

Bioinformatics Describes Novel Loci for High Resolution Discrimination of *Leptospira* Isolates

Gustavo M. Cerqueira^{1*}, Alan J. A. McBride^{2,9}, Rudy A. Hartskeerl³, Niyaz Ahmed⁴, Odir A. Dellagostin⁵, Marcus R. Eslabão⁵, Ana L. T. O. Nascimento^{1,6}

1 Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil, **2** Gonçalo Moniz Institute, Oswaldo Cruz Foundation, Salvador, Brazil, **3** Department of Biomedical Research, Royal Tropical Institute, World Health Organization/Food and Agriculture Organization of the United Nations/World Organisation for Animal Health (WHO/FAO/OIE) and National Leptospirosis Reference Centre, Amsterdam, The Netherlands, **4** Pathogen Biology Laboratory, Department of Biotechnology, University of Hyderabad, Hyderabad, India, **5** Centro de Biotecnologia, Universidade Federal de Pelotas, Pelotas, Brazil, **6** Interunidades em Biotecnologia, Instituto de Ciências Biomédicas, USP, São Paulo, Brazil

Abstract

Background: Leptospirosis is one of the most widespread zoonoses in the world and with over 260 pathogenic serovars there is an urgent need for a molecular system of classification. The development of multilocus sequence typing (MLST) schemes for *Leptospira* spp. is addressing this issue. The aim of this study was to identify loci with potential to enhance *Leptospira* strain discrimination by sequencing-based methods.

Methodology and Principal Findings: We used bioinformatics to evaluate pre-existing loci with the potential to increase the discrimination of outbreak strains. Previously deposited sequence data were evaluated by phylogenetic analyses using either single or concatenated sequences. We identified and evaluated the applicability of the *ligB*, *secY*, *rpoB* and *lipL41* loci, individually and in combination, to discriminate between 38 pathogenic *Leptospira* strains and to cluster them according to the species they belonged to. Pairwise identity among the loci ranged from 82.0–92.0%, while interspecies identity was 97.7–98.5%. Using the *ligB-secY-rpoB-lipL41* superlocus it was possible to discriminate 34/38 strains, which belong to six pathogenic *Leptospira* species. In addition, the sequences were concatenated with the superloci from 16 sequence types from a previous MLST scheme employed to study the association of a leptospiral clone with an outbreak of human leptospirosis in Thailand. Their use enhanced the discriminative power of the existing scheme. The *lipL41* and *rpoB* loci raised the resolution from 81.0–100%, but the enhanced scheme still remains limited to the *L. interrogans* and *L. kirschneri* species.

Conclusions: As the first aim of our study, the *ligB-secY-rpoB-lipL41* superlocus demonstrated a satisfactory level of discrimination among the strains evaluated. Second, the inclusion of the *rpoB* and *lipL41* loci to a MLST scheme provided high resolution for discrimination of strains within *L. interrogans* and *L. kirschneri* and might be useful in future epidemiological studies.

Citation: Cerqueira GM, McBride AJA, Hartskeerl RA, Ahmed N, Dellagostin OA, et al. (2010) Bioinformatics Describes Novel Loci for High Resolution Discrimination of *Leptospira* Isolates. PLoS ONE 5(10): e15335. doi:10.1371/journal.pone.0015335

Editor: Stefan Bereswill, Charité-Universitätsmedizin Berlin, Germany

Received: July 28, 2010; **Accepted:** August 9, 2010; **Published:** October 15, 2010

Copyright: © 2010 Cerqueira, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study did not receive any designated funding. Gustavo M. Cerqueira holds a postdoctoral fellowship from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Alan J.A. McBride would like to acknowledge the support from the Research Support Foundation for the State of Bahia (FAPESB), grant PES-0092/2008. Ana L.T.O. Nascimento would like to acknowledge the support from the FAPESP, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação Butantan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Niyaz Ahmed and Rudy Hartskeerl are members of the PLoS ONE Editorial Board. Niyaz Ahmed is a member of the PLoS International Advisory Group.

* E-mail: cerqueiragm@yahoo.com.br

⁹ These authors contributed equally to this work.

Introduction

Leptospirosis is a zoonotic disease caused by pathogenic *Leptospira* spp. and is considered an emerging global public health problem [1,2]. Furthermore, the impact of leptospirosis has increased, particularly in poverty stricken regions of the world, due to the high mortality (>50%) associated with the recent increase of severe pulmonary haemorrhage syndrome (SPHS) in patients with severe leptospirosis [3,4]. Based on serology, *Leptospira* spp. are traditionally classified into 29 serogroups and over 300 serovars [5–7]. More recently, genetic methods have

attempted to replace the traditional classification methods and DNA-DNA hybridization studies have identified 20 *Leptospira* spp. to date [5,8–11]. Several typing methods have been employed to classify isolates with differing degrees of success [12]. However, a major limitation is the lack of correlation between the serologic and genotypic classification methods [5,12,13].

Multilocus sequence typing (MLST) was originally developed for bacteria using *Neisseria meningitidis* isolates [14] and, so far, it has been successfully applied to over 30 bacteria [15,16]. In the field of leptospirosis, efforts to develop a typing method have focused on MLST [17,18]. MLST allows the adoption of a universal format

for a particular bacterial species and permits the sequence data generated to be easily exchanged over the Internet. Traditionally, the loci chosen for MLST analyses are based on 6–10 housekeeping genes that are under selection for metabolic functionality [15]. Since this group is comprised of slowly evolving genes they are likely to be more conserved and stable within a particular species [19]. Unfortunately it has not been possible to identify a set of housekeeping genes with universal applicability to all bacterial pathogens. Rather, MLST loci are chosen empirically and evaluated for each individual pathogen [15]. Ahmed and colleagues presented the first MLST scheme based on loci from four housekeeping genes and two genes encoding outer-membrane proteins for typing *L. alexanderi*, *L. borgpetersenii*, *L. interrogans*, *L. kirschneri*, *L. noguchii* and *L. santarosai* isolates [18]. An alternative MLST scheme using loci from seven housekeeping genes was used to type *L. interrogans* and *L. kirschneri* isolates and is available on the Internet (<http://leptospira.mlst.net/>). The database contains 109 sequence types (ST) and sequences from 263 isolates at time of writing [17]. Although this evidently moved the field forward, a limitation of this database is that it only applies to isolates from two *Leptospira* species, *L. interrogans* and *L. kirschneri*. The ideal MLST scheme should be valid for all *Leptospira* spp. or at least the pathogenic species [20], and provide discrimination beyond the species level [21].

