Title
Genomic insights into the pathogenesis of Leptospira

Permalink
https://escholarship.org/uc/item/906898vz

Author
Ricaldi Camahualí, Jessica Nancy

Publication Date
2009

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Genomic insights into the pathogenesis of Leptospira

A dissertation submitted in partial satisfaction of the requirements for the
degree Doctor of Philosophy

in
Molecular Pathology

by
Jessica Nancy Ricaldi Camahuali

Committee in charge:

Joseph Vinetz, Chair
Lynette Corbeil
Donald Guiney
Mark Kamps
Victor Nizet
Elizabeth Winzeler

2009
The Dissertation of Jessica Nancy Ricaldi Camahuali is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

______________________________________________

______________________________________________

______________________________________________

______________________________________________

______________________________________________

Chair

University of California, San Diego

2009
Dedication

To my parents Javier and Nancy.

To my husband Mark.
Table of Contents

Signature page ........................................................................................................................... iii

Dedication .................................................................................................................................. iv

List of Tables ............................................................................................................................... x

List of Figures .............................................................................................................................. xi

Acknowledgments ....................................................................................................................... xiv

Vita ................................................................................................................................................ xvi

Abstract of the Dissertation ...................................................................................................... xviii

Chapter 1

Introduction ................................................................................................................................. 1

  1.1 Introduction ......................................................................................................................... 1
  1.2 History ................................................................................................................................. 1
  1.3 Epidemiology ....................................................................................................................... 4
  1.4 Leptospira taxonomy ........................................................................................................... 8
  1.5 Clinical presentation ........................................................................................................... 11
  1.6 Microbiological diagnosis .................................................................................................. 12
  1.7 Pathogenesis and virulence ................................................................................................ 14

Chapter 2

Isolation and characterization of a new intermediate specie of Leptospira .................. 23


2.1 Introduction...........................................................................................................23

2.2 Materials and Methods ..........................................................................................24

Humans: Enrollment, Sampling and Culture ..............................................................24

Human Serological Analysis and Interpretation Criteria ..........................................25

Animals: Trapping and Culture for Leptospira .........................................................27

Pulsed Field Gel Electrophoresis (PFGE) Characterization of Isolates.................27

Characterization of isolates by 16S rRNA Gene Sequencing ..............................28

DNA-DNA Hybridization Analysis ....................................................................29

Determination of leptospiral serogroup ...............................................................30

Biological Characterization .................................................................................30

Growth Characteristics .........................................................................................30

*LipL32* PCR ........................................................................................................32

LipL32 Western Blot ...............................................................................................32

Experimental challenge infections by Leptospira ..................................................33

DNA Preparation and Real-Time qPCR ...............................................................34

LigA Southern Blot .................................................................................................35

2.3 Results ......................................................................................................................36

Patient Isolates of Leptospira .................................................................................36

Case Descriptions of Patients with Novel Leptospires .........................................36

Patient VAR10 ........................................................................................................36

Patient HAI029 ........................................................................................................37

Follow Up ...............................................................................................................37

Characterization of Isolates ..................................................................................38

16S rRNA Gene Sequencing ...............................................................................38
Serotyping .................................................................................................................42
DNA-DNA Hybridization .........................................................................................43
Biological characterization ......................................................................................43
Demonstration of a LipL32-related protein in “L. licerasiae”, L. fainei and
L. biflexa ..................................................................................................................45
Experimental animal infections ..............................................................................47
Prevalence of Leptospira licerasiae serovar Varillal seropositivity in the
Iquitos region .............................................................................................................49
Specificity of MAT for anti-Leptospira licerasiae serovar Varillal
antibodies ..................................................................................................................52

2.4 Discussion ..........................................................................................................54

Chapter 3

Genomic features of L. licerasiae ..............................................................................62

3.1 Introduction ..........................................................................................................62
3.2 Methods ................................................................................................................63

Bacterial strains .......................................................................................................63
Sequencing ...............................................................................................................64
Annotation ..............................................................................................................65
Gene, protein and genome comparisons ..............................................................66

3.3 Results ................................................................................................................67

General genome features .......................................................................................67
Origin of replication ...............................................................................................70
Transcription, translation and gene expression regulation .............................72
Replication, repair and recombination ...............................................................76
Metabolism ..............................................................................................................76
Comparative genomics .........................................................................................77
Potential virulence factors ......................................................... 78
Lipoproteins ................................................................. 81
Iron metabolism ............................................................... 82
3.4 Discussion .................................................................. 85

Chapter 4

Experimental evidence for virulence factors ........................................ 89

4.1 Introduction .................................................................... 89

4.2 Methods ......................................................................... 93
Mild acid hydrolysis and DMB-derivatization ................................. 93
Solid phase lectin binding ...................................................... 93
Mass spectrometry ................................................................ 94
Glycoprotein staining ................................................................ 94
Metabolic staining .................................................................. 94
LPS extraction ....................................................................... 95
PCR of sialic acid cluster genes ............................................. 95
NeuA2 southern blot ........................................................... 96
Palmitate labeling .................................................................. 96
Western blots ........................................................................ 97
Live immunofluorescence assay ............................................. 97

4.3 Results ........................................................................... 99
Genome analysis ................................................................. 99
DMB-derivatization and HPLS analysis ................................... 99
Solid-phase binding assay .................................................. 105
Mass spectrometry analysis ................................................ 108
LPS extraction and analysis ................................................................. 108
PCR of sialic acid cluster genes ............................................................ 110
NeuA2 southern blot ........................................................................ 110
Palmitate labeling .............................................................................. 110
Live immunofluorescence .................................................................. 116

4.4 Discussion ...................................................................................... 118

Bibliography ......................................................................................... 124
List of Tables

Table 2.1 Comprehensive panel of leptospiral serogroup antisera used to characterize

*L. licerasiae* strain Var10 ........................................................................................................ 31

Table 2.2 DNA relatedness of *L. licerasiae* to *Leptospira* species, *L. bromii* 5399,

*L. inadai* Lyme and *L. interrogans* RGA ........................................................................ 44

Table 2.3 Seroreactivity to *L. licerasiae* in acute leptospirosis patients from different

regions of Peru ......................................................................................................................... 53

Table 3.1 Sequencing statistic .................................................................................................. 68

Table 3.2 Alignment consensus results .................................................................................. 68

Table 3.3 Comparison general genome data between *L. licerasiae* and other leptospiral

species ..................................................................................................................................... 69

Table 4.1 List of proteins unique to in *L. interrogans* L1130 .............................................. 100
List of Figures

Figure 1.1  Phylogram of the three main *Leptospira* clades ........................................ 10

Figure 2.1  Pulsed field gel electrophoresis analysis of leptospiral isolates obtained from rats and humans in Iquitos, Peru .............................................................. 40

Figure 2.2  Phylogram of leptospiral 16s RNA gene sequences generated by Bayesian phylogenetic analysis ................................................................. 41

Figure 2.3  Southern blot to determine the presence of *LigA* in *L. licerasiae*. ........ 46

