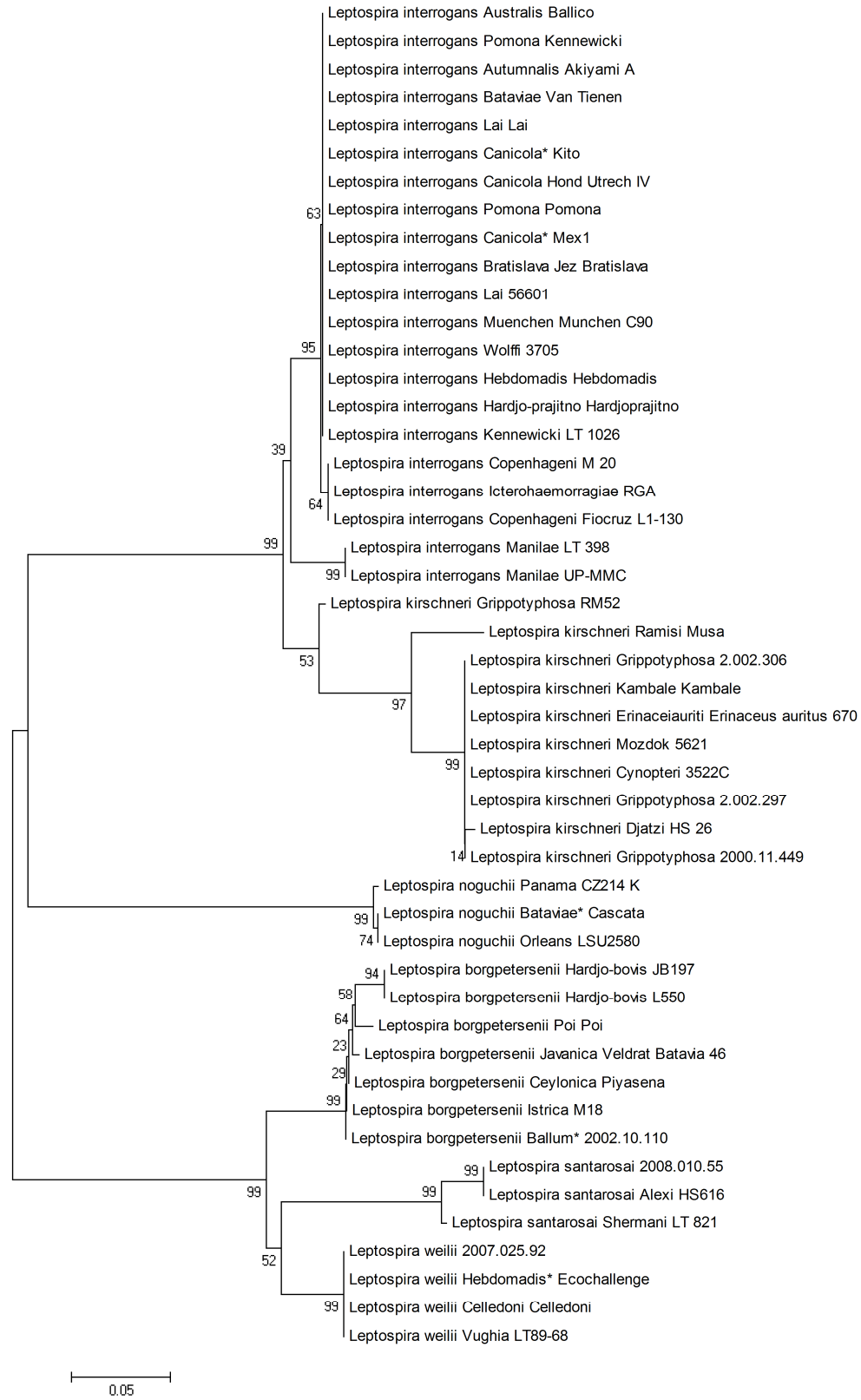
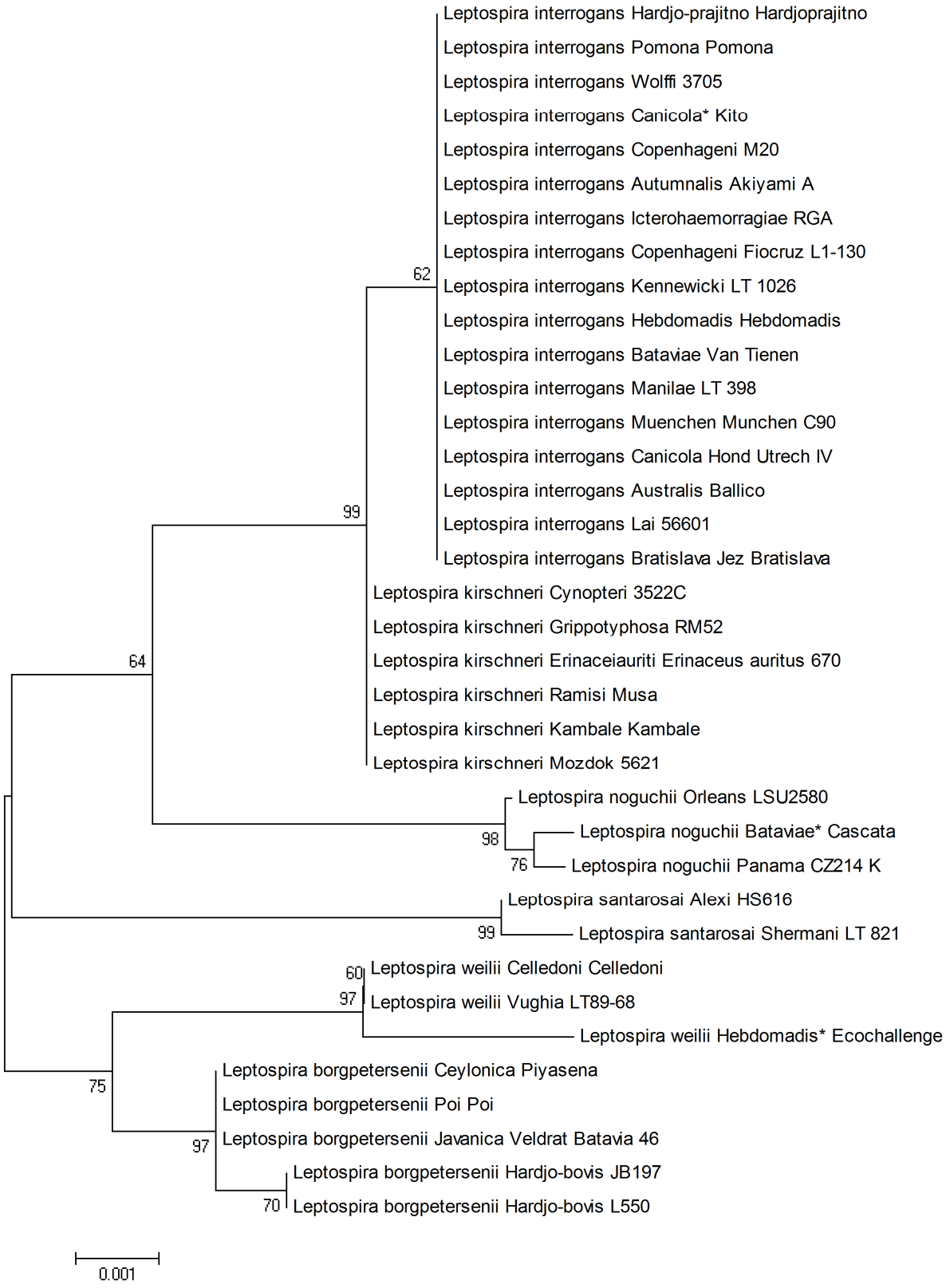

45
40
35
30
25
20
15
10
5
0
% variability
16S
rpoB
ligA
ligB
ligC
16S
rpoB
ligA
ligB
ligC
(36)
(19)
(20)
(48)
(22)
P + + +/- + +/-
I + + - - -
S + + - - -

Figure 3

(A)


Leptospira interrogans Australis Ballico
Leptospira interrogans Pomona Kennewicki
Leptospira interrogans Autumnalis Akiyami A
Leptospira interrogans Bataviae Van Tienen
Leptospira interrogans Lai Lai
Leptospira interrogans Canicola* Kito
Leptospira interrogans Canicola Hond Utrech IV
63
Leptospira interrogans Pomona Pomona
Leptospira interrogans Canicola* Mex1
Leptospira interrogans Bratislava Jez Bratislava
Leptospira interrogans Lai 56601
Leptospira interrogans Muenchen Munchen C90
95
Leptospira interrogans Wolffi 3705
Leptospira interrogans Hebdomadis Hebdomadis
Leptospira interrogans Hardjo-prajitno Hardjoprajitno
Leptospira interrogans Kennewicki LT 1026
39
Leptospira interrogans Copenhageni M 20
Leptospira interrogans Icterohaemorragiae RGA
64
Leptospira interrogans Copenhageni Fiocruz L1-130
99
Leptospira interrogans Manilae LT 398
99
Leptospira interrogans Manilae UP-MMC
Leptospira kirschneri Grippotyphosa RM52
Leptospira kirschneri Ramisi Musa
53
Leptospira kirschneri Grippotyphosa 2.002.306
Leptospira kirschneri Kambale Kambale
97
Leptospira kirschneri Erinaceiauriti Erinaceus auritus 670
Leptospira kirschneri Mozdok 5621
99
Leptospira kirschneri Cynopteri 3522C
Leptospira kirschneri Grippotyphosa 2.002.297
Leptospira kirschneri Djatzi HS 26
14
Leptospira kirschneri Grippotyphosa 2000.11.449
Leptospira noguchii Panama CZ214 K
99
Leptospira noguchii Bataviae* Cascata
74
Leptospira noguchii Orleans LSU2580
94
Leptospira borgpetersenii Hardjo-bovis JB197
58
Leptospira borgpetersenii Hardjo-bovis L550
64
Leptospira borgpetersenii Poi Poi
23
Leptospira borgpetersenii Javanica Veldrat Batavia 46
29
Leptospira borgpetersenii Ceylonica Piyasena
99
Leptospira borgpetersenii Istrica M18
Leptospira borgpetersenii Ballum* 2002.10.110
99
Leptospira santarosai 2008.010.55
99
Leptospira santarosai Alexi HS616
99
Leptospira santarosai Shermani LT 821
52
Leptospira weilii 2007.025.92
Leptospira weilii Hebdomadis* Ecochallenge
99
Leptospira weilii Celledoni Celledoni
Leptospira weilii Vughia LT89-68
0.05

(B)


Leptospira interrogans Hardjo-prajitno Hardjoprajitno
Leptospira interrogans Pomona Pomona
Leptospira interrogans Wolffi 3705
Leptospira interrogans Canicola* Kito
Leptospira interrogans Copenhageni M20
Leptospira interrogans Autumnalis Akiyami A
Leptospira interrogans Icterohaemorragiae RGA
Leptospira interrogans Copenhageni Fiocruz L1-130
62
Leptospira interrogans Kennewicki LT 1026
Leptospira interrogans Hebdomadis Hebdomadis
Leptospira interrogans Bataviae Van Tienen
Leptospira interrogans Manilae LT 398
Leptospira interrogans Muenchen Munchen C90
Leptospira interrogans Canicola Hond Utrech IV
Leptospira interrogans Australis Ballico
99
Leptospira interrogans Lai 56601
Leptospira interrogans Bratislava Jez Bratislava
Leptospira kirschneri Cynopteri 3522C
Leptospira kirschneri Grippotyphosa RM52
64
Leptospira kirschneri Erinaceiauriti Erinaceus auritus 670
Leptospira kirschneri Ramisi Musa
Leptospira kirschneri Kambale Kambale
Leptospira kirschneri Mozdok 5621
Leptospira noguchii Orleans LSU2580
98
Leptospira noguchii Bataviae* Cascata
76
Leptospira noguchii Panama CZ214 K
Leptospira santarosai Alexi HS616
99
Leptospira santarosai Shermani LT 821
60
Leptospira weilii Celledoni Celledoni
97
Leptospira weilii Vughia LT89-68
Leptospira weilii Hebdomadis* Ecochallenge
75
Leptospira borgpetersenii Ceylonica Piyasena
Leptospira borgpetersenii Poi Poi
97
Leptospira borgpetersenii Javanica Veldrat Batavia 46
Leptospira borgpetersenii Hardjo-bovis JB197
70
Leptospira borgpetersenii Hardjo-bovis L550
0.001

7. CONCLUSÕES GERAIS

Artigo 2

- O transposon do tipo mariner, Himar 1, é uma ferramenta eficaz na transformação genética de *Leptospira* spp., conforme comprovado pela criação de um banco de mutantes;
- O método de transposon mutagênese pode ser aplicado para a identificação de fatores de virulência e processos biológicos ligados à patogenicidade de leptospiras patogênicas, mediante a avaliação dos mutantes gerados em modelos animais;

Artigo 3

- Um sistema de expressão induzível, baseado no sistema Lac, permite a expressão controlada de antígenos e complementação de mutantes em *L. biflexa*.
- A proteína GFP não é tóxica para *L. biflexa* e pode ser utilizada como um repórter para sua manipulação genética;

Artigo 4

- Os genes *lig* estão presentes apenas em sorovares patogênicos de *Leptospira*, e *ligB* é o único presente entre todas as cepas e isolados;
- O gene *ligA* está presente apenas entre os sorovares das espécies *L. interrogans* e *L. kirschneri*, enquanto *ligC* aparece, além destas, também entre sorovares das espécies *L. noguchii* e *L. weilii*;
- A sequência do fragmento amplificado do gene *ligB* pode ser usada na diferenciação das espécies patogênicas de *Leptospira*.

8. REFERÊNCIAS

BARBOSA, A. S.; ABREU, P.A.; NEVES, F.O.; ATZINGEN, M.V.; WATANABE, M.M.; VIEIRA, M.L.; MORAIS, Z.M.; VASCONCELLOS, S.A.; NASCIMENTO, A.L. A newly identified leptospiral adhesin mediates attachment to laminin. **Infection and Immunity**, v.74, p. 6356–6364, 2006.

BHARTI, A. R.; NALLY, J. E.; RICALDI, J. N.; MATTHIAS, M. A.; DIAZ, M. M.; LOVETT, M. A.; LEVETT, P. N.; GILMAN, R. H.; WILLIG, M. R.; GOTUZZO, E.; VINETZ, J. M. Leptospirosis: a zoonotic disease of global importance. **The Lancet Infectious Diseases***,* v.3, n.12, p.757-771, 2003.

BOURHY, P.; LOUVEL, H.; SAINT GIRONS, I.; PICARDEAU, M. Random insertional mutagenesis of *Leptospira interrogans*, the agent of leptospirosis, using a *mariner* transposon. **Journal of Bacteriology**, v. 187, p. 3255–3258, 2005..

BRANGER, C.; SONRIER, C.; CHATRENET, B.; KLONJKOWSKI, B.; RUVOEN-CLOUET, N.; AUBERT, A.; ANDRE-FONTAINE, G.; ELOIT, M. Identification of the hemolysis-associated protein 1 as a cross-protective immunogen of *Leptospira interrogans* by adenovirus-mediated vaccination. **Infection and Immunity***,* v. 69, n.11, p.6831-6838, 2001.

BRANGER, C.; CHATRENET, B.; GAUVRIT, A.; AVIAT, F.; AUBERT, A.; BACH, J. M.; ANDRE-FONTAINE, G. Protection against *Leptospira interrogans* sensu lato challenge by DNA immunization with the gene encoding hemolysin-associated protein 1. **Infection and Immunity***,* v. 73, n.7, p.4062-4069, 2005.

BULACH, D. M.; ZUERNER, R. L.; WILSON, P.; SEEMANN, T.; MCGRATH, A.; CULLEN, P. A.; DAVIS, J.; JOHNSON, M.; KUCZEK, E.; ALT, D. P.; PETERSON-BURCH, B.; COPPEL, R. L.; ROOD, J. I.; DAVIES, J. K.; ADLER, B. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. **Proceedings of National Academy of Science of USA**, v. 103, p. 14560–14565, 2006.

CHOY, H.A., KELLEY, M.M., CHEN, T.L., MOLLER, A.K., MATSUNAGA, J., HAAKE, D.A. Physiological osmotic induction of Leptospira interrogans adhesion: LigA and LigB bind extracellular matrix proteins and fibrinogen. **Infection and Immunity**, v. 75, p. 2441–2450, 2007.

CRODA, J.; FIGUEIRA, C. P.; WUNDER, E. A. J.; SANTOS, C. S.; REIS, M. G.; KO, A. I.; PICARDEAU, M. Targeted mutagenesis in pathogenic *Leptospira*: disruption of the *ligB* gene does not affect virulence in animal models of leptospirosis. **Infection and Immunity**. v. 76, p. 5826–5833, 2008.

Cullen, P.A.; Xu, X.; Matsunaga, J.; Sanchez, Y.; Ko, A.I.; Haake, D.A.; Adler, B. Surfaceome of *Leptospira* spp. **Infection and Immunity**, v. 73, p. 4853–4863, 2005.

(a) FAISAL, S. M.; YAN, W.; CHEN, C. S.; PALANIAPPAN, R. U.; MCDONOUGH, S. P.; CHANG, Y. F.. Evaluation of protective immunity of *Leptospira* immunoglobulin like protein A (LigA) DNA vaccine against challenge in hamsters. **Vaccine**, v. 26, p. 277-287, 2008.

(b) FAISAL, S. M.; YAN, W.; MCDONOUGH, S. P.; CHANG, Y. F. *Leptospira* immunoglobulin-like protein A variable region (LigAvar) incorporated in liposomes and PLGA microspheres produces a robust immune response correlating to protective immunity. **Vaccine**, 2008 *in press.*

HAAKE, D. A.; MAZEL, M. K.; MCCOY, A. M.; MILWARD, F.; CHAO, G.; MATSUNAGA, J.; WAGAR, E. A. Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection. **Infection and Immunity***,* v. 67, n.12, p.6572-6582, 1999.

KOIZUMI, N.; WATANABE, H. Leptospiral immunoglobulin-like proteins elicit protective immunity. **Vaccine***,* v. 22, n.11-12, p.1545-1552, 2004.

Lee, S. H.; Kim, S.; Park, S. C.; Kim, M. J. Cytotoxic activities of *Leptospira interrogans* hemolysin SphH as a pore-forming protein on mammalian cells. **Infection and Immunity**, v. 70, n.1, p.315-22, 2002.

LEVETT, Paul N. Leptospirosis. **Clinical Microbiology Reviews***,* v. 14, n.2, p.296-326, 2001.

MATSUNAGA, J.; SANCHEZ, Y.; XU, X.; HAAKE, D.A. Osmolarity, a key environmental signal controlling expression of leptospiral proteins LigA and LigB and the extracellular release of LigA. **Infection and Immunity**, v. 73, p. 70–78, 2005.

MCBRIDE, A. J.; ATHANAZIO, D. A.; REIS, M. G.; KO, A. I. Leptospirosis. **Current Opinion in Infectious Diseases***,* v. 18, n.5, p.376-386, 2005.

MERIEN, F.; TRUCCOLO, J.; BARANTON, G.; PEROLAT, P. Identification of a 36-kDa fibronectin-binding protein expressed by a virulent variant of *Leptospira interrogans* serovar icterohaemorrhagiae. **FEMS Microbiology Letters**, v. 185, n. 1, p. 17-22, 2000.

MOREY, R. E.; GALLOWAY, R. L.; BRAGG, S. L.; STEIGERWALT, A. G.; MAYER, L. W.; LEVETT, P. N.. Species-specific identification of Leptospiraceae by 16S rRNA gene sequencing. **Journal of Clinical Microbiology**, v. 44, p. 3510–3516, 2006.

(a) MURRAY, G. L.; ELLIS, K. M.; LO, M.; ADLER, B. *Leptospira interrogans* requires a functional heme oxygenase to scavenge iron from hemogglobin. **Microbes Infection**, v. 10, p. 791–797, 2008.

(b) MURRAY, G. L.; MOREL, V.; CERQUEIRA, G. M.; CRODA, J.; SRIKRAM, A.; HENRY, R.; KO, A. I.; DELLAGOSTIN, O. A.; BULACH, D. M.; SERMSWAN, R. W.; ADLER, B.; PICARDEAU, M. Genome-wide transposon mutagenesis in pathogenic *Leptospira* species. **Infection and Immunity**, v. 77, n. 2, p. 810-816, 2009.

(c) MURRAY, G. L.; SRIKRAM, A.; HOKE, D. E.; WUNDER, E. A. JR.; HENRY, R.; LO, M.; ZHANG, K.; SERMSWAN, R. W.; KO, A. I.; ADLER, B. The major surface protein LipL32 is not required for either acute or chronic infection with *Leptospira* interrogans. **Infection and Immunity**, 2008, *in press*.

NASCIMENTO, A. L.; KO, A. I.; MARTINS, E. A.; MONTEIRO-VITORELLO, C. B.; O, P. L.; HAAKE, D. A.; VERJOVSKI-ALMEIDA, S.; HARTSKEERL, R. A.; MARQUES, M. V.; OLIVEIRA, M. C.; MENCK, C. F.; LEITE, L. C.; CARRER, H.; COUTINHO, L. L.; DEGRAVE, W. M.; DELLAGOSTIN, O. A.; EL-DORRY, H.; FERRO, E. S.; FERRO, M. I.; FURLAN, L. R.; GAMBERINI, M.; GIGLIOTI, E. A.; GOES-NETO, A.; GOLDMAN, G. H.; GOLDMAN, M. H.; HARAKAVA, R.; JERONIMO, S. M.; JUNQUEIRA-DE-AZEVEDO, I. L.; KIMURA, E. T.; KURAMAE, E. E.; LEMOS, E. G.; LEMOS, M. V.; MARINO, C. L.; NUNES, L. R.; DE OLIVEIRA, R. C.; PEREIRA, G. G.; REIS, M. S.; SCHRIEFER, A.; SIQUEIRA, W. J.; SOMMER, P.; TSAI, S. M.; SIMPSON, A. J.; FERRO, J. A.; CAMARGO, L. E.; KITAJIMA, J. P.; SETUBAL, J. C.; VAN SLUYS, M. A. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. **Journal of Bacteriology**, v. 186, p. 2164–2172, 2004.

PALANIAPPAN, R. U., MCDONOUGH, S. P., DIVERS, T. J., CHEN, C. S., PAN, M. J.; MATSUMOTO, M.; CHANG, Y. F. Immunoprotection of recombinant leptospiral immunoglobulin-like protein A against *Leptospira interrogans* serovar Pomona infection. **Infection and Immunity**, v. 74, n. 3, p. 1745-1750, 2006.

REN, S.; FU, G.; JIANG, X.; ZENG, R.; XIONG, H.; LU, G.; JIANG, H. Q.; MIAO, XU, H.; ZHANG, Y.; GUO, X.; SHEN, Y.; QIANG, B. Q.; Q. X., DANCHIN, A.; SAINT GIRONS, I.; SOMERVILLE, R. L.; WENG, Y. M.; SHI, M.; CHEN, Z.; XU, J. G.; ZHAO, G. P. Unique and physiological and pathogenic features of *Leptospira interrogans* revealed by whole genome sequencing. **Nature**. v. 422, p.888–893, 2003.

RISTOW, P.; BOURHY, P.; DA CRUZ MCBRIDE, F. W.; FIGUEIRA, C. P.; HUERRE, M.; AVE, P.; GIRONS, I. S.; KO, A. I.; PICARDEAU, M. The OmpA-like protein Loa22 is essential for leptospiral virulence. **PLoS Pathogens**, v. 3, e97, 2007.

SILVA, E. F.; MEDEIROS, M. A.; MCBRIDE, A. J.; MATSUNAGA, J.; ESTEVES, G. S.; RAMOS, J. G.; SANTOS, C. S.; CRODA, J.; HOMMA, A.; DELLAGOSTIN, O. A.; HAAKE, D. A.; REIS, M. G.; KO, A. I. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. **Vaccine**, v. 25, 6277-6286, 2007.

SLACK, A. T.; KALAMBAHETI, T.; SYMONDS, M. L.; DOHNT, M. F.; GALLOWAY, R. L.; STEIGERWALT, A. G.; CHAICUMPA, W.; BUNYARAKSYOTIN, G.; CRAIG, S.; HARROWER, B. J.; SMYTHE, L. D. *Leptospira wolffii* sp. nov., isolated from a human with suspected leptospirosis in Thailand. **International Journal of Systematics and Evolutionary Microbiology**, v. 58, n. 10, 2305-2308, 2008.

STEVENSON, B.; CHOY, H. A.; PINNE, M.; ROTONDI, M. L.; MILLER, M. C.; DEMOLL, E.; KRAICZY, P.; COOLEY, A. E.; CREAMER, T. P.; SUCHARD, M. A.; BRISSETTE, C. A.; VERMA, A.; HAAKE, D. A. *Leptospira interrogans* endostatin-like outer membrane proteins bind host fibronectin, laminin and regulators of complement. **PLoS ONE**, v. 2, p. 1188, 2007.

YAN, W.; FAISAL, S. M.; MCDONOUGH, S. P.; DIVERS, T. J.; BARR, S. C.; CHANG, C. F.; PAN, M. J.; CHANG, Y. F. Immunogenicity and protective efficacy of recombinant Leptospira immunoglobulin-like protein B (rLigB) in a hamster challenge model. **Microbes and Infecttion**, 2008, *in press*.

ANEXOS

ANEXO I

O ANEXO I é o material suplementar do Artigo 2. Neste anexo estão apresentados os dados de identificação dos genes interrompidos pelo transposon Himar 1 nos sorovares Lai, Manilae, Copenhageni, Canicola, Bataviae e Hebdomadis. Nesta tabela constam a identificação dos mutantes, suas coordenadas de localização com base no genoma de *L. interrogans* Lai, o nome do gene interrompido, sua função, a existência de ortólogos em outras espécies de *Leptospira* e comentários gerais. Os mutantes marcados em amarelo correspondem àqueles interrompidos mais de uma vez no mesmo gene. Mutantes em negrito foram utilizados para infecção experimental no modelo de hamster.