High-resolution typing, such as that required during outbreak investigations, usually requires the inclusion of genes with greater diversity, e.g. antigen genes, rather than housekeeping genes [15]. The objective of this study was to carry out a bioinformatics-based analysis of *Leptospira* genes available in GenBank to identify potential targets for improved *Leptospira* discrimination. The genes *ligB*, *secY*, *lipL41* and *rpoB* were identified as potential genes for use in an improved typing scheme.

Methods

DNA sequences

The DNA sequences for the *ligB*, *secY*, *rpoB* and *lipL41* loci used in this study were obtained from GenBank and LepBank [22] (Table 1) or from the authors of the original *Leptospira* MLST scheme [18]. Most of these sequences were generated by the authors during previous studies and they belong to different reference strains and clinical isolates. The sizes of the loci analyzed were 214 bp (*ligB*), 245 bp (*secY*), 541 bp (*rpoB*) and 884 bp (*lipL41*) and correspond to nucleotide positions 2236–2449 (*ligB*), 771–1015 (*secY*), 1922–2462 (*rpoB*) and 73–956 (*lipL41*). Note that nucleotide positions are based on the *L. interrogans* Copenhageni Fiocruz L1-130 genome (AE016823). These genes can be amplified by using the primers *ligB* (PSBF: 5'-ACWRVHVHR-GYWDCCTGGTCTCTTC-3'; PSBR: 5'-TARRHDGCYB-TAATATYCGRWYYTCCTAA-3'), [23]; *secY* (SeqYII: 5'-GAA-TTTCTCTTTTGTCTTCG-3'; SeqYIV: 5'-GAGTTAGAG-CTCAAATCTAAG-3'), [24]; *rpoB* (Lept 1900f: 5'-CCTCATG-GGTTCCAACATGCA-3'; Lept 2500r: 5'-CGCATCCTCRA-AGTTGTAWCCTT-3'), [25] and *lipL41* (*lipL41F*: 5'-TAG-GAAATTGCGCAGCTACA-3'; *lipL41R*: 5'-GCATCGAGAG-GAATTAACATCA-3'), [18]. The DNA sequences corresponding to the *glmU* (18 alleles), *pntA* (24 alleles), *sucA* (20 alleles), *fadD* (20 alleles), *tpiA* (30 alleles), *pfkB* (35 alleles) and *mreA* (23 alleles) loci were downloaded from the *Leptospira* MLST database at <http://leptospira.mlst.net/> [26].

Phylogenetic analysis

DNA sequences were aligned using ClustalW at the default settings (<http://www.ebi.ac.uk/clustalw>). The phylogenetic anal-

yses were performed with Mega 4.1 [27] or Geneious Pro ver 4.7 [28], and the neighbour-joining method with no outgroup. The Tamura-Nei genetic distance model was selected and all trees were resampled using the bootstrap method and 1000 replicates. The phylogenetic trees constructed using sequences of 38 reference strains (Table 1) was based a 1884 bp superlocus composed of the concatenated sequences of the loci for each strain in the following order: *ligB-secY-rpoB-lipL41*.

Results

Phylogenetic analysis of the *ligB*, *secY*, *rpoB* and *lipL41* loci

The main criteria to select genes for the presented MLST scheme were their ability to separate one or more species in different clusters and to discriminate the strains and clinical isolates within them. To identify candidate loci with these properties we searched those previously characterized by the authors and mined public databases to obtain additional representative sequences. The genes used to constitute this study (*ligB*, *secY*, *rpoB* and *lipL41*) showed an optimal discriminative power. Following alignment, the percentage of identical pairwise amino acids residues (PI) for the *lipL41* locus was 92.0% and the percentage of identical sites (IS) among the DNA sequences was 77.9%, the *rpoB* locus was 91.8 and 75.2%, the *secY* locus was 87.8 and 71.4%, and the *ligB* locus was 82.0 and 48.6%, respectively. When all four loci for each strain were concatenated and when the resulting superloci were aligned the overall PI was 90.3% and the IS was 72.9%. Of note, intraspecies identity was considerably higher, 97.7–98.5±0.2% (Figure 1). The phylogenetic tree formed two major clusters representing *L. kirschneri* and the other species: *L. interrogans*, *L. noguchii*, *L. borgpetersenii*, *L. santarosai* and *L. weilii* (Figure 2).

The 15 *L. interrogans* strains, two *L. noguchii* strains, four *L. weilii* strains and three *L. santarosai* strains, listed in Table 1, could all be discriminated using the superlocus and of the seven *L. kirschneri* strains, only two proved to be identical at the sequence level (strains 5621 and 3522C). Among the seven *L. borgpetersenii* strains, the superlocus sequence was capable of discriminating all but two strains (JB197 and L550). Polymorphic sites, where one sampled sequence exhibits a unique base relative to the common nucleotide of the others were observed in 13 serovars and all species except *L. weilii*. The *L. interrogans* species included the largest number of polymorphic-containing serovars, six, followed by Pomona with two, then Autumnalis, Lai, Manilae, Muenchen and Pyrogenes with one. The *L. noguchii* serovars contained the largest number of unique polymorphic sites per serovar, Orleans and Panama had five each. Inclusion of the low polymorphic genes *ms2* and *lipL32* as in the scheme of Ahmed and colleagues [18] was assessed but showed no advantage to the presented scheme (results not shown).