Figure 2.4  Western immunoblot of *L. interrogans* Copenhageni, *L. licerasiae* and *L. fainei*, using polyclonal antiserum to rLipL32 ............................................................... 48

Figure 2.5  Real time quantitative PCR analysis of experimental infection of Hamsters ................................................................................................................. 50

Figure 2.6  Seroprevalence of *L. licerasiae* in acute febrile patients in Iquitos ...... 51

Figure 3.1  Diagram of the chromosome initiation site of *L. interrogans* L1130, *L. interrogans* Lai, *L. licerasiae* and *L. biflexa* ......................................................... 71

Figure 3.2  Nucleotide sequence and translation of orf00124 contig 0012 in *L. licerasiae* ............................................................................................................. 74
Figure 3.3 Diagram of the dnaK cluster in *L. interrogans* L1130, *L. interrogans* Lai, *L. licerasiae* and *L. biflexa* ................................................................. 75

Figure 3.4 Two way comparison between all analysed genomes...................... 79

Figure 3.5 Four way comparison between genomes and number of genes unique to *L. interrogans* L1130 ........................................................................................................ 80

Figure 3.6 Diagram of ExbB/ExbD/TonB cluster in *L. interrogans* L1130 and *L. licerasiae* .................................................................................................................. 84

Figure 4.1 Diagram of potential sialic acid metabolism pathway in *L. interrogans* L1130 ..................................................................................................................... 102

Figure 4.2 Sialic acid cluster in *L. interrogans* L1130 and *L. interrogans* Lai ....... 103

Figure 4.3 Comparison of sialic acid cluster genes present in various species of leptospires ...................................................................................................................... 104

Figure 4.4 Schematic of DMB-derivatization and HPLC/MS analysis of leptospires and media .................................................................................................................. 106

Figure 4.5 Composite figure of : silver stain of affinity purified proteins, glycoprotein stain, sialic acid metabolic precursor and Loa22 western blot................. 107

Figure 4.6 DMB-derivatization analysis of lipopolysaccharide .............................. 109
Figure 4.7  PCR of sialic acid genes of different leptospiral species and Southern blot using a DIG-labelled neuA2 probe in various leptospiral species .......... 111

Figure 4.8 Comparison of signal peptide sequences of lipoproteins in *L. interrogans* L1130 and *L. licerasiae* ................................................................. 112

Figure 4.9 Palmitate labeling and LipL32 western blot in *L. interrogans* L1130 and *L. licerasiae* ................................................................. 113

Figure 4.10 Palmitate labeling and Loa22 western blot in *L. interrogans* L1130 and *L. licerasiae* ................................................................. 114

Figure 4.11 Comparison of Loa22 and LipL32 western blots and palmitate labeling in *L. interrogans* L1130 and *L. licerasiae* ................................. 115

Figure 4.12 Live immunofluorescence of surface lipoproteins in *L. interrogans* L1130 and *L. licerasiae* ................................................................. 117
Acknowledgments

I would like to acknowledge Professor Joseph Vinetz for his support and guidance as the chair of my committee.

I would like to thank Mike Matthias and Mandy Lewis for their invaluable guidance and support during this project, Mark Swancutt, Christian Ganoza and Kailash Patra for their encouragement and very helpful discussions, Andrew Borkosky and all present and past members of the Vinetz lab, without whom research would not have been as fun as it has been.

I would also like to thank Dr. Humberto Guerra for inspiring me to pursue translational research in tropical diseases and giving me continuous support and encouragement during all these years.

I would like to acknowledge Dr. Lee Smythe, Conrad Estrada, James-Platt Mills, Patrick LaRochelle, Kalina Campos, Eddy Segura, Nahir Chuquipiondo and Sonia Torres for their support in the isolation, identification and characterization of L. licerasiae and Hans Trüper for his kind advice regarding correct Latin etymology.

Chapter 2, in full, is a reprint of the material as it appears in PLOS Neglected Diseases, 2008, Ricaldi, JN, Matthias MA, Cespedes M, Diaz, MM, Galloway RL, Saito M, Steigerwalt, AG, Patra, KP, Vidal Ore C, Gotuzzo E, Gilman RH, Levett PN, Vinetz JM. The dissertation author was the co-primary investigator and author of this paper.
Chapter 3, in part is currently being prepared for submission for publication of the material. Ricaldi, JN, Matthias MA, Vinetz, JM. The dissertation author was the primary investigator and author of this material.

Chapter 4, in part is currently being prepared for submission for publication of the material. Ricaldi, JN, Borkowsky A, Matthias MA, Nizet V, Vinetz, JM, Lewis AL. The dissertation author was the primary investigator and author of this material.
Vita

2001 Medical Doctor, Physician Surgeon
Universidad Peruana Cayetano Heredia, School of Medicine.

2001- 2002 Master’s degree in Tropical Medicine and Infectious Diseases.
Universidad Peruana Cayetano Heredia (coursework completed)

2002-2003 Field Project Coordinator
Asociación Benefica Prisma/ University of Texas Medical Branch/
University of California San Diego. NIH funded project: Leptospirosis
transmission in the Peruvian Amazon, Iquitos, Peru.

2003-2004 Post Graduate Researcher
Division of Infectious Diseases, University of California San Diego.

2009 Doctor of Philosophy. Molecular Pathology Program, Microbiology track
University of California San Diego.

Publications

Segura ER, Ganoza CA, Campos K, Ricaldi JN, Torres S, Silva H, Cespedes MJ,
Matthias MA, Swancutt MA, Lopez Linan R, Gotuzzo E, Guerra H, Gilman RH, Vinetz
JM; Peru-United States Leptospirosis Consortium. Clinical spectrum of pulmonary
involvement in leptospirosis in a region of endemcity, with quantification of leptospiral


ABSTRACT OF THE DISSERTATION

Genomic insights into the pathogenesis of *Leptospira*

by

Jessica Nancy Ricaldi Camahuali

Doctor of Philosophy in Molecular Pathology

University of California, San Diego, 2009

Professor Joseph Vinetz, Chair

Leptospirosis is a zoonosis worldwide distribution, difficult diagnosis and underestimated burden. Most of its virulence factors remain unknown. The identification and characterization of a novel species of intermediate *Leptospira* isolated from humans and rodents from the Peruvian Amazon are described in this dissertation. This new specie has been denominated *L. licerasiae*. The genome of this new species was sequenced and analyzed. It has characteristics that are closer to those of pathogens than to those of saprophytic leptospires.

This new genome was compared with other known leptospiral genomes to predict virulence factor by identifying genes unique to pathogenic leptospires. Enzymes required for the synthesis of sialic acid were found among the genes unique to pathogens. Various pathogenic bacteria use sialic acids as a mechanism for immune evasion by promoting complement factor H binding and therefore preventing activation of the alternative pathway response.
A biochemical approach was used to demonstrate the synthesis of sialic acids by pathogenic leptospires. Potential localizations for sialic acid identified by MS/peptide identification and HPLC include the virulence factor Loa22, the lipoprotein LipL32 and lipopolysaccharide.