Supplemental Table: Location of transposon insertions in the L. interrogans chromosomes
yellow: multiple Himar1 integration in individual gene/bold letters: mutants tested in animal model

mutant strain	Himar1 insertion coordinate (Lai)	Putative interrupted CDS coordinates		Potential function	leptospiral orthologs	comments
FLaiS320	2418515/3065176	16S ribosomal RNA	ND			
FLai-K223	2418098/3064759	16S ribosomal RNA	ND			
FLaiK152	2418602/3065263	16S ribosomal RNA	ND			
FLaiS288	2418515/3065176	16S ribosomal RNA	ND			
FKit23	2418850 /3065511	16S ribosomal RNA	ND			
AMan756	2417550/3064211	16S ribosomal RNA	ND			
AMan1089	2417902/3064565	16S ribosomal RNA	ND			
AMan1243	2417998/3064659	16S ribosomal RNA	ND			
AMan822	2418092/3064753	16S ribosomal RNA	ND			
AMan790	2418103/3064764	16S ribosomal RNA	ND			
AMan1312	2418582/3065243	16S ribosomal RNA	ND			
AMan1019	3064653/2417992	16S ribosomal RNA	ND			
AMan764	3064785/2418124	16S ribosomal RNA	ND			
AMan798	3065069/2418408	16S ribosomal RNA	ND			
AMan1015	2417769	16S ribosomal RNA	ND			
AMan1091	2417879	16S ribosomal RNA	ND			
AMan1084	2418092	16S ribosomal RNA	ND			
AMan1034	2418423	16S ribosomal RNA	ND			
AMan994	516506	23S ribosomal RNA	ND			
AMan1112	516801	23S ribosomal RNA	ND			
AMan882	516862	23S ribosomal RNA	ND			
AMan964	517419	23S ribosomal RNA	ND			
AMan879	517517	23S ribosomal RNA	ND			
AMan814	517752	23S ribosomal RNA	ND			
FLaiTK4	517518	23S ribosomal RNA	ND			
FKit188	6711	LA0005 (+)	6626-8545	gyrB1 DNA gyrase subunit B	LEPBIa0005/LIC10005/LBJ_0008/LBL_0008	
FCan4	6976	LA0005 (+)	6626-8545	gyrB1 DNA gyrase subunit B	LEPBIa0005/LBJ_0008/LBL_0008/LIC10005	
AMan1310	7467	LA0005 (+)	6626-8545	gyrB1 DNA gyrase subunit B	LIC10005/LEPBIa0005/LBJ_0008/LBL_0008	
AMan1252	8541	LA0005 (+)	6626-8545	gyrB1 DNA gyrase subunit B	LIC10005/LEPBIa0005/LBJ_0008/LBL_0008	
ALaiC1	9671	LA0006 (+)	8564 - 11068	gyrA1 DNA gyrase subunit A	LIC10006/LEPBIa0006/LBJ_0009/LBL_0009	
AMan151	12105	LA0008 (+)	12069 -12413	hypothetical protein	LIC10008	
FLaiS304	12424	LA0008-LA0009				
AMan1305	13934	LA0010 (-)	13744-14745	hypothetical protein	LIC10010	
AMan1139	14291	LA0010 (-)	13744-14745	hypothetical protein	LIC10010	
ALaiC12	14377	LA0010 (-)	13744-14745	hypothetical protein	LIC10010	

FLaiS250	19457	LA0014- LA0015				
FLai22	20202	LA0017 (-)	20125-20688	Maf-like protein	LIC10016/LEPBIa0014/LBL_2976/LBJ_3012	Septum formation protein Maf
FLai58	28400	LA0025 (+)	28064-29077	fliG1	flagellar motor switch protein FliG	LIC10023/LBJ_0019/LBL_0019/LEPBIa3423
FLai193S	28774	LA0025 (+)	28064-29077	fliG1	flagellar motor switch protein FliG	LIC10023/LBJ_0019/LBL_0019/LEPBIa3423
AMan298	35038	LA0032 (+)	34705- 35064	hypothetical protein	LIC10028/LBL_0080/LBJ_2985	Cytochrome c, mono- and diheme variants
FLaiK3E	38980	LA0035 (-)	37450-39426	beta-galactosidase	LIC10031 LBJ_2988 LBL_0077 LEPBIa0024	
FLaiS50	39506	LA0035- LA0036				
AMan1311	44686	LA0042 (+)	44654-46267	GMC oxidoreductase	LEPBIa3414/LIC10037	FAD/NAD(P)-binding domain
AMan809	44688	LA0042 (+)	44654-46267	GMC oxidoreductase	LEPBIa3414/LIC10037	FAD/NAD(P)-binding domain
AMan800	51008	LA0048-LA0049				
AMan1163	51819	LA0048-LA0049				LA0049 (methyl accepting chemotaxis protein) pro
ALaiC4	52428	LA0049(+)	51872-53569	aerotaxis sensor receptor, flavoprotein	LIC10043/LBL_2991/LBJ_0037/LEPBIa0595	methyl accepting chemotaxis protein
AMan17	54960	LA0050 (+)	54175-55053	hypothetical protein	LEPBIa3231/LBL_2992/LBJ_0038/LIC10044	Permeases of the major facilitator superfamily 6TM
ALaiC6	57646	LA0053 (-)	57190-57726	hypothetical protein	LIC10047/LBJ_0041/LBL_2995/LEPBIa3235	3TM Putative membrane-associated phospholipid p
AMan778	60147	LA0055 (-)	59049-60197	hypothetical protein	LEPBIa3237/LBJ_0043/LBL_2997/LIC10049	
AMan653	65483	LA0060 (-)	65283-66185	TPR-repeat-containing protein	LIC10053/LBJ_0051/LBL_3005/LEPBIa3241	
AMan936	70759	LA0064 (+)	69702 -71117	hypothetical protein	LEPBIa0035/LBL_3009/LBJ_0055/LIC10057	
AMan987	74352	LA0066 (-)	73311-74594	Proton/sodium-glutamate symport protein	LIC10059/LBL_3011/LBJ_0057/LEPBIa3441	8TM Sodium:dicarboxylate symporter
AMan988	76752	LA0067 (+)	74734-77919	hypothetical protein	LIC10060/LBJ_0058/LBL_3012/LEPBIa3243	8TM Putative chemotaxis sensory transducer; putat
AMan1346	77797	LA0067 (+)	74734-77919	hypothetical protein	LIC10060/LBJ_0058/LBL_3012/LEPBIa3243	8TM Putative chemotaxis sensory transducer; putat
AMan972	82772	LA0070 (+)	79673-82879	hypothetical protein	LIC10063/LBJ_0060/LBL_3014/LEPBIa3243	10TM Putative chemotaxis sensory transducer; putz
FKitGK48	93555	LA0084 (-)	93349-94059	acyltransferase	LBJ_0072/LBL_3026/LIC10075/LEPBIa0586	Phospholipid/glycerol acyltransferase
FLaiK150	97392	LA0091 (+)	97163-97510	hypothetical protein	LIC10077/LBJ_0074/LBL_3028/LEPBIa3042	
AMan142	99326	LA0094 (-)	99185- 100006	hypothetical protein	LIC10080	
AMan1013	105419	LA0100 (+)	104945-106225	hypothetical protein	LEPBIa1252/LBL_0024/LBJ_0111/LIC10088	SP
AMan930	111993	LA0105 (-)	111132 -112058	hypothetical protein	LIC10093	6TM
AMan751	124153	LA0115 (+)	123968-124573	hypothetical protein	LIC10103/LBL_0041/LBJ_0094/LEPBIa3283	
AMan833	126808	LA0117 (-)	126245-127267	hypothetical protein	LIC10105/LBJ_0096/LBL_0039	
FLai5	141354	LA0133 (-)	140801-142630	Sensory transduction histidine kinase	LIC10120/LBJ_0116/LBL_2963/LEPBIa2901	
AMan807	141831	LA0133 (-)	140801 -142630	Sensory transduction histidine kinase	LIC10120/LBL_2963/LBJ_0116/LEPBIa2901	Two-component sensor histidine kinase
AMan843	142559	LA0133 (-)	140801 -142630	Sensory transduction histidine kinase	LIC10120/LBL_2963/LBJ_0116/LEPBIa2901	Two-component sensor histidine kinase
FLaiPCRK11	144538	LA0135 (+)	143601-144902	hypothetical protein	LIC10122/LBJ_0118/LBL_2965/LEPBIa1560	
AMan1140	144564	LA0135 (+)	143601 -144902	hypothetical protein	LIC10122/LBJ_0118/LBL_2965/LEPBIa1560	
AMan845	145819	LA0136 (+)	144972 - 145964	lipoprotein qlp42	LIC10123/LBJ_0119/LBL_2966/LEPBIa2895	LipL45-like protein
AMan1225	**145820**	**LA0136 (+)**	**144972 - 145964**	**lipoprotein qlp42**	**LIC10123/LBJ_0119/LBL_2966/LEPBIa28!**	**LipL45-like protein**
AMan1368	146292	LA0137 (+)	146075 - 147730	hypothetical protein	LIC10124/LBJ_0120/LBL_2967	
AMan253	146584	LA0137 (+)	146075 - 147730	hypothetical protein	LIC10124/LBJ_0120/LBL_2967	
AMan1073	**146655**	**LA0137 (+)**	**146075 - 147730**	**hypothetical protein**	**LIC10124/LBJ_0120/LBL_2967**	
AMan888	149929	LA0138 (+)	147727 -151320	TPR-repeat-containing proteins	LIC10125/LBJ_0121/LBL_2968/LEPBIa2897	
FLaiS60	151081	LA0138 (+)	147727-151320	TPR-repeat-containing proteins	LIC10125/LBJ_0121/LBL_2968/LEPBIa2897	

AMan851	154948	LA0143 (+)	154281-154949	hypothetical protein	LIC10129/LBJ_0126/LBL_2957/LEPBIa0561	
FLaiTK78	154973	LA0143-LA0144				
FLaiPCR35	158128	LA0146 (+)	157912-160011	nifA | Nif-specific regulatory protein	LIC10132/LBJ_0129/LBL_2954/LEPBIa0566	Transcriptional regulator
FLaiS59	161192	LEPIN0145 (+)	160928-161056	hypothetical protein; putative exported protein		
FLaiK143	162318	LA0150 (-)	162352-164172	hlyB | Hemolysin secretion ATP-binding pr	LIC10136/LBJ_0132/LBL_2951	ABC transporter ATP-binding protein
FKitGK42	163911	LA0150 (-)	162352-164172	hlyB | Hemolysin secretion ATP-binding pr	LIC10136/LBJ_0132/LBL_2951	ABC transporter ATP-binding protein
AMan469	164858	LA0151 (+)	164345 -164959	ACT family protein	LIC10137/LBJ_0133/LBL_2950/LEPBIa0066	[protein-PII] uridylyltransferase
AMan632	165909	LA0153 (+)	165126 -166595	hypothetical protein	LEPBIa1137/LBJ_0134/LBL_2949/LIC10138	HD-GYP hydrolase domain protein
FPC3	167610	LA0153-LA0156				
AMan1365	173500	LA0161-LA0162				
AMan621	174060	LA0163-LA0164				
AMan1076	174170	LA0164-LA0165				
FLaiPCR42	174250	LA0164-LA0165				
AMan631	174978	LA0166 (+)	174924 -176297	hypothetical protein	LIC10145	
FLai46	177449	LA0168 (-)	177350-177457	hypothetical protein	LIC11322	
FLaiS65	179927	LA0171 (+)	179516-180778	aminotransferase	LIC10150/LEPBIa0695	
AMan891	182695	LA0172 (+)	180854 -183208	hypothetical protein	LEPBIa0694/LBJ_0143/LBL_2940/LIC10151	
AMan909	189878	LA0181 (+)	189124 - 190437	NADP-dependent malic enzyme	LIC10160/LBL_4280/LBL_2932/LEPBIa0143	
AMani2	192132	LA0184 (+)	191138 - 193216	hypothetical protein	LIC10161/LBL_2931/LBJ_0152	
AMan1001	199127	LA0197-LA0198				
FLaiK-235	207438	LA0209 (+)	205534-207741	phage-like protein	LIC10178	Probable bacteriophage tail protein, membrane prot
FLaiS1	220548	LA0222 (-)	220446-221033	outer membrane protein OmpA family	LIC10191/LBJ_0158/LBL_2925/LEPBIa3149	Loa22
FLai7	257504	LA0262 (-)	257323-258222	Peptidase family M23/M37	LIC10224/LBJ_2782/LBL_0289/LEPBIa2933	
AMan1071	257620	LA0262 (-)	257323 -258222	Peptidase family M23/M37	LIC10224/LBJ_2782/LBL_0289/LEPBIa2933	Membrane proteins related to metalloendopeptidase
AMan821	257906	LA0262 (-)	257323 -258222	Peptidase family M23/M37	LIC10224/LBJ_2782/LBL_0289/LEPBIa2933	Membrane proteins related to metalloendopeptidase
FKitGK62	259285	LA0263 (+)	259201-259929	hypothetical protein	LIC10226/LBJ_2780/LBL_0291/LEPBIa2930	
AMan478	259869	LA0263 (+)	259201 -259929	hypothetical protein	LIC10226/LBJ_2780/LBL_0291/LEPBIa2930	
AMan1152	265903	LA0269-LA0270				
AMan704	273517	LA0279 (+)	273468 -274106	hypothetical protein	LIC10236/LBJ_0664/LBL_2415/LEPBIa0169	
AMan993	275796	LA0281 (+)	275430 -276116	DNA-3-methyladenine glycosylase	LIC10238/LBJ_0662/LBL_2417	
FLaiPCRK10	276429	LA0282 (+)	276107-276214	hypothetical protein		
AMan772	276456	LA0282_LA0283				
AMan1042	280617	LA0286 (+)	280400 -281158	hypothetical protein	LIC10244/LBJ_2606/LBL_0506	
AMan1189	280743	LA0286 (+)	280400 -281158	hypothetical protein	LIC10244/LBJ_2606/LBL_0506	
FPom04	284102	LA0289-LA0290				
AMan711	288689	LA0293 (-)	287046-289361	purine NTPase	LIC10251/LBL_2421/LBJ_0658/LEPBIa2905	Rad50-like protein
AMan494	289980	LA0294 (-)	289363 -289989	hypothetical protein	LIC10252/LBL_2420/LBJ_0659/LEPBIa2906	DNA repair exonuclease
AMan1167	307635	LA0310 (-)	306815-307870	probable catabolite gene activator	LIC10269/LBJ_0378/LBL_2699/LEPBIa1150	Putative cAMP binding protein
AMan859	307748	LA0310 (-)	306815-307870	probable catabolite gene activator	LIC10269/LBJ_0378/LBL_2699/LEPBIa1150	Putative cAMP binding protein
AMan978	308349	LA0311 (-)	307881 -309131	probable chlorohydrolase	LEPBIa1151/LBJ_0379/LBL_2698/LIC10270	Amidohydrolase

FWei8	308905	LA0311 (-)	307881-309131	probable chlorohydrolase	LIC10270/LBJ_0379/LBL_2698/LEPBIa1151	Putative deaminase or related metal-dependent hyd
AMan832	316141	LA0318-LA0319				
AMan796	318056	LA0320 (+)	318007 -318510	Thiol-disulfide isomerase and thioredoxins	LIC10278/LBJ_0384/LBL_2693/LEPBIa1157	
FKit76	318541	LA0320- LA0321				
FLaiK-211	330122	LA0330-LA0331				
AMan1298	351712	LA0359-LA0360				
FLaiS348	367510	LA0374 (-)	365935-368025	hypothetical protein	LBJ_2760/LBL_0311/LIC10321	coiled-coil
FLaiTK113	382847	LA0388 (-)	382316-383629	hypothetical protein	LIC10336/LBJ_2767/LBL_0304/LEPBIa2571	2-methylthioadenine synthetase
AMan889	387802	LA0391-LA0392				
FLaiS43	396594	LA0400-LA0401				
FLai21	396670	LA0400-LA0401				
AMan1011	401672	LA0405 (-)	401414 -402355	dimethyladenosine transferase	LIC10354/LBL_2401/LBJ_0678/LEPBIa1126	mutations in ksgA causes resistance to the translatic
AMan1304	419913	LA0422-LA0423				
AMan1197	420736	LA0423 (-)	420312-421754	hypothetical protein	LIC10371/LBL_1162/LBJ_1108	putative lipoprotein
AMan168	429556	LA0430 (+)	429451-430209	hypothetical protein	LIC10377	putative lipoprotein
FLai40	441469	LA0442 (-)	441059-441922	hypothetical protein	LIC10385/LBJ_2870/LBL_0201	Sm-like ribonucleoproteins
AMan1030	442995	LA0444 (+)	442683 -443762	hypothetical protein	LIC10387/LBJ_2867/LBL_0204	
AMan797	445967	LA0446-LA0447				
AMan1157	447220	LA0448 (+)	447243 -449183	ABC transporter, ATP-binding protein	LIC10389/LBJ_2864/LBL_0207/LEPBIa0274	
FLaiPCRK19	462662	LA0465 (+)	462142-464208	TPR-repeat-containing proteins	LIC10405/LBJ_2506/LBL_0635/LEPBIa2138	
FLaiS8	467990	LA0468-LA0469				
AMan706	475198	LA0477-LA0478				
AMan1020	**498008**	**LA0505 (-)**	**497374-498474**	**probable glycosyl hydrolase**	**LBL_0375/LBJ_2697/LIC13050/LEPBIa2670**	
FLai44	499073	LA0506 (+)	498559-499236	hypothetical protein	LIC13049/LEPBIa2671/LBJ_2698/LBL_0374	
FPom45	502325	LA0513 (-)	502279-502461	hypothetical protein		
AMan1265	503479	LA0512-LA0514				
AMan847	515507	LA0526-23S ribosomal RNA				
AMan757	515508	LA0526-23S ribosomal RNA				
FLaiPCR40	519109	LA0528 (-)	519055-519297	hypothetical protein	LIC13034	
AMan864	537381	LA0544 (-)	536808-537983	hypothetical protein	LIC13022/LBL_2261/LBJ_0820/LEPBIa2385	
AMan1056	538576	LA0546 (+)	538108- 540633	hypothetical protein	LIC13021/LEPBIa2599/LBJ_0819/LBL_2262	
AMan1240	538613	LA0546 (+)	538108 -540633	hypothetical protein	LIC13021/LEPBIa2599/LBJ_0819/LBL_2262	
FFCJ1	538823	LA0546 (+)	538108-540633	hypothetical protein	LIC13021/LEPBIa2599/LBJ_0819/LBL_2262	
FLai51	538855	LA0546 (+)	538108-540633	hypothetical protein	LIC13021/LEPBIa2599/LBJ_0819/LBL_2262	cAMP-binding domain-like
Aman648	542498	LA0548 (+)	542037 -542501	hypothetical protein	LIC13019/LEPBIa2602/LBJ_0811/LBL_2269	
AMan1328	562775	LA0562 (+)	562637 -563119	hypothetical protein	LIC13007/LBJ_2406/LBL_0702	
FLaiS258	571017	LA0568 (-)	570430-571851	hypothetical protein	LIC13002	
FLai34	580530	LA0575 (+)	580408-581415	hypothetical protein	LIC12995/LBL_2500/LBJ_0580	
AMan1078	582004	LA0576 (-)	581690 -582829	hypothetical protein	LIC12994/LBL_2499/LBJ_0581/LEPBIa2723	
AMan1277	582032	LA0576 (-)	581690 -582829	hypothetical protein	LIC12994/LBL_2499/LBJ_0581/LEPBIa2723	

AMan607	598396	LA0589 (+)	597157- 599055	hypothetical protein	LIC12986/LBJ_4195/LBJ_1339	
AMan846	603582	LA0594 (+)	601542 -603761	heavy metal-transporting ATPase	LIC12982/LBL_2491/LBJ_0589/LEPBIa2738	7TM
FLai1	608021	LA0600 (-)	607461-608054	Peptide methionine sulfoxide reductase	LIC12978	
AMan79	608860	LA0602 (+)	608788 -610044	lipoprotein	LIC12976/LBL_2486/LBJ_0594/LEPBIa0852	Predicted NAD/FAD-binding protein
AMan874	**619339**	**LA0615 (-)**	**619238 -619621**	**hypothetical protein**	**LIC12967/LBJ_0318/LBL_2758**	
AMan1168	642997	LA0636-LA0637				
FLai32	653095	LA0648 (+)	652906-653574	hypothetical protein	LIC12942/LEPBIa0844/LBJ_2352/LBL_0756	Anti-sigma regulatory factor
FLaiTK2	653987	LA0649 (+)	653581-655581	uvrB | Excinuclease ABC subunit B	LIC12941/LEPBIa0845/LBJ_2353/LBL_0755	nucleotide excision repair subunit B
AMan1082	656245	LA0650- LA0651				
FLai29	656668	LA0650-LA0651				
FLai3	665633	LA0663 (+)	665172-666497	hypothetical protein	LIC12930	Fimh-like protein
FLai37	668374	LA0666 (+)	668375-669397	Cytochrome c peroxidase	LIC12927/LEPBIa2430/LBL_0579/LBJ_2533	
AMan856	673403	LA0671-LA0672				
AMan725	681657	LA0677-LA0678				
AMan1115	**682991**	**LA0678 (-)**	**681902 -683980**	**Methyl-accepting chemotaxis protein mcp**	**LIC12921/LBJ_0536/LBL_2543/LEPBIa2436**	
AMan926	683432	LA0678 (-)	681902 -683980	Methyl-accepting chemotaxis protein mcpB	LIC12921/LBJ_0536/LBL_2543/LEPBIa2436	
AMan1227	683646	LA0678 (-)	681902 -683980	Methyl-accepting chemotaxis protein mcpB	LIC12921/LBJ_0536/LBL_2543/LEPBIa2436	
AMan766	684067	LA0678-LA0679				
AMan784	697548	LA0693-LA0694				
AMan1061	**723772**	**LA0709 (+)**	**722546-731971**	**hypothetical protein**	**LIC12896**	
AMan1107	723784	LA0709 (+)	722546-731971	hypothetical protein	LIC12896	
AMan1245	724524	LA0709 (+)	722546-731971	hypothetical protein	LIC12896	
FFC8	730245	LA0709 (+)	722546-731971	hypothetical protein	LIC12896	
FLaiS180	731529	LA0709 (+)	722546-731971	hypothetical protein	LIC12896	coiled-coil /340kDa
AMan1148	735772	LA0713 (+)	734684 -737362	hypothetical protein	LIC12894	
AMan848	738497	LA0715 (+)	738344 -738910	lipoprotein	LIC12891	
AMan1217	756621	LA0734 (+)	756386 -756937	hypothetical protein	LIC12878/LBL_2297/LBJ_0782	Thioredoxin-like
AMan1012	770793	LA0758 (+)	770607 -771149	50S ribosomal protein L15 rplO	LIC12854/LBJ_2640/LBL_0432/LEPBIa1946	
AMan1366	770841	LA0758 (+)	770607 -771149	50S ribosomal protein L15 rplO	LIC12854/LBJ_2640/LBL_0432/LEPBIa1946	
FLaiTK65	770908	LA0758 (+)	770607-771149	rplO | ribosomal protein L15	LIC12854/LBJ_2640/LBL_0432/LEPBIa1946	
AMan1024	783094	LA0773-LA0774				
FLaiS325	790547	LA0781 (-)	790374-790682	hypothetical protein	LIC12837	
AMan493	794110	LA0786 (-)	793868 -794992	Lipid-A-disaccharide synthetase(lpxB) like	LIC12833/LBJ_0743/LBL_2335/LEPBIa1039	UDP-Glycosyltransferase/glycogen phosphorylase
AMan976	806807	LA0798 (+)	806685 -807452	hypothetical protein	LIC12822/LBJ_0751/LBL_2327/LEPBIa1046	
AMan604	813790	LA0808 (+)	813245 -813868	hypothetical protein	LIC12813/LBJ_0759/LBL_2319	Peptidase, trypsin-like serine and cysteine proteases
AMan1232	825434	LA0823 (-)	825197 -826108	alpha/beta hydrolase fold	LIC12801/LBJ_0546/LBL_2534	alpha/beta-Hydrolases
AMan1077	830310	LA0829 (+)	829998-830489	hypothetical protein	LIC12794	
AMan1344	834865	LA0835 (-)	833787 -835682	hypothetical protein	LIC12791/LBJ_1339/LBL_1564	
FLaiS15	858722	LA0861 (+)	857994-858398	probable anti-sigma factor antagonist	LIC12766/LBJ_0359	
AMan1208	867610	LA0869-LA0870				

AMan617	875497	LA0876-LA0877				promoter sigX Probable RNA polymerase ECF-typ
FLaiS160	885200	LA0886 (-)	883565-885526	NuoL	NADH dehydrogenase I chain L	LIC12750/LBJ_0510/LBL_2569/LEPBIa1305
AMan734	893116	LA0898 (+)	892717 -893118	hypothetical protein		LIC12740/LBL_2579/LBJ_0500
AMan898	910231	LA0915- LA0916				
FLaiS175	911320	LA0916 (+)	910393-911598	hypothetical protein		LIC12726/LBJ_2212/LBL_2205/LEPBIa0779
AMan1201	913590	LA0919 (-)	913498-914025	hypothetical protein		LIC12724/LBJ_2210/LBL_2203/LEPBIa0869
AMan1329	914237	LA0919-LA0920				
AMan802	914335	LA0919-LA0920				
AMan947	914338	LA0919-LA0920				
AMan816	919216	LA0927 (-)	918803 -920044	hypothetical protein		LIC12719/LBJ_2208/LBL_2201/LEPBIa0867 TPR repeat
AMan715	922176	LA0930 (-)	921642-923039	Sugar and other transporter		LIC12717/LBJ_2206/LBL_2199/LEPBIa2961 12 TM Transporter, MFS superfamily; putative mes
AMan749	926951	LA0934 (+)	925932-927848	hypothetical protein		LIC12715/LBJ_4195/LBL_4210
AMan969	940235	LA0945 (+)	940128 -940979	tRNA	pseudouridine synthase B (tRNA pse	LIC12703/LBJ_0948/LBL_2085/LEPBIa1527 Pseudouridine synthase
AMan773	940288	LA0945 (+)	940128 -940979	tRNA	pseudouridine synthase B (tRNA pse	LIC12703/LBJ_0948/LBL_2085/LEPBIa1527 Pseudouridine synthase
AMan768	940619	LA0945 (+)	940128 -940979	tRNA	pseudouridine synthase B (tRNA pse	LIC12703/LBJ_0948/LBL_2085/LEPBIa1527 Pseudouridine synthase
AMan1147	940942	LA0945 (+)	940128 -940979	tRNA	pseudouridine synthase B (tRNA pse	LIC12703/LBJ_0948/LBL_2085/LEPBIa1527 Pseudouridine synthase
AMan2	946564	LA0951 (+)	945989 -946852	hypothetical protein		LIC12698 3TM
AMan638	946643	LA0951 (+)	945989 -946852	hypothetical protein		LIC12698 3TM
FLai45	963874	LA0962 (-)	961648-964173	Putative lipoprotein		LIC12690
AMan602	980190	LA0975 (+)	979812 -981239	Fimh-like protein		LIC12680/LBJ_0461
AMan925	995914	LA0990 (+)	995677 -996375	hypothetical protein		LIC12673/LBJ_0710/LBL_2369 4TM NADH-ubiquinone oxidoreductase subunit 5 (
AMan645	998195	LA0992 (-)	997253 -998263	Mechanosensitive ion channel		LIC12671/LBJ_0712/LBL_2367 5TM Small-conductance mechanosensitive channel
FKit46	1021009	LA1013 (+)	1020602-1021606	hypothetical protein		LIC12644/LBL_2612/LBJ_0467
AMan900	1022523	LA1015 (-)	1022455 - 1023072	hypothetical protein		LIC12642 SP
AMan967	1023463	LA1016 (-)	1023077 - 1025245	hypothetical protein		LIC12641 SP coiled-coil, down-regulated at physiologic osmo
AMan1321	1024107	LA1016 (-)	1023077 - 1025245	hypothetical protein		LIC12641 SP coiled-coil, down-regulated at physiologic osmo
FLaiS339	1024221	LA1016 (-)	1023077-1025245	hypothetical protein		LIC12641 SP coiled-coil, down-regulated at physiologic osmo
AMan763	1025792	LA1017 (-)	1025242 - 1026075	TPR-repeat-containing proteins		LIC12640
FLai43	1027356	LA1019 (-)	1026573-1027364	NifU-like protein		LIC12638/LBJ_0726/LBL_2353/LEPBIa1034
AMan875	1031432	LA1023 (+)	1030904 - 1034392	hypothetical protein		LIC12634
ALaiC7	1039593	LA1029 (-)	1038898-1040769	sph2	Sphingomyelinase C 2 precursor	LIC12631/LBJ_0291/LBL_2785
AMan1340	1044167	LA1035-LA1036				
AMan1307	1047161	LA1037 (-)	1047045-1047215	hypothetical protein		LEPIC2671
AMan927	1047341	LA1038 (+)	1047269 - 1047568	hypothetical protein		LIC12626/LBL_2790/LBJ_0286
AMan829	1052635	LA1045 (-)	1051918 - 1054401	Penicillin-binding protein 1A		LIC12620/LBJ_0283/LBL_2793/LEPBIa1432 Membrane carboxypeptidase (penicillin-binding pr
AMan975	1057575	LA1048-LA1049				
AMan957	1058010	LA1049-LA1050				
FLaiS345	1074433	LA1070 (-)	1073851-1074951	adenylate cyclase		LEPBIa2128/LBJ_0708/LBL_2371/LIC12598 intracellular signaling cascade
AMan1029	1089268	LA1085 (+)	1087834 - 1089858	ATP-dependent DNA helicase pcrA		LIC12588/LBJ_2305/LBL_0802/LEPBIa2066 Superfamily I DNA and RNA helicases
AMan1273	1105172	LA1102-LA1103				