Increased discrimination using the *ligB*, *secY*, *rpoB* and *lipL41* loci in an existing MLST scheme

Based on analysis of the loci sequences from 38 *Leptospira* reference strains, *ligB* contained 15 alleles, *secY* 16 alleles, *rpoB* 19 alleles and *lipL41* 28 alleles, Figure 3 and Figure S1 B–D, respectively. Using the online *Leptospira* MLST scheme (<http://leptospira.mlst.net>, [17]) we identified 16 sequence types (ST), out of 109 ST, where the corresponding *ligB*, *secY*, *rpoB* and *lipL41* sequences for each strain were readily available. Using the concatenated sequences that corresponded to the *glmU*, *pntA*, *sucA*, *fadD*, *tpiA*, *pfkB* and *mreA* loci a phylogenetic tree was constructed that identified 13 unique ST among the 16 different strains

Table 1. *Leptospira* serovars and the candidate alleles for MLST.

Species	Serogroup	Serovar	Strain	Accession numbers				
				<i>ligB</i>	<i>secY</i>	<i>rpoB</i>	<i>lipL41</i>	
<i>L. borgpetersenii</i>	Javanica	Ceylonica	Piyasena	EU938500 ^a	EU358041 ^g	DQ296134 ^j	AY461936 ^m	
	Javanica	Javanica	Veldrat Batavia 46	EU938501 ^a	EU358040 ^g	DQ296134 ^j	AY461938 ^m	
	Javanica	Poi	Poi	EU938502 ^a	EU358007 ^g	DQ296134 ^j	n/a ⁿ	
	Mini	Mini	Sari	n/a	EU358032 ^g	DQ296134 ^j	n/a ⁿ	
	Sejroe	Hardjo-bovis	JB197	CP000350 ^b	CP000350 ^b	CP000350 ^b	CP000350 ^b	
	Sejroe	Hardjo-bovis	L550	CP000348 ^b	CP000348 ^b	CP000348 ^b	CP000348 ^b	
	Tarassovi	Tarassovi	Perepelitsin	n/a	EU358057 ^g	EU747307 ^k	AY461937 ^m	
<i>L. interrogans</i>	Australis	Australis	Ballico	EU938484 ^a	DQ882850 ^h	DQ296144 ^j	n/a ⁿ	
	Australis	Bratislava	Jez Bratislava	EU938487 ^a	EU357939 ^g	EU747300 ^k	AY461939 ^m	
	Australis	Muenchen	Muenchen C90	EU938497 ^a	EU357938 ^g	DQ296133 ^j	n/a ⁿ	
	Autumnalis	Autumnalis	Akiyami A	EU938485 ^a	EU357943 ^g	DQ296145 ^j	AY461940 ^m	
	Bataviae	Bataviae	Van Tienen	EU938486 ^a	EU357956 ^g	DQ296146 ^j	AY461941 ^m	
	Canicola	Canicola	Hond Utrecht IV	EU938488 ^a	EU357961 ^g	EU747299 ^k	AY461942 ^m	
	Icterohaemorrhagiae	Copenhageni	Fiocruz L1-130	AE016823 ^c	AE016823 ^c	AE016823 ^c	AE016823 ^c	
	Icterohaemorrhagiae	Icterohaemorrhagiae	RGA	EU938493 ^a	EU365950 ^g	DQ296133 ^j	AY461947 ^m	
	Icterohaemorrhagiae	Lai	56601	AE010300 ^d	AE010300 ^d	AE010300 ^d	AE010300 ^d	
	Pyrogenes	Manilae	LT398	EU938496 ^a	EU358049 ^g	DQ296133 ^j	n/a ⁿ	
	Pyrogenes	Pyrogenes	Salinem	n/a	DQ882863 ^h	DQ296147 ^j	n/a ⁿ	
	Sejroe	Hardjo-prajitno	Hardjoprajitno	EU938491 ^a	EU357983 ^g	EU747303 ^k	AY461943 ^m	
	Sejroe	Wolffi	3705	EU938499 ^a	EU357985 ^g	EU747308 ^k	n/a ⁿ	
	Hebdomadis	Hebdomadis	Hebdomadis	EU938492 ^a	EU357974 ^g	EU747304 ^k	n/a ⁿ	
	Pomona	Pomona	Pomona	EU938498 ^a	EU358013 ^g	EU747306 ^k	AY461948 ^m	
	<i>L. kirschneri</i>	Australis	Ramisi	Musa	EU938507 ^a	EU358020 ^g	DQ296139 ^j	AY461949 ^m
		Autumnalis	Erinaceiauriti	Erinaceus Auritus 670	EU938504 ^a	EU358021 ^g	DQ296139 ^j	AY461950 ^m
Bataviae		Djatzi	HS 26	EU938503 ^a	EU358027 ^g	DQ296139 ^j	AY461951 ^m	
Cynopteri		Cynopteri	3522 C	EU938508 ^a	EU358027 ^g	DQ296139 ^j	n/a ⁿ	
Grippotyphosa		Grippotyphosa	RM52	AY190126 ^e	EU358027 ^g	EU747301 ^k	AY461953 ^m	
Hebdomadis		Kambale	Kambale	EU938505 ^a	EU358030 ^g	DQ296139 ^j	AY461954 ^m	
Pomona		Mozdok	5621	EU938506 ^a	EU358015 ^g	DQ296139 ^j	AY461955 ^m	
<i>L. noguchii</i>	Louisiana	Orleans	LSU 2580	EU938509 ^a	EU365958 ^g	EU349500 ^l	AY461957 ^m	
	Panama	Panama	CZ 214 K	EU938510 ^a	EU365958 ^g	DQ296141 ^j	n/a ⁿ	
<i>L. santarosai</i>	Pyrogenes	Alexi	HS 616	EU938512 ^a	EU358047 ^g	DQ296131 ^j	AY461964 ^m	
	Sejroe	Trinidad	TRVL 34056	n/a	EU358035 ^g	DQ296131 ^j	n/a ⁿ	
	Shermani	Shermani	LT 821	EU938511 ^a	DQ882866 ^h	DQ296131 ^j	AY461965 ^m	
<i>L. weilii</i>	Celledoni	Celledoni	Celledoni	EU938514 ^a	EU365960 ^g	DQ296132 ^j	n/a ⁿ	
	Hebdomadis	n/a	EcoChallenge	EU700274 ^f	AY034036 ⁱ	DQ296132 ^j	n/a ⁿ	
	Javanica	Coxi	Cox	n/a	EU358009 ^g	DQ296132 ^j	AY461967 ^m	
	Tarassovi	Vughia	LT 89–68	EU938515 ^a	EU365960 ^g	DQ296132 ^j	AY461968 ^m	