The sequences of known leptospiral lipoproteins were identified in *L. licerasiae* and compared with the sequences of lipoproteins from pathogenic leptospires. Differences were found in the sequence of the signal peptide, particularly in the lipobox region that determines cleavage by signal peptidases, membrane export and lipidation.

It was found that the virulence factor Loa22, among other lipoproteins in *L. licerasiae*, has lost the characteristics necessary for a functional lipobox.

Palmitate labeling shows differences in the pattern of protein lipidation between *L. licerasiae* and the pathogen *L. interrogans*. Live immunofluorescence studies show that Loa 2 is not expressed in the surface of *L. licerasiae* or expressed in much lower levels than in the pathogen *L. interrogans*, while LipL71 and LipL32 continue to be detected in the surface.

Loss of the lipobox and therefore loss of membrane-localization and post-translational modifications including sialylation could be mechanisms responsible for the less significant virulence of *L. licerasiae*. 
Chapter 1

Introduction

1.1 Introduction

Leptospirosis is a zoonotic disease of worldwide distribution (Levett 2001; Vinetz 2001; Bharti, Nally et al. 2003). This disease cannot be diagnosed on clinical grounds alone because its clinical presentation is diverse, ranging from undifferentiated fever to fulminant disease typified by various combinations of jaundice, renal failure, hemorrhage, and shock. Involvement of other organs such as gallbladder, pancreas, myocardium, and central nervous system has also been reported. The determinants of virulence are mostly unknown.

The aim of this dissertation is to try to infer potential virulence mechanisms by using translational research, bringing together what can be learned from field work in the Peruvian Amazon to data obtained with the use of new sequencing technologies and a genomics approach to analyze these findings.

1.2 History

There are many historical references describing the existence of fever and jaundice syndromes. These include references in the Talmud to jaundice associated with “withholding of urination” and epidemics of jaundice mentioned in the Bible (Rosner 1972). There are references to 34 epidemics of jaundice in Europe from 1745 to 1885.
There are also descriptions of epidemics of infectious jaundice in Alexandria and cases in Cairo, Malta, Athens, Constantinople, Syria and the Balkans (Sandwith 1905).

Different names have been used to describe infectious jaundice that we now know are caused by leptospirosis. Leptospirosis synonyms include “nanukayami”, cane-cutter’s disease, rice-field jaundice, “Schalmmfieber”, seven-day fever, rat fever and mud fever. In several instances these illnesses have been associated with working in a certain trade (like butcher’s) or exposure to environments like swamps or mud. The described symptoms could be associated with various etiologies, including yellow fever and both the mild and severe forms of leptospirosis. The history of these two diseases is intertwined, as the severe forms of leptospirosis can be indistinguishable from yellow fever causing misidentification of the causative agents of both diseases.

In 1886 Adolf Weil described the clinical features of four patients presenting with fever, jaundice and renal failure in the article: “A peculiar form of acute infectious disease characterized by jaundice, swelling of the spleen and nephritis” (Weil 1886). Since then, this syndrome became known as Weil’s disease.

During the late 1800’s and the early 1900’s there was an ongoing international effort to identify microbial etiological agents of disease. Some of the largest efforts were aimed at identifying and understanding the transmission of the causative agent of yellow fever.

In 1907 Stimson reported finding organisms resembling the spirochetes in silver stained kidney tissue of patients with a diagnosis of “yellow fever” (Stimson 1907).

In 1914 Inada and Ito identified the causative agent of Weil’s disease and Odaneki in Japan by replicating the human disease in guinea pigs inoculated with blood from
patients and later identifying similar organisms in both the human and guinea pig tissue. This agent was named *Spirochæta icterohæmorrhagiae* (Inada, Ido et al. 1916). Later they described isolating a similar organism from patients with seven-day fever, which they called *Spirochæta hebdomadis* (Ido, Ito et al. 1918). It is now known that both are leptospiral diseases and represent different ends of the spectrum of the same disease.

At the same time, research in Europe was being done to determine the etiology of Weil’s disease, which was common among soldiers in World War I. An organism similar to *S. icterohæmorrhagiae* was identified in samples from soldiers in Britain, France and Italy (Noguchi 1917).

Hideyo Noguchi performed studies to compare the various organisms associated with Weil’s disease that had been isolated from rats in Japan, Europe and the United States. Through evaluation of cross reactivity of serum from infected guinea pigs he concluded that the Japanese, Belgian and North American strains were “probably identical”. He suggested designating them as a new genus, called *Leptospira*. His studies also found evidence supporting the role of wild rats in transmission of this disease (Noguchi 1917).

In addition to performing research in Colombia and Ecuador trying to identify the causative agent of yellow fever, Noguchi reported isolating another spirochete from patients with yellow fever. He denominated this organism *Leptospira icteroides* and reported isolation not only from yellow fever patients but also from guinea pigs inoculated with the patients’ blood (Noguchi 1919; Noguchi 1920).

In 1930 it was established that “*Clinically yellow fever is similar to infectious jaundice. The differences existing between the two diseases appear to be chiefly those of*”
degree. There is more marked jaundice and less hemorrhage in yellow fever than in infectious jaundice" (Sawyer, Kitchen et al. 1930).

Based on clinical studies of yellow fever in West Africa where no leptospires were isolated from patients, it was highly questioned that this was the etiologic agent of yellow fever. Furthermore, no serological evidence of leptospiiral infection was found in a yellow fever outbreak in Brazil. Finally, experiments done by different researchers trying to show the transmission of leptospires via Aedes aegypti failed. It was soon concluded that Weil’s disease and yellow fever had similar clinical presentations but different etiologic agents and that only yellow fever could be transmitted through mosquito bites. (Sawyer, Kitchen et al. 1930).

1.3 Epidemiology

Starting shortly after the initial description by Professor Weil, there have been many reports of epidemic and endemic occurrences of Weil’s disease. The wide distribution of leptospirosis was known as early as 1937 and it was suspected to be “universal” (Alston and Brown 1937). Some of the epidemiological patterns present until today were described, such as a high incidence in sewage workers in London (Fairley 1934), fish workers in Aberdeen (Davidson and Smith 1936) and the fact that up to 60% of serologically confirmed cases do not present with jaundice (Alston and Brown 1937).

Leptospirosis has been found everywhere it has been looked for. While some authors consider it an emerging disease there is also the notion that it is “persistent and often under recognized” (Brandling-Bennett 1996; Levett 1999; Meites, Jay et al. 2004; Vijayachari, Sugunan et al. 2008). The need for laborious and specific diagnostic
tests for a proper diagnostic of leptospirosis supports the latter statement (Bharti, Nally et al. 2003). This fact associated to the varied clinical forms leptospirosis can present make it difficult to known the true prevalence and worldwide burden of this disease (Hotez, Bottazzi et al. 2008). Although the best conditions for the persistence and transmission of *Leptospira* are found in tropical areas, leptospirosis cases have also been reported in subtropical and temperate climates, albeit differing epidemiological patterns. The transmission of leptospirosis occurs directly through contact with animal reservoirs and their blood or urine or indirectly through water contaminated with leptospires shed from reservoir mammals through urine. Humans come into contact with contaminated surface waters through occupational and recreational activitie. In endemic tropical areas everyday activities can lead to exposure to contaminated water sources. Infection occurs through mucous membranes or non-intact skin (Tassinari, Pellegrini et al. 2008).