AMan732	1107354	LA1104 (-)	1106786-1107415	chaperone DnaJ	LIC12571/LBJ_0870/LBL_0881/LEPBIa1816	
AMan970	1107361	LA1104 (-)	1106786-1107415	chaperone DnaJ	LIC12571/LBJ_0870/LBL_0881/LEPBIa1816	
AMan1004	1110793	LA1106 (+)	1108861-1110756	hypothetical protein	LIC12569/LBJ_0868/LBL_0883/LEPBIa1814	6TM Putative permease; putative membrane protei
AMan1299	1111104	LA1107 (+)	1110788 - 1111624	hypothetical protein	LIC12568/LBJ_0867/LBL_0884/LEPBIa1813	
AMan949	1116245	LA1112 (+)	1115799 - 1117808	Glycerol-3-phosphate dehydrogenase	LIC12563/LBJ_2684/LBL_0389/LEPBIa1936	
FLai-K233	1119103	LA1114 (+)	1118414-1119547	Rieske 2Fe-2S family protein	LIC12562	electron transport
AMan603	1123319	LA1119- LA1120				
AMan1170	1132875	LA1128 (-)	1132502 - 1133404	hypothetical protein	LIC12553/LBL_0384/LBJ_2689/	
AMan276	1139164	LA1135 (+)	1139061 - 1140236	hypothetical protein	LIC12546/LBL_0407/LBJ_2681/LEPBIa2723	
AMan986	1139372	LA1135 (+)	1139061 - 1140236	hypothetical protein	LIC12546/LBL_0407/LBJ_2681/LEPBIa2723	
AMan876	1140670	LA1135-LA1136				LA1136 promoter
FLaiS242	1155817	LA1151 (+)	1155459-1155734	hypothetical protein	LIC12532	
AMan1096	1156108	LA1152 (+)	1155958-1156644	hypothetical protein	LIC12531	
AMan921	1156191	LA1152 (+)	1155958-1156644	hypothetical protein	LIC12531	
AMan1219	1156476	LA1152 (+)	1155958-1156644	hypothetical protein	LIC12531	
AMan932	1162822	LA1158- LA1159				
AMan18	1163485	LA1159 (-)	1162939 - 1164147	hypothetical protein	LIC12525/LBJ_2488/LBL_0617	putative lipoprotein
FLaiB1	1166054	LA1161 (-)	1165784-1167241	Fatty acid transport / FadL	LIC12524/LBJ_2145/LBL_2139/LEPBIa0390	Outer membrane fatty-acid transport protein
AMan1184	1179781	LA1173 (+)	1177835 - 1179817	acetyl-CoA synthetase	LIC12516/LBJ_2139	Firefly luciferase-like
AMan1181	1179831	LA1173-LA1174				
AMan250	1186239	LA1180 (-)	1186084 - 1186401	hypothetical protein	LIC12509/LEPJB12766/LEPL512724	
AMan1131	1188629	LA1184 (-)	1188498 - 1190090	adenylate cyclase	LIC12506	
AMan1054	1188683	LA1184 (-)	1188498 - 1190090	adenylate cyclase	LIC12506	
FLai-K221	1190744	LA1185 (+)	1190272-1191981	Response regulator receiver domain	LIC12505	
AMan1047	1191930	LA1185 (+)	1190272-1191981	two-component response regulator	LIC12505/LBJ_4069/LBL_4069/LEPBIa0930	Response regulator receiver
AMan1216	1191933	LA1185 (+)	1190272-1191981	two-component response regulator	LIC12505/LBJ_4069/LBL_4069/LEPBIa0930	Response regulator receiver
AMan719	1194641	LA1190 (+)	1194600 - 1195541	hypothetical protein	LIC12501/LBJ_2216/LBL_2209	
FLaiPCR33	1197935	LA1192 (-)	1197671-1199242	hypothetical protein	LIC12499/LEPBIa0923	SP
FLai-S71	1207750	LA1205 (+)	1207444-1208064	rpoE1 | Probable RNA polymerase ECF-typ	LIC12490/LBL_0908/LBJ_0893/LEPBIa2150	
AMan1370	1208559	LA1206 (+)	1208045-1208632	hypothetical protein	LIC12489/LBL_0909/LBJ_0894/LEPBIa2149	
AMan767	1217568	LA1214-LA1215				
AMan714	1220635	LA1220 (+)	1220575 - 1221531	Metallo-hydrolase/oxidoreductase	LIC12478/LBJ_2127/LBL_2124/LEPBIa1009	
AMan1338	1226743	LA1223-LA1224				LA1224 promoter
FPom09	1229646	LA1224-LA1225				
AMan1162	1232331	LA1230 (-)	1232167 - 1233282	hypothetical protein	LIC12470	
FLaiK7E	1232770	LA1230 (-)	1232167-1233282	hypothetical protein	LIC12470	
AMan629	1250065	LA1251 (+)	1249643-1252855	Chemotaxis protein histidine kinase and rel	LIC12456/LBL_2106/LBJ_0927/LEPBIa2392	CheW-like protein
FLaiS6	1252285	LA1251 (+)	1249643-1252855	cheA1 | Chemotaxis protein histidine kinase	LEPBIa2392/LBJ_0927/LBL_2106/LIC12456	che operon
AMan1065	**1254060**	**LA1252 (+)**	**1252861 - 1253934**	**Chemotaxis response regulator protein-g**	**LIC12455/LBL_2105/LBJ_0928/LEPBIa23**	**Response regulator receiver**
FLaiS39	1257512	LA1257 (+)	1256712-1257629	tyrA | Prephenate dehydrogenase	LIC12450/LBL_2100/LBJ_0933/LEPBIa2386	tyrosine biosynthesis

AMan1220	1263341	LA1264 (+)	1263237 - 1264244	Fatty acid/phospholipid synthesis protein pl	LIC12443/LBL_2093/LBJ_0940/LEPBIa2591	
AMan713	1269841	LA1276 (+)	1269050-1270963	hypothetical protein	LIC12436/LBL_2011/LBJ_1023/	5TM Alkaline phosphatase-like
AMan852	1270481	LA1276 (+)	1269050-1270963	hypothetical protein	LIC12436/LBL_2011/LBJ_1023/	5TM Alkaline phosphatase-like
FLaiS327	1290944	LA1299 (-)	1290285-1292543	Ankyrin repeat proteins	LIC12419	
FLai9	1291386	LA1299 (-)	1290285-1292543	Ankyrin repeat protein	LIC12419	
FLaiS78	1295824	LA1304 (+)	1295266-1295916	hypothetical protein	LIC12416/LBJ_1035/LBL_1999/LEPBIa1413	putative membrane protein
AMan1172	1296875	LA1306 (+)	1296780 - 1296992	hypothetical protein	LIC12414/LBJ_1037/LBL_1997	
FLai39	1308319	LA1315 (-)	1308257-1308376	hypothetical protein	LBL_1490/LBJ_1266	
AMan1337	1308771	LA1317 (+)	1308724-1309281	hypothetical protein	LIC12406	
AMan628	1308879	LA1317 (+)	1308724-1309281	hypothetical protein	LIC12406	
AMan1254	1312921	LA1321-LA1322				
AMan1258	1320285	LA1326 (-)	1320187 - 1320927	hypothetical protein	LIC12399/LBJ_2269/LBL_0838/LEPBIa1037	
AMan249	1320918	LA1326 (-)	1320187 - 1320927	hypothetical protein	LIC12399/LBJ_2269/LBL_0838/LEPBIa1037	
AMan1231	1322152	LA1328 (+)	1321399-1323243	hypothetical protein	LIC12397/LBJ_2267/LBL_0840/LEPBIa1748	
AMan743	1325864	LA1332 (-)	1325508-1326491	ankyrin-like protein	LIC12393/LBJ_2265/LBL_0842	
AMan1333	1326648	LA1333 (+)	1326606-1328108	amidase	LIC12392	6-aminohexanoate-cyclic-dimer hydrolase
AMan712	1326704	LA1333 (+)	1326606-1328108	amidase	LIC12392	6-aminohexanoate-cyclic-dimer hydrolase
AMan1238	1326768	LA1333 (+)	1326606-1328108	amidase	LIC12392	6-aminohexanoate-cyclic-dimer hydrolase
AMan842	1365357	LA1372 (+)	1365106-1366335	permease	LIC12361/LBJ_0880/LBL_0870	13TM Tetracycline resistance protein
AMan1123	1365642	LA1372 (+)	1365106-1366335	permease	LIC12361/LBJ_0880/LBL_0870	13TM Tetracycline resistance protein
AMan948	1367439	LA1374 (+)	1367191-1368198	Pirin-related protein	LIC12360/LBJ_0878/LBL_0873	
AMan532	1383365	LA1389 (+)	1382816 - 1384540	hypothetical protein	LIC12349/LBL_2226/LBJ_2233/LEPBIa0891	6TM TolA/TonB C-terminal domain
AMan1297	1385880	LA1391 (+)	1385286-1386188	Ribosomal protein L11 methyltransferase (I	LIC12347/LBJ_2231/LBL_2224/LEPBIa3084	
AMan1249	1391800	LA1397 (-)	1389836-1392790	Protein export membrane protein SecD/Sec	LIC12342/LBJ_2228/LBL_2221	Predicted exporters of the RND superfamily
AMan733	1397251	LA1400 (+)	1395464-1397410	hypothetical protein	LIC12340/LBL_2503/LBJ_0577	
AMan916	1399565	LA1402 (+)	1398328-1400319	hypothetical protein	LIC12339/LBL_2503/LBJ_0577	
AMan1325	1406063	LA1406 (-)	1404927-1406255	sterol desaturase-related protein	LIC12336/LEPBIa1125/LBJ_1635/LBL_1853	8TM
AMan1199	1419045	LA1422 (-)	1418327-1423669	Serine/threonine protein kinase	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	
AMan1248	1419151	LA1422 (-)	1418327-1423669	Serine/threonine protein kinase	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	
FLaiTK108	1419830	LA1422 (-)	1418327-1423669	Serine/threonine protein kinases	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	Signal transduction protein with multiple domains
AMan1049	1420435	LA1422 (-)	1418327-1423669	Serine/threonine protein kinase	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	
AMan1040	1421124	LA1422 (-)	1418327-1423669	Serine/threonine protein kinase	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	
AMan997	1421855	LA1422 (-)	1418327-1423669	Serine/threonine protein kinase	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	
AMan1045	1423509	LA1422 (-)	1418327-1423669	Serine/threonine protein kinase	LIC12324/LBL_2179/LBJ_2185/LEPBIa2646	
AMan1051	1424128	LA1423 (-)	1423694-1424662	3-oxoacyl-[acyl-carrier-protein] synthase II	LIC12323/LBL_2178/LBJ_2184	
AMan1146	1424519	LA1423 (-)	1423694-1424662	3-oxoacyl-[acyl-carrier-protein] synthase II	LIC12323/LBL_2178/LBJ_2184	
AMan1295	1424712	LA1424 (-)	1424696-1426477	Probable succinyl-CoA:3-ketoacid-coenzyn	LIC12322/LBL_2177/LBJ_2183	
AMan1101	1425230	LA1424 (-)	1424696-1426477	Probable succinyl-CoA:3-ketoacid-coenzyn	LIC12322/LBL_2177/LBJ_2183	
AMan1143	1425409	LA1424 (-)	1424696-1426477	Probable succinyl-CoA:3-ketoacid-coenzyn	LIC12322/LBL_2177/LBJ_2183	
AMan705	1425588	LA1424 (-)	1424696-1426477	Probable succinyl-CoA:3-ketoacid-coenzyn	LIC12322/LBL_2177/LBJ_2183	

AMan1135	1425650	LA1424 (-)	1424696-1426477	Probable succinyl-CoA:3-ketoacid-coenzyn	LIC12322/LBL_2177/LBJ_2183	
AMan6	1427628	LA1425 (+)	1426897-1427760	hypothetical protein	LIC12321/LBL_2176/LBJ_2182	
AMan1022	1428766	LA1427 (+)	1428115-1428885	3-oxoacyl-(acyl carrier protein) reductase	LIC12320/LBL_2175/LBJ_2181/LEPBIa2647	
FLaiK148	1429455	LA1428 (+)	1429583-1430392	protein phosphatase	LIC12319/LBL_2174/LBJ_2180	
AMan871	1429649	LA1428 (+)	1429583 - 1430392	protein phosphatase	LIC12319/LBL_2174/LBJ_2180	
AMan1039	1431020	LA1430 (-)	1430905 - 143220(	3-oxoacyl-[acyl-carrier-protein] synthase	LIC12317/LBL_2172/LBJ_2178	
AMan1032	1431037	LA1430 (-)	1430905 - 143220(	3-oxoacyl-[acyl-carrier-protein] synthase	LIC12317/LBL_2172/LBJ_2178	
AMan1050	1432069	LA1430 (-)	1430905 - 143220(	3-oxoacyl-[acyl-carrier-protein] synthase	LIC12317/LBL_2172/LBJ_2178	
FFC6	1433973	LA1432-LA1433				LA1432 promoter
FLaiTK101	1434970	LA1433 (-)	1434255-1435568	hypothetical protein	LIC12315	putative lipoprotein/LenD
AMan934	1440952	LA1440 (+)	1440832-1442391	hypothetical protein	LIC12309/LBL_2164/LBJ_2170/LEPBIa1499	TPR repeat
FPom33	1448159	LA1446 (+)	1447019-1448200	Cobalt-zinc-cadmium resistance protein CZ	LBJ_2132/LBL_2129/LIC12306/LEPBIa1315	
FLaiK144	1452371	LA1451 (+)	1451589-1452353	psuA 1Phosphatidylglycerophosphate synth	LIC12301/LEPBIa1320/LBJ_1981/LBL_1067	phospholipid biosynthesis
AMan652	1455887	LA1455 (+)	1455724-1455984	hypothetical protein	LIC12298/LBJ_0955/LBL_2078/LEPBIa0934	
AMan747	1467917	LA1466-LA1467				
FLaiS52	1472754	LA1471 (+)	1471781-1473895	Pyrophosphate-energized vacuolar membra	LIC12285/LBJ_0973/LBL_2060/LEPBIa3448	Inorganic pyrophosphatase
AMan980	1479811	LA1478-LA1479				LA1478 promoter
AMan069	1482020	LA1481-LA1482				
AMan837	1494128	LA1494 (+)	1494125-1494358	hypothetical protein	LIC12264	
AMan1224	**1498966**	**LA1499 (-)**	**1498207-1501698**	**hypothetical protein**	**LIC12259/LBL_4283/LBJ_0994/LEPBIa14**	**Integrin alpha N-terminal domain**
FLaiS275	1499629	LA1499 (-)	1498207-1501698	hypothetical protein	LIC12259/LBJ_0994/LBL_4283/LEPBIa1446	Integrin alpha N-terminal domain
AMan140	1500208	LA1499 (-)	1498207-1501698	hypothetical protein	LIC12259/LBL_4283/LBJ_0994/LEPBIa1446	Integrin alpha N-terminal domain
FKitGK36	1508885	LA1507 (+)	1507737-1509305	hypothetical protein	LIC12254/LBL_1930/LBJ_1711/LEPBIa1510	SPI outer membrane protein /Bacterial surface anti
AMan1347	1525124	LA1523 (-)	1524677 - 1525192	lipoprotein	LIC12238/LBJ_1696/LBL_1915/LEPBIa0490	
AMan1160	1526005	LA1526 (-)	1525564-1526004	hypothetical protein	LIC12236/LBJ_1694/LBL_1913/LEPBIa2312	CBS domain protein
FLaiS313	1526632	LA1528 (+)	1526376-1527842	two-component response regulator	LBJ_1693/LBL_1912/LIC12235/LEPBIa0081	Signal transduction histidine kinase
AMan1267	1528608	LA1530 (-)	1528407-1528910	hypothetical protein	LIC12234	Leucine-rich repeat
AMan831	1529925	LA1531 (-)	1529847 - 1529981	hypothetical protein		
AMan1118	1542219	LA1543 (-)	1541152 - 1544448	Putative cation efflux system protein	LIC12224/LBJ_1796/LBL_1078/LEPBIa3142	12TM
AMan1121	1542266	LA1543 (-)	1541152 - 1544448	Putative cation efflux system protein	LIC12224/LBJ_1796/LBL_1078/LEPBIa3142	12TM
AMan1182	1543396	LA1543 (-)	1541152 - 1544448	Putative cation efflux system protein	LIC12224/LBJ_1796/LBL_1078/LEPBIa3142	12TM
FLaiS38	1547461	LA1548 (-)	1547159-1549411	napA 1Na(+)/H(+) antiporter	LIC12221/LBJ_1793/LBL_1081/LEPBIa0085	cation:proton antiporter
AMan1320	1562818	LA1567 (-)	1562735 - 1563499	hypothetical protein	LIC12209/LBL_1163/LBJ_1109	putative lipoprotein
AMan1102	1591382	LA1599 (+)	1590161 - 1591414	hypothetical protein	LIC12183	
AMan1255	1613714	LA1618-LA1619				LA1619 promoter
FLaiPCR48	1613893	LA1619 (+)	1613744-1615552	probable carbamoyl transferase	LIC12163/LBJ_1117/LBL_1171	
AMan895	**1638223**	**LA1641 (+)**	**1637205 - 1638494**	**hypothetical protein**	**LIC12143**	**11TM integral membrane protein**
AMan943	1670282	LA1669 (+)	1670065-1670499	hypothetical protein	LIC12116	
AMan730	1670719	LA1670 (+)	1670557 - 1670736	hypothetical protein		
AMan641	1673472	LA1674 (+)	1673131 - 167436(	hypothetical protein	LIC12114/LBJ_1192/LBL_1244/LEPBIa2102	Putative enzyme of poly-gamma-glutamate biosynt

AMan1113	1682431	LA1683 (+)	1680107 - 1682521	hypothetical protein	LIC12106/LEPBIa2093/LBJ_1200/LBL_1252	6TM Putative metal-dependent phosphohydrolase,
AMan1221	1684402	LA1687 (+)	1683986 - 168440	hypothetical protein	LIC12103	
AMan1035	1691975	LA1693 (-)	1691757 - 16928	Sensory transduction histidine kinase	LIC12097/LEPBa2079/LBJ_1215/LBL_1266	Histidine kinase
AMan1179	1693347	LA1694 (-)	1692897 - 1693553	hypothetical protein	LIC12096/LEPBIa2084/LBJ_1216/LBL_1267	
AMan709	1701702	LA1705-LA1706				
AMan1359	1704886	LA1711 (-)	1704173-170489	histidine kinase sensor protein	LIC12087/LBJ_1223/LBL_1274	PAS/PAC domain containing protein
AMan372	1706294	LA1713-LA1715				
AMan646	1714389	LA1724 (-)	1713813-1714625	ABC transporter ATP-binding protein	LIC12079/LBJ_1229/LBL_1280/LEPBIa2866	
AMan1117	1714988	LA1726 (-)	1714622-1716199	hypothetical protein	LIC12078/LBJ_1230/LBL_1281/LEPBIa2865	Metal-dependent phosphohydrolase, HD subdomair
FLaiS285	1720191	LA1733 (+)	1720116-1721279	hypothetical protein	LBJ_1237/ LBL_1288 /LIC12071	
FPom03	1731077	LA1741 (-)	1730975-1731724	acyltransferase	LIC12063/LBJ_1247/LBL_1296	
AMan1264	1736097	LA1745 (+)	1733955-1737578	two-component hybrid sensor and regulator	LIC12060/LBJ_1250/LBL_1299/	ATPase domain of HSP90 chaperone/DNA topoiso
AMan1372	1767170	LA1770 (-)	1766629-1767801	Putative transcriptional regulator, AraC fam	LBL_0695/LBJ_2414/	6TM
FLaiS280	1767368	LA1770 (-)	1766629-1767801	Putative transcriptional regulator, AraC fam	LBJ_2414 /LBL_0695	Genomic island
AMan722	1768439	LA1772 (+)	1768259-1768684	Putative regulator, lambda repressor-like	LBJ_0614/LBL_2465	
AMan1098	1772138	LA1775-LA1776				
FLaiS168	1774655	LA1779 (+)	1773866-1774657	hypothetical protein		SP, Genomic island
AMan803	1805798	LA1828-LA1829				
FLai380	1806280	LA1828-LA1829				
FLaiPCR49	1827938	LA1856- LA1857				
AMan776	**1828342**	**LA1857 (-)**	**1827966-1828403**	**Ferric uptake regulation protein**	**LIC12034/LBL_1818/LBJ_1600/LEPBIa2461**	
FPom43	1829320	LA1858 (+)	1828892-1829437	ankyrin like protein	LIC12033/LBL_1817/LBJ_1599	
AMan468	1829332	LA1858 (+)	1828892-1829437	ankyrin like protein	LIC12033/LBL_1817/LBJ_1599	
AMan1142	1830986	LA1859-LA1860				
FLai14	1831539	LA1860 (+)	1831286-1833505	two-component hybrid sensor and regulator	LIC12031	hybrid sensor and regulator
AMan1108	1831944	LA1860 (+)	1831286-1833505	two-component hybrid sensor and regulator	LIC12031	
AMan1288	1840045	LA1866 (+)	1839820-1841127	Asparaginyl-tRNA synthetase (Asparagine-	LIC12025/LBJ_1650/LBL_1869/LEPBIa1722	
AMan1275	1840610	LA1866 (+)	1839820-1841127	Asparaginyl-tRNA synthetase (Asparagine-	LIC12025/LBJ_1650/LBL_1869/LEPBIa1722	
FLaiS384	1848571	LA1875 (-)	1848186 - 1850162	ATP-dependent DNA helicase	LIC12019/LEPBIa1719/LBJ_1613/LBL_1831	
AMan998	1852876	LA1879 (+)	1852280-1854862	Chaperone protein clpB	LIC12017/LBJ_1611/LBL_1829/LEPBIa2574	
AMan855	1860143	LA1885 (-)	1859656 - 1860318	hypothetical protein	LIC12013/LBJ_1511/LBL_1735	4TM
Aman1303	1867996	LA1894 (+)	1867854-1868303	hypothetical protein	LIC12005	
AMan1087	1876992	LA1905 (+)	1876606-1877943	hypothetical protein	LIC11996	
FLaiS263	1878905	LA1908(-)	1878626-1879072	hypothetical protein	LIC11994/LBJ_1498/LBL_1722	
FLaiK15E	1884488	LA1915 (-)	1882896-1886495	TPR-repeat-containing proteins	LIC11990/ LBJ_1494/ LBL_1718 /LEPBIa2964	
FKitGK17	1888609	LA1918 (+)	1888546-1888815	hypothetical protein	LIC11987	
AMan902	1891061	LA1919-LA1920				
AMan823	1891909	LA1919-LA1920				
AMan721	1893134	LA1921-LA1922				
AMan853	1893812	LA1922 (+)	1893183-1893920	hypothetical protein	LIC11983/LBL_1811/LBJ_1593	protein with MORN repeat

AMan1025	1894053	LA1922-LA1924				
FLai53	1900702	LA1929 (+)	1898684-1900843	Cyclic nucleotide binding protein /Cgs	LIC11977/LBJ_1584/LBL_1802/LEPBIa2301	CAP family transcription factor
AMan465	1901770	LA1930 (+)	1900861-1901991	Acyl-CoA dehydrogenase	LIC11976/LBJ_1583/LBL_1801/LEPBIa2302	
AMan952	1907104	LA1936 (-)	1907103-1908974	hypothetical protein	LIC11970/LEPBIa3301/LBJ_1577/LBL_1795	alpha/beta-Hydrolases
AMan1031	1908915	LA1936 (-)	1907103-1908974	hypothetical protein	LIC11970/LEPBIa3301/LBJ_1577/LBL_1795	alpha/beta-Hydrolases
AMan982	1912904	LA1940-LA1941				
AMan1259	1918155	LA1948 (-)	1917732-1918220	two-component response regulator	LIC11957/LBJ_1564/LBL_1788/LEPBIa0930	Response regulator receiver
AMan737	1919587	LA1949 (-)	1918291-1919403	hypothetical protein	LIC11956/LBJ_1563/LBL_1787/LEPBIa2863	
AMan1218	1923655	LA1954 (+)	1923576-1923932	hypothetical protein	LIC11950/LBJ_1558/LBL_1782	
FPom59	1924727	LA1955 (+)	1923956-1924966	hydrolase or acyltransferase	LIC11949/LBJ_1557/LBL_1781/LEPBIa1707	alpha/beta hydrolase superfamily
AManp	1924833	LA1955 (+)	1923956-1924966	Putative hydrolase	LIC11949/LBJ_1557/LBL_1781/LEPBIa1707	alpha/beta-Hydrolases
AMan141	1930600	LA1961 (+)	1930526-1930963	hypothetical protein	LIC11943/LBJ_1551/LBL_1775/LEPBIa2853	Integral outer membrane protein TolC, efflux pump
FLaiA2	1930761	LA1961 (+)	1930526-1930963	hypothetical protein	LIC11943/LBJ_1551/LBL_1775/LEPBIa2853	
ALaiC19	1930954	LA1961 (+)	1930526-1930963	hypothetical protein	LIC11943/LBL_1775/LBJ_1551/LEPBIa2853	
AManc2	1932298	LA1963 (+)	1931541-1932995	heavy metal efflux pump	LIC11941/LBJ_1549/LBL_1773/LEPBIa0575	
AMan1175	1941333	LA1967 (+)	1938620-1941940	cation efflux system protein	LIC11937/LBJ_1546/LBL_1770/LEPBIa0572	12TM
FLai-S42	1948177	LA1973 (-)	1947965-1948282	hypothetical protein	LIC11931/LBJ_1539/LBL_1763	putative membrane protein
AMan951	1955888	LA1980-LA1981				
AMan605	1963321	LA1983- LA1985				
AMan622	1976293	LA1997 (-)	1974104-1976284	glycosyltransferase	LIC11911/LBJ_1521/LBL_1745	7TM
AMan1154	1981319	LA2003 (+)	1981077-1982321	hypothetical protein	LIC11904/LBJ_1973/LBL_1311	lipoprotein, down-regulated at physiologic osmolar
FLai-S34	1981703	LA2003 (+)	1981077-1982321	hypothetical protein	LIC11904/LBJ_1973/LBL_1311	lipoprotein, down-regulated at physiologic osmolar
AMan886	**1982261**	**LA2003 (+)**	**1981077-1982321**	**hypothetical protein**	**LIC11904/LBJ_1973/LBL_1311**	**lipoprotein, down-regulated at physiologic osmo**
AMan1192	1993546	LA2016 (+)	1993425-1993799	hypothetical protein	LIC11891	
AMan1053	1994502	LEPIN1786 (+)	1994349-1994507	hypothetical protein	LEPIC1930	
AMan779	2000525	LA2023-LA2024				
AMan519	2001560	LA2024 (-)	2001025 -2002263	hypothetical protein	LIC11885/LBJ_1958/LBL_1326/LEPBIa2117	putative lipoprotein
FLaiS299	2002382	LA2025 (+)	2002293-2003591	hypothetical protein	LBJ_1957 LEPBIa2116 LBL_1327 LIC11884	Bacterial extracellular solute-binding protein, famil
FLai-S57	2036148	LA2062 (-)	2034932-2036236	metY | O-acetylhomoserine sulfhydrylase	LIC11852/LEPBIa1593	
AMan723	2036224	LA2062 (-)	2034932 -2036236	metY | O-acetylhomoserine sulfhydrylase	LIC11852/LEPBIa1593	
AMan1266	2036699	LA2063 (+)	2036536-2038179	hypothetical protein	LIC11851/LBJ_1943/LBL_1341/LEPBIa1594	impL63 / cytoplasmic membrane protein
AMan1095	2042293	LA2069 (+)	2041972-2042289	Flagellar motor switch protein fliN	LIC11846/LBJ_1938/LBL_1346/LEPBIa1599	
FLai-K237	2042301	LA2069-LA2070				
FKit44	2042822	LA2071 (-)	2042764-2043696	hypothetical protein	LIC11845/LBL_1348/LBJ_1936/LEPBIa1602	Membrane protease subunits, stomatin/prohibitin h
FPom13	2043533	LA2071 (-)	2042764-2043696	hypothetical protein	LIC11845/LBL_1348/LBJ_1936/LEPBIa1602	Membrane protease subunits, stomatin/prohibitin h
AMan1215	2044561	LA2072 (-)	2043698-2044645	hypothetical protein	LIC11844/LBJ_1935/LBL_1349/LEPBIa1603	Membrane protease subunits, stomatin/prohibitin h
AMan1322	2044588	LA2072 (-)	2043698-2044645	hypothetical protein	LIC11844/LBJ_1935/LBL_1349/LEPBIa1603	Membrane protease subunits, stomatin/prohibitin h
FKit38	2044628	LA2072 (-)	2043698-2044645	hypothetical protein	LIC11844/LBJ_1935/LBL_1349/LEPBIa1603	Membrane protease subunits, stomatin/prohibitin h
FKit12	2044992	LA2073-LA2074				
FLaiK-224	2052062	LA2082-LA2083				LA2083= LipL45 (-52),down-regulated at physiolo