n/a - Not applicable. Source of sequence data: [26]^a; [32]^b; [33]^c; [34]^d; [35]^e; [30]^f; [28]^g; Riediger^h, unpublished data; [36]ⁱ; [29]^j; Bomfim^k, unpublished data; [37]^j; [27]^m; [18]ⁿ.

doi:10.1371/journal.pone.0015335.t001

included in the analysis (Figure 4). The *ligB*, *secY*, *rpoB* and *lipL41* loci were added to the MLST superloci from each strain to determine their impact on the level of discrimination. While inclusion of the *ligB* locus did not improve upon the original scheme (Figure S2 B), the *secY* and *lipL41* loci resolved 14 ST (Figures S2 C & D, respectively) and the *rpoB* locus discriminated between 15 ST (Figure S2 E). All possible combinations of the candidate loci were used to create additional superloci to determine the assembly with the greatest discriminatory power

(data not shown). Complete resolution of the 16 ST was achieved by inclusion of both *lipL41* and *rpoB* loci in the original concatenated sequence of each strain, Figure S2 F.

Discussion

The MLST scheme proposed by Thaipadungpanit and colleagues is based on loci from seven housekeeping genes and identified 109 unique ST among 263 isolates from either *L.*

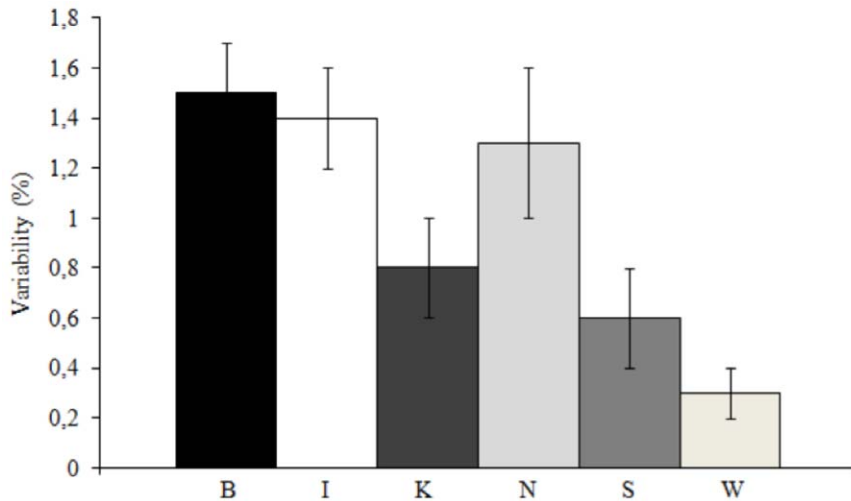


Figure 1. Alignment analysis of the concatenated superlocus. Comparison of the variability of the DNA sequences from the *ligB-secY-rpoB-lipL41* superlocus among the *Leptospira* species included in this study, where B – *L. borgpetersenii*, I – *L. interrogans*, K – *L. kirschneri*, N – *L. noguchii*, S – *L. santarosai* and W – *L. weilii*. Results are shown as means \pm SD. The nucleotide positions used during the alignment analysis were: nt 2236–2449 (*ligB*), 771–1015 (*secY*), 1922–2462 (*rpoB*) and 73–956 (*lipL41*) and refer to the *L. interrogans* serovar Copenhageni L1-130 strain. doi:10.1371/journal.pone.0015335.g001

interrogans or *L. kirschneri* [17]. This represented a major advance in the molecular epidemiology of *Leptospira* isolates. Unfortunately, as noted by the authors themselves, a major limitation of this scheme is that it does not allow for the inclusion of the other common

pathogens associated with human leptospirosis [20]. Previously, it was shown that the antigen encoding genes *ligB*, *secY*, *lipL41* and the *rpoB* gene are potentially useful for the molecular discrimination of *Leptospira* strains and that they can be readily amplified from