Three distinct epidemiological patterns have been described for leptospirosis: endemic, epidemic, and sporadic or urban. Occupational leptospirosis could be considered a fourth pattern that in recent years has decreased in the developed world (A.M. Monahan 2009).

Serological studies have been used to demonstrate the endemicity of leptospiral infection in various tropical areas of the world, including Africa (Bertherat, Renaut et al. 1999), Asia (Sehgal, Vijayachari et al. 1999) and Central and South America (Sulzer, Sulzer et al. 1978;Leal-Castellanos, Garcia-Suarez et al. 2003). Prevalence tends to be underestimated, mostly due to the high percentage of asymptomatic infections. The data from endemic areas show that asymptomatic infection is common, with rates as high as 70% reported during an outbreak (Ashford, Kaiser et al. 2000).
The endemicity of leptospiral infection in these tropical areas can be attributed to prevalent geographic, climatic and socioeconomic factors. A temperate climate with frequent rain, and urban slum or rural environments that lack proper sanitation and water services provide the ideal setting for the transmission of leptospirosis.

Leptospires can survive up to weeks and months in warm, moist soil and water (Levett 2001). The close relationship of environment and transmission has been described in the Peruvian Amazon, where all of these conditions occur. In this area, a high percentage of the population, both from rural areas and urban slums (16-28%) was seropositive and those who were seronegative had a high rate of seroconversion (12% in <6 months). No difference in seroprevalence was found between males and females and the presentation of severe disease is uncommon (Johnson, Smith et al. 2004; Segura, Ganoza et al. 2005). Isolates obtained from patients, carrier animals and water samples in these areas represent diverse leptospiral serovars (Ganoza, Matthias et al. 2006).

Epidemic leptospirosis occurs in many different settings, sometimes in a cyclical manner associated with rainy seasons and particularly when rainfall surpasses a certain threshold (Tassinari, Pellegrini et al. 2008). One of the most studied settings is the yearly outbreak of severe leptospirosis in the favelas or urban slums of Sao Paulo, Brazil. Isolates obtained from patients from different locations in this area and at different time points have been shown to be part of a clonal subpopulation (Pereira, Matsuo et al. 2000). Epidemic leptospirosis is also often associated with flooding or other natural disasters, such as the outbreak in Nicaragua in 1995 (Centers for Disease Control and Prevention 1995; Munoz, Jarquin et al. 1995; Zaki and Shieh 1996). Risk factors associated with infection during the Nicaragua outbreak in were living in a rural household, washing
clothes or bathing in a river or creek, protective factors included signs of relative wealth including owning television or radio and having an indoor water source.

Outbreaks with attack rates ranging from 16 to 50% have been associated with recreational exposure such as water sports or adventure races, military training and occupational exposure such as farming (Hahn 2001; Stone and McNutt 2001; Russell, Montiel Gonzalez et al. 2003; Yoder, Hlavsa et al. 2008; Desai, van Treeck et al. 2009). In these cases the main risk factors are contact of mucous membranes with contaminated water or contact with rodents. Clinical disease was mild in most cases.

Endemic and epidemic leptospirosis may occur simultaneously, as in Nicaragua, Brazil and India (Kuriakose, Eapen et al. 1997; Bharadwaj, Bal et al. 2002). There is evidence that epidemic urban leptospirosis occurs in urban slums where rodents are abundant and contribute to the maintenance of leptospires. Endemic transmission in these areas seem to be associated with factors such as lower income and disposal of waste in vacant land, rivers or the ocean, which could be a correlate of poor access to sanitation services (Oliveira, Guimaraes et al. 2009). Urban outbreaks tend to present in clusters that can be associated temporally with factors such as high average daily rainfall (Tassinari, Pellegrini et al. 2008).

Sporadic cases of urban leptospirosis occur in inner cities in developed countries mostly in the homeless and other disadvantaged urban populations (Vinetz, Glass et al. 1996; Shaukat, Pohlel et al. 2004; Hotez 2008). Serological studies have found reservoirs in many cities around the world: rodents and shrews in Zurich, opossums in Sydney, wild boars in Berlin, rats in Baltimore and Detroit, coyotes in Tucson, raccoons in St. Louis and dogs in northern California. Seroprevalence ranges from 10 to 30% in most cases.
The order Spirochaetales was proposed as a result of an existing tendency to attribute taxonomic value to morphological features. Spirochetes share the presence of a helical protoplasmic cylinder bound to axial filaments inserted to either ends of the protoplasmic cylinder and that vary widely in number depending on the species. The protoplasmic cylinders are surrounded by an outer membrane. At the same time the spirochetes can vary widely not only in size but also physiologically, including energy-yielding specialization (Canale-Parola 1977) and ability to survive in different environments. These variations reflect the differences that can be found in the genotypes of these bacteria, including G+C contents varying from 36 to 66%.

The genus *Leptospira* Noguchi (1917), type species *Leptospira interrogans* (Stimson, 1907), type strain serovar Icterohaemorrhagiae, Ictero No. I (ATCC 43782) is made up by both host-associated and free living (saprophyte) bacteria. The taxonomy of *Leptospira* is complicated and has often changed over time, until the recent use of molecular techniques (Adler and Faine 2006). Classically, leptospires have been classified into serovars, and the approximately 240 serovars were divided into 24 antigenically related serogroups. This classification is achieved by using antisera raised in rabbits specific to each serovar. These antisera react mainly against the lipo-polysaccharide on the surface of *Leptospira*. The serologic classification is also the basis for the microagglutination test (MAT), which remains the gold standard for the diagnosis of leptospirosis. Currently, all pathogenic leptospires are classified as *L. interrogans*,

sensu lato and all saprophytes as *L. biflexa* sensu lato. They are designated by the specie name followed by the serovar.

Molecular analysis allows for the classification of leptospires into genomospecies. DNA hybridization and 16S ribosomal RNA gene sequencing have been used to develop a clearer taxonomic classification. (Perolat, Lecuyer et al. 1993). Twenty species of *Leptospira* are currently recognized (Figure 1.1). These are grouped into 3 genomic clades: one that comprises pathogens, the second that comprises non pathogens and the third that comprises species of unclear pathogenicity, usually referred to as intermediates (Morey, Galloway et al. 2006). There is poor correlation between the serological and the DNA based taxonomy as some members of the same serovar belong to different species.

Other methods that have been used to classify leptospires include DNA fingerprinting based on pulsed-field electrophoresis of *Not*I restriction digests of DNA. These fingerprint patterns correlate well with serovar grouping (Herrmann, Bellenger et al. 1992; Galloway and Levett 2008).
Figure 1.1. Phylogram showing the three main leptospiral clades: pathogens, intermediates and saprophytes (Adapted from Creative Commons: Evolutionary relationships of Leptospira and Leptonema).
1.5 Clinical presentation

The clinical presentations of leptospirosis are very diverse. In most cases serological conversion will be the only evidence of infection, however the data are limited due to the lack of prospective studies. The most frequent presentation is an undifferentiated febrile syndrome with malaise, headache and sometimes gastrointestinal manifestations such as nausea vomiting or diarrhea, or a maculo-papular rash or a non productive cough (Katz, Ansdell et al. 2001). Complicating diagnosis even more, severe leptospirosis does not always present as the classic triad of fever, jaundice and renal failure in a biphasic illness, but also pulmonary hemorrhage, aseptic meningitis, refractory shock and myocarditis have been reported (Panicker, Mammachan et al. 2001).