AMan1081	2052434	LA2083 (+)	2051826-2052536	hypothetical protein	LIC11834/LBJ_1926/LBL_1358	putative lipoprotein
AMan985	2053178	LA2084 (-)	2052915-2053205	hypothetical protein	LIC11833/LBJ_1925/LBL_1359/LEPBIa1613	3TM, MtN3 and saliva related transmembrane prot
AMan1191	2056477	LA2088 (-)	2056439-2057110	hypothetical protein	LIC11830/LBJ_1922/LBL_1362	
AMan941	2067325	LA2097 (+)	2067003-2068004	cysteine synthase B	LIC11820/LBJ_1911/LBL_1373/LEPBIa1628	
AMan1315	2067533	LA2097 (+)	2067003-2068004	cysteine synthase B	LIC11820/LBJ_1911/LBL_1373/LEPBIa1628	
FLaiS170	2074494	LA2104 (+)	2073650-2075890	clpA | ATPases with chaperone activity, AT	LIC11814/LBJ_1905 /LBL_1379/LEPBIa1852	
FLaiPCR32	2082217	LA2111-LA2112				
AMan1130	2087355	LA2117 (-)	2087204-2087518	Anti-sigma B factor antagonist	LIC11801/LBJ_1893/LBL_1391/LEPBIa1741	
AMan1300	2087452	LA2117 (-)	2087204-2087518	Anti-sigma B factor antagonist	LIC11801/LBJ_1893/LBL_1391/LEPBIa1741	
AMan917	2087453	LA2117 (-)	2087204-2087518	Anti-sigma B factor antagonist	LIC11801/LBJ_1893/LBL_1391/LEPBIa1741	
FLaiK10B	2092588	LA2122 (-)	2091947-2092927	putative sigma-B regulator	LIC11796 LBJ_1888 LBL_1396	Sigma factor PP2C-like phosphatase
FLaiTK103	2096711	LA2127 (+)	2096320-2097225	hypothetical protein	LIC11792/LBL_140/LBJ_1884/LEPBa1014	Galactose mutarotase and related enzymes
FCan13	2096813	LA2127 (+)	2096320-2097225	hypothetical protein	LIC11792/LBL_1400/LBJ_1884/LEPBa1014	Galactose mutarotase and related enzymes
AMan1033	2104454	LA2133 (-)	2102914-2104536	sodium:solute symporter family protein	LEPBIa1109/LIC11787/LBJ_1879/LBL_1405	12TM Na+/solute symporter
FLaiPCR45	2107054	LA2137 (-)	2106766-2107071	hypothetical protein	LIC11783/LBJ_1876 /LBL_1408/LEPBIa1142	
FKit41	2118243	LA2146 (-)	2115765-2118311	MutS proteinDNA mismatch repair protein	LIC11774/LBJ_1867/LBL_1417/LEPBIa0833	Mismatch repair ATPase (MutS family)
AMan1114	2119481	LA2149 (+)	2119398-2119907	hypothetical protein	LIC11771	
FLaiK149	2129949	LA2162 (+)	2129703-2129987	rpmB | ribosomal protein L28	LIC11760/LBJ_1857/LBL_1427/LEPBIa1020	
AMan1226	2137795	LA2168 (+)	2137054-2138364	hypothetical protein	LEPIC1793/LBJ_1851/LBL_1433/LEPBIa2074	
AMan826	2143035	LA2172 (+)	2141722-2143032	glycosyl transferase	LIC11752/LBJ_1847/LBL_1437/LEPBIa1000	
AMan1306	2145337	LA2175 (+)	2145242-2147023	hypothetical protein	LEPBIa0156/LIC11749/LBJ_1843/LBL_1441	
AMan1093	2163471	LA2186 (+)	2157784-2163588	hypothetical protein	LEPBIa1279/LBJ_1350/LBL_1575/LIC11739	
AMan1027	2163773	LA2187 (+)	2163751-2165880	Penicillin-binding protein 1F	LEPBIa0534/LBJ_1351/LBL_1576/LIC11738	
AMan703	2171793	LA2193 (-)	2170218-2172236	ATP-dependent DNA helicase uvrD	LIC11732/LBJ_1356/LBL_1581/LEPBIa1736	
AMan726	2173073	LA2195 (-)	2172771-2173124	hypothetical protein	LEPIC1767/LBJ_1358/LBL_1583/LEPBIa1734	
AMan1198	2173117	LA2195 (-)	2172771-2173124	hypothetical protein	LEPIC1767/LBJ_1358/LBL_1583/LEPBIa1734	
FLaiS324	2178141	LA2197 (-)	2176665-2178674	fadH | 2,4-dienoyl-CoA reductase (NADPH	LEPBIa2481/LBJ_1360/LBL_1585/LIC11729	
FLaiS340	2178146	LA2197 (-)	2176665-2178674	fadH | 2,4-dienoyl-CoA reductase (NADPH	LEPBIa2481/LBJ_1360/LBL_1585/LIC11729	
FCan12	2185118	LA2207 (-)	2184830-2185606	hypothetical protein	LIC11721/LEPBIa1841/LBJ_1399/LBL_1625	putative membrane protein (6TMS)
AMan708	2190709	LA2211 (-)	2189875-2191614	signal peptide peptidase sppA	LIC11718/LEPBIa1838/LBL_1628/LBJ_1402	Peptidase
AMan608	2191684	LA2211-LA2212				
FLaiTK75	2193492	LA2212 (+)	2191724-2194570	uvrA | Excinuclease ABC subunit A	LIC11717/LEPBIa1837/LBJ_1403/LBL_1629	
AMan1083	2194727	LA2213 (+)	2194662-2195636	hypothetical protein	LIC11716/LEPBIa1836/LBJ_1404/LBL_1630	Acyl-CoA dehydrogenase, middle and N-terminal
AMan1014	**2197023**	**LA2215 (-)**	**2196745-2197764**	**Chemotaxis motB protein**	**LIC11713/LBJ_1365/LBL_1590/LEPBIa1834**	
AMan1066	2197447	LA2215 (-)	2196745-2197764	Chemotaxis motB protein	LIC11713/LBJ_1365/LBL_1590/LEPBIa1834	
AMan1190	2204347	LA2222 (+)	2203802-2206300	two-component hybrid sensor and regulator	LIC11709/LBJ_1368/LBL_1593	
AMan11917	2207150	LA2223 (+)	2206339-2207487	two-component hybrid sensor and regulator	LIC11708/LBJ_1369/LBL_1594	
FFCJ2	2209628	LA2227 (+)	2209464-2209907	hypothetical protein	LIC11706/LBL_1596/LBJ_1371/LEPBIa1563	
AMan1222	2214860	LA2232-LA2235				
AMan1207	2218419	LA2238 (+)	2218290-2219486	polysaccharide deacetylase	LIC11697/LBJ_1382/LEPBIa1571/LBL_1607	

FPom23	2226794	LA2246 (+)	2226654-2228162	Methyl-accepting chemotaxis protein	LIC11691/LBL_0142/LBJ_2921/LEPBIa0109	Putative methyl-accepting chemotaxis protein; puta
FLai-S51	2230515	LA2249 (+)	2229957-2230838	hypothetical protein	LIC11688	
AMan1171	2231457	LA2250 (-)	2231152-2232072	Nuclease S1	LIC11687/LBJ_1405/LBL_1631	SP
AMan841	2248985	LA2267 (-)	2248192-2250180	hypothetical protein	LIC11670/LBJ_1386/LBL_1612/LEPBIa2104	Putative TPR-repeat-containing protein
AMan1349	2259361	LA2280-LA2281				
AMan1209	2259432	LA2280-LA2281				LA2281 promoter
AMan981	2260066	LA2282 (+)	2259613-2260239	Phosphoribosylglycinamide formyltransferi	LIC11656/LBJ_1330/LBL_1555/LEPBIa1470	
FLaiS291	2260207	LA2282 (+)	2259613-2260239	purN | Phosphoribosylglycinamide formyltr	LEPBIa1470 LBJ_1330 LBL_1555 LIC11656	
AMan1116	2261968	LA2284 (+)	2261781-2262407	hypothetical protein	LIC11654/LBJ_1328/LBL_1553/LEPBIa1468	5TM putative membrane protein
FLaiTK63	2262299	LA2284 (+)	2261781-2262407	hypothetical protein	LIC11654/LBJ_1328/LBL_1553/LEPBIa1468	5TM putative membrane protein
FLaiTK8	2262479	LA2285 (+)	2262438-2263169	hypothetical protein	LIC11653/LBJ_1327/LBL_1552/LEPBIa1467	SP
FLaiS266	2268549	LA2291(-)	2268055-2268795	rluB | ribosomal large subunit pseudouridin	LIC11647/LBJ_1323/LBL_1548/LEPBIa1463	
FLai31	2284270	LA2307 (+)	2283350-2285182	Phosphotransferase system, enzyme I	LIC11632/LBJ_1314/LBL_1539/LEPBIa1869	
FLai-K232	2284619	LA2307 (+)	2283350-2285182	Phosphotransferase system, enzyme I	LIC11632/LBJ_1314/LBL_1539/LEPBIa1869	
AMan1336	2290180	LA2311-LA2314				
AMan973	2300455	LA2321 (-)	2299250-2300956	DNA repair protein RecN	LIC11620/LBJ_1302/LBL_1527/LEPBIa1630	
AMan1223	2301306	LA2321-LA2322				
FLaiS352	2311859	LA2332-LA233				
AMan271	2322037	LA2345 (-)	2321603-2323042	ATP-dependent hsl protease ATP-binding s	LIC11601/LBJ_1284/LBL_1509/LEPBIa2354	
AMan1253	2332744	LA2356 (+)	2332238-2333506	Exonuclease VII, large subunit	LEPBIa1860/LBJ_1274/LBL_1499/LIC11591	
AMan762	2354054	LA2378-LA2379				
FLaiS391	2355004	LA2379 (+)	2354228-2355712	hypothetical protein	LIC11566/LBJ_1425/LBL_1649/LEPBIa1671	Mg chelatase-related protein
FLaiTK102	2357731	LA2383 (-)	2357091-2358275	hypothetical protein	LIC11563/LBJ_1428/LBL_1652 /LEPBIa1668	Putative metal-dependent phosphohydrolase
FLaiS295	2358191	LA2383 (-)	2357091-2358275	hypothetical protein	LBJ_1428 LBL_1652 LIC11563 LEPBIa1668	Putative metal-dependent phosphohydrolase
FLaiPCR34	2358574	LA2385 (-)	2358536-2359279	hypothetical protein	LIC11561/LBJ_1430/LBL_1654	
AMan1301	2386028	LA2417 (+)	2385507-2386364	Possible hook-associated protein, FlaB	LIC11532/LBJ_1259/LBL_1484/LEPBIa1872	flagellin family
FLai-S72	2388183	LA2418-LA2419				
FCan15	2393941	LA2422 (+)	2392707-2393987	two-component response regulator	LIC11527/LBL_1474/LBJ_1809/LEPBIa1586	Response regulator receiver /Putative transcriptiona
FKit28	2393941	LA2423 (+)	2393950-2394363	two-component response regulator	LIC11526/LBL_1473/LBJ_1810/LEPBIa1585	Response regulator receiver Chemotaxis protein Ch
AMan477	2402433	LA2430 (-)	2401934-2402068	hypothetical protein		
FLaiS292	2408044	LA2435 (-)	2407557-2410373	Sigma factor sigB regulation protein rsbU	LBJ_1820 LBL_1463 LEPBIa1575 LIC11515	phosphatase
AMan1237	2411046	LA2437 (+)	2410986-2411666	hypothetical protein	LIC11512	
FLaiTK96	2412974	LA2439 (+)	2412225-2415467	heavy metal efflux pump	LIC11510/LBL_1459/LBJ_1824/LEPBIa1572	11TM
AMan881	2413553	LA2439 (+)	2412225-2415467	heavy metal efflux pump	LIC11510/LBL_1459/LBJ_1824/LEPBIa1572	11TM
FLaiS306	2413694	LA2439 (+)	2412225-2415467	heavy metal efflux pump	LIC11510/LBL_1459/LBJ_1824/LEPBIa1572	11TM
AMan889	2416832	LA2440-LA2441				
FLaiS342	2420871	LA2443 (-)	2420583-2422163	hypothetical protein	LIC11507	SP, Leucine-rich repeat
AMan159	2421204	LA2443 (-)	2420583-2422163	hypothetical protein	LIC11507	SP, Leucine-rich repeat
AMan409	2421703	LA2443 (-)	2420583-2422163	hypothetical protein	LIC11507	SP, Leucine-rich repeat
AMan1211	2421770	LA2443 (-)	2420583-2422163	hypothetical protein	LIC11507	SP, Leucine-rich repeat

AMan1021	2434826	LA2458-LA2459				
AMan1048	2434916	LA2458-LA2459				
FLaiS252	2440675	LA2462 (+)	2439818-2440756	bcrA | Bacitracin transport ATP-binding pro	LIC11495/LBJ_1463/LBL_1687/LEPBIa0504	gldA/ ABC transporter ATP-binding protein /glidin
AMan780	**2448336**	**LA2469 (+)**	**2448158-2448622**	**chemotaxis protein CheX**	**LIC11488/LBJ_1470/LBL_1694/LEPBIa2661**	
AMan1043	2450250	LA2471 (-)	2449877-2450872	hypothetical protein	LIC11486/LBL_1695/LBJ_1471/LEPBIa2662	TPR-like
AMan1269	2450200	LA2471 (-)	2449877-2450872	hypothetical protein	LIC11486/LBL_1695/LBJ_1471/LEPBIa2662	
FKit58	2459788	LA2487-LA2488				
FLaiS293	2498369	LA2519 (-)	2498414-2498917	hypothetical protein	LIC11447 LBL_1889 LBJ_1670	
FLaiS335	2498369	LA2518- LA2519				
AMan808	2500985	LA2523-LA2524				
AMan869	2506542	LA2528 (-)	2506149-2507141	Putative GGDEF/response regulator receive	LIC11444/LBJ_1672/LBL_1891/pLEPBI0053	
AMan979	2506809	LA2528 (-)	2506149-2507141	Putative GGDEF/response regulator receive	LIC11444/LBJ_1672/LBL_1891/pLEPBI0053	
AMan920	2507719	LA2530 (+)	2507724-2508182	hypothetical protein	LIC11442/LBJ_1674/LBL_1893/LEPBIb0077	
AMan1134	2511247	LA2532 (-)	2509850-2511301	two-component hybrid sensor and regulator	LEPBIa2597/LBJ_1676/LBL_1895/LIC11440	
AMan1141	2513991	LA2534-LA2535				
FLaiS41	2514202	LA2535 (-)	2514161-2514619	ppiA | Probable peptidyl-prolyl cis-trans iso	LIC11438/LBJ_1679/LBL_1909/LEPBIa2121	protein folding
FPom6	2514248	LA2535 (-)	2514161-2514619	Probable peptidyl-prolyl cis-trans isomerase	LIC11438/LBL_1909/LBJ_1679/LEPBIa2121	
AMan817	2514890	LA2536 (+)	2514820-2517702	Adenylate cyclase	LIC11437/LBJ_1680/LBL_1908/LEPBIa0546	
AMan1151	2523622	LA2540 (-)	2523475-2524944	two-component hybrid sensor and regulator	LIC11433/LBJ_1684/LBL_1904/LEPBIa2442	
FLai-K215	2523939	LA2540 (-)	2523475-2524944	two-component hybrid sensor and regulator	LIC11433/LBL_1904/LBJ_1684	
AMan1314	2528320	LA2542 (-)	2528310-2528786	probable cyclic nucleotide binding protein	LEPBIa2338/LBJ_1686/LBL_1902/LIC11431	cAMP-binding domain-like 9
FLaiS311	2536598	LA2551 (-)	2536337-2536732	glycine-rich RNA binding protein	LEPBIa1541/LBJ_1691/LBL_1897/LIC11423	RNA-binding proteins
FKit170	2537290	LA2552 (+)	2536843-2537889	predicted kinase	LIC11422/LBJ_1088/LBL_1145	
AMan1100	2545268	LA2560-LA2562				
AMan1129	2545464	LA2562 (-)	2545411-2546043	acyl-CoA thioesterase	LIC11414/LBJ_1081/LBL_1138/LEPBIa0966	
FPom28	2552725	LA2570 (-)	2551653-2553431	acetolactate synthase large subunit Ilv	LIC11409/LBJ_1076/LBL_1133/LEPBIa2362	
AMan746	2554021	LA2572 (+)	2553792-2554901	TPR-repeat-containing protein	LIC11408/LBJ_1075/LBL_1132/LEPBIa2363	
AMan340	2554365	LA2572 (+)	2553792-2554901	TPR-repeat-containing protein	LIC11408/LBJ_1075/LBL_1132/LEPBIa2363	
AMan1313	2554736	LA2572 (+)	2553792-2554901	TPR-repeat-containing protein	LIC11408/LBJ_1075/LBL_1132/LEPBIa2363	
AMan787	2558127	LA2575 (+)	2558071-2558340	hypothetical protein	LIC11406/LBJ_1073/LBL_1130/LEPBIa0883	
AMan887	2558288	LA2575 (+)	2558071-2558340	hypothetical protein	LIC11406/LBJ_1073/LBL_1130/LEPBIa0883	
AMan34	2567587	LA2582 (+)	2567456-2568598	Peptidase family M23/M37	LIC11399/LBL_1123/LBJ_1066	Membrane proteins related to metalloendopeptidase
AMan1327	2572583	LA2587-LA2588				
FLaiS44	2578764	LA2592 (+)	2578811-2580175	fliI | flagellum-specific ATP synthase fliI	LIC11391/LEPBIa0955/LBJ_1054/LBL_1115	
AMan736	2583541	LA2597-LA2598				
AMan1341	2590585	LA2605 (-)	2589985-2590899	minD-related ATP-binding protein	LIC11378l4:1/LBJ_1622/LBL_1840/LEPBIa	ATPases involved in chromosome partitioning
ALaiC24	2590585	LA2605 (-)	2589985-2590899	minD-related ATP-binding protein	LIC11378l4:1/LBJ_1622/LBL_1840/LEPBIa	ATPases involved in chromosome partitioning
AMan1125	2603843	LA2617 (+)	2603634-2604500	hypothetical protein	LIC11366/LBJ_1633/LBL_1851/LEPBIa2555	
AMan794	2607915	LA2624 (+)	2607745 -2607991	hypothetical protein	LIC11362/LBL_1856/LEPBIa2556	
AMan933	2620244	LA2637 (+)	2620070-2620888	LipL32 protein	LIC11352/LBJ_1647/LBL_1866	

AMan1326	2624015	LA2639 (+)	2622335-2624095	acyl-CoA dehydrogenase	LIC11350/LBJ_1648/LBL_1867/LEPBIa1259	
AMan1323	2626506	LA2641 (+)	2625561-2627801	TonB-dependent receptor	LIC11345/LBL_1111/LBJ_1058/	
AMan146	2632820	LA2649 (-)	2632602-2634152	Deoxyribodipyrimidine photolyase	LIC11339/LBL_1105/LBJ_1769/LEPBIa2409	
AMan1000	2637474	LA2653-LA2654				
AMan1104	2637992	LA2653-LA2654				
FLaiS163	2691330	LA2711 (-)	2690674-2692608	hypothetical protein	LIC11295	12TM putative membrane protein
FLaiK139	2691489	LA2711 (-)	2690674-2692608	hypothetical protein	LIC11295	12TM putative membrane protein
AMan1256	2697792	LA2715 (-)	2697281-2698225	Sensory transduction histidine kinase	LIC11292/LBL_1032/LBJ_2018	
AMan275	2698357	LA2715 (-)	2697281-2698225	Sensory transduction histidine kinase	LIC11292/LBL_1032/LBJ_2018	
AMan801	2705971	LA2723 (-)	2705381-2706211	hypothetical protein	LIC11287/LEPBIa1260	Putative metal-dependent hydrolase
AMan914	2707127	LA2725-LA2727				
AMan1345	2709094	LA2728 (+)	2708865-2709284	hypothetical protein	LIC11283/LBJ_1723/LBL_1942	
AMan897	2718985	LA2737 (-)	2718397-2719194	acyltransferase	LIC11276/LBJ_1714/LBL_1933/LEPBIa0552	
FLaiK9E	2719988	LA2738 (-)	2719428-2720537	hypothetical protein	LIC11274LBJ_1715LBL_1934LEPBIa0551	
AMan1205	2720613	LA2738-LA2739				
AMan1038	2720887	LA2739 (+)	2720682-2721563	hypothetical protein	LIC11273/LBJ_1716/LBJ_1716/LEPBIa0550	
ALaiC5	2721117	LA2739 (+)	2720682- 2721563	hypothetical protein	LEPBIa0550/LBL_1935/LBJ_1716/LIC11273	
FPom47	2724574	LA2743 (-)	2724504-2724650	hypothetical protein	LEPBIa1193	
FLaiTK107	2726347	LA2746 (+)	2726334-2728121	hypothetical protein	LIC11268/LBJ_1729/LBL_1948/LEPBIa3035	
FLaiTK99	2741494	LA2761 (+)	2740183-2741745	Exopolyphosphatase	LIC11257/LBJ_1742/LBL_196 /LEPBIa0447	
AMan867	2742494	LA2762 (+)	2742198-2742539	hypothetical protein	LIC11256/LEPJB12150/LEPL512399	
AMan1068	2749993	LA2770 (-)	2749078-2750214	hypothetical protein	LIC11249/LBJ_1746/LBL_1965/LEPBIa0813	
FLaiS324	2750350	LA2770-LA2771				
AMan620	2765614	LA2786 (-)	2765179-2765964	hypothetical protein	LIC11235/LBL_1979/LBJ_1760	
AMan839	2771120	LA2791 (-)	2769638-2771827	hypothetical protein	LIC11230	putative lipoprotein
FLaiK155	2791571	LA2809 (-)	2791009-2791590	2-Cys thioredoxin peroxidase	LIC11219/LBJ_1015/LBL_2019/LEPBIa1358	
FKit55	2791806	LA2809-LA2810				
FLai63	2791902	LA2810 (+)	2791918-2792031	hypothetical protein	LEPIC1247	
AMan1060	2796486	LA2814-LA2816				
AMan919	2799810	LA2819-LA2820				
FLaiA3	2805059	LA2822-LA2823				
AMan508	2806473	LA2823 (+)	2805796-2806887	lipoprotein	LIC11207/LBJ_1570/LBL_0972	SP
AMan1356	2818674	LA2834 (-)	2818641-2819912	adenylate cyclase	LIC11198	
FLaiS344	2821078	LA2839 (+)	2821061-2821234	hypothetical protein		
FLaiPCRK13	2827413	LA2847 (-)	2827338-2829320	cGMP-dependent 3',5'-cyclic phosphodieste	LIC11189/LBJ_0788/LBL_2291/LEPBIa1194	Response regulator with HD-GYP domain
AMan788	2829485	LA2847-LA2848				
AMan1360	2833807	LA2850-LA2852				
FLaiS268	2836768	LA2854-LA2855				
AMan1234	2840649	LA2858-LA2859				
AMan1241	2841490	LA2860 (-)	2841280-2842416	hypothetical protein	LIC11181	Predicted amidohydrolase

AMan782	2848199	LA2867 (-)	2848084-2849151	rhomboid-like protein	LIC11176/LBJ_2025/LBL_1025/LEPBIa0872	7TM
FPC2	2852466	LA2871 (-)	2851461-2852960	hypothetical protein	LIC11173/LBL_1023/LBJ_2027/LEPBIa2924	
AMan611	2852549	LA2871 (-)	2851461-2852960	hypothetical protein	LIC11173/LBJ_2027/LBL_1023/LEPBIa2924	
AMan906	2852874	LA2871 (-)	2851461-2852960	hypothetical protein	LIC11173/LBJ_2027/LBL_1023/LEPBIa2924	
AMan1006	2779113	LA2880 (+)	2860414-2860025	hypothetical protein	LIC11164/LBL_0575/LBJ_2537	
AMan786	2866365	LA2887-LA2886				
FKit32	2889407	LA2918 (+)	2889260-2891200	Alpha-galactosidase	LIC11140/LBJ_2282/LBL_0825/LEPBIa0772	Alpha-glucosidase
FLaiPCR55	2891506	LA2919 (+)	2891197-2891739	hypothetical protein	LIC11139/LBJ_2283/LBL_0824/LEPBIa0770	
AMan830	2896114	LA2922 (-)	2895394-2896209	hypothetical protein	LIC11135/LBJ_2286/LBL_0821/LEPBIa0766	
AMan720	2897248	LA2924 (+)	2897214-2898650	pyruvate kinase	LIC11132/LBJ_2288/LBL_0819/LEPBIa0764	
AMan1212	2900821	LA2926 (+)	2900067-2901026	sensory box/GGDEF family protein	LIC11131/LBJ_2289/LBL_0818	
AMan1003	2901706	LA2927 (+)	2901147-2902094	sensory box/GGDEF family protein	LIC11130/LBJ_2289/LBL_0818/	
AMan724	2903190	LA2930 (+)	2903091-2904017	sensory box/GGDEF family protein	LIC11128/LBJ_2289/LBL_0818	
FLaiS367	2919860	LA2948 (+)	2919607-2920398	hypothetical protein	LIC11113/LBL_0998/LBJ_2052/LEPBIa3204	
FLai-K234	2928394	LA2955 (+)	2928309-2929106	putative acyltransferase	LIC11106/LBL_0991/LBJ_2059/LEPBI3197	
AMan1193	2938755	LA2960 (-)	2938615-2938758	hypothetical protein	LIC11100	
FLaiS338	2945027	LA2968 (-)	2943766-2945091	cyaA1 1 adenylate cyclase	LBJ_2066/LBL_0984/LIC11095	intracellular signaling cascade, TMS>6
FLai55	2946263	LA2969 (-)	2945895-2946788	Putative glycosyl transferase	LIC11093/LBL_0983 /LBJ_2067/LEPBIa17(	Glycosyltransferases involved in cell wall biogenes
FKitGK56	2949502	LA2971 (-)	2947162-2949222	hypothetical protein	LBJ_2068/LBL_0982/LIC11092	SP 13TM Lipid A core - O-antigen ligase and relate
AMan1276	2954262	LA2976-LA2977				
FLai41	2974467	LA2997 (+)	2974649-2975383	Glycosyl hydrolase/lysozyme	LIC11068	cell wall catabolism
FLaiS64	2979817	LA3002 (-)	2979090-2980553	pcnB 1 Probable poly(A) polymerase	LIC11065/LBJ_1990/LBL_1060/LEPBIa1386	RNA processing
AMan1251	2983905	LA3005 (-)	2983647-2984270	hypothetical protein	LIC11062/LEPBIa3357	Putative methyltransferase
FLaiTK88	2984070	LA3005 (-)	2983647-2984270	hypothetical protein	LIC11062/LEPBIa3357	Putative methyltransferase
AMan789	**3001826**	**LA3028 (+)**	**3000858-3002915**	**hypothetical protein**	**LIC11051/LBJ_2033/LBL_1017**	**Leucine-rich repeat containing protein**
AMan1257	3009390	LA3039 (-)	3008548-3009720	hypothetical protein	LIC11046	
FKit69	3019796	LA3054 (-)	3019796-3020080	serine protease DO	LIC11037/LBJ_2294/LBL_0813/LEPBIa0911	Peptidase S1C, hrtA/degQ
FLaiS267	3021688	LA3055 (+)	3021440-3021817	hypothetical protein	LIC11036/LBL_0812/LBJ_2295/LEPBIa0910	
FLaiS389	3039066	LA3069 (-)	3036390-3039398	hypothetical protein	LIC11026/LBJ_2298/LBL_0809/LEPBIa0906	
AMan968	**3046361**	**LA3075 (+)**	**3044798-3050668**	**surface protein Lk90-like protein LigC**	**LIC10464/LBJ_0609/LBL_2470**	**Bacterial Ig-like**
FLaiK2E	3046600	LA3075 (+)	3044798-3050668	surface protein Lk90-like protein LigC		LigC,pseudogene in Lai
FLaiS368	3054659	LA3079 (-)	3054623-3055405	hypothetical protein	LIC11017/LBJ_0846/LBL_2236	
FLai2	3060492	LA3085 (-)	3060298-3062340	ppGpp synthetase /RelA	LIC11012/LBJ_0842/LBL_2240/LEPBIa2324	
AMan1064	3061578	LA3085 (-)	3060298-3062340	ppGpp synthetase /RelA	LIC11012/LBJ_0842/LBL_2240/LEPBIa2324	
FLaiS392	3063405	LA3087-16S ribosomal RNA				
FLaiS47	3069467	LA3093 (-)	3069370-3071262	Peptidase family M23/M37	LIC11008/LBJ_0835/LBL_2247/LEPBIa2332	proteolysis and peptidolysis
AMan1092	3072573	LA3095 (+)	3072283-3073407	Queuine tRNA-ribosyltransferase (tRNA-gu	LIC11005/LBJ_0838/LBL_2244/LEPBIa2329	
AMan1247	3072655	LA3095 (+)	3072283-3073407	Queuine tRNA-ribosyltransferase (tRNA-gu	LIC11005/LBJ_0838/LBL_2244/LEPBIa2329	
AMan1122	3073201	LA3095 (+)	3072283-3073407	Queuine tRNA-ribosyltransferase (tRNA-gu	LIC11005/LBJ_0838/LBL_2244/LEPBIa2329	
AMan754	3074835	LA3097 (+)	3073861-3075528	LipL71	LIC11003/LBJ_0840/LBL_2242/LEPBIa2327	