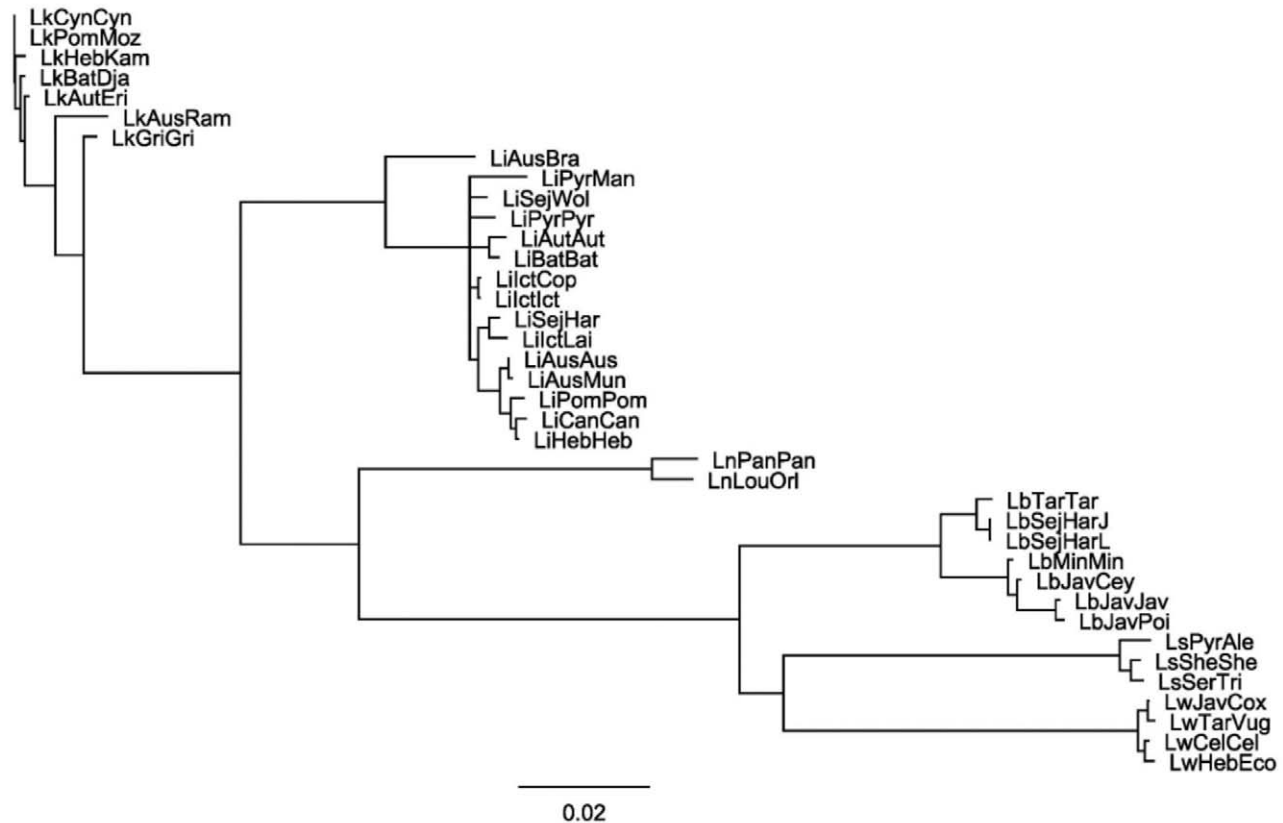


Figure 2. Phylogenetic analysis of 38 *Leptospira* serovars. The tree was constructed based on the *ligB-secY-rpoB-lipL41* superlocus sequences. The loci were analyzed using the Neighbor-Joining method as implemented in Geneious Pro 4.7.5 [25]. The samples are represented by the serovar followed by the strain designations. Confidence in the topology of this tree was gauged by bootstrap resampling (1,000 times). doi:10.1371/journal.pone.0015335.g002