In the initial reports of Weil’s disease leptospirosis was described as being biphasic. It is now known that symptoms start developing between 3-14 days after infection. The initial septicemic phase lasts for of up to seven days. Leptospires can be found in the bloodstream during this phase. This period is characterized by fever of sudden onset, malaise and anorexia. Conjuntival suffusion and myalgia (muscle pain) mostly located in the calves are described as classical signs of this early phase of leptospirosis, but these are not specific.

The immune phase of the disease begins after a period of apparent resolution of the illness lasting 3-4 days. During this phase leptospires cannot be cultured from blood, but start being detected in urine. Is during this period that the severe manifestations of disease may present, suggesting an immune system dependant pathogenesis.
Severe leptospirosis presenting as a fulminant monophasic illness associated with pulmonary hemorrhage has been reported (Marotto, Nascimento et al. 1999; Spichler, Vilaca et al. 2008).

Although most cases of mild disease resolve spontaneously, it is important to consider leptospirosis a diagnostic possibility in the appropriate clinical and epidemiological context, since early antibiotic treatment will decrease mortality in the 5-10% of patients that develop the severe form of disease. In the patients that present severe disease mortality seems to be increased in older patients and those who present with pulmonary involvement and oliguria (Spichler, Vilaca et al. 2008).

1.6 Microbiological diagnosis

Laboratory confirmation is necessary for a conclusive diagnosis of leptospirosis. This can be made through direct visualization, isolation, molecular or serological methods. Choosing the best diagnostic method to use depends on the time in the course of illness. During the early phase of the disease leptospires might be seen in a blood smear examined with dark field microscopy. Unfortunately, this requires expertise and training to be able to differentiate leptospires from fibrin strands and other artifacts in the slide, thus making this method unreliable.

Isolation and identification of the infectious agent is the ideal form of diagnosis of any infectious disease. Leptospira is a fastidious organism that has specific media requirements to grow properly. Its nutritional needs include long chain fatty acids and mineral salts. To obtain isolates in leptospirosis one requires not only special media, but the blood sample should be taken during the early phase of the disease. Even in ideal
conditions isolation rates are low. Later in the course of illness isolation can be attempted from urine or CSF. Due to the slow growth rate of the organism (up to 3 months until detection) a positive result is often obtained only when it no longer has relevance for the treatment of the patient. However, isolates are needed to know the prevalence of serovars in a region, to identify new species and serovars and for research purposes.

To this the day the microagglutination test (MAT) remains the gold standard for the diagnosis of leptospirosis. In this test, one to three cultured leptospire strains of each serogroup are incubated with patient serum for 2 hours in a microtiter plate and then observed for microagglutination under a dark field microscope. To make a confirmed diagnosis, paired sera showing a four-fold rise in titer or one sample with a high titer (>1:800) are needed; therefore its usefulness in a clinical setting is limited. This test is not only laborious and time consuming, but requires regular maintenance of leptospiral cultures representative of each serogroup, prohibiting its widespread use.

Non-MAT serological methods are most often used for the diagnosis of leptospirosis. Several ELISA based assays have been developed, some using whole non-pathogenic leptospires, others using recombinant leptospiral antigens (Terpstra, Ligthart et al. 1985;Natarajaseenivasan, Vijayachari et al. 2008). Several tests are commercially available. In a comparison study done in Hawaii including a dot ELISA, 3 different ELISA kits, commercially available IHA and IFA kits, a dipstick and a latex agglutination test, all except the latex test had similar results. Sensitivities ranged from 26 to 51% when using samples from confirmed and probable cases while specificity was above 80% for all except the latex test (Effler, Bogard et al. 2002). In an outbreak study in
Pichanaqui, Peru a commercially available ELISA test had a sensitivity of 26% and a specificity of 60% when compared with MAT (Russell, Montiel Gonzalez et al. 2003). More recently DNA methods based on the amplification of genes *hap1* and *rrs5* have been evaluated (Branger, Blanchard et al. 2005). Real time PCR has been used in a research setting for quantification of leptospiral burden (Manu, Roy et al. 2009), but these methods are not yet used for routine diagnosis (Segura, Ganoza et al. 2005; Céspedes, Tapia et al. 2007). These methods might be limited by the low levels of leptospiremia commonly found in patients.

1.7 Pathogenesis and virulence

The genomes of two strains of each of two pathogenic species of *Leptospira*, *L. interrogans* serovars (Lai and Copenhageni strain L1130), two strains of *L. borgpetersenii* and one of the saprophyte *L. biflexa* have been published, but still little is known about the underlying mechanisms of pathogenicity and virulence. Both host and pathogen specific factors seem to guide the course of infection, determining the occurrence of mild self-limiting flu-like disease, severe leptospirosis with liver and kidney failure, or pulmonary hemorrhage or anywhere in between.

In order to cause disease, leptospires need to trespass host skin or mucous membrane barriers, find their way to the bloodstream, evade opsonization and complement deposition, find nutrition sources while in the bloodstream, replicate and make it to the liver and kidneys. Factors that mediate each one of these steps must be present to produce disease.

The first step in invasion must include attachment to host cells. It has been
reported that virulent leptospires are able to attach to mammalian cells. (Thomas and Higbie 1990; Dobrina, Nardon et al. 1995; Barbosa, Abreu et al. 2006). Leptospiral proteins LigA and LigB contain immunoglobulin-like repeats and these repeats have been associated with adhesion to host cells in other gram-negative bacteria. Highly passaged pathogenic leptospires that are unable to infect animals do not express these proteins. It has been reported that in pathogenic leptospires exposure to physiologic osmolarity upregulates their expression. Increased expression in a higher osmolarity has been associated in vitro with an increase of 1.9 times the adhesion to fibronectin, and increased binding to type IV collagen, fibrinogen and laminin. Fibrinogen binding may have a role in binding during tissue colonization. At the same time laminin binding to exposed basal lamina could have a role in colonization of the glomeruli. Invasion of the glomeruli has been proven by the finding of a high density of leptospires in infected hamsters (Matsunaga, Werneid et al. 2006). Exogenous recombinant Lig proteins partially inhibit binding to fibrinogen suggesting that more than one factor mediates adhesion (Marangoni, Aldini et al. 2000; Choy, Kelley et al. 2007). LigA is also released during infection potentially promoting dissemination. Other proteins that have been shown to have laminin binding activity in-vitro include a 36Kd protein and recombinant LenA/LfhA/Lsf24 and its related proteins LenBCDEF. LenA seems to bind factor H of complement so it might have more that one role in virulence. (Merien, Truccolo et al. 2000; Barbosa, Abreu et al. 2006; Stevenson, Choy et al. 2007).