AMan765	**3075429**	**LA3097 (+)**	**3073861-3075528**	**LipL71**	**LIC11003/LBJ_0840/LBL_2242/LEPBIa2327**	
AMan977	**3083805**	**LA3103 (-)**	**3083485-3083883**	**hypothetical protein**	**LIC10997/LBJ_2226/LBL_2219**	
AMan41	3087730	LA3108 (-)	3087641-3088324	Regulatory protein KdpE	LIC10994/	
FLaiS246	3090259	LA3109(-)	3088328-3090940	kdpD I two-component sensor KdpD	LIC10993	4TM histidine kinase sensor protein
FLaiTK84	3127739	LA3144 (-)	3127708-3128880	Alpha-methylacyl-CoA racemase	LIC10969/LBJ_0567/LBL_2513/LEPBIa2907	metabolism
FLaiS273	3128776	LA3144 (-)	3127708-3128880	Alpha-methylacyl-CoA racemase	LIC10969/LBJ_0567/LBL_2513/LEPBIa2907	metabolism
AMan738	3130902	LA3145-LA3146				
AMan923	3131802	LA3147 (+)	3131340-3132203	Putative hydrolase, alpha/beta hydrolase sup	LIC10966/LBJ_0564/LBL_2516/LEPBIa2911	
FLaiK-209	3136525	LA3152 (-)	3136320-3136877	predicted transcriptional regulator, copG fan	LIC10962	
AMan1058	3140789	LA3156 (+)	3139596-3141002	flavin-containing monooxygenase	LIC10959/LEPBIa0245	
FLaiS79	3153797	LA3172 (+)	3153463-3154299	aldehyde reductase	LIC10949/LEPBIa2448/LBL_0189	Aldo/keto reductase
AMan755	3176835	LA3197 (+)	3176524-3178065	Type I restriction enzyme	LIC10932/LBL_2771/LBJ_0305	
AMan840	3180790	LA3200-LA3201				
FLaiTK111	3181118	LA3200-LA3201				
FLaiPCRK3	3193309	LA3217 (-)	3193282-3194268	hypothetical protein	LIC10915/LBL_0492/LBJ_2614/LEPBIa0645	
AMan1185	3193564	LA3217 (-)	3193282-3194268	hypothetical protein	LIC10915/LBL_0492/LBJ_2614/LEPBIa0645	
AMan1260	3201760	LA3230 (-)	3200979-3201887	hypothetical protein	LIC10906/LBJ_0615/LBL_2464	7TM
AMan955	3201823	LA3230 (-)	3200979-3201887	hypothetical protein	LIC10906/LBJ_0615/LBL_2464	7TM
FLaiK137	3208080	LA3236-LA3237				
AMan630	3211609	LA3242 (+)	3211269-3214244	TonB-dependent receptor	LIC10896/LBJ_0718/LBL_2361/LEPBIa3362	Outer membrane receptor proteins, mostly Fe trans
AMan894	**3224989**	**LA3258 (-)**	**3224914-3227739**	**TonB-dependent receptor**	**LIC10882/LBJ_2189/LBL_2721**	
FLaiS370	3228503	LA3260 (-)	3228494-3228961	hypothetical protein	LIC10880	
FLaiK6B	3229572	LA3262 (-)	3229121-3229600	putative lipoprotein	LIC10879 LBJ_0807 LBL_2273 LEPBIa1374	
AMan781	3239548	LA3270-LA3271				
AMan792	3243490	LA3274 (-)	3243370-3243813	hypothetical protein	LIC10866	
AMan844	3246912	LA3277 (+)	3246174-3246986	hypothetical protein	LIC10869	
AMan760	3254297	LA3286 (-)	3253540-3255852	adenylate cyclase	LIC10862/LBJ_0916/LBL_0931/LEPBIa2606	6TM Adenylyl and guanylyl cyclase catalytic doma
AMan761	3268386	LA3301 (-)	3268226-3268447	hypothetical protein	LIC10847/LBJ_0901/LBL_0916	TPR repeat
AMan1079	3275629	LA3312 (-)	3275416-3275841	glyoxalase	LIC10838	Glyoxalase/Bleomycin resistance protein/Dioxygen
AMan1111	3277621	LA3314 (-)	3277082-3277984	cell-division inhibitor	LIC10837	
FLaiS355	3283181	LA3320 (+)	3282382-3283515	Leucine-rich repeat containing protein	LBJ_1970/LBL_1314/LIC10831	Phosphoprotein phosphatase
AMan1088	3296767	LA3333 (-)	3294139-3301059	hypothetical protein	LIC10825	3TM Integrin alpha N-terminal domain
AMan1145	3301010	LA3333 (-)	3294139-3301059	hypothetical protein	LIC10825	3TM Integrin alpha N-terminal domain
AMan1174	3301025	LA3333 (-)	3294139-3301059	hypothetical protein	LIC10825	3TM Integrin alpha N-terminal domain
FKitGK2	3302161	LA3334 (+)	3301526-3302410	hypothetical protein	LIC10824 / LBL_1742 /LBJ_1517	3TM
AMan956	3309884	LA3344 (-)	3309848-3310132	hypothetical protein	LIC10816/LBJ_2317/LBL_0790/	3TM
AMan373	3311676	LA3345 (+)	3311349-3312119	hypothetical protein	LIC10815/LBJ_0560/LBL_2520/LEPBIa0726	
AMan635	3320079	LA3356 (+)	3319731-3320405	glutathione S-transferase	LIC10807/LBL_0650/LBJ_2462/LEPBIa2267	
FLaiTK70	3320150	LA3356 (+)	3319731-3320405	glutathione S-transferase	LIC10807/LEPIC0823 /LBL_0650/LBJ_2462	
AMan1294	3323982	LA3359-LA3360				

AMan1229	**3331158**	**LA3367 (-)**	**3331017-3331448**	**hypothetical protein**	**LIC10798**	
AMan865	3342344	LA3379 (-)	3342172-3343149	flagellar filament outer layer protein A	LIC10788/LBJ_0704/LBL_2375/LEPBIa2335	
AMan1067	3342900	LA3379 (-)	3342172-3343149	flagellar filament outer layer protein A	LIC10788/LBJ_0704/LBL_2375/LEPBIa2335	
AMan718	3342957	LA3379 (-)	3342172-3343149	flagellar filament outer layer protein A	LIC10788/LBJ_0704/LBL_2375/LEPBIa2335	
ALaiC2	3343151	LA3380 (-)	3343143-3343862	flaA2 flagellar filament outer layer protein	LIC10787/LBJ_0703/LBL_2376/LEPBIa2336	
AMan908	3345014	LA3383 (-)	3344859-3345113	hypothetical protein	LEPIC0802	
FLai47	3358434	LA3397 (-)	3358111-3358935	hypothetical protein	LIC10772/LBJ_2382/LBL_0726/LEPBIa1352	
AMan1036	3364768	LA3403 (+)	3364728-3365162	hypothetical protein	LIC10766/LBJ_2388/LBL_0720	
AMan878	3376669	LA3418-LA3419				
AMan866	3376825	LA3418-LA3419				
AMan707	3388245	LA3427-LA3428				
AMan1002	3392068	LA3432 (-)	3391359-3392957	hypothetical protein	LIC10741/LBJ_2580/LBL_0532	
AMan623	3397999	LA3438 (+)	3397796-3398308	hypothetical protein	LIC10735/LBJ_2573/LBL_0539/LEPBIa1851	4TM
AMan614	3408523	LA3450 (+)	3408274-3410070	regulation protein	LIC10727/LBJ_0644/LBL_2435/	7TM Protein phosphatase 2C-like
FLaiS278	3420414	LA3459 (+)	3420315-3422453	polyphosphate kinase	LIC10719/LEPBIa0579/LBJ_2339/LBL_0769	
AMan1239	3428577	LA3466 (-)	3427074-3429476	thermolysin	LIC10715	Peptidase M4, thermolysin
FLaiS388	3438034	LA3471-LA3472				
AMan1169	3449561	LA3484 (-)	3449172-3449951	3-oxoacyl-(acyl-carrier protein) reductase	LIC10700/LBJ_0451/LBL_2627/LEPBIa0050	
ALaiC9	3451957	LA3487(+)	3451740-3451976	hypothetical protein	LIC10697	
ALaiC25	3451962	LA3487 (+)	3451740 -3451976	hypothetical protein	LIC10697	
AMan1180	3461836	LA3497 (-)	3460682-3462904	hypothetical protein	LIC10690/LBJ_0456/LBL_2622	
FLaiTK81	3470716	LA3506 (-)	3470528-3472240	Methyl-accepting chemotaxis protein	LIC10685/LBJ_2374/LBL_0734	Bacterial chemotaxis sensory transducer
AMan935	3475049	LA3510-LA3511				
AMan872	3481487	LA3521 (+)	3481487-3482623	TPR-repeat-containing protein	LIC10673/LBJ_0693/LBL_2386/LEPBIa1930	
FLai189S	3487406	LA3527-LA3528				
Aman910	3501586	LA3543 (-)	3501472-3501921	hypothetical protein	LIC10655	
AMan903	3509526	LA3553 (+)	3509303-3510301	tRNA-dihydrouridine synthase A	LIC10646/LBJ_0299/LBL_2777/LEPBIa2545	
AMan1150	3510954	LA3553-LA3554				
AMan774	3518356	LA3563-LA3564				
AMan1334	3518508	LA3563-LA3564				
AMan1302	3531020	LA3580 (-)	3530592-3531110	hypothetical protein	LIC10619	
AMan901	3538261	LA3586 (-)	3538040-3538387	hypothetical protein	LIC10613/LBL_2748/LBJ_0328	3TM
AMan966	3547080	LA3597-LA3598				
AMan1228	3573890	LA3623 (-)	3573827-3574603	hypothetical protein	LIC10588/LBJ_2562/LBL_0550	glutamine amidotransferase
FLai-K208	3574162	LA3623 (-)	3573827-3574603	hypothetical protein	LIC10588/LBL_0550/LBJ_2562	glutamine amidotransferase
FKitGK60	3599103	LA3644 (-)	3598797-3599600	enoyl-CoA hydratase	LIC10565/LEPBIa0052/LBJ_0255/LBL_2825	
FLaiS270	3599601	LA3644 (-)	3598797-3599600	enoyl-CoA hydratase	LIC10565/LEPBIa0052/LBJ_0255/LBL_2825	
AMan990	3599601	LA3644 (-)	3598797-3599600	enoyl-CoA hydratase	LIC10565/LEPBIa0052/LBJ_0255/LBL_2825	
AMan626	3601508	LA3646 (-)	3600928-3601704	hypothetical protein	LIC10563/LEPBIa2728/	Uncharacterized protein SCO1/SenC/PrrC, involve
AMan297	3604150	LA3649 (-)	3603250-3604167	hypothetical protein	LIC10561/LEPBIa2726	

AMan1052	3616686	LA3664-LA3665				
AMan1178	3616960	LA3664-LA3665				
AMan1080	3617839	LA3666 (+)	3617702-3618295	Nucleoside-triphosphatase (Nucleoside tripl	LIC10549/LBJ_2875/LBL_0196/LEPBIa2667	
Aman1244	3617966	LA3666 (+)	3617702-3618295	Nucleoside-triphosphatase (Nucleoside tripl	LIC10549/LBJ_2875/LBL_0196/LEPBIa2667	
AMan804	3624364	LA3675 (-)	3623915-3625405	hypothetical protein	LIC10544/LBJ_2708/LBL_0366/LEPBIa0571	
AMan868	3634069	LA3684 (-)	3633806-3635440	ABC transporter, ATP-binding protein	LIC10538/LBJ_2874/LBL_0197/LEPBIa2665	
AMan1133	3635424	LA3684 (-)	3633806-3635440	ABC transporter, ATP-binding protein	LIC10538/LBJ_2874/LBL_0197/LEPBIa2665	
FLaiPCR54	3641601	LA3692 (+)	3641402-3643828	Penicillin-binding protein 1A	LIC10534/LBJ_2676/LBL_0402/LEPBIa0847	
FLaiS55	3664609	LA3712 (-)	3664471-3667008	unknown protein	LIC10519/LBJ_2539/LBL_0573/LEPBIa0266	Permease component of an ABC transporter compl
FLaiK-219	3687140	LA3725 (-)	3679212-3687569	probable phenazine biosynthesis family pro	LIC10508-LIC10510	
AManf2	3688121	LA3726 (-)	3687710-3689242	hypothetical protein	LIC10501	putative lipoprotein
AMan1263	3694539	LA3730-LA3731				
AMan1155	3695781	LA3731 (-)	3695022-3698219	hypothetical protein	LIC10497/LEPBIa2988	
AMan1250	3697539	LA3731 (-)	3695022-3698219	hypothetical protein	LIC10497/LEPBIa2988	
AMan929	3707018	LA3738 (+)	3705231-3708311	Multidrug efflux transporter AcrB transmem	LIC10491/LBJ_2094/LBL_0956/	12TM
AMan913	3715203	LA3745 (-)	3714602-3715327	hypothetical protein	LIC10485/LBJ_2548/LBL_0564/LEPBIa0260	
AMan974	3715289	LA3745 (-)	3714602-3715327	hypothetical protein	LIC10485/LBJ_2548/LBL_0564/LEPBIa0260	
AMan944	3729043	LA3762 (+)	3728686-3729114	hypothetical protein	LIC10474/LBJ_2496/LBL_0625	SP 4TM
FLaiS376	3749696	LA3783 (-)	3749420-3750748	Adenosine deaminase	LIC10459/LBJ_0617/LBL_2462/LEPBIa0038	
FLaiS68	3750797	LA3783- LA3784				
AMan1188	3761381	LA3798 (-)	3761202-3761672	hypothetical protein	LIC10449/LBL_2713/LBJ_0363	
AMan937	3771321	LA3806 (+)	3770039-3771331	ammonium transporter	LIC10441/LBJ_0371/LBL_2706/LEPBIa0794	SP 12TM
FLaiTK72	3782046	LA3819-LA3820				
AMan922	3788755	LA3825-LA3826				
FLaiPCR27	3789743	LEPIN3351 (+)	3788850-3790295	hypothetical protein	LIC10424/LBJ_2582/LBL_0530	putative membrane protein
AMan905	3808085	LA3843-LA3844				
FLaiPCRK23	3820462	LA3855 (-)	3819673-3820578	Pirin-related protein	LIC13080/LBJ_2665/LBL_0391/LEPBIa3351	
FKit59	3823842	LA3861 (-)	3823530-3824177	hypothetical protein	LIC13083/LBJ_0655/LBL_2424	
AMan854	3829153	LA3868 (+)	3828548-3829609	Sensory transduction histidine kinase	LIC13087/LBJ_0353/LBL_2724/	
AMan813	3829990	LA3869 (+)	3829606-3830277	response regulator	LIC13088/LBJ_0352/LBL_2725	
AMan861	3844437	LA3881 (-)	3844076-3847681	hypothetical protein	LIC13101/LBJ_0994/LBL_4283/	SP Integrin alpha N-terminal domain
AMan834	**3845285**	**LA3881 (-)**	**3844076-3847681**	**hypothetical protein**	**LIC13101/LBJ_0994/LBL_4283/**	**SP Integrin alpha N-terminal domain**
FLai-K216	3847449	LA3881 (-)	3844076-3847681	hypothetical protein	LIC13101/LBJ_0994/LBL_4283/	SP Integrin alpha N-terminal domain
AMan1124	3850641	LA3883 (-)	3850548-3851165	Uracil-DNA glycosylase	LIC13102/LBJ_0598/LBL_2482	
FLaiK15	3863670	LA3895 (-)	3861636-3863960	two-component hybrid sensor and regulator	LIC13111/LBJ_0606/LBL_2474/LEPBIa0078	signal transduction/pair genes
FLaiS372	3872046	LA3905-LA3906				
FKit71	3877233	LA3909 (+)	3875890-3877404	hypothetical protein	LIC13120/LBL_0602/LBJ_2510	EAL domain, diguanylate phosphodiesterase
AMann2	3879649	LA3912 (+)	3879560-3880435	hypothetical protein	LIC13123/LBL_0599/LBJ_2513/LEPBIa1161	SP FOG:HEAT repeat
FLaiL33	3879890	LA3912 (+)	3879560-3880435	hypothetical protein	LIC13123/LBL_0599/LBJ_2513/LEPBIa1161	SP FOG:HEAT repeat
AMan806	3881834	LA3913-LA3915				

FLaiS306	3890122	LA3920-LA3921				
FLaiS294	3890527	LA3920-LA3921				
FLaiS245	3903522	LA3928 (+)	3902801-3904588	fadE2 | acyl-CoA dehydrogenase	LIC13136/LEPBIa1259/LBJ_0263/LBL_2817	
FLai35	3904171	LA3928 (+)	3902801-3904588	fadE2 | acyl-CoA dehydrogenase	LIC13136/LBL_2817/LBJ_0263/LEPBIa1259	
AMan1195	3917628	LA3943 (+)	3917353-3918030	hypothetical protein	LIC13148/LBL_2679/LBJ_0398	
FLai190S	3918589	LA3945 (+)	3918427-3920601	recG | ATP-dependent DNA helicase recG	LIC13150/LBL_2677/LBL_2673/LEPBIa0349	
FLaiS244	3940548	LA3965- LA3966				
AMan1017	3943529	LA3969-LA3970				LA3970 promoter
AMan1161	3947520	LA3973 (-)	3946980-3948047	putative aldo/keto reductase-family protein	LIC13176/LEPBIa3328	
FKit11	3951030	LA3977-LA3978				
AMan1138	3953148	LA3980 (-)	3952359-3953147	glycerophosphoryl diester phosphodiesteras	LIC13182/LBJ_0420/LBL_2657/LEPBIa1842	
FLaiPCRK9	3953566	LA3982 (-)	3953282-3953665	hypothetical protein	LIC13183/LBJ_0421/LBL_2656/LEPBIa0144	
AMan791	3958953	LA3986-LA3987				
AMan945	3959160	LA3986-LA3987				
AMan480	3967446	LA3996 (-)	3967396-3969858	two-component hybrid sensor and regulator	LIC13192/LBL_2811/LBJ_2667/LEPBIa0901	SP 6TM
AMan777	**3979064**	**LA4002-LA4003**				
FLaiPCR56	3980787	LA4004 (+)	3980227-3981903	sph3 | Sphingomyelinase C precursor	LIC13198/LBJ_0291/LBJ_0527	
FPom36	3994772	LA4016 (-)	3994551-3995018	hypothetical protein	LIC13206/LBL_2802/LBJ_0274	
FLaiS14	4014083	LA4036-LA4038				
FLai4	4025365	LA4049 (+)	4024617-4026182	ATP-dependent RNA helicase	LIC13230/LBJ_0225/LBL_2858/LEPBIb0208	
AMan1343	4037848	LA4064 (+)	4037645-4038052	hypothetical protein	LIC13241/LBJ_0213/LBL_2869	
FLaiK16E	4039501	LA4066-LA4067				
AMan928	4063823	LA4093 (+)	4063746-4064417	Putative hydrolase	LIC13264/LBL_2891/LBJ_0192/LEPBIa0171	MutT/nudix family protein
AMan742	4064033	LA4093 (+)	4063746-4064417	Putative hydrolase	LIC13264/LBL_2891/LBJ_0192/LEPBIa0171	MutT/nudix family protein
AMan727	4069552	LA4102 (+)	4069396-4070097	Putative response regulator receiver, transcr	LIC13269/LBL_2897/LBJ_0186/LEPBIa3226	
AMan1126	4076618	LA4107-LA4108				
AMan1109	4077101	LA4108 (-)	4076718-4077404	hypothetical protein	LIC13274/LBL_2902/LBJ_0181	
FLaiPCR43	4084752	LA4117 (-)	4083775-4085379	Acyl-CoA dehydrogenase	LIC13281/LEPBIa1112	
AMan280	4090178	LA4122 (-)	4090101-4090610	RNA polymerase ECF-type sigma factor	LIC13285/LBL_2910/LBJ_0173/LEPBIb0053	
FPom44	4090426	LA4122 (-)	4090101-4090610	RNA polymerase ECF-type sigma factor	LIC13285/LBL_2910/LBJ_0173/LEPBIb0053	
AMan1059	**4090592**	**LA4122 (-)**	**4090101-4090610**	**RNA polymerase ECF-type sigma factor**	**LIC13285/LBL_2910/LBJ_0173/LEPBIb0053**	
AMan1086	**4096975**	**LA4129 (-)**	**4096524-4097036**	**Ankyrin repeat proteins**	**LIC13291**	
FLai56	4098837	LA4131 (+)	4098661-4100640	Peptidase family M48	LIC13293	SP Zn-dependent protease with chaperone function
AMan907	4099404	LA4131 (+)	4098661-4100640	Peptidase family M48	LIC13293	SP Zn-dependent protease with chaperone function
FPom38	4100227	LA4131 (+)	4098661-4100640	Peptidase family M48	LIC13293	SP Zn-dependent protease with chaperone function
AMan1105	4103996	LA4135 (-)	4102858-4105446	hypothetical protein	LIC13296/LBJ_1775/LBL_1099	putative lipoprotein
AMan870	4127134	LA4158 (-)	4127119-4127682	hypothetical protein	LIC13317	
AMan1236	4131796	LA4161 (-)	4130892-4133237	thermolysin	LIC13320	Peptidase M4
AMan880	**4132301**	**LA4161 (-)**	**4130892-4133237**	**thermolysin**	**LIC13320**	**Peptidase M4**
FFC18	4132419	LA4161 (-)	4130892-4133237	thermolysin	LIC13320	Peptidase M4

FLaiPCR39	4132057	LA4161 (-)	4130892-4133237	thermolysin	LIC13320	Peptidase M4
FLaiK218	4151828	LA4181 (-)	4151843-4152835	carboxypeptidase, putative	LIC13337	
AMan1309	4156451	LA4187 (-)	4156160- 4156687	hypothetical protein	LIC13342/LBJ_1829/LBJ_2572	
AMan1136	4160081	LA4191 (-)	4159908-4160822	hypothetical protein	LIC13346/LBJ_2816/LBL_0255	
AMan996	4160831	LA4191-LA4193				
FLaiS251	4167287	LA4198(-)	4166365-4167336	hypothetical protein	LIC13351	
FLaiS336	4176807	LA4208 (-)	4176002-4177837	hypothetical protein	LBJ_2830/LBL_0241/LIC13360	coiled-coil
AMan911	**4178550**	**LA4209 (-)**	**4177839-4179926**	**hypothetical protein**	**LIC13361/LBJ_2831/LBL_0240**	**SP**
AMan507	4186843	LA4215 (+)	4186230-4187306	putative strictosidine synthase	LIC13366/LBJ_2836/LBL_0235/LEPBIa2637	SP
FLaiS253	4211068	LA4240 (-)	4211011-4212633	unknown protein	LIC13392/LBJ_2913/LBL_0150/LEPBIa0312	polysaccharide deacetylase
FKitGK3	4216454	LA4243 (-)	4215230-4216822	Methyl-accepting chemotaxis protein	LIC13394/ LEPBIa1258	
FPom10	4218121	LA4244- LA4245				
FLai200S	4219117	LA4245 (+)	4218339-4219772	putative flavin-containing monooxygenase	LIC13396/LEPBIa0245	electron transport
FWei3	4224204	LA4249-LA4250				
AMan1318	4231298	LA4254 (-)	4229746-4231731	Acetyl-coenzyme A synthetase (Acetate--C	LIC13404/LBJ_2938/LBL_0125/LEPBIa3450	
FPC4	4239388	LA4264 (+)	4239308-4241737	gpsA3 | Glycerol-3-phosphate dehydrogena	LIC13412/LBJ_2953/LBL_0111/LEPBIa0227	
AMan1144	4240146	LA4264 (+)	4239308-4241737	gpsA3 | Glycerol-3-phosphate dehydrogena	LIC13412/LBJ_2953/LBL_0111/LEPBIa0227	
AMan904	4241579	LA4264 (+)	4239308-4241737	gpsA3 | Glycerol-3-phosphate dehydrogena	LIC13412/LBJ_2953/LBL_0111/LEPBIa0227	
AMan896	4246429	LA4269 (-)	4245479-4247110	transcriptional regulator AraC family	LIC13415/LBJ_2951/LBL_0113/LEPBIa3033	SP 8TM
AMan619	4252210	LA4273 (+)	4250901-4252622	hydrogenase-4 component B	LIC13419/LBJ_2943/LBL_0120/LEPBIa2423	SP 13 TM
AMan1075	4254873	LA4276 (+)	4254112-4255320	Hydrogenase-4 component F	LIC13422/LBL_0117/LBJ_2946/LEPBIa2420	SP 12TM
AMan1308	4261380	LA4283-LA4284				
AMan745	4267525	LA4290 (+)	4266389-4267537	dihydroorotate dehydrogenase	LIC13433/LBJ_2928/LBL_0135/LEPBIa0092	
FKit72	4268619	LA4291 (+)	4267641-4269083	hypothetical protein	LIC13434/LBJ_2929/LBL_0134/LEPBIa0128	Metalloproteases
AMan1028	4273684	LA4296 (+)	4272264-4274036	Putative sodium/hydrogen exchanger	LIC13439/LBJ_2934/LBL_0129/LEPBIa0200	12TM
FLaiPCRK22	4276673	LA4300 (+)	4276374-4277138	ThiJ/PfpI family protein	LIC13443/LBJ_2959/LBL_0105	Putative intracellular protease/amidase
FLaiS162	4278337	LA4301 (+)	4277149-4278945	kefB | Glutathione-regulated potassium-eff	LIC13444/LBJ_2960/LBL_0104/LEPBIa2123	
AMan770	4288334	LA4311 (+)	4287461-4288567	metallopeptidase	LIC13453/LBJ_2969/LBL_0095/LEPBIa3428	
AMan421	**4298996**	**LA4324 (+)**	**4298618-4300036**	**hypothetical protein**	**LIC13467**	**C-type lectin-like**
FLaiS283	4314334	LA4340 (-)	4313561-4314370	deoxyribonuclease, tatD family	LIC13481/LBL_2979/LBJ_0025/LEPBIa3468	Metallo-dependent hydrolases
FKitGK6	4320084	LA4348 (-)	4319808-4320572	Glucose inhibited division protein gidB	LIC13490/LEPBIa3475/LBJ_0031/LBL_2985	
AMand	4320118	LA4348 (-)	4319808-4320572	Glucose inhibited division protein gidB	LIC13490/LEPBIa3475/LBJ_0031/LBL_2985	
FLaiTK25	3866	LB003 (+)	3779-3886	hypothetical protein		
FKit47	5110	LB006-LB007				
AMan1158	5568	LB007 (+)	5239-5754	hypothetical protein	LIC20006/LBJ_4006/LBL_4006	Predicted hydrolases of the HAD superfamily
FLaiS302	26867	LB033 (-)	26903-27112	hypothetical protein		
AMan615	45799	LB056 (+)	45467 -46210	TPR-repeat-containing protein	LIC20042/LBL_4047/LBJ_4047	
AMan1233	47891	LB058 (+)	47847 - 49670	heat shock protein 90	LIC20044/LBL_4049/LBJ_4049/LEPBIb0005	Molecular chaperone, HSP90 family
AMan783	54118	LB065 (+)	54100 -54918	hypothetical protein	LIC20050	
AMan422	61498	LB072 (+)	61294 -61851	hypothetical protein	LIC20056/LBL_4060/LBJ_4060	