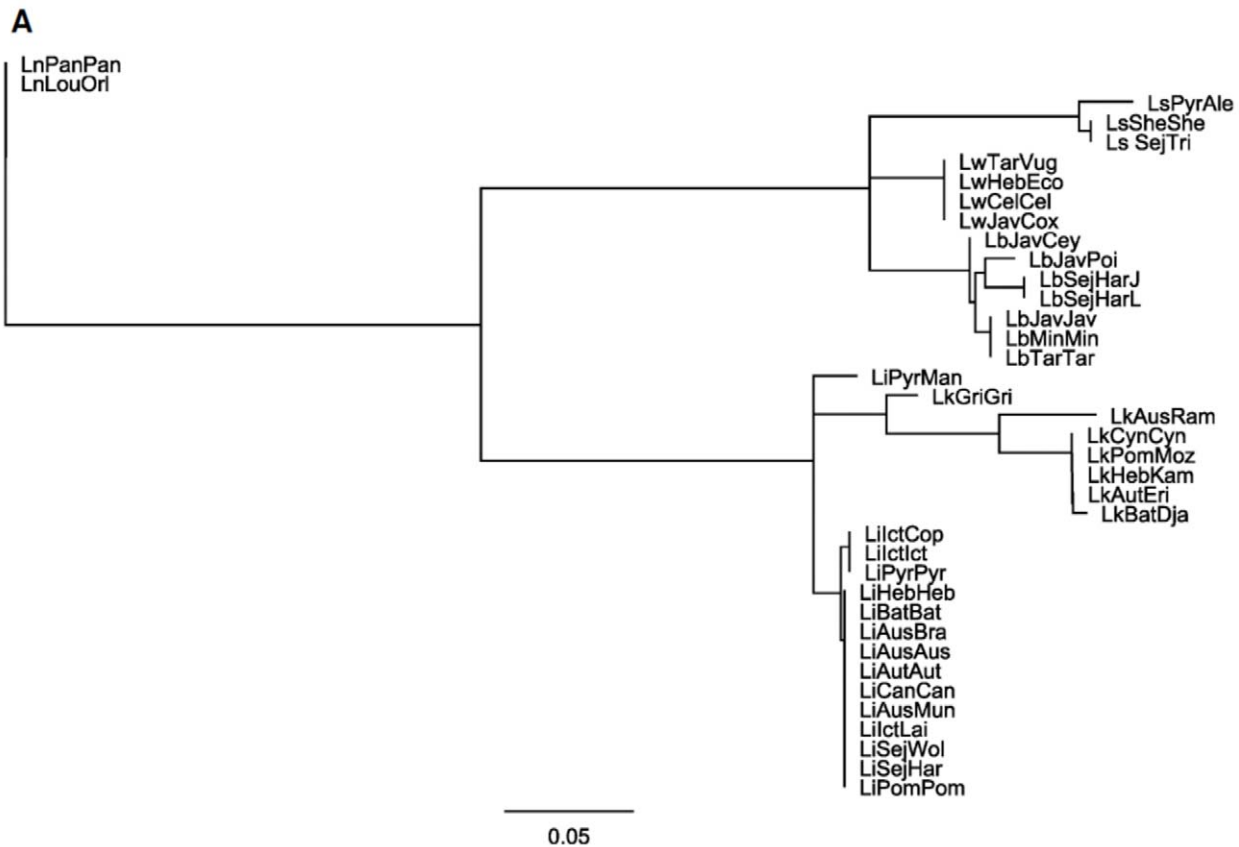


Figure 3. Analysis of the loci sequences from 38 *Leptospira* reference strains. The tree was constructed based on the individual loci sequences. The loci were analyzed using the Neighbor-Joining method as implemented in Geneious Pro 4.7.5 [25]. The samples are represented by the serovar followed by the strain designations. Confidence in the topology of this tree was gauged by bootstrap resampling (1,000 times). A: *ligB* loci; B: *secY* loci; C: *rpoB* loci; D: *lipL41* loci. Phylogenetic analysis was used to demonstrate the number of alleles that were distinguished for each loci A: *ligB* (15); B: *secY* (16); C: *rpoB* (19); and D: *lipL41* (27). Parts B–D can be found as supporting information in Figure S1. doi:10.1371/journal.pone.0015335.g003

all known pathogenic leptospires [23–25,29,30]. Furthermore, the authors of the original bacterial MLST scheme recommend the inclusion of loci from antigen coding genes to improve discrimination, especially during outbreak investigations [31].

To determine the potential benefits of using *ligB*, *secY*, *rpoB* and *lipL41* loci in an MLST typing scheme focused on strain differentiation we identified the corresponding sequences in a reference collection containing 38 *Leptospira* strains (Table 1). The overall level of pairwise identity ranged from 82–92% among the individual loci while the intraspecies identity was even higher (Figure 1). Furthermore, when the sequences were concatenated to create a superlocus for each strain and analysed, the overall pairwise identity was >90%. Following the modelling of phylogenetic trees and in agreement with previous studies, two distinct evolutionary branches were observed, the first contained *L. kirschneri*, *L. interrogans* and *L. noguchii* strains and the second the *L. borgpetersenii*, *L. santarosai* and *L. weilii* strains (Figure 4 and Figure S2) [23,29]. There was some evidence that serovars from the same serogroup clustered together, serogroups: Icterohaemorrhagiae: *L. interrogans* Icterohaemorrhagiae RGA and Copenhageni Fiocruz L1-130; Australis: *L. interrogans* Australis Ballico and Muenchen Muenchen C90; and Javanica: *L. borgpetersenii* Javanica Veldrat Batavia 46, Poi Poi and Ceylonica Piyasena (Figure 2). This may indicate homoplasy (similarity due to convergent evolution) of the genes from these serovars. This analysis showed a separation

of serovar Bratislava from the other *L. interrogans* serovars (Figure 2). Alignment analyses demonstrated this is probably due to the high similarity of the *L. interrogans* Jez Bratislava strain *rpoB* gene sequence with those from the *L. borgpetersenii* strains (data not shown). Despite this, serovar Bratislava was correctly located within the *L. interrogans* clade. We did not determine however, whether this was due to sequence mosaicism or horizontal gene transfer. The phylogenetic organisation of the *Leptospira* genus based on the superlocus supports the theory [29] that *L. interrogans* is more recently evolved from *L. kirschneri*, more recent evolutionary subdivisions resulted in the separation of *L. borgpetersenii* followed by *L. santarosai* and *L. weilii* clades.

Analysis of the discriminatory power of the *ligB-secY-rpoB-lipL41* superlocus found that the *L. interrogans*, *L. noguchii*, *L. santarosai* and *L. weilii* strains could be separated into individual ST. However, among the *L. kirschneri* and *L. borgpetersenii* strains, two could not be resolved by the superlocus. Serological analysis of the two *L. borgpetersenii* strains indicated that both belonged to serogroup Sejroe serovar Hardjo-bovis and their genomes were found to be highly conserved. Yet they are distinct clonal subtypes, both strains established chronic infections in cattle yet differed in their ability to cause lethal infections in hamsters [32]. Despite this no specific polymorphisms were observed in either of the *L. borgpetersenii* L550 or JB197 strains. These polymorphic regions are normally useful for surveillance purposes, to monitor outbreaks or for epidemiological studies. Thirteen serovars, out of