Leptospires have been reported to rapidly achieve hematogenous dissemination. Studies done in guinea pigs have found that once in the bloodstream the number of leptospires initially decreases and then achieves logarithmic growth with a generation
time of 8.3 hours in vivo (Faine 1957). This logarithmic growth could either suggest that a selection process occurs with only the most virulent leptospires replicating and invading host tissues, or it could be reflecting growth happening in the recently infected tissue. Leptospires can be found in the tissue and blood vessels of kidneys and liver 2 hours after intra-peritoneal infection (Faine 1957).

Data published about 20 years ago shows evidence for differential attachment and migration through cultured cells when using pathogenic strains, compared to non-pathogen ones. These experiments used endothelial (HUVEC) and epithelial cells (MDCK) and found that attachment occurred in a dose- and time-dependent manner and decreased with pre-treatment with heat, proteases or rabbit immune serum (Thomas and Higbie 1990). Experiments performed more recently using Madin-Darby canine kidney cells (MDCK-II) have shown that both pathogenic and non-pathogenic leptospires have the ability to translocate through the cellular barrier with no evidence of cytoskeleton rearrangement or monolayer integrity, albeit at different rates and with no evidence of intracellular multiplication or cell to cell spread (Barocchi, Ko et al. 2002). This mechanism could be used to go from tissues to bloodstream and vice versa.

Pathogenic *Leptospira* species are capable of multiplying in the host bloodstream and penetrating tissues. In order to survive in the bloodstream leptospires must evade different components of the immune response including the classic, alternative and lectin pathways of complement. In a non-immune host pathogenic leptospires avoid being destroyed by the alternative pathway of complement during the initial leptospiremic phase that can last up to 10 days, while non-pathogenic leptospires are readily killed by serum components or engulfed by phagocytes. Intermediately virulent leptospires are
able to resist serum killing in a lesser degree than highly virulent ones (Meri, Murgia et al. 2005). Data shows that C3 deposition occurs in both serum resistant or serum sensitive strains, but binding of the terminal components C5, C6 and C8, and the cytolytic MAC (membrane attack complex) seems to be markedly diminished in pathogenic and intermediate strains, as seen in ELISA testing. It has also been known that factor H (fH), the soluble protein of the alternative pathway prevents binding of factor B to C3b, therefore stopping the formation of C3bBb and leading to termination of the alternative pathway cascade. It has been shown that fH does not bind to serum sensitive strains but only to intermediate and resistant strains (Meri, Murgia et al. 2005).

The soluble protein of the classical pathway of complement, C4BP a 570Kd spider shaped protein interferes with the assembly and decay of C4bC2a and is a cofactor of the proteolytic inactivation of C4b. Both fH and C4BP down-regulate complement activation, prevent opsonization and the formation of the membrane attack complex. It has been recently described that non-pathogenic strains undergo serum killing through the alternative pathway. Pathogens bind C4BP and avoid killing by the classical pathway, non-pathogens do not bind C4BP (Barbosa, Abreu et al. 2009).

Mannose-binding lectin (MBL) can modulate inflammatory responses and enhance activation of the complement cascade. MBL mediates opsonophagocytosis by binding to mannan on the surface of bacteria and both activating the lectin pathway and directly. MBL level in humans depend on genetic factors, including polymorphism of the both the promoter or exon 1 on the MBL2 gene. A retrospective serologic study looking at patients with a confirmed diagnosis of leptospirosis found an association between levels of MBL higher than 1000 ng/mL and an increased severity of disease, including
hemorrhage and respiratory, renal and cardiac complications (Miranda, Vasconcelos et al. 2009).

In addition to evading complement and opsonization in the blood stream leptospires must also evade the other components of the cellular immune responses.

Naturally acquired immunity has been shown to occur in animal models, but its protective effect is serovar specific. In vaccination studies in cattle it has been shown that a type 1 immune response (Th 1) occurs. PBMC cultures from unvaccinated animals produce interferon-gamma (IFN-γ) in lower levels than vaccinated animals. Vaccinated animals also produced IFN-γ cells including CD4(+) and WC1(+) gamma-delta T cells. Similar levels of immunoglobulin G1 (IgG1) are found in vaccinated and naïve animals, but with a two-fold increase in IgG2 in vaccinated over unvaccinated animals. It is suggested that the natural type 1 response induced is not strong enough to prevent chronic infection (Naiman, Blumerman et al. 2002)

Antibodies are directed mostly against leptospiral LPS and are thought to be mediated by humoral immunity. At the same time LPS stimulates the innate immune system through Toll like receptors (TLRs).

TLRs are membrane proteins that contain an extracellular domain rich in Leucine repeats and a cytoplasmic domain with Toll/IL-1R homology that mediates inflammatory signaling. TLRs recognize elements of self versus foreign elements. During enterobacterial infections the lipid transferase LPS-binding protein (LBP) attaches to the lipid A moiety of LPS in the surface of the microorganism. LBP detaches LPS from the surface and binds it to CD14. CD14 is a glycoprotein that can be found both as a GPI anchor protein attached to myelomonocytic and in soluble form. CD14 delivers LPS to a
TLR4-MD2 complex and also has a role in LPS clearance. (Cullen 2004; Miyake 2004; Nahori, Fournie-Amazouz et al. 2005; Schröder, Eckert et al. 2008).

LPS from gram negative bacteria that cause systemic disease seem to be less recognized by TLR4-MD2 than that from bacteria that cause localized infection (Munford 2008). Consistent to this finding it has been reported that leptospiral LPS has much lower toxicity than other gram negative LPS and that it activates TLR2 \textit{in vitro} (Werts C 2001). Leptospiral LPS has not been thoroughly characterized; only the structure of the lipid A portion has been elucidated and is described as having some unique characteristics. Leptospiral lipid A contains four N-linked acyl chains, two unsaturated secondary chains and has been found to have only one phosphate group instead of two. This phosphate group is also methylated, which is a very rare occurrence in biology (Que-Gewirth 2004). The low recognition of leptospiral LPS by the TLR4-MD2 complex could be explained by these changes, since it is known that variants of lipid A structure seem to have different effects in the pro-inflammatory response. The recently reported structure of the TLR4-MD2 complex highlights the need of the two phosphate groups present in enterobacterial LPS for TLR4-MD2 recognition (Park, Song et al. 2009). The diminished recognition of leptospiral LPS could contribute to its survival in the bloodstream and dissemination.

Other components of the leptospiral surface can elicit a response from different components of the immune system. Lipoproteins are abundant in the leptospiral outer membrane. They play an important role in activating the host inflammatory response. In \textit{Borrelia} and \textit{Treponema} it has been shown that TLR2 is needed for lipoproteins to activate innate immunity (Cullen 2004; Schröder, Eckert et al. 2008). Interestingly
enough, there are data that suggest that leptospiral LPS can induce NF-κB activation through CD14 and TLR2 and that LPS from avirulent leptospires is more potent than that from virulent leptospires (Werts C 2001).