AMan647	66025	LB075 (+)	65076-66413	hypothetical protein	LBJ_4063/LBL_4063/LIC20059/LEPBIb00457	TM ABC transporter
AMan165	70416	LB080 (+)	70349-72280	hypothetical protein	LIC20063/LBJ_0472/LBL_2607	9TM
AMan1044	72467	LB080-LB081				
AMan1230	72478	LB080-LB081				
AMan702	80393	LB089 (+)	79179 -80735	Uncharacterized FAD-dependent dehydroge	LEPBIa0364/LBJ_4076/LBL_4093/LIC20071	
AMan1159	80529	LB089 (+)	79179 -80735	Uncharacterized FAD-dependent dehydroge	LEPBIa0364/LBJ_4076/LBL_4093/LIC20071	
AMan769	82303	LB092 (-)	81925-82335	hypothetical protein	LIC20073/LBJ_4078/LBL_4095	4TM
FLaiS48	91840	LB102 (+)	91672-92481	phospholysine phosphohistidine inorganic p	LIC20079/LBJ_4085/LBL_4102	Sugar phosphatase
AMan918	105763	LB112 (-)	104517-106262	putative regulatory protein contains GAF dc	LBJ_4095/LBL_4111/LIC20089/LEPBIb0171	Sigma factor SigB regulation protein RsbU; putativ
AMan873	110760	LB116 (-)	109072 -110814	hypothetical protein	LIC20092/LBJ_4098/LBL_4114/LEPBIb0195	Predicted protease with the C-terminal PDZ domair
AMan1037	115124	LB119 (+)	114803 -115147	hypothetical protein	LIC20095/LBJ_4100/LBL_4116/LEPBIa3129	SP 4TM
AMan827	115176	LB119-LB120				LB120 promoter
AMan1119	115210	LB120 (+)	115194-116291	hypothetical protein	LBL_4117/LBJ_4101/LIC20096	Zn-dependent hydrolases, including glyoxylases
FKitGK5	115216	LB120 (+)	115194-116291	hypothetical protein	LBJ_4101/LBL_4117/LIC20096	Zn-dependent hydrolases, including glyoxylases
Aman319	117966	LB122 (-)	117169-118068	site-specific integrase/recombinase XerD re	LIC20098/LBL_4120/LBJ_4104	phage-related integrase/recombinase
AMan739	**119588**	**LB125 (-)**	**119577 -120128**	**Probable chemoreceptor glutamine deam**	**LEPBIb0074/LBL_4123/LBJ_4107/LIC2010**	**Chemotaxis protein; stimulates methylation of M**
AMan606	119736	LB125 (-)	119577 -120128	Probable chemoreceptor glutamine deamidz	LEPBIb0074/LBL_4123/LBJ_4107/LIC20100	Chemotaxis protein; stimulates methylation of MCI
AMan1348	127119	LB133 (+)	126928 -128532	probable protein containing EAL domain	LIC20106	
FPCJ4	128332	LB133 (+)	126928-128532	probable protein containing EAL domain	LIC20106	putative diguanylate phosphodiesterase
AMan893	132936	LB138 (-)	132172 -133629	hypothetical protein	LIC20110/LBL_4084/LEPBIa1742	
FLaiS303	137877	LB142 (-)	137699-138250	hypothetical protein	LBL_4217 LBJ_4203 LEPBIb0141 LIC20113	Methyl-accepting chemotaxis protein
AMan785	138044	LB142 (-)	137699 - 138250	hypothetical protein	LIC20113/LBJ_4203/LBL_4217/LEPBIb0141	Methyl-accepting chemotaxis protein
AMan1007	138095	LB142 (-)	137699 - 138250	hypothetical protein	LIC20113/LBJ_4203/LBL_4217/LEPBIb0141	Methyl-accepting chemotaxis protein
AMan154	145266	LB150 (-)	144510 -145457	Cobalamin biosynthesis protein cobD	LIC20120/LBJ_4192/LBL_4207	
AMan613	146933	LB151 (-)	145461- 146933	Cobyric acid synthase cobQ	LIC20121/LBJ_4191/LBL_4206	
ALaiC22	152732	LB157 (-)	151967-153460	cbiH/ precorrin methylase	LIC20127/LBJ_4185/LBL_4200	
AMan862	159746	LB165 (-)	159516 - 160685	cbiX	putative cbiX protein	LIC20135/LBJ_4177/LBL_4192	ferredoxin/cob operon
FLaiS4	159837	LB165 (-)	159516-160685	cbiX	putative cbiX protein	LIC20135/LBJ_4177/LBL_4192	ferredoxin/cob operon
AMan912	163269	LB167 (+)	163042-165165	hypothetical protein	LIC20136/LBJ_4176/LBL_4191	
AMan890	168635	LB172 (+)	168134 -169744	hypothetical protein	LIC20140/LBJ_4172/LBL_4187/LEPBIb0126	14TM Permeases
FLaiPCR44	169521	LB172 (+)	168134-169744	hypothetical protein	LIC20140/LBJ_4172/LBL_4187/LEPBIb0126	14TM Permeases
AMan1316	171626	LB176 (+)	171595- 172128	Adenine phosphoribosyltransferase (APRT)	LIC20142/LBJ_4170/LBL_4185/LEPBIb0127	
AMan728	172337	LB177 (+)	172166 -173680	putative serine protease	LIC20143/LBJ_4169/LBL_4184/LEPBIb0125	Peptidase S1C, hrtA/degQ
AMan610	174715	LB178 (+)	173695-175278	putative serine protease	LIC20144/LBJ_4168/LBL_4183/LEPBIb0124	Peptidase S1C, hrtA/degQ
AMan484	181753	LB186 (+)	181530 -182207	heme oxygenase	LEPBIa0669/LBL_4180/LBJ_4165/LIC20148	
FLai-K220	181837	LB186 (+)	181530-182207	heme oxygenase	LIC20148/LBL_4180/LBJ_4165/LEPBIa0669	
FLai38	184452	LB190 (-)	183575-184105	Hypothetical protein	LIC20056/LBL_4060/LBJ_4060	
FLai194S	185347	LB191 (-)	184996-187128	TonB-dependent outer membrane receptor p	LIC20151/LBJ_4163/LBL_4178/LEPBIa3432	Hemin receptor
AMan601	185907	LB191 (-)	184996-187128	TonB-dependent outer membrane receptor p	LIC20151/LBJ_4163/LBL_4178/LEPBIa3432	Hemin receptor
AMan655	188384	LB194 (-)	187833 - 188411	hypothetical protein	LBL_4176/LBJ_4161/LIC20153	

AMan420	193578	LB201 (-)	192835-194457	lonA putative ATP-dependent protease LA	LIC20158/LBJ_4156/LBL_4171	Predicted ATP-dependent protease
FLaiS349	198975	LB205-LB206				
AMan1272	210604	LB216 (+)	209350 -211110	hypothetical protein	LIC20172/LBL_4151/LBJ_4136/LEPBIb0156	
FKitGK4	215189	LB222 (-)	216332-216784	hypothetical protein	LIC20175/LBJ_4139/LBL_4154/LEPBIb0163	Putative DnaK suppressor; putative DksA/TraR zin
AMan1062	220280	LB225 (+)	219779 -222361	hypothetical protein	LIC20176	lipoprotein
AMan915	234541	LB240 (-)	234521 -236095	GGDEF family protein	LIC20182/LBL_4068/LBJ_4068	
FLaiS329	236667	LB241 (+)	236543-239152	two-component hybrid sensor and regulator	LEPBIa1069/LIC20183/LBJ_2667/LBL_2811	Response regulator receiver , TMS>8
AMan835	247800	LB251 (-)	247350 -249314	hypothetical protein	LIC20191/LBJ_4208/LBL_4222/LEPBIa0998	(Trans)glycosidases
AMan924	247900	LB251 (-)	247350 -249314	hypothetical protein	LIC20191/LBJ_4208/LBL_4222/LEPBIa0998	(Trans)glycosidases
ALaiC3	250806	LB253-LB254				
FLaiB3	251220	LB254 (-)	250911-252125	Sugar transport protein	LIC20194/LBJ_4210/LBL_4224/LEPBIa3254	
AMan828	255436	LB258 (+)	254727 -257126	Cysteine protease	LIC20197/LBJ_4115/LBL_4131/LEPBIb0082	
FLaiS287	255843	LB258 (+)	254727-257126	Cysteine protease	LIC20197 LBJ_4115 LBL_4131 LEPBIa3193	
AMan295	256062	LB258 (+)	254727 -257126	Cysteine protease	LIC20197/LBJ_4115/LBL_4131/LEPBIb0082	
Aman989	269710	LB272 (+)	269486 -270124	hypothetical protein	LIC20207/LBJ_4128/LBL_4144	5TM Integral membrane protein TerC family
AMan1023	276054	LB275-LB276				
AMan1041	276362	LB275-LB276				
AMan609	276972	LB276-LB277				
FLaiPCRK53	279113	LB279 (+)	278133-280793	hypothetical protein	LIC20214/LBJ_4121/LBL_4137/LEPBIa3017	TonB dependent receptor
FLaiS256	279329	LB279 (+)	278133-280793	hypothetical protein	LIC20214/LBJ_4121/LBL_4137/LEPBIa3017	TonB dependent receptor
AMan164	279608	LB279 (+)	278133 -280793	hypothetical protein	LIC20214/LBJ_4121/LBL_4137/LEPBIa3017	TonB dependent receptor
AMan1246	280734	LB279 (+)	278133 -280793	hypothetical protein	LIC20214/LBJ_4121/LBL_4137/LEPBIa3017	TonB dependent receptor
AMan1279	285501	LB286 (+)	285144 -286796	glutamate synthase	LIC20220/LBJ_4213/LBL_4227/LEPBIa2978	
FLaiTK71	286266	LB286 (+)	285144-286796	glutamate synthase	LIC20220/LBJ_4213/LBL_4227/LEPBIa2978	
AMan1094	295736	LB295-LB296				
FLai13	300276	LB300 (+)	300168-300389	Hypothetical protein	LEPIC0243/LBJ_4224/LBL_4238/LEPBIb0204	
FLaiPCR47	306198	LB309 (+)	305503-306762	similar to Fimh-like protein	LIC20235	
FLaiS298	315707	LB320-LB321				
AMan516	319440	LB322 (-)	316778-319093	two-component hybrid sensor and regulator	LIC20246/LBJ_4237/LBL_4251/LEPBIb0008	Signal transduction histidine kinase
AMan984	323971	LB327-LB328				
AMan775	**324378**	**LB328 (-)**	**324060-325211**	**peptidoglycan-associated cytoplasmic me**	**LIC20250/LBJ_4241/LBL_4255**	**OmpA/MotB domain**
AMan452	325981	LB329 (-)	325387-326403	lysyl-tRNA synthetase	LIC20251/LBJ_4242/LBL_4256/LEPBIb0234	
AMan1137	328462	LB330 (+)	326607-326936	hypothetical protein	LIC20252/LBJ_4243/LBL_4257/LEPBIb0235	
FPom02	328686	LB333-LB334				
AMan1110	350313	LB359-LB360				
AMan942	350414	LB359-LB360				
FLaiS58	350804	LB359-LB360				
AMan799	351250	LB360-LB361				LB360 promoter
FLaiTK105	352466	LB362 (-)	352294-353268	hypothetical protein	LIC20271	
AMan983	**353190**	**LB362 (-)**	**352294- 353268**	**hypothetical protein**	**LIC20271**	

AMan1330	354030	LB363 (-)	353351 -354448	hypothetical protein	LIC20272/LBJ_4262/LBL_4276/LEPBIb0280
AMan759	354383	LB364 (-)	354363 -355733	transcriptional regulator (FIS family)	LIC20273/LBJ_4263/LBL_4277/LEPBIb0281 Response regulator containing CheY-like receiver,
FLai-S75	87924	LB095-LB096			
FLaiS353	3591722	LEPIN3193 (-)	3591527-3591817	transposase	
FLai4	301832	LB303 (-)	301790-302635	transposase / IS*1501*	
FKit80	multiple locations	ND	ND	transposase / ISlin1	
FLai17	multiple locations	ND	ND	transposase / IS*1501*	
FLaiS240	multiple locations	ND	ND	transposase	
AMan836	multiple locations	ND	ND	transposase / IS*1500* (part)	
AMan892	multiple locations	ND	ND	transposase	
Aman953	multiple locations	ND	ND	transposase / IS*1500*	
AMan1057	multiple locations	ND	ND	transposase / IS*1500*	
FLaiL16	multiple locations	ND	ND	transposase / IS*1500*	
FLaiS69	multiple locations	ND	ND	transposase	

A, France; A, Australia; Man, L. interrogans serovar Manilae strain L495; Lai, L. interrogans serovar Lai strain 56601; Kit, L. interrogans serovar Canicola strain Kito EFS; P

L.int+L.weil	L.int
931 mutants	929 mutants
11 Tn	11 Tn
25 rRNA	25 rRNA
1732 intergenic	172
722 genes	721
551 ind genes	551 ind genes

ANEXO II

Artigo “Genetic diversity of the Leptospiral immunoglobulin-like (Lig) genes in pathogenic *Leptospira* spp.” de autoria de McBride, A. J. A.; Cerqueira, G. M.; Suchard, M. A.; Moreira, A. N.; Zuerner, R. L.; Reis, M. G.; Haake, D. A.; Ko, A. I.; Dellagostin, O. A, publicado no periódico Infection Genetics and Evolution, 2009, v. 9, p. 196-205.

Este artigo foi derivado de uma colaboração entre o Centro de Pesquisas Gonçalo Moniz e a Universidade Federal de Pelotas para sequenciar completamente os genes *ligA*, *B* e *C* em sorovares patogênicos de *Leptospira* spp.

A minha participação foi na construção das bibliotecas e sequenciamento dos genes, além de realizar ensaios de PCR e Southern blot para confirmar a presença ou ausência dos genes nos diferentes sorovares testados.

Infection, Genetics and Evolution 9 (2009) 196–205

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Genetic diversity of the Leptospiral immunoglobulin-like (Lig) genes in pathogenic *Leptospira* spp.

Alan J.A. McBride [a,1], Gustavo M. Cerqueira [b,1], Marc A. Suchard [c,d,e], Angela N. Moreira [b], Richard L. Zuerner [f], Mitermayer G. Reis [a], David A. Haake [g,h], Albert I. Ko [a,i], Odir A. Dellagostin [b,*]

[a] Centro de Pesquisa Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, BA, Brazil
[b] Centro de Biotecnologia, Universidade Federal de Pelotas, Pelotas, RS, Brazil
[c] Department of Biomathematics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
[d] Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
[e] Department of Biostatistics, UCLA School of Public Health, Los Angeles, CA, USA
[f] National Animal Disease Centre, Agricultural Research Service, U.S. Department of Agriculture, Ames, Iowa, USA
[g] Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, CA, USA
[h] Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
[i] Division of International Medicine and Infectious Disease, Weill Medical College of Cornell University, New York, NY, USA

ARTICLE INFO

Article history:
Received 27 August 2008
Received in revised form 23 October 2008
Accepted 28 October 2008
Available online 6 November 2008

Keywords:
Leptospirosis
Lig
Pathogenesis
Molecular evolution
Sequence analysis

ABSTRACT

Recent serologic, immunoprotection, and pathogenesis studies identified the Lig proteins as key virulence determinants in interactions of leptospiral pathogens with the mammalian host. We examined the sequence variation and recombination patterns of *ligA*, *ligB*, and *ligC* among 10 pathogenic strains from five *Leptospira* species. All strains were found to have intact *ligB* genes and genetic drift accounting for most of the *ligB* genetic diversity observed. The *ligA* gene was found exclusively in *L. interrogans* and *L. kirschneri* strains, and was created from *ligB* by a two-step partial gene duplication process. The aminoterminal domain of LigB and the LigA paralog were essentially identical (98.5 ± 0.8% mean identity) in strains with both genes. Like *ligB*, *ligC* gene variation also followed phylogenetic patterns, suggesting an early gene duplication event. However, *ligC* is a pseudogene in several strains, suggesting that LigC is not essential for virulence. Two *ligB* genes and one *ligC* gene had mosaic compositions and evidence for recombination events between related *Leptospira* species was also found for some *ligA* genes. In conclusion, the results presented here indicate that Lig diversity has important ramifications for the selection of Lig polypeptides for use in diagnosis and as vaccine candidates. This sequence information will aid the identification of highly conserved regions within the Lig proteins and improve upon the performance characteristics of the Lig proteins in diagnostic assays and in subunit vaccine formulations with the potential to confer heterologous protection.

1. Introduction

Pathogenic spirochaetes belonging to the genus *Leptospira* are the agents of leptospirosis, which is considered to be the most widespread zoonosis in the world (Faine et al., 1999; Levett, 2001; Bharti et al., 2003). Susceptible animals, including humans, are infected by direct contact with urine from a reservoir host, usually rats or other rodents, or indirectly through contaminated water. Transmission occurs via dermal abrasions or inoculation of the mucous or conjunctival membranes (Faine et al., 1999). In the majority of infected individuals, leptospirosis is a self-limited disease characterized by flu-like symptoms (Faine et al., 1999). However, hepatorenal manifestations, as observed in Weil's disease, are frequent complications and are associated with significant (10–15%) mortality (Bharti et al., 2003; McBride et al., 2005). In addition, leptospirosis causes severe pulmonary haemorrhage syndrome (SPHS), for which case fatality is >50% (Segura et al., 2005; Gouveia et al., 2008). Leptospirosis is considered to be an emerging infectious disease in endemic regions of Asia (Karande et al., 2003, 2005; LaRocque et al., 2005; Yanagihara et al., 2007; Peacock and Newton, 2008) and Latin America (Ko et al., 1999; Sarkar et al., 2002; Romero et al., 2003; Johnson et al., 2004) and is a major public health concern in

* Corresponding author. Tel.: +55 53 3275 7350; fax: +55 53 3275 7555.
E-mail address: odir@ufpel.edu.br (O.A. Dellagostin).
[1] The first two authors contributed equally to this work.

1567-1348/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2008.10.012

poverty stricken regions of the world (McBride et al., 2005; Ganoza et al., 2006; Riley et al., 2007).

The *Leptospira* genus is sub-classified into 18 genomospecies that includes both saprophytic and pathogenic species (Levett, 2001; Levett et al., 2006; Matthias et al., 2008). Classification based on serologic methods has identified ~300 serovars, of which more than 200 are considered to be pathogenic (Faine et al., 1999; Levett, 2001; Bharti et al., 2003). The availability of genomic sequence data from five *Leptospira* strains, *L. interrogans* serovars Lai (Ren et al., 2003) and Copenhageni (Nascimento et al., 2004), *L. borgpetersenii* serovar Hardjo strains L550 and JB197 (Bulach et al., 2006), and the saprophyte *L. biflexa* serovar Patoc I (Picardeau et al., 2008), is driving the discovery of new diagnostic tools and vaccines for leptospirosis. Considerable effort has been expended towards identifying conserved surface-exposed antigenic determinants that could improve diagnosis and provide heterologous protection via subunit or DNA vaccines.

A number of leptospiral outer membrane proteins (OMPs) have been characterized (Cullen et al., 2005), including OmpL1 (Haake et al., 1993), LipL41 (Shang et al., 1996), LipL36 (Haake et al., 1998), the major outer membrane protein, LipL32 (Haake et al., 2000), LipL21 (Cullen et al., 2003), LipL46 (Matsunaga et al., 2006), LenA (Verma et al., 2006), and the OmpA-like proteins Loa22 (Koizumi and Watanabe, 2003) and Omp52 (Hsieh et al., 2005). However, their performance in diagnostic assays for acute leptospirosis or as vaccine candidates has been problematic (Haake et al., 1999; Branger et al., 2001; Flannery et al., 2001; Guerreiro et al., 2001). LigA and LigB, belonging to a family of leptospiral immunoglobulin-like (Lig) proteins, appear to be promising antigens (Palaniappan et al., 2002; Matsunaga et al., 2003). The gene encoding a third Lig protein, *lig*C, was identified as a pseudogene in *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa (Matsunaga et al., 2003), but was found to be intact in *L. interrogans* serovar Lai (Ren et al., 2003). The Lig proteins contain a series of bacterial immunoglobulin-like (Big) repeat domains that were originally identified in virulence determinants from *Escherichia coli* and *Yersina pseudotuberculosis* (Hamburger et al., 1999; Luo et al., 2000).

The *lig* genes are of great interest because emerging serologic, vaccine, and pathogenesis studies indicate that Lig proteins are key virulence determinants involved in host–pathogen interactions. Lig proteins mediate interaction with multiple host extracellular matrix proteins, including fibronectin, fibrinogen, collagen, and laminin (Choy et al., 2007). Several studies have provided evidence that the Lig proteins are protective immunogens in animal models of leptospirosis (Koizumi and Watanabe, 2004; Palaniappan et al., 2006; Silva et al., 2007). In addition, we recently demonstrated that a recombinant polypeptide containing Big domains 2–6 from LigB was able to protect hamsters against homologous challenge by *L. interrogans* serovar Copenhageni (unpublished data). Virulent forms of *L. interrogans* and *L. kirschneri* strains express higher levels of Lig proteins than culture-attenuated forms (Matsunaga et al., 2003). Lig expression is strongly induced by shifting the osmolarity from low levels used in EMJH culture medium to osmolarity levels found in host tissues (Matsunaga et al., 2005). Up-regulation during early host infection is consistent with the strong serologic response to Lig proteins observed during acute leptospirosis (Croda et al., 2007).

Considering the large number of pathogenic *Leptospira* serovars and the broad distribution of leptospiral host reservoirs, the potential effect of selective pressure on the genetic diversity of the Lig proteins was unclear. Given the potential of the Lig proteins as diagnostic antigens and vaccine candidates, we examined their sequence diversity in the serovars most often associated with leptospirosis.

2. Materials and methods

2.1. Leptospira strains and culture conditions

Virulent leptospiral strains (Table 1) were obtained from culture collections maintained by the authors. The isolation conditions of a number of the strains used in this study were previously described (Ko et al., 1999; Haake et al., 2002; Silva et al., 2008). The identity of each of the strains used in this study was confirmed by 16S rRNA gene sequencing (Hookey et al., 1993) and serogrouping based on the microscopic agglutination test (MAT) (Cole et al., 1973). Strains were cultured in liquid Ellinghausen–McCullough–Johnson–Harris modified tween 80-bovine albumin medium (Ellinghausen and McCullough, 1965; Johnson and Harris,

Table 1
Leptospira strains, the status of their *lig* genes and level of identity compared to that of serovar Copenhageni.

Leptospira species	Serogroup	Serovar	Strain	*lig* gene status (PCR/SB/Seq)			% *lig* DNA sequence identity vs. Copenhageni		
				ligA	*ligB*	*ligC*	*ligA*	*ligB*	*ligC*
L. borgpetersenii		Hardjo	L550[a]	ND/ND/−	+/ND/+	ND/ND/−	NA	68.7	NA
		Hardjo	JB197[a]	ND/ND/−	+/ND/+	ND/ND/−	NA	67.9	NA
L. interrogans	Icterohaemorrhagiae	Copenhageni	Fiocruz L1-130[b]	+/+/+	+/+/+	+/+/+	100	100	100
	Icterohaemorrhagiae	Lai	56601[c]	−/ND/−	+/+/+	+/+/+	NA	97.1	99.5
	Canicola	ND	Kito[d]	+/+/+	+/+/+	+/ND/+	90.3	96.7	98.5
	Pomona	Pomona	Kennewicki Cornell[e]	+/ND/+	+/+/+	+/ND/+	90.1	97.1	99.5
	Pomona	Pomona	Kennewicki PO-06-047[f]	+/+/+	+/+/+	+/ND/+	90.3	96.7	98.8
L. kirschneri	Grippotyphosa	Grippotyphosa	RM52[g]	+/+/+	+/+/+	+/+/+	91.4	93.2	90.5
L. noguchii	Bataviae	ND	Cascata[h]	−/−/−	+/+/+	−/ND/−	NA	74.3	NA
L. weilii	Hebdomadis	ND	Eco-Challenge[i]	−/−/−	+/+/+	+/ND/+	NA	69.6	78

PCR: PCR screening assay; SB: Southern blot; Seq: gene sequenced; ND: not determined; NA: not applicable.
[a] Isolated from a chronically infected bovine, Australia, genome sequenced (Bulach et al., 2006).
[b] Isolated from a patient during a leptospirosis epidemic, Brazil (Ko et al., 1999), genome sequenced (Nascimento et al., 2004).
[c] Causative agent of rural leptospirosis in China, genome sequenced (Ren et al., 2003).
[d] Isolated from an infected canine, Brazil (Silva et al., 2008).
[e] Isolated from an aborted equine foetus, USA (Palaniappan et al., 2002).
[f] Isolated from an aborted swine foetus, USA.
[g] Isolated from an infected swine, USA (Shang et al., 1996).
[h] Isolated from a patient, Brazil (Silva et al., 2008).
[i] Isolated from a triathlete diagnosed with leptospirosis, Malaysia (Haake et al., 2002).

1967) at 30 °C. Virulence of each *Leptospira* strain was evaluated in the hamster model of lethal leptospirosis as previously described (Silva et al., 2008).

2.2. Genomic DNA extraction

Leptospiral genomic DNA was extracted from 7-day-old cultures using the GFX Genomic Blood DNA Purification Kit (GE Healthcare) according to the instructions provided for Gram-negative bacteria. The concentration was determined by absorbance at 260 nm and the quality of the genomic DNA was confirmed by agarose gel electrophoresis.

2.3. PCR detection of lig genes

A PCR-based screening method for the detection of the *lig* genes was developed. Degenerate primers (Table S4) were designed to amplify *ligA*, *ligB* and *ligC* based on sequence information deposited in GenBank. Additional primers were designed to amplify *ligB* from *L. weilii* strains due to the heterogeneity of *ligB* from this species compared to the other leptospiral species. The DNA sequence of each *lig* gene was aligned using AlignX (Vector NTI, Invitrogen) and, based on homology between the sequences, suitable regions were identified and primers, degenerate where appropriate, were designed. PCR (30 cycles) was carried out using recombinant *Taq* DNA polymerase (Invitrogen) following the protocol provided. PCR products were analysed by agarose gel electrophoresis. The expected product size for *ligA* was estimated to be 211 bp, 536 bp for *ligB* (1076 bp for *ligB* from *L. noguchii* and 1625 bp for *ligB* from *L. weilii*) and 248 bp for *ligC*. Quality control PCR amplification of the 16S rDNA and *lipL32* genes was used to verify the quality of the genomic DNA.