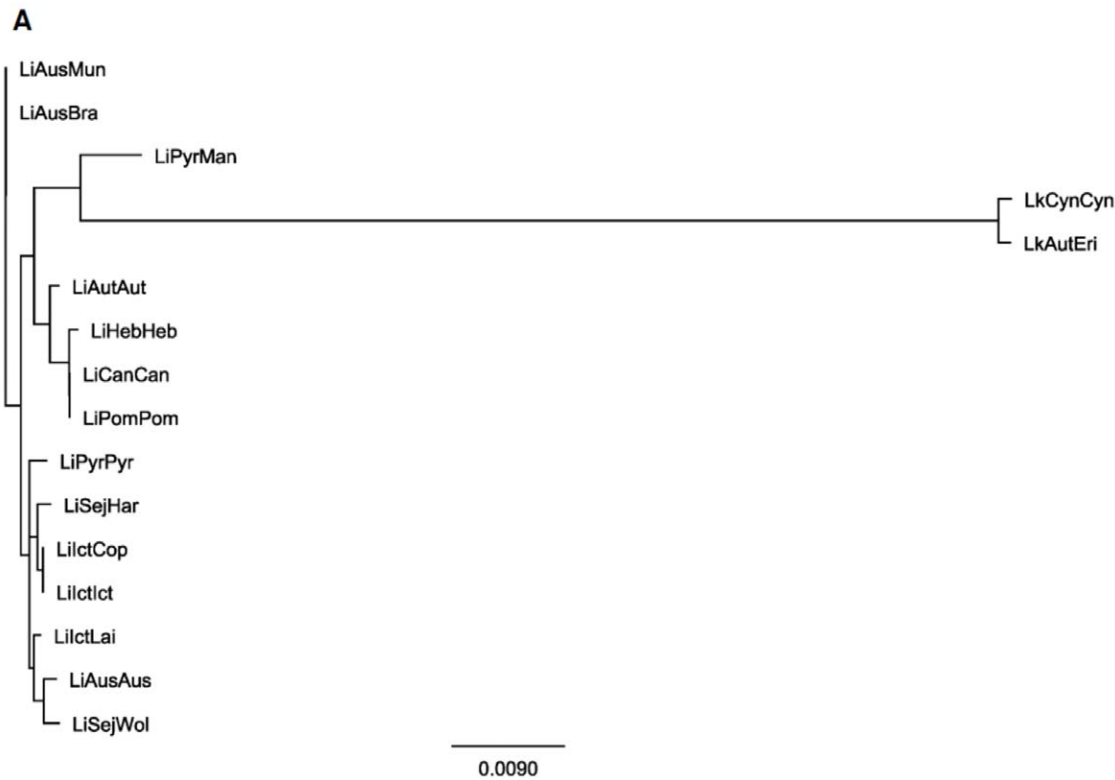


Figure 4. Phylogenetic analysis of the 16 ST from *L. interrogans* and *L. kirschneri*. The STs determined by Thaipadungpanit and colleagues [17] were concatenated with the loci analyzed during this study to test the effect of different loci combinations over serovars discrimination: (A) ST; (B) ST+*ligB*; (C) ST+*secY*; (D) ST+*rpoB*; (E) ST+*lipL41* and (F) ST+*rpoB*+*lipL41*. The tree was constructed by the Geneious Pro 4.7.5 software [25] as previously described. Confidence in the topology of this tree was gauged by bootstrap resampling (1,000 times). Parts B–F can be found as supporting information in Figure S2. doi:10.1371/journal.pone.0015335.g004

38 in this study, exhibited unique polymorphic sites. These findings highlight the efficiency of the proposed *ligB-secY-rpoB-lipL41* superlocus to discriminate *Leptospira* strains. The study previously performed by Ahmed and colleagues [18] was the pioneer in the use of a concatenated superlocus to discriminate among the *Leptospira* species. However, this work was intended as a step towards the study of pathogen evolution rather than strain discrimination. Inclusion of low polymorphic genes such as *ns2* and *lipL32* used by Ahmed et al., [18] did not contribute to enhance the discriminative power of the MLST scheme presented here. In the present work, we included some of those sequences in combination with recently sequenced polymorphic genes to increase the resolution, and observed the occurrence of discrimination to the subspecies level. Although a reduced number of strains and isolates were included in our study we believe the proposed superlocus presents a solid basis for discriminating within large panels of *Leptospira* strains and isolates.

An analysis of the *ligB*, *secY*, *rpoB* and *lipL41* loci found that the *ligB* locus was the most conserved with 15 alleles, followed by *secY* with 16 alleles, *rpoB* with 19 alleles and *lipL41* with 28 alleles out of a potential 38 (Figure 3 and Figure S1). Following analysis of the 263 isolates (109 ST) contained in the *Leptospira* MLST database (leptospora.mlst.net), 16 strains (corresponding to 13 ST) were identified as having the corresponding *ligB*, *secY*, *rpoB* and *lipL41* loci sequences available (Figure 4). To determine the utility of the loci proposed in this study the relevant sequence for each individual locus was concatenated to the *glmU*, *pntA*, *sucA*, *fadD*, *tpiA*, *pfkB* and *mreA* superlocus of each strain (Figures S2 B–E). All possible variables were evaluated in order to identify the most useful

additional loci. The combination of the original superlocus together with the *rpoB* and the *lipL41* loci was found to be the simplest superlocus that could discriminate between all 16 of the strains (Figure S2 F). This is in agreement with the ability of the superlocus determined by Ahmed and colleagues [18], which includes the *lipL41* locus, to discriminate the *Leptospira* spp. in study.

The phylogenetic analysis of our sequences showed a great diversity of ST and no clustering, due to the use of epidemiologically unrelated strains. Thus, when the two new loci sequences were concatenated to the original ST sequences we observed the complete discrimination of the strains, although our adapted scheme remains limited to *L. interrogans* and *L. kirschneri*. We recommend that the *rpoB* and *lipL41* loci be evaluated in existing or future MLST schemes to enhance their typing power during outbreak investigations.

Supporting Information

Figure S1 Continued from figure 3.

Figure S2 Continued from figure 4.

Acknowledgments

This publication made use of the Multi Locus Sequence Typing website (<http://www.mlst.net>) at Imperial College London developed by David Aanensen and funded by the Wellcome Trust.

Author Contributions

Conceived and designed the experiments: GMC AJAM. Performed the experiments: GMC AJAM MRE. Analyzed the data: GMC AJAM. Wrote the manuscript: GMC AJAM RAH NA OAD MRE ALTON.