Data from experiments in TLR4-deficient mice suggest that the response to LPS is different than in humans. TLR4-deficient mice infected with *L. interrogans* present disease similar to the severe forms of leptospirosis in humans (Viriyakosol, Matthias et al. 2006). These data suggest that murine TLR4 recognizes leptospiral LPS and it prevents dissemination of *Leptospira* in mice. A similar difference in response to LPS has been reported in *Pseudomonas aeruginosa* and *Coxiella burnettii* (Munford 2008). Whether this is true in humans remains to be shown.

Dendritic cells (DC) form a link between the innate and the adaptive immune response. They express pattern recognition receptors (PPR) to recognize pathogen molecular patterns. PPRs include C-type lectins such as DC-specific ICAM-3 grabbing integrin (DC-SIGN). DC-SIGN recognizes high mannose glycans and fucose containing antigens. *L. interrogans* induces DC through DC-SIGN maturation *in vitro* but there are inter-serovar differences in cytokine production independent of virulence (Gaudart, Ekpo et al. 2008).

Clinical data shows that levels of soluble IL-2R, IL-6, and TNF-α are elevated in patients with leptospirosis (Klimpel, Matthias et al. 2003). Comparing PBMC from healthy donors and patients with acute disease it has been shown that the ratio of IL-10 to TNF-α is associated with disease progression, a high ratio of IL-10/TNF-α predicts a good outcome.

Besides the factors involved in the pathogenesis mentioned previously, there are
other potential virulence factors for which more evidence remains to be published. Motility and chemotaxis are thought to be necessary for invasion of the host. Early reports of chemotaxis towards hemoglobin can be found in the literature (Yuri, Takamoto et al. 1993). It was reported that only virulent strains of pathogenic leptospires showed migration towards 10-fold dilutions of hemoglobin. Chemotaxis genes have been shown to be upregulated in *L. interrogans* Lai in temperature shift experiments, suggesting that the temperature inside the host might elicit similar changes in expression (Lo 2006). Evidence of a more complex chemotaxis pathway involving a family of *Che* genes has been reported. Using gene complementation with leptospiral proteins in mutants of *E. coli*, Li and colleagues showed that CheW1 and CheW2 could replace *E. coli* proteins in the phospho-relay pathway but do not restore swarming induced by chemotaxis (Li, Dong et al. 2006). Using the same system, it was shown that CheY genes were able to restore swarming in some degree (Li, Dong et al. 2006). Still, the exact role of chemotaxis is pathogenesis in not known.

The role of apoptosis in leptospiral infection has not been thoroughly studied. It has been reported that pathogenic leptospires and some of their components induce NFκB and p38 phosphorylation in a dose-dependent manner in cultured murine microglial cell lines (Blasi, Ardizzoni et al. 2007). TNFα and nitric oxide were also induced in a dose dependent manner, also lipoproteins were the cellular component that induced the highest responses. Only very recently has a report on the potential apoptosis-inducing pathway in a murine model been published. It has been shown that leptospires reduce the viability of macrophage-like murine cells J774A.1 in a dose-dependant manner. This reduction of viability appears to be caused by apoptosis, which was seen even at a MOI of 10 and as
early as 30 min after infection. The same study used caspase-inhibitor assays to show that apoptosis is induced via the intrinsic death pathway not only by caspase-3, but also by caspase-6. A role was also found for caspase-8 and -9 (Jin, Ojcius et al. 2009).

The goal of this dissertation is to elucidate new virulence mechanisms for leptospiral pathogenesis. To achieve this, I will first describe the isolation and characterization of a novel intermediate Leptospira specie. I will then analyze its genome sequence with emphasis on functions present in other leptospires, components present on the surface and previously suggested virulence factors. I will compare this genome to currently known pathogenic and saprophytic leptospire genomes. Finally, I will present experimental evidence supporting the role in virulence of proteins that are unique to the pathogen L. interrogans  L1130.
Chapter 2

Isolation and characterization of a new intermediate specie of

*Leptospira*

2.1 Introduction

In the Iquitos region in the Peruvian Amazon, as is the case in developing countries around the world, many patients with undifferentiated febrile illnesses do not have an etiology identified, even in comprehensive, prospective studies (Watts, Callahan et al. 1998). The Peruvian Amazon region of Iquitos and its surrounding areas provide an ideal ecological setting for the maintenance and transmission of leptospirosis. Diverse fauna, tropical climate and the lack of proper sanitation (Levett 2001; Vinetz 2001; Bharti, Nally et al. 2003). The diagnosis of leptospirosis is made even more difficult by the lack of sensitive and readily accessible diagnostics. Clinical leptospirosis has neither been commonly recognized nor reported in Iquitos, so that it has been mostly ignored as a cause of febrile illness. Malaria and dengue are important causes of acute febrile illness in the Iquitos region but leptospirosis has only been reported there when research studies have specifically looked for it (Segura, Ganoza et al. 2005). Renal carriage among wild animals in Iquitos is common (Bunnell, Hice et al. 2000), yet comparatively few strains have been isolated in the Peruvian Amazon region of Iquitos (Liceras de Hidalgo and Mejia 1981). In the context of a prospective study to determine the proportion of acute,
differentiated febrile illnesses caused by acute leptospirosis, we isolated a new leptospiral species and serovar. We have provisionally named this isolate “\textit{Leptospira licerasiae}” serovar Varillal strain VAR 010, determined its major mammalian reservoir, and shown its importance in regional diagnosis of acute leptospirosis.

2.2 Materials and Methods

\textit{Humans: Enrollment, Sampling and Culture}

Patients presenting at the Belen, Moralillo, Varillal, Padrecocha and Zungarococha Ministry of Health health posts and the Hospital de Apoyo in the Iquitos region of the Peruvian Amazon with complaint of fever were enrolled in a prospective study after oral assent for adults (after reading a detailed script of what participation would consist of along with potential risks and benefits) or written informed consent from parents or legal guardians for children. Included in the informed consent process was a request to administer a questionnaire that asked for personal, medical, demographic and economic information, and requests for serial samples of blood and urine. Specific dates of the study periods are as follows: Belen, January 2003 to September 2005; Hospital Apoyo de Iquitos, May 2003-April 2006; Zungarococha, November 2002 to July 2005; Moralillo, January 2003 to January 2005; Varillal, November 2002 to July 2005; Padre Cocha February 2004 to May 2005.

Inclusion criteria were a self-reported undifferentiated febrile illness of \( \leq \) 2 weeks duration with a negative malaria smear. Clinical and demographic data were collected from the patients using a questionnaire. Seven milliliters of whole blood were collected by venipuncture at the time of presentation for culture and serological analysis.
up blood samples for serological analysis and mid-stream urine samples for leptospiral culture were collected 10-70 days after enrollment. For urine culture, the pH of samples was adjusted to ~7.4 with 10 M NaOH at the time of collection. Two tubes of 5 ml semisolid EMJH (Difco, BD Biosciences, Sparks, MD) containing 0.01% (w/v) 5-fluorouracil (5-FU) and 300 μg/ml neomycin were inoculated on site with 2 and 4 drops of whole blood, respectively, using strict aseptic techniques. Urine samples were centrifuged briefly at ~ 800 rpm and 2 tubes of semisolid EMJH medium (BD Biosciences) containing the same antibiotics and concentrations were inoculated with 2 and 4 drops of clarified urine, respectively. Cultures were examined weekly by darkfield microscopy and classified as negative if no organisms typical of *Leptospira* were observed by 12 weeks. A high level of care was taken to avoid contamination by water-borne saprophytic *Leptospira*; no saprophytes were obtained during the course of the study (as determined by 16S rRNA gene sequencing).