2.4. Southern blotting

Southern blotting was carried out as described previously (Sambrook and Russell, 2001). Briefly, 3 μg of genomic DNA was digested with 20 units of BamHI (Invitrogen) and separated by agarose gel electrophoresis. DNA was transferred from the gel to a positively charged Hybond-N nylon membrane (GE Healthcare) with a vacuum blotter (Bio-Rad). Probes to each of the *lig* genes were based on pooled PCR products amplified using the primers described in Table S4 and labelled using the ECL Direct Nucleic Acid Labelling and Detection System (GE Healthcare) as described in the protocol provided. Prehybridization was carried out at 42 °C for 1 h in hybridization buffer supplemented with 0.5 M NaCl and 5% blocking agent. Hybridization was carried out overnight at 42 °C in roller bottles. Following hybridization, the membrane was washed twice for 10 min at 55 °C in wash solution (0.4% SDS, 0.5× SSC). Finally, the membrane was washed twice in 2× SSC, 5 min per wash at room temperature. After incubation with ECL detection reagents, hybridization products were detected by exposure of the membrane to Hyperfilm ECL X-ray film (GE Healthcare).

2.5. DNA sequencing

Dependant on their presence, full-length *ligA* (~3.7 kb), *ligB* (~5.7 kb) and *ligC* (~5.9 kb) genes from *L. interrogans* serovars Pomona and Canicola, *L. noguchii* serogroup Bataviae and *L. weilii* serogroup Hebdomadis were amplified using Elongase Mix (Invitrogen), which contains a proof-reading polymerase, and subsequently cloned using the TOPO-TA cloning kit (Invitrogen). Each gene sequence was determined by direct sequencing of PCR products amplified using the cloned genes and internal primers (Table S4). Ambiguous bases or those different to previously published sequences were resolved by direct sequencing of PCR products amplified from genomic DNA. Each base was sequenced on both strands a minimum of two times resulting in each base being sequenced a minimum of four times. Previously unpublished sequences were submitted to GenBank and assigned accession numbers EU700267 to EU700275.

2.6. DNA and protein sequence analysis

Raw DNA sequences were analysed for the quality of base calling using Phred (Ewing et al., 1998), the 5′ and 3′ ends were trimmed accordingly using FinchTV (Geospiza Inc.) and each contig was assembled using Contig Express (Vector NTI, Invitrogen) at the default settings. Consensus DNA sequences were exported and coding sequences (CDS) identified using Vector NTI. CDS were aligned using AlignX (Vector NTI, Invitrogen), based on the ClustalW algorithm (Thompson et al., 1994), at the default settings. Protein sequences were generated by translation of the CDS and were aligned using AlignX at the default settings. The number of observed synonymous (S_d), and nonsynonymous (N_d) substitutions in the *lig* genes were calculated using Syn-SCAN (Gonzales et al., 2002). The evolutionary history relating the *ligA*, *ligB*, and *ligC* genes from a total of 14 strains was examined using a statistical phylogenetic approach. Alignments of the genes were based on their amino acid translations to maintain reading frames and then reverse translated back to their nucleotide sequences using the Geneious Pro 3.0.1 software package (Drummond et al., 2007). Possible intragenic recombination was examined using the DualBrothers plugin for Geneious. Default settings were used with the following exceptions: chain length 220,000, subsampling frequency 100 and burn-in length 10,000. DualBrothers is a recombination detection algorithm based on a phylogenetic dual multiple change-point model (MCP) (Suchard et al., 2002, 2003; Minin et al., 2005). The MCP model allows for changes in evolutionary relationships and rates across sites in a multiple sequence alignment by assuming that the sites separate into an unknown number of contiguous segments, each with possibly different topologies or mutation processes. Differing evolutionary topologies on either side of a break-point suggests recombination (Li et al., 1988). The DualBrothers implementation takes a Bayesian approach that employs Markov chain Monte Carlo to simulate from the posterior distribution of model parameters. One strength of the Bayesian model is that it can measure uncertainty in break-point locations, determine the most likely parental sequences of the putative recombinant and assess the statistical significance of recombination simultaneously; this simultaneous approach avoids the pitfall inherent in sequential testing for recombination found in many recombination detection programs (Suchard et al., 2002). For each putative recombinant, break-points were inferred and parental representatives and the significance of recombination using the approach of Suchard et al. (2003) were assessed.

Statistical analysis was carried out using Graphpad Prism 5.01 (GraphPad Software, San Diego, CA, USA). The Mann–Whitney test was used to calculate differences between populations and two-tailed P values <0.05 were considered significant.

3. Results

3.1. The lig genes in pathogenic Leptospira spp.

A PCR screening assay demonstrated that *ligB* was present in all 10 strains studied, *ligA* in five strains and *ligC* in seven strains (Table 1). Using the *lig* sequences deposited in GenBank, primers were designed to screen for the presence of *lig* genes in the leptospiral species most commonly associated with human

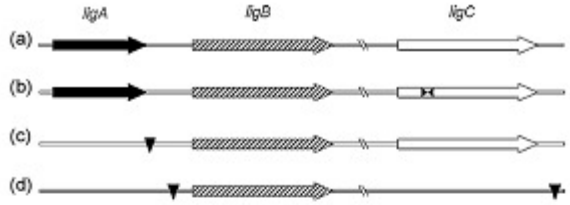


Fig. 1. Genomic organization of the *lig* genes in *Leptospira* spp. The individual *lig* genes are represented by arrows, the presence of a stop codon is indicated by the ▶◀ symbol and transposase genes are designated by the ▼ symbol. The *ligA* (solid arrow) and *ligB* (hatched arrow) genes are located within 0.9–1.4 kb of each other in the same orientation within the genome in an operon-like arrangement. The *ligC* gene (open arrow) is distally located on the opposite side of the genome (~715 kb distant). Genomic organization in the *Leptospira* strains fell into one of four distinct groups: (a) full-length intact *lig* genes: *L. interrogans* serovar Canicola and serovar Pomona; (b) intact *ligA*, *ligB* and *ligC* pseudogene: *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa; (c) intact *ligB* and *ligC*: *L. interrogans* serovar Lai and *L. weilii* serogroup Hebdomadis; (d) intact *ligB* only: *L. borgpetersenii* serovar Hardjo and *L. noguchii* serogroup Bataviae strains.

leptospirosis (Table 1). *L. interrogans* serovars Copenhageni Fiocruz L1-130, Pomona Kennewicki, Canicola Kito and *L. kirschneri* serovar Grippotyphosa RM52 were found to contain three *lig* genes (Fig. 1). *L. interrogans* serovar Lai 56601 and *L. weilii* serogroup Hebdomadis Eco-Challenge were found to contain *ligB* and *ligC* while *L. noguchii* serogroup Bataviae Cascata and *L. borgpetersenii* serovar Hardjo strains JB197 and L550 contained only *ligB* (Fig. 1). Of note, in place of the *ligA* gene in *L. interrogans* serovar Lai and the two *L. borgpetersenii* serovar Hardjo genomes there were transposases (Fig. 1). In serovar Lai, upstream of *ligB* and where *ligA* would normally be located, two transposase genes were identified, Tn*8* and an integrase, and a *ccrB* ortholog (encodes camphor resistance in bacteria). *L. borgpetersenii* serovar Hardjo strain JB197 contained a transposase, IS*1533*, and the C-terminal of the *ccrB* ortholog. *L. borgpetersenii* serovar Hardjo L550 contained an IS*1533* pseudogene and the C-terminal region of the *ccrB* gene. The JB197 strain has a genome layout similar to that of serovar Lai with a genome inversion (Ren et al., 2003), while the genome organization of strain L550 more closely resembles that of the serovars Copenhageni, Pomona and Canicola.

To confirm the integrity of the *lig* genes in pathogenic *Leptospira* spp. they were cloned and sequenced from *L. interrogans* serovar Canicola, serovar Pomona, *L. noguchii* serogroup Bataviae and *L. weilii* serogroup Hebdomadis. The presence or absence of the *lig* genes was also confirmed by Southern blotting (data not shown). The *ligA* and *ligB* genes in serovars Pomona and Canicola were of the expected sizes and genome locations, with *ligA* situated upstream of *ligB* in both strains. We previously reported that the *ligC* gene was a pseudogene in *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa (Matsunaga et al., 2003). In this study we found *ligC* to be present in *L. interrogans* serovars Pomona and Canicola, and *L. weilii* serogroup Hebdomadis, but absent in *L. noguchii* serogroup Bataviae.

3.2. Lig sequence conservation

LigB was the only *lig* gene ortholog found in all 10 pathogenic strains of *Leptospira* spp. studied (Table 1) and was compared by estimating pair-wise sequence distances. LigB was found to be significantly more variable than *ligA* ($P < 0.05$) and *ligC* ($P < 0.005$), ranging from 67.9 to 99.9% (mean 80.2 ± 13.1%) and 62.6 to 99.9% (mean 77.3 ± 15.3%) for pair-wise DNA and amino acid sequence identity, respectively (Table S1). The base substitutions in *ligB* were found throughout the gene, although there was increased variability within a region of the carboxyterminal domain (amino acids 1518–1669). When phylogenetic groupings were analysed, *L. interrogans* serovars Copenhageni, Canicola, Lai and Pomona had a 95.4 and 94.4% mean pair-wise DNA and amino acid sequence identity, respectively. The mean number of observed synonymous substitutions (S_d) was 77.8, the mean number of observed nonsynonymous substitutions (N_d) was 65.5 and the ratio of synonymous to nonsynonymous substitutions (d_S/d_N) was 4.01. Inclusion of *L. kirschneri* serovar Grippotyphosa in this grouping reduced, but not significantly, the mean identity observed to 91.0 and 90.6% for DNA and amino acid sequences, respectively ($d_S/d_N = 4.88$). The two *L. borgpetersenii* serovar Hardjo strains and the *L. weilii* Eco-Challenge strain had a 90.9 and 90.2% mean DNA and amino acid sequence identity, respectively, and the majority of the base substitutions were synonymous (mean $S_d = 182.7$, mean $N_d = 125.7$, mean $d_S/d_N = 4.24$). LigB from the *L. noguchii* strain did not fall within either of these two groupings and inclusion in either group increased the levels of sequence variability by 20–30% (Table S1).

LigA from five strains (Table 1) was highly conserved, ranging from 85.4 to 99.8% (mean 91.8 ± 5.9%) and 80.5 to 99.8% (mean 88.9 ± 6.8%) DNA and amino acid sequence identity, respectively (Table S2). The carboxyterminal region of LigA, comprising Big domain repeats 11–13 (see Fig. 2a), exhibited the highest level of variability, with 78.1 and 71.2% mean pair-wise DNA and amino acid identity, respectively. Although LigA sequence conservation was high, the majority of the single-nucleotide polymorphisms resulted in nonsynonymous amino acid substitutions (mean $S_d = 119.7$, mean $N_d = 174.9$, mean $d_S/d_N = 2.89$). Of note, the serovar Pomona LigA amino acid sequence previously deposited in GenBank ((Palaniappan et al., 2002), accession number AAN52495), contained a region (amino acids 740–774) of very low identity (<10%) compared to the other LigA sequences, specifically, point-insertions at nucleotide positions 2218, 2315 and 2322. We resequenced the *ligA* gene from the serovar Pomona Kennewicki PO-06-047 strain, accession number EU700270, and did not find the point-insertions.

The *ligC* gene was present in 7 of the 10 strains evaluated in this study. Furthermore, it was the most conserved Lig protein. LigC DNA and amino acid sequence identity ranged from 77.0 to 100% (mean 91.2 ± 8.9%) and 83.6 to 99.9% (mean 94.1 ± 6.2%), respectively (Table S3). Similar to *ligB*, the base substitutions are spread throughout the gene. The majority of the base substitutions were nonsynonymous (mean $S_d = 96.1$, mean $N_d = 342.9$, mean $d_S/d_N = 1.03$). Previously we found that the *ligC* genes from *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa appear to be pseudogenes. Serovar Copenhageni contained a point mutation while serovar Grippotyphosa contained a frame-shift mutation, both of which resulted in stop codons (Matsunaga et al., 2003). The major difference between the various *ligC* sequences was a gap towards the middle of *ligC* (nucleotides 2839–3104) from the Eco-Challenge strain, as there was one less Big repeat domain in the *L. weilii ligC*.

3.3. Phylogenetic analysis of Lig domains

Translation-based alignment of nucleotide sequences of the immunoglobulin-like domains of LigA, LigB, and LigC of *L. interrogans* serovar Copenhageni revealed clustering of LigB domains 5–10 with LigA domains 8–13 (Fig. 2b). The same relatedness pattern was found for the Big domains of LigA, B, and C of *L. kirschneri* serovar Grippotyphosa (data not shown). This result suggests that the *ligA* gene was derived from the *ligB* gene in a partial gene duplication event (Fig. 2a). The *ligA* genes were found only in *L. interrogans* and *L. kirschneri* strains, indicating that the partial gene duplication event occurred in a progenitor of the *L. interrogans*–*L. kirschneri* branch of the leptospiral evolutionary

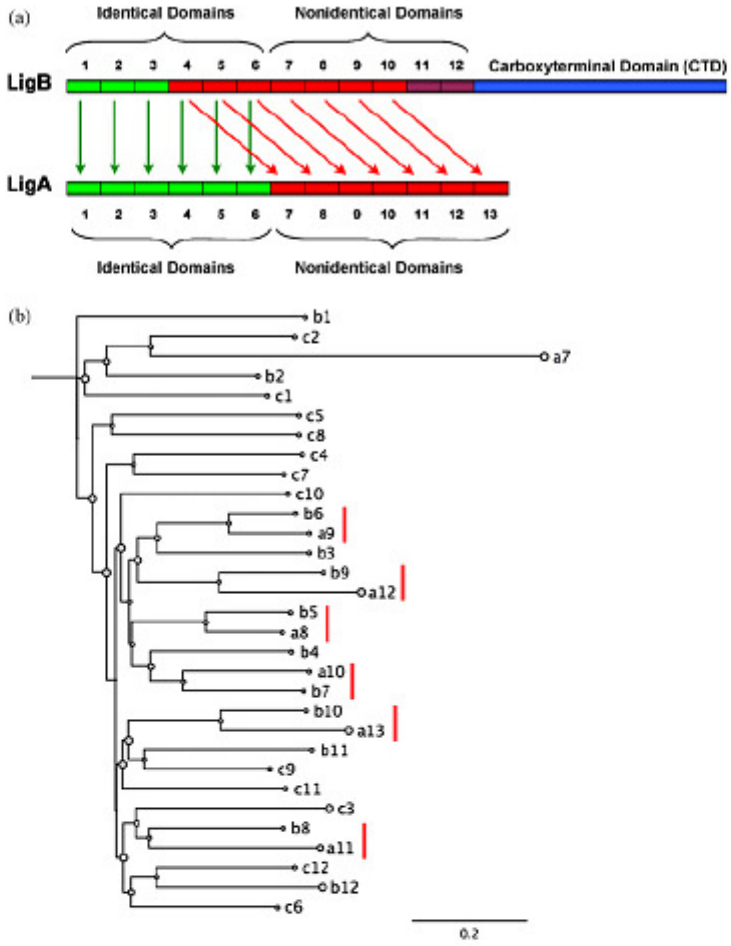


Fig. 2. (a) Proposed mechanism of *ligA* gene creation. A two-step partial gene duplication process is proposed: in the first step (green arrows), the first six Big domains of *ligA* were derived from the first six domains of *ligB*. In the second step (red arrows), Big domains 7–13 of *ligA* were derived from the domains 4–10 of *ligB*. Ordering of steps remains unknown. The individual Big domains comprising the *lig* genes are numbered, and the carboxyterminal domain of *ligB* is indicated. (b) Phylogenetic tree of the Lig domains. The tree is built from a translation-based alignment of nucleotide sequences of the immunoglobulin-like domains of LigA (a1–13), LigB (b1–12), and LigC (c1–12) of *L. interrogans* serovar Copenhageni. Vertical red lines indicate clustering of corresponding LigB domains 5–10 with LigA domains 8–13 in a "+3" relationship. LigC domains are more irregularly related than the LigA and LigB domains (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

tree. LigC domains were more distantly related than the LigA and LigB domains, consistent with the more ancient origin of *ligC* (see below).

3.4. Phylogenetic analysis of ligB

The relatedness of the 10 full-length *ligB* sequences is presented in Fig. 3a. The *ligB* sequences of *L. borgpetersenii* and *L. weilii* cluster together in one region of the tree while the *L. interrogans* and *L. kirschneri* sequences cluster together in a separate region of the tree, with the *ligB* sequence of *L. noguchii* occupying an intermediate position. This relatedness pattern is consistent with a phylogenetic tree based on 16S rDNA sequences (Haake et al., 2004). Analysis using the multiple change-point (MCP) model revealed that all *ligB* genes except two were phylogenetically clonal (no evidence of rearrangements). The *ligB* sequence of *L. interrogans* serovar Copenhageni was found to have two *L. kirschneri*-like regions: a 153 nucleotide region in Big domain 11, and a 500 nucleotide region in the carboxyterminal domain (Fig. 3b). The *ligB* sequence of *L. kirschneri* serovar Grippotyphosa was found to have a large 1300 nucleotide rearrangement containing an *L. interrogans* serovar Lai-like sequence in the region coding for the carboxyterminal domain (Fig. 3c).

3.5. Phylogenetic analysis of ligA

There were two different phylogenetic trees for the 5′ (nucleotides 1–2820) and 3′ regions (nucleotides 2821–3675) of the *ligA* gene sequences. The phylogenetic tree for the 5′ region (Fig. 4a) showed low levels of sequence non-identity (ranging from 1 to 9%) for the five strains. In contrast, the tree for the 3′ region (Fig. 4b) revealed that while the *ligA* sequences of the Canicola and Pomona strains are 100% identical, their *ligA* sequences were 33% non-identical to the Grippotyphosa and Copenhageni *ligA* sequences, a 4–10-fold increase in sequence non-identity. One possible interpretation of this difference in sequence variation for

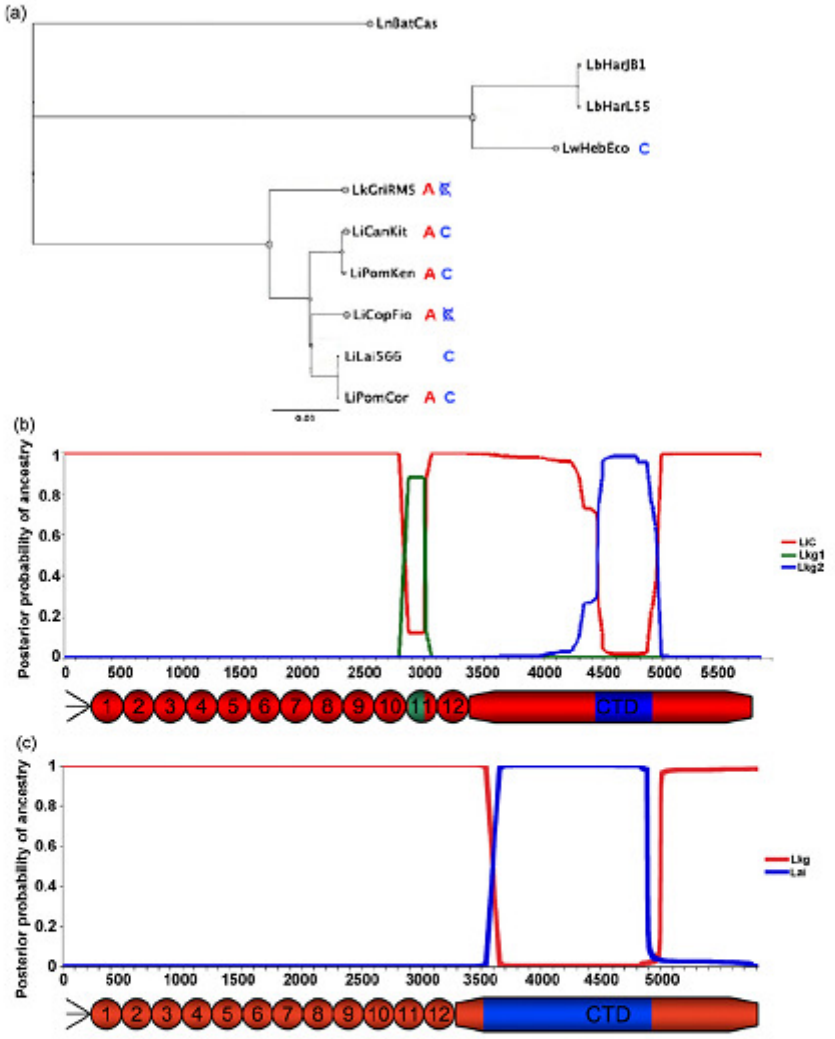


Fig. 3. (a) Phylogenetic tree of *ligB* sequences, the tree is built from the ten available full-length *ligB* sequences, and shows the same branching relationships as would be anticipated from the 16S-based phylogenetic tree for the five species represented. Strains with *ligA* (red) and *ligC* (blue) genes are shown. The *ligC* genes of *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa are pseudogenes. Mosaicism in the *ligB* genes, MCP analysis of the *ligB* genes of *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa revealed two distinct mosaic patterns. (b) In the first pattern, the *ligB* gene of *L. interrogans* serovar Copenhageni contains two regions of *L. kirschneri*-like sequences: a 153 nucleotide region in Big domain 11 (green), and a 500 nucleotide region in the carboxyterminal domain (blue). (c) In the second pattern, the *ligB* gene of *L. kirschneri* serovar Grippotyphosa contains a large 1300 nucleotide region (blue) containing an *L. interrogans* Lai-like sequence in the region coding for the carboxyterminal domain (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

the two different regions of the *ligA* gene is the recent acquisition the 3′ *ligA* region, encoding the repeats 11–13 (Fig. 4c).

3.6. Phylogenetic analysis of ligC

In contrast to the distribution of the *ligA* genes, the *ligC* genes were found not only in *L. interrogans* and *L. kirschneri*, but also in *L. weilii*. The sequence identity between the *ligB* and *ligC* sequences of *L. interrogans* and *L. weilii* was similar (69 and 78%, respectively), suggesting that the *ligB* and *ligC* genes coevolved with leptospiral evolution, rather than the *L. weilii ligC* gene representing a more recent horizontal acquisition. MCP analysis of the seven *ligC* genes revealed a single tree structure consistent with leptospiral evolution throughout, except for an *L. interrogans*-like 640 nucleotide insertion (4660–5300) into the carboxyterminal domain coding region of the *ligC* gene of *L. weilii* serovar Eco-Challenge (Fig. 5). In addition, the *ligC* gene of *L. weilii* serovar Eco-Challenge was found to lack one Big repeat domain (domain 11).

4. Discussion

Genome sequencing studies have demonstrated that pathogenic *Leptospira* spp. contain *ligB* together with up to two *lig* paralogs, *ligA* and *ligC* (Ren et al., 2003; Nascimento et al., 2004; Bulach et al., 2006), while they are absent from the non-pathogenic saprophyte, *L. biflexa* (Matsunaga et al., 2003; Picardeau et al.,

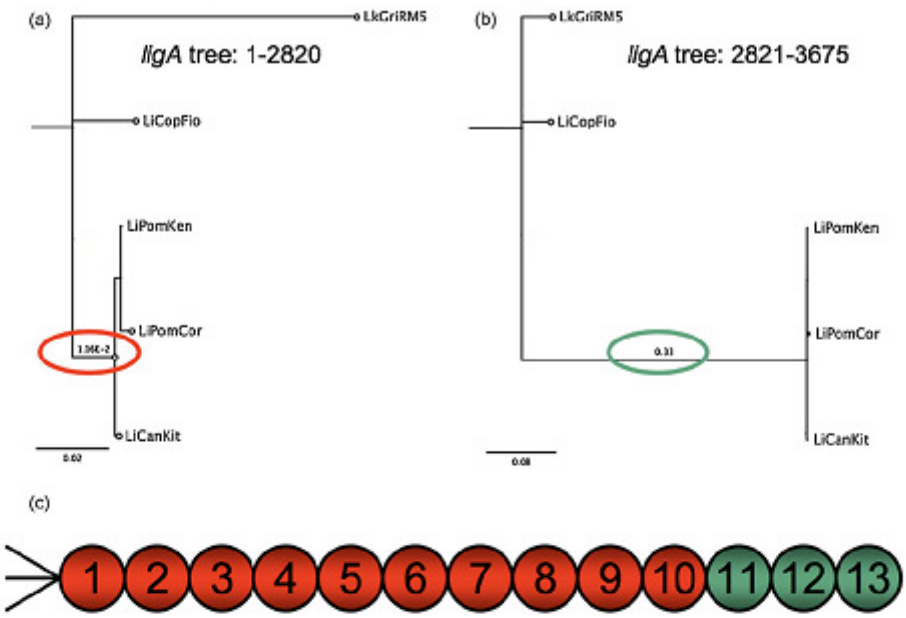


Fig. 4. Evidence for recombination in *ligA*. Comparison of five *ligA* genes revealed two different evolutionary rates for the 5′ and 3′ regions. (a) The phylogenetic tree for the 5′ region showed low levels of sequence non-identity (ranging from 1 to 9%) for the five strains (key branch length circled in red). (b) In contrast, the tree for the 3′ region revealed that while the *ligA* sequences of the Canicola and Pomona strains are 100% identical, their *ligA* sequences were 33% non-identical to the Grippotyphosa and Copenhageni *ligA* sequences (key branch length circled in green). (c) One interpretation of this result is that the last three Big domains of Canicola and Pomona strains were acquired by horizontal DNA transfer (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

2008). However, little was known about the distribution of the *lig* genes among pathogenic leptospiral strains or their interrelationships. Our study demonstrates that an intact *ligB* gene is found in all leptospiral pathogens studied to date, suggesting an important, or perhaps essential, role in virulence. The *ligA* gene appears to have been derived from *ligB* by a two-step partial gene duplication process. LigC is structurally similar to LigB and *ligC* gene variation also follows phylogenetic patterns, suggesting an early gene duplication event. However, the role of LigC in virulence is less clear, as some strains have lost *ligC*, while in others, such as *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa, *ligC* is a pseudogene. Of note, we were unable to demonstrate any association between the number of *lig* genes, their diversity, or the hosts from which they were isolated and the degree of virulence in the hamster model (data not shown). There is strong evidence that LigA is expressed during infection (Palaniappan et al., 2002; Matsunaga et al., 2003; Koizumi and Watanabe, 2004; Silva et al., 2007; Srimanote et al., 2008) and yet several virulent pathogenic *Leptospira* strains do not contain *ligA* (Table 1). A possible explanation is that *ligA* and *ligB* are involved in virulence but that both copies are not required, an example of gene redundancy. Indeed, a recent report showed that a *ligB* knockout in *L. interrogans* strain Fiocruz L1-130 did not alter the virulence of the *ligB*$^{-}$ strain (Croda et al., 2008). As LigA was expressed in the *ligB*$^{-}$ strain it would appear that LigA can replace LigB during infection, although the role of the Lig proteins in virulence is not yet clear.

Phylogenetic analysis based on LigB, which is conserved in all strains, sorted the strains into three distinct groups: (i) *L. kirschneri*

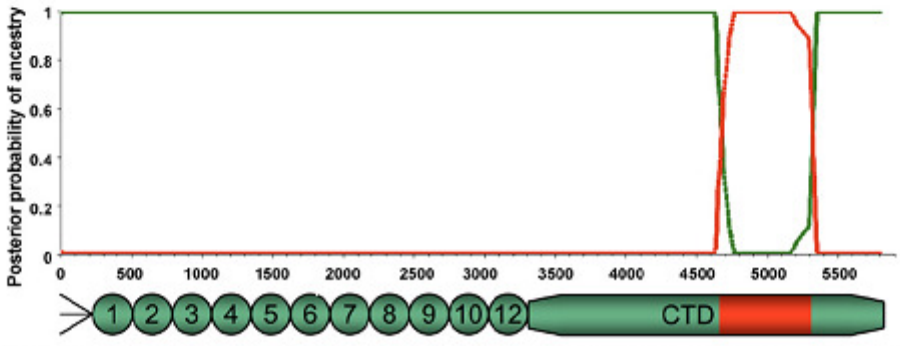


Fig. 5. Mosaicism in the *ligC* gene. MCP analysis of the *ligC* gene from the *L. weilii* serogroup Hebdomadis Eco-Challenge strain. The *ligC* gene of *L. weilii* has a 640 nucleotide interrogans-like insertion in the carboxyterminal domain (red). In addition, the *L. weilii* Eco-Challenge strain lacks Big domain 11 (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

and *L. interrogans*; (ii) *L. borgpetersenii* and *L. weilii*; and (iii) *L. noguchii* (Fig. 1). Amino acid sequence alignment of LigB reveals that overall the carboxyterminal domain is the most conserved region (>60% mean identity), although there are short, highly variable regions within this domain. However, when aligned by phylogenetic group the aminoterminal region (Big domains 1–12) exhibited a similar level of identity when compared with the carboxyterminal domain (>90% mean identity). LigA, although only present in *L. interrogans* and *L. kirschneri* strains, demonstrated a high level of conservation of Big domains 1–10 (89% mean identity), while the full-length LigA had a mean identity of 80%. LigC was the most conserved of the Lig proteins (90% mean identity) however, it appears to be a pseudogene in *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa (Matsunaga et al., 2003) and is absent in *L. borgpetersenii* and *L. noguchii* strains (this work). The most widely used method for identifying non-functional genes is the d_S/d_N test, which compares the rate of synonymous to nonsynonymous mutations (Nei and Kumar, 2000). Pseudogenes have no functional restrictions as they are not expressed and are therefore expected to have a d_S/d_N ratio that does not differ significantly from one. Analysis of *ligC* revealed a mean d_S/d_N ratio of 1.03 suggesting that *ligC* may have lost or is in the process of losing its role in virulence and strains that contain this gene may be subject to a genome reduction event in the future.