References

- Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, et al. (2003) Leptospirosis: a zoonotic disease of global importance. *Lancet Infect Dis* 3: 757–771.
- McBride AJ, Athanazio DA, Reis MG, Ko AI (2005) Leptospirosis. *Curr Opin Infect Dis* 18: 376–386.
- Gouveia EL, Metcalfe J, de Carvalho AL, Aires TS, Villasboas-Bisneto JC, et al. (2008) Leptospirosis-associated Severe Pulmonary Hemorrhagic Syndrome, Salvador, Brazil. *Emerg Infect Dis* 14: 505–508.
- Segura ER, Ganoza CA, Campos K, Ricaldi JN, Torres S, et al. (2005) Clinical spectrum of pulmonary involvement in leptospirosis in a region of endemicity, with quantification of leptospiral burden. *Clin Infect Dis* 40: 343–351.
- Levett PN (2001) Leptospirosis. *Clin Microbiol Rev* 14: 296–326.
- Kmetz E, Dikken H (1993) Classification of the species *Leptospira interrogans* and history of its serovars. Groningen, The Netherlands: University Press. pp 17–22.
- Faine SB, Adler B, Bolin C, Perolat P (1999) *Leptospira* and leptospirosis. Melbourne, Australia: MediSci.
- Levett PN, Morey RE, Galloway RL, Steigerwalt AG (2006) *Leptospira broomii* sp. nov., isolated from humans with leptospirosis. *Int J Syst Evol Microbiol* 56: 671–673.
- Matthias MA, Ricaldi JN, Cespedes M, Diaz MM, Galloway RL, et al. (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.
- Slack AT, Khairani-Bejo S, Symonds ML, Dohnt MF, Galloway RL, et al. (2009) *Leptospira kmetyi* sp. nov., isolated from an environmental source in Malaysia. *Int J Syst Evol Microbiol* 59: 705–708.
- Brenner DJ, Kaufmann AF, Sulzer KR, Steigerwalt AG, Rogers FC, et al. (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(Pt 2): 839–858.
- Cerqueira GM, Picardeau M (2009) A century of *Leptospira* strain typing. *Infect Genet Evol* 9: 760–768.
- Yasuda PH, Steigerwalt AG, Sulzer KR, Kaufman AF, Rogers FC, et al. (1987) Deoxyribonucleic acid relatedness between serogroups and serovars in the family Leptospiraceae with proposals for seven new *Leptospira* species. *Int J Syst Bacteriol* 37: 9.
- Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, et al. (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 95: 3140–3145.
- Maiden MC (2006) Multilocus sequence typing of bacteria. *Annu Rev Microbiol* 60: 561–588.
- Ibarz Pavon AB, Maiden MC (2009) Multilocus sequence typing. *Methods Mol Biol* 551: 129–140.
- Thaipadungpanit J, Wuthiekanun V, Chierakul W, Smythe LD, Petkanchanapong W, et al. (2007) A dominant clone of *Leptospira interrogans* associated with an outbreak of human leptospirosis in Thailand. *PLoS Negl Trop Dis* 1: e56.
- Ahmed N, Devi SM, Valverde Mde L, Vijayachari P, Machang'u RS, et al. (2006) Multilocus sequence typing method for identification and genotypic classification of pathogenic *Leptospira* species. *Ann Clin Microbiol Antimicrob* 5: 28.
- Enright MC, Spratt BG (1999) Multilocus sequence typing. *Trends Microbiol* 7: 482–487.
- Levett PN (2007) Sequence-based typing of *Leptospira*: epidemiology in the genomic era. *PLoS Negl Trop Dis* 1: e120.
- Morey RE, Galloway RL, Bragg SL, Steigerwalt AG, Mayer LW, et al. (2006) Species-specific identification of *Leptospiraceae* by 16S rRNA gene sequencing. *J Clin Microbiol* 44: 3510–3516.
- Eslabão MR, Dellagostin OA, Cerqueira GM (2010) LepBank: A *Leptospira* sequence repository and a portal for phylogenetic studies. *Infection Genetics and Evolution* 10: 5.
- Cerqueira GM, McBride AJ, Picardeau M, Ribeiro SG, Moreira AN, et al. (2009) Distribution of the leptospiral immunoglobulin-like (*lig*) genes in pathogenic *Leptospira* species and application of *ligB* to typing leptospiral isolates. *J Med Microbiol* 58: 1173–1181.
- Victoria B, Ahmed A, Zuerner RL, Ahmed N, Bulach DM, et al. (2008) Conservation of the *S10-spec-alpha* locus within otherwise highly plastic genomes provides phylogenetic insight into the genus *Leptospira*. *PLoS One* 3: e2752.
- La Scola B, Bui LT, Baranton G, Khamis A, Raoult D (2006) Partial *spoB* gene sequencing for identification of *Leptospira* species. *FEMS Microbiol Lett* 263: 142–147.
- Aanensen DM, Spratt BG (2005) The multilocus sequence typing network: mlst.net. *Nucleic Acids Res* 33: W728–733.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 4.
- Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, et al. (2007) Geneious v3.5 Available from <http://www.geneious.com/>. 3.5 ed.
- Haake DA, Suchard MA, Kelley MM, Dundoo M, Alt DP, et al. (2004) Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer. *J Bacteriol* 186: 2818–2828.
- McBride AJA, Cerqueira GM, Suchard MA, Moreira AN, Zuerner RL, et al. (2009) Genetic diversity of the Leptospiral immunoglobulin-like (*lig*) genes in pathogenic *Leptospira* spp. *Infection Genetics and Evolution* 9: 10.
- Jolley KA, Brehony C, Maiden MC (2007) Molecular typing of meningococci: recommendations for target choice and nomenclature. *FEMS Microbiol Rev* 31: 89–96.
- Bulach DM, Zuerner RL, Wilson P, Seemann T, McGrath A, et al. (2006) Genome reduction in *Leptospira borgpeterseni* reflects limited transmission potential. *Proc Natl Acad Sci U S A* 103: 14560–14565.
- Nascimento AL, Ko AI, Martins EA, Monteiro-Vitorello CB, Ho PL, et al. (2004) Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *Journal of Bacteriology* 186: 2164–2172.
- Ren SX, Fu G, Jiang XG, Zeng R, Miao YG, et al. (2003) Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature* 422: 888–893.
- Matsunaga J, Barocchi MA, Croda J, Young TA, Sanchez Y, et al. (2003) Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. *Molecular Microbiology* 49: 929–945.
- Haake DA, Dundoo M, Cader R, Kubak BM, Hartskeerl RA, et al. (2002) Leptospirosis, water sports, and chemoprophylaxis. *Clinical Infectious Diseases* 34: e40–e43.
- Silva EF, Cerqueira GM, Seyffert N, Seixas FK, Hartwig DD, et al. (2009) *Leptospira noguchii* and human and animal leptospirosis, Southern Brazil. *Emerging Infectious Diseases* 15: 3.