Evidence is presented for horizontal recombination events affecting all three *lig* genes. The results indicate that several types of evolutionary mechanisms have been acting on the *lig* genes, including genetic drift, gene duplication, and horizontal gene transfer. We find that genetic drift accounts for most of *ligB* genetic diversity, suggesting that *ligB* was acquired early during the evolution of leptospires from free-living saprophytes to colonizers of host tissues. Although the level of DNA and amino acid sequence diversity for the LigB was similar to what had been observed previously for the gene encoding the porin, OmpL1, it would have been difficult to anticipate what level of diversity to expect without a sequence analysis study of this type. Our previous study found surprisingly large differences in the rates of sequence variation among genes encoding surface-exposed leptospiral outer membrane proteins (Haake et al., 2004). Sequences of *lipL32* genes encoding the major outer membrane protein were highly conserved (3.1% DNA and 0.9% amino acid sequence non-identity). Strong *lipL32* sequence conservation was not anticipated because LipL32 is highly immunogenic; >95% of patients with leptospirosis have an antibody response to LipL32 (Flannery et al., 2001). In contrast to *lipL32*, genes encoding OmpL1 exhibited significantly higher rates of sequence variation (14.9% DNA and 9.6% amino acid sequence non-identity). Rates of sequence variation and recombination for the *lig* genes were comparable to those observed for the *ompL1* genes. However, there were some notable differences compared to the earlier study. For example, DNA sequence variability was higher than amino acid sequence variability for *lipL32*, *lipL41*, and *ompL1*. In contrast, the rates of DNA and amino acid sequence variation for the *lig* genes were comparable, indicating that *lig* genes had a higher overall rate of nonsynonymous sequence changes.

Our analysis revealed evidence for a second mechanism of sequence diversity for all three *lig* genes: recombination events and horizontal DNA transfer between related bacterial species. Two *ligB* genes and one *ligC* gene were found to be mosaics. The *ligB* gene of *L. interrogans* serovar Copenhageni was found to contain two *L. kirschneri*-like insertions: one insertion in Big domain 11 and a second insertion in the carboxyterminal domain (relative probability of recombination event >1000:1). The *ligB* gene of *L. kirschneri* strain RM52 and the *ligC* gene of *L. weilii* strain Ecochallenge included *L. interrogans*-like insertions in the regions encoding their carboxyterminal domains (relative probability of recombination event >1000:1). We previously reported mosaicism for 20% of genes encoding the outer membrane protein, OmpL1. Mosaicism does not affect all leptospiral outer membrane proteins, as no evidence of recombination events was found for the gene encoding the major outer membrane lipoprotein, LipL32, and only one recombination event affecting a second outer membrane lipoprotein, LipL41, from 38 different strains representing six pathogenic *Leptospira* spp. (Haake et al., 2004). In the case of *ligA*, three of five strains appear to have acquired the same DNA encoding their last three *ligA* Big domains. Phylogenetic comparison of the transferred DNA encoding the exogenous *ligA* Big domains 11–13 with all known *lig* Big domains shows that they are most closely related to their *ligA* orthologs in *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa (data not shown). However, insufficient *ligA* sequence data is available to determine the phylogenetic origin of this exogenous *ligA* DNA.

Of note, we found that the first six Big domains from *ligB* and the first six Big domains from *ligA* were essentially identical (98.5 ± 0.8% mean identity). This is an important observation for the future development of diagnostic reagents. Recombinant LigB polypeptides containing Big domains 2–6 from *L. interrogans* serovar Copenhageni and *L. kirschneri* serovar Grippotyphosa were evaluated as antigens for the diagnosis of leptospirosis. Sensitivity and specificity were reported to be >90 and >97%, respectively, during the acute-phase of leptospirosis (Croda et al., 2007). Furthermore, there was no evidence of major genetic rearrangements in this region. The lack of genetic drift within the identical regions of *ligB* and *ligA* is evidence of selective pressure, intragenic recombination or gene conversion. The remaining Big domains (7–12 in *ligB* and 7–13 in *ligA*) were considerably more variable (34.2 ± 1.6% mean identity). This is an important observation, as our previous findings indicate that these Big domains are involved in binding extracellular matrix proteins and fibrinogen (Choy et al., 2007). Together these findings support a role for the Lig proteins during the transmission of leptospirosis. The carboxyterminal Big domains may have evolved to recognise specific host extracellular matrix proteins. This region was subject to horizontal recombination between *Leptospira* spp. suggesting that increased variability in the carboxyterminal Big domains of *ligA* and *ligB* may have contributed towards the adaption to novel hosts and potentially accounting for the extensive serovar-host specificity that typifies leptospirosis.

Our study had several potential limitations. The status of the *ligB* gene of *L. interrogans* serovar Manilae remains to be clarified since the *ligB* gene sequence has been reported to lack the carboxyterminal domain due to a point mutation that created a stop codon (Koizumi and Watanabe, 2004). Due to the difficulties inherent in sequencing the *lig* genes it is possible that the *ligB* sequence for serovar Manilae contains an erroneous stop codon in the carboxyterminal domain. Bacterial genes containing indels resulting in premature stop codons are defined as pseudogenes and are not functional (Ochman and Davalos, 2006). Yet, experimental data suggested that LigB was expressed in *L. interrogans* serovar Manilae. Koizumi and Watanabe (2004) showed that sera from leptospirosis patients specifically recognized recombinant LigB cloned from serovar Manilae. The partial *ligB* sequence (accession number AB098517) demonstrated >95% identity with the same region in the other *ligB* orthologs (data not shown). However, we were unable to obtain the serovar Manilae strain to evaluate this possibility and therefore excluded the serovar Manilae *lig* gene sequences from further analyses. Nevertheless, sequence data from other serovars from the same Icterohaemorrhagiae serogroup (serovars Lai and Copenhageni) as well as other strains from the same species (serovars Canicola and Pomona) provided consistent

findings for *ligB* (see Table 1). Sequence alignment suggests that the previously sequenced *ligA* from *L. interrogans* serovar Pomona strain Kennewicki (GenBank sequence AAN52495) contains a highly variable region (<10% amino acid identity) due to three indels that alter the coding sequence of this region. These indels are either representative of strain-to-strain variation within serovar Pomona Kennewicki or of PCR artefacts. The *ligA* sequence data in this study was derived from three independent PCR products sequenced in both directions for the region in question. Each base was therefore sequenced a minimum of six times to rule out the possibility of PCR artefacts and sequence errors. As the strains are probably different, our strain was isolated from an aborted swine foetus while the other was isolated from a case of equine recurrent uveitis (Palaniappan et al., 2002), more serovar Pomona strains would need to be sequenced to resolve this issue. A further limitation is that we were only able to include 10 *Leptospira* strains in our analysis however, all of these strains were considered virulent.

To date, Lig research has focused mainly on the Big domains, which are highly antigenic and are the focus of most of the anti-Lig antibody response during leptospirosis infection. LigA and LigB repeats have been shown to provide protective immunity in animal models of leptospirosis (Koizumi and Watanabe, 2004; Palaniappan et al., 2006; Silva et al., 2007). Lastly, Big domains 7–13 of LigA and domains 7–12 of LigB have been shown to function in binding to host extracellular matrix proteins (Choy et al., 2007). The results presented here serve to highlight the potential importance of the carboxyterminal domains of LigB and to prompt studies investigating its cellular location and function. Sequence analysis of *lig* genes from multiple strains is important in the ongoing efforts to develop new Lig-based serologic tests and vaccines. LigB has been found in all strains studied, making it an ideal candidate for vaccine and diagnostic applications. Furthermore, sequence conservation of LigB from *L. interrogans*, which contains the serovars most important to public health, was high, >96%. The sequence information provided here will improve the performance characteristics of the Lig proteins in diagnostic and subunit vaccine formulations. Knowledge of sequence variations among Lig serodiagnostic antigens should prompt cross-reaction studies using sera from different human and animal populations and possibly inclusion of homologous antigens from antigenically distinct strains. Future Lig vaccine studies should evaluate cross-protection using antigens and challenge strains with different *lig* sequences in order to assess the significance of sequence variation on immunoprotection.

Acknowledgements

This work was supported by Bio-Manguinhos, Oswaldo Cruz Foundation, Brazilian Ministry of Health (grant 09224-7), Brazilian National Research Council (grants 01.06.0298.00 3773/2005, 420067/2005, 554788/2006 and 473006/2006-5); Public Health Service grants AI-34431 (to D.A.H.) and R01 AI034431 (to A.I.K.) from the National Institute of Allergy and Infectious Diseases, the VA Medical Research Funds (to D.A.H.) and grant D43 TW00919 (to A.I.K.) from the Fogarty International Center, National Institutes of Health. M.A.S. is supported by an Alfred P. Sloane Research Fellowship and a John Simon Guggenheim Memorial Fellowship.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2008.10.012.

References

Bharti, A.R., Nally, J.E., Ricaldi, J.N., Matthias, M.A., Diaz, M.M., Lovett, M.A., Levett, P.N., Gilman, R.H., Willig, M.R., Gotuzzo, E., Vinetz, J.M., 2003. Leptospirosis: a zoonotic disease of global importance. Lancet Infect. Dis. 3, 757–771.

Branger, C., Sonrier, C., Chatrenet, B., Klonjkowski, B., Ruvoen-Clouet, N., Aubert, A., Andre-Fontaine, G., Eloit, M., 2001. Identification of the hemolysis-associated protein 1 as a cross-protective immunogen of *Leptospira interrogans* by adenovirus-mediated vaccination. Infect. Immun. 69, 6831–6838.

Bulach, D.M., Zuerner, R.L., Wilson, P., Seemann, T., McGrath, A., Cullen, P.A., Davis, J., Johnson, M., Kuczek, E., Alt, D.P., Peterson-Burch, B., Coppel, R.L., Rood, J.I., Davies, J.K., Adler, B., 2006. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. Proc. Natl. Acad. Sci. U.S.A. 103, 14560–14565.

Choy, H.A., Kelley, M.M., Chen, T.L., Moller, A.K., Matsunaga, J., Haake, D.A., 2007. Physiological osmotic induction of *Leptospira interrogans* adhesion: LigA and LigB bind extracellular matrix proteins and fibrinogen. Infect. Immun. 75, 2441–2450.

Cole Jr., J.R., Sulzer, C.R., Pursell, A.R., 1973. Improved microtechnique for the leptospiral microscopic agglutination test. Appl. Microbiol. 25, 976–980.

Croda, J., Ramos, J.G., Matsunaga, J., Queiroz, A., Homma, A., Riley, L.W., Haake, D.A., Reis, M.G., Ko, A.I., 2007. *Leptospira* immunoglobulin-like proteins as a serodiagnostic marker for acute leptospirosis. J. Clin. Microbiol. 45, 1528–1534.

Croda, J., Figueira, C.P., Wunder Jr., E.A., Santos, C.S., Reis, M.G., Ko, A.I., Picardeau, M., 2008. Targeted mutagenesis in pathogenic leptospira: disruption of the LigB gene does not affect virulence in animal models of leptospirosis. Infect. Immun. 76, 5826–5833.

Cullen, P.A., Haake, D.A., Bulach, D.M., Zuerner, R.L., Adler, B., 2003. LipL21 is a novel surface-exposed lipoprotein of pathogenic *Leptospira* species. Infect. Immun. 71, 2414–2421.

Cullen, P.A., Xu, X., Matsunaga, J., Sanchez, Y., Ko, A.I., Haake, D.A., Adler, B., 2005. Surfaceome of Leptospira spp. Infect. Immun. 73, 4853–4863.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2007. Geneious v3.5. Available from http://www.geneious.com/.

Ellinghausen Jr., H.C., McCullough, W.G., 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.

Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175–185.

Faine, S.B., Adler, B., Bolin, C., Perolat, P., 1999. *Leptospira* and Leptospirosis. MediSci, Melbourne, Australia.

Flannery, B., Costa, D., Carvalho, F.P., Guerreiro, H., Matsunaga, J., da Silva, E.D., Ferreira, A.G.P., Riley, L.W., Reis, M.G., Haake, D.A., Ko, A.I., 2001. Evaluation of recombinant *Leptospira* antigen-based enzyme-linked immunosorbent assays for the serodiagnosis of leptospirosis. J. Clin. Microbiol. 39, 3303–3310.

Ganoza, C.A., Matthias, M.A., Collins-Richards, D., Brouwer, K.C., Cunningham, C.B., Segura, E.R., Gilman, R.H., Gotuzzo, E., Vinetz, J.M., 2006. Determining risk for severe leptospirosis by molecular analysis of environmental surface waters for pathogenic *Leptospira*. PLoS Med. 3, e308.

Gonzales, M.J., Dugan, J.M., Shafer, R.W., 2002. Synonymous-non-synonymous mutation rates between sequences containing ambiguous nucleotides (SynSCAN). Bioinformatics 18, 886–887.

Gouveia, E.L., Metcalfe, J., de Carvalho, A.L., Aires, T.S., Villasboas-Bisneto, J.C., Queirroz, A., Santos, A.C., Salgado, K., Reis, M.G., Ko, A.I., 2008. Leptospirosis-associated severe pulmonary hemorrhagic syndrome, Salvador, Brazil. Emerg. Infect. Dis. 14, 505–508.

Guerreiro, H., Croda, J., Flannery, B., Mazel, M., Matsunaga, J., Galvao, R.M., Levett, P.N., Ko, A.I., Haake, D.A., 2001. Leptospiral proteins recognized during the humoral immune response to leptospirosis in humans. Infect. Immun. 69, 4958–4968.

Haake, D.A., Champion, C.I., Martinich, C., Shang, E.S., Blanco, D.R., Miller, J.N., Lovett, M.A., 1993. Molecular cloning and sequence analysis of the gene encoding OmpL1, a transmembrane outer membrane protein of pathogenic *Leptospira* spp. J. Bacteriol. 175, 4225–4234.

Haake, D.A., Martinich, C., Summers, T.A., Shang, E.S., Pruetz, J.D., McCoy, A.M., Mazel, M.K., Bolin, C.A., 1998. Characterization of leptospiral outer membrane lipoprotein LipL36: downregulation associated with late-log-phase growth and mammalian infection. Infect. Immun. 66, 1579–1587.

Haake, D.A., Mazel, M.K., McCoy, A.M., Milward, F., Chao, G., Matsunaga, J., Wagar, E.A., 1999. Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection. Infect. Immun. 67, 6572–6582.

Haake, D.A., Chao, G., Zuerner, R.L., Barnett, J.K., Barnett, D., Mazel, M., Matsunaga, J., Levett, P.N., Bolin, C.A., 2000. The leptospiral major outer membrane protein LipL32 is a lipoprotein expressed during mammalian infection. Infect. Immun. 68, 2276–2285.

Haake, D.A., Dundoo, M., Cader, R., Kubak, B.M., Hartskeerl, R.A., Sejvar, J.J., Ashford, D.A., 2002. Leptospirosis, water sports, and chemoprophylaxis. Clin. Infect. Dis. 34, e40–43.

Haake, D.A., Suchard, M.A., Kelley, M.M., Dundoo, M., Alt, D.P., Zuerner, R.L., 2004. Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer. J. Bacteriol. 186, 2818–2828.

Hamburger, Z.A., Brown, M.S., Isberg, R.R., Bjorkman, P.J., 1999. Crystal structure of invasin: a bacterial integrin-binding protein. Science 286, 291–295.

Hookey, J.V., Bryden, J., Gatehouse, L., 1993. The use of 16S rDNA sequence analysis to investigate the phylogeny of Leptospiraceae and related spirochaetes. J. Gen. Microbiol. 139, 2585–2590.
Hsieh, W.J., Chang, Y.F., Chen, C.S., Pan, M.J., 2005. Omp52 is a growth-phase-regulated outer membrane protein of Leptospira santarosai serovar Shermani. FEMS Microbiol. Lett. 243, 339–345.
Johnson, M.A., Smith, H., Joeph, P., Gilman, R.H., Bautista, C.T., Campos, K.J., Cespedes, M., Klatsky, P., Vidal, C., Terry, H., Calderon, M.M., Coral, C., Cabrera, L., Parmar, P.S., Vinetz, J.M., 2004. Environmental exposure and leptospirosis, Peru. Emerg. Infect. Dis. 10, 1016–1022.
Johnson, R.C., Harris, V.G., 1967. Differentiation of pathogenic and saprophytic leptospires. I. Growth at low temperatures. J. Bacteriol. 94, 27–31.
Karande, S., Bhatt, M., Kelkar, A., Kulkarni, M., De, A., Varaiya, A., 2003. An observational study to detect leptospirosis in Mumbai, India, 2000. Arch. Dis. Child. 88, 1070–1075.
Karande, S., Gandhi, D., Kulkarni, M., Bharadwaj, R., Pol, S., Thakare, J., De, A., 2005. Concurrent outbreak of leptospirosis and dengue in Mumbai, India, 2002. J. Trop. Pediatr. 51, 174–181.
Ko, A.I., Galvao Reis, M., Ribeiro Dourado, C.M., Johnson Jr., W.D., Riley, L.W., Salvador Leptospirosis Study Group, 1999. Urban epidemic of severe leptospirosis in Brazil. Lancet 354, 820–825.
Koizumi, N., Watanabe, H., 2003. Molecular cloning and characterization of a novel leptospiral lipoprotein with OmpA domain. FEMS Microbiol. Lett. 226, 215–219.
Koizumi, N., Watanabe, H., 2004. Leptospiral immunoglobulin-like proteins elicit protective immunity. Vaccine 22, 1545–1552.
LaRocque, R.C., Breiman, R.F., Ari, M.D., Morey, R.E., Janan, F.A., Hayes, J.M., Hossain, M.A., Brooks, W.A., Levett, P.N., 2005. Leptospirosis during dengue outbreak, Bangladesh. Emerg. Infect. Dis. 11, 766–769.
Levett, P.N., 2001. Leptospirosis. Clin. Microbiol. Rev. 14, 296–326.
Levett, P.N., Morey, R.E., Galloway, R.L., Steigerwalt, A.G., 2006. *Leptospira broomii* sp. nov., isolated from humans with leptospirosis. Int. J. Syst. Evol. Microbiol. 56, 671–673.
Li, W.H., Tanimura, M., Sharp, P.M., 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. 5, 313–330.
Luo, Y., Frey, E.A., Pfuetzner, R.A., Creagh, A.L., Knoechel, D.G., Haynes, C.A., Finlay, B.B., Strynadka, N.C., 2000. Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. Nature 405, 1073–1077.
Matsunaga, J., Barocchi, M.A., Croda, J., Young, T.A., Sanchez, Y., Siqueira, I., Bolin, C.A., Reis, M.G., Riley, L.W., Haake, D.A., Ko, A.I., 2003. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. Mol. Microbiol. 49, 929–945.
Matsunaga, J., Sanchez, Y., Xu, X., Haake, D.A., 2005. Osmolarity, a key environmental signal controlling expression of leptospiral proteins LigA and LigB and the extracellular release of LigA. Infect. Immun. 73, 70–78.
Matsunaga, J., Werneid, K., Zuerner, R.L., Frank, A., Haake, D.A., 2006. LipL46 is a novel surface-exposed lipoprotein expressed during leptospiral dissemination in the mammalian host. Microbiology 152, 3777–3786.
Matthias, M.A., Ricaldi, J.N., Cespedes, M., Diaz, M.M., Galloway, R.L., Saito, M., Steigerwalt, A.G., Patra, K.P., Ore, C.V., Gotuzzo, E., Gilman, R.H., Levett, P.N., Vinetz, J.M., 2008. Human leptospirosis caused by a new, antigenically unique *Leptospira* associated with a *Rattus* species reservoir in the Peruvian Amazon. PLoS Negl. Trop. Dis. 2, e213.
McBride, A.J., Athanazio, D.A., Reis, M.G., Ko, A.I., 2005. Leptospirosis. Curr. Opin. Infect. Dis. 18, 376–386.
Minin, V.N., Dorman, K.S., Fang, F., Suchard, M.A., 2005. Dual multiple change-point model leads to more accurate recombination detection. Bioinformatics 21, 3034–3042.
Nascimento, A.L., Ko, A.I., Martins, E.A., Monteiro-Vitorello, C.B., Ho, P.L., Haake, D.A., Verjovski-Almeida, S., Hartskeerl, R.A., Marques, M.V., Oliveira, M.C., Menck, C.F., Leite, L.C., Carrer, H., Coutinho, L.L., Degrave, W.M., Dellagostin, O.A., El-Dorry, H., Ferro, E.S., Ferro, M.I., Furlan, L.R., Gamberini, M., Giglioti, E.A., Goes-Neto, A., Goldman, G.H., Goldman, M.H., Harakava, R., Jeronimo, S.M., Junqueira-de-Azevedo, I.L., Kimura, E.T., Kuramae, E.E., Lemos, E.G., Lemos, M.V., Marino, C.L., Nunes, L.R., de, O.R.C., Pereira, G.G., Reis, M.S., Schriefer, A., Siqueira, W.J., Sommer, P., Tsai, S.M., Simpson, A.J., Ferro, J.A., Camargo, L.E., Kitajima, J.P., Setubal, J.C., Van, S.M.A., 2004. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. J. Bacteriol. 186, 2164–2172.
Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press US, New York.
Ochman, H., Davalos, L.M., 2006. The nature and dynamics of bacterial genomes. Science 311, 1730–1733.
Palaniappan, R.U., Chang, Y.F., Jusuf, S.S., Artiushin, S., Timoney, J.F., McDonough, S.P., Barr, S.C., Divers, T.J., Simpson, K.W., McDonough, P.L., Mohammed, H.O., 2002. Cloning and molecular characterization of an immunogenic LigA protein of *Leptospira interrogans*. Infect. Immun. 70, 5924–5930.
Palaniappan, R.U., McDonough, S.P., Divers, T.J., Chen, C.S., Pan, M.J., Matsumoto, M., Chang, Y.F., 2006. Immunoprotection of recombinant leptospiral immunoglobulin-like protein A against *Leptospira interrogans* serovar Pomona infection. Infect. Immun. 74, 1745–1750.
Peacock, S.J., Newton, P.N., 2008. Public health impact of establishing the cause of bacterial infections in rural Asia. Trans. R. Soc. Trop. Med. Hyg. 102, 5–6.
Picardeau, M., Bulach, D.M., Bouchier, C., Zuerner, R.L., Zidane, N., Wilson, P.J., Creno, S., Kuczek, E.S., Bommezzadri, S., Davis, J.C., McGrath, A., Johnson, M.J., Boursaux-Eude, C., Seemann, T., Rouy, Z., Coppel, R.L., Rood, J.I., Lajus, A., Davies, J.K., Medigue, C., Adler, B., 2008. Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis. PLoS ONE 3, e1607.
Ren, S.X., Fu, G., Jiang, X.G., Zeng, R., Miao, Y.G., Xu, H., Zhang, Y.X., Xiong, H., Lu, G., Lu, L.F., Jiang, H.Q., Jia, J., Tu, Y.F., Jiang, J.X., Gu, W.Y., Zhang, Y.Q., Cai, Z., Sheng, H.H., Yin, H.F., Zhang, Y., Zhu, G.F., Wan, M., Huang, H.L., Qian, Z., Wang, S.Y., Ma, W., Yao, Z.J., Shen, Y., Qiang, B.Q., Xia, Q.C., Guo, X.K., Danchin, A., Saint, G.I., Somerville, R.L., Wen, Y.M., Shi, M.H., Chen, Z., Xu, J.G., Zhao, G.P., 2003. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. Nature 422, 888–893.
Riley, L.W., Ko, A.I., Unger, A., Reis, M.G., 2007. Slum health: diseases of neglected populations. BMC Int. Health Hum. Rights 7, 2.
Romero, E.C., Bernardo, C.C., Yasuda, P.H., 2003. Human leptospirosis: a twenty-nine-year serological study in Sao Paulo, Brazil. Rev. Inst. Med. Trop. Sao Paulo 45, 245–248.
Samhrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sarkar, U., Nascimento, S.F., Barbosa, R., Martins, R., Nuevo, H., Kalafanos, I., Grunstein, I., Flannery, B., Dias, J., Riley, L.W., Reis, M.G., Ko, A.I., 2002. Population-based case-control investigation of risk factors for leptospirosis during an urban epidemic. Am. J. Trop. Med. Hyg. 66, 605–610.
Segura, E.R., Ganoza, C.A., Campos, K., Ricaldi, J.N., Torres, S., Silva, H., Cespedes, M.J., Matthias, M.A., Swancutt, M.A., Lopez Linan, R., Gotuzzo, E., Guerra, H., Gilman, R.H., Vinetz, J.M., 2005. Clinical spectrum of pulmonary involvement in leptospirosis in a region of endemicity, with quantification of leptospiral burden. Clin. Infect. Dis. 40, 343–351.
Shang, E.S., Summers, T.A., Haake, D.A., 1996. Molecular cloning and sequence analysis of the gene encoding LipL41, a surface-exposed lipoprotein of pathogenic *Leptospira* species. Infect. Immun. 64, 2322–2330.
Silva, E.F., Medeiros, M.A., McBride, A.J., Matsunaga, J., Esteves, G.S., Ramos, J.G., Santos, C.S., Croda, J., Homma, A., Dellagostin, O.A., Haake, D.A., Reis, M.G., Ko, A.I., 2007. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. Vaccine 25, 6277–6286.
Silva, E.F., Santos, C.S., Athanazio, D.A., Seyffert, N., Seixas, F.K., Cerqueira, G.M., Fagundes, M.Q., Brod, C.S., Reis, M.G., Dellagostin, O.A., Ko, A.I., 2008. Characterization of virulence of Leptospira isolates in a hamster model. Vaccine 26, 3892–3896.
Srimanote, P., Wongdeethai, N., Jieanampunkul, P., Samonkiert, S., Leepiyasakulchai, C., Kalambaheti, T., Prachayasittikul, V., 2008. Recombinant LigA for leptospirosis diagnosis and *ligA* among the *Leptospira* spp. clinical isolates. J. Microbiol. Methods 72, 73–81.
Suchard, M.A., Weiss, R.E., Dorman, K.S., Sinsheimer, J.S., 2002. Oh brother, where art thou? A Bayes factor test for recombination with uncertain heritage. Syst. Biol. 51, 715–728.
Suchard, M.A., Weiss, R.E., Dorman, K.S., Sinsheimer, J.S., 2003. Inferring spatial phylogenetic variation along nucleotide sequences: a multiple change-point model. J. Am. Stat. Assoc. 98, 427–437.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Verma, A., Hellwage, J., Artiushin, S., Zipfel, P.F., Kraiczy, P., Timoney, J.F., Stevenson, B., 2006. LfhA, a novel factor H-binding protein of *Leptospira interrogans*. Infect. Immun. 74, 2659–2666.
Yanagihara, Y., Villanueva, S.Y., Yoshida, S., Okamoto, Y., Masuzawa, T., 2007. Current status of leptospirosis in Japan and Philippines. Comp. Immunol. Microbiol. Infect. Dis. 30, 399–413.