This study was approved by the Human Subjects Protection Program, University of California San Diego, and the Ethical Committees of Asociacion Benefica PRISMA, Lima, Peru, and Universidad Peruana Cayetano Heredia, Lima, Peru.

*Human Serological Analysis and Interpretation Criteria*

Serologic testing of patient samples was performed at the Instituto Nacional de Salud in Lima, Peru using an in-house IgM ELISA (Cespedes, Glenny et al. 2002) (which includes as antigens serovars Icterohaemorrhagiae, Bratislava, Ballum, Canicola, Cynopteri, Grippotyphosa but not “*L. licerasiae*” serovar Varillal). Microscopic agglutination testing (MAT) was done using the following antigens (serogroup followed
by serovar in parentheses): serogroup Andamana (serovar Andamana), Australis (Australis and Bratislava), Ballum (Ballum), Bataviae (Bataviae), Canicola (Canicola), Celledoni (Celledoni), Cynopteri (Cynopteri), Djasiman (Djasiman), Grippotyphosa (Grippotyphosa), Hebdomadis (Borincana), Icterohaemorrhagiae (Copenhageni, Icterohaemorrhagiae and Mankarso), Javanica (Javanica), Mini (Georgia), Pomona (Pomona), Pyrogenes (Alexi and Pyrogenes), Sejroe (Hardjo and Wolffi), and Tarassovi (Tarassovi). Sera were screened at a dilution of 1:100 and positive sera were titrated to endpoint using standard methods (Faine 1982).

Clinical criteria for submitting sera on patients (both in Iquitos and nationwide) for serological diagnosis were undifferentiated fever for 2 weeks or less, malaria smear negative, and no alternative explanation for fever.

Serological criteria for diagnosing acute leptospirosis in all areas of Peru other than Iquitos included any one of the following: seroconversion in IgM by ELISA from acute to convalescent sera; seroconversion in MAT from negative to 1:100 or greater; 4-fold rise in titer between acute and convalescent sera; or a single titer of 1:400 or greater. The single titer of 1:400 in non-Iquitos regions was chosen as the seropositivity cutoff because of national data indicating that titers at this level or lower were background in the population in asymptomatic individuals (M. Cespedes, Instituto Nacional de Salud, Lima, Peru, unpublished observations).

Serological criteria for stating that a specific MAT titer was associated with acute leptospirosis were made more stringent in Iquitos than in other parts of Peru because of the high prevalence of low level (1:400 or less) anti-“L. licerasiae” serovar Varillal antibodies in asymptomatic individuals (data not shown). Serological criteria to assign a
diagnosis of acute leptospirosis in Iquitos included any one of the following: IgM positive by ELISA in either acute or convalescent sera; seroconversion in MAT from negative to 1:100 or greater; 4-fold rise in titer between acute and convalescent sera; or a single titer of 1:800 or greater.

Animals: Trapping and Culture for Leptospira

Rats were caught live in baited wire-mesh traps (Tomahawk, USA) left overnight near dwellings in the urban area of San Juan near the Iquitos airport, in the urban slum of Belen in Iquitos, or the rural area of Moralillo 15 km outside Iquitos removed 1 km from the Iquitos-Nauta road. Animals were anesthetized with isoflurane and the kidneys were removed aseptically; blood was collected by cardiac puncture. Excised kidney material was minced using a clean, sterile scalpel blade and cultured in semisolid PLM-5 containing antibiotics. All cultures were incubated at 30°C for up to 12 weeks and checked bi-weekly for growth. Positive cultures were sub-cultured into liquid EMJH for serological and molecular typing. Animal trapping and use was approved by the Instituto Nacional de Recursos Naturales of Peru (INRENA) and the Institutional Animal Care and Use Committee, University of California San Diego.

Pulsed Field Gel Electrophoresis (PFGE) Characterization of Isolates

Agarose blocks containing leptospiral DNA were prepared and then digested with 30 units of NotI restriction enzyme (New England Biolabs, USA) for 2 hours at 37°C. Plug slices containing the digested DNA were placed in the wells of a 1% agarose gel and electrophoresed in a Bio-Rad CHEF-DRIII for 18 hours at 14°C with recirculating TBE
buffer. Initial and final switch times of 2.16 and 35.07s, respectively, were employed, and voltage was 6 V/cm. *Salmonella* serotype Braenderup H9812 was digested with 50 U XbaI (New England Biolabs) for use as a DNA size standard (Ribot, Fair et al. 2006). Gels were stained with ethidium bromide and then photographed under UV transillumination using the Gel Doc 2000 system (Bio-Rad). PFGE fingerprints were analyzed using the BioNumerics software package (Applied Maths, Belgium) and a database of PFGE profiles from reference strains and clinical isolates (Galloway and Levett, unpublished data). The Dice band-based coefficient was used for cluster analysis (Carrico, Pinto et al. 2005).

*Characterization of isolates by 16S rRNA Gene Sequencing*

Total genomic DNA was extracted from 7 day cultures (2 x 10^8 leptospires/mL) using the QIAamp DNA extraction kit (QIAGEN, USA). Initial PCR amplification was performed using the eubacterial rDNA primers fD1/rD1 as described previously for leptospiral 16S rRNA gene sequencing (Morey, Galloway et al. 2006). PCR products were purified using the Qiaquick PCR purification kit (QIAGEN, USA). Sequencing was performed on an ABI 3100 automated sequencer (Perkin Elmer, USA). Since the most distinguishing 16S sequence is found in the middle of the leptospiral 16S rRNA gene, base pairs from ~32 to 1355 were sequenced, using the following internal sequencing primers (Supplementary Figure 1, which includes a schematic of the PCR and sequencing primer locations and an example of one such sequence assembly): lepto16S11f, a 20 bp forward primer located at position 11 (5´- GGC GGC GCG TCT TAA ACA TGC - 3´); and lepto16S1388r, a 20 bp reverse primer located at position 1388, (5´-TGT GTA CAA CAA CAA GCA CCG GCG TCG TCG TCG TCG TG - 3´).
GGT CCG GGA AC - 3’). Additional internal sequencing was done using specific forward primers beginning at position 505 (5’- TCA TTG GGC GTA AAG GGT G – 3’) and position 1006 (5’ - TCA GCT CGT GTC GTG AGA TG – 3’). Reads of 650-700 bp were routinely obtained. 16S rRNA gene segments were sequenced 8 times in both directions. Given that informative sequence cannot include the PCR primers, 1355 bp of informative primary sequence was obtained for each isolate. Reaction conditions for cycle sequencing were according to manufacturer’s directions. Sequences were edited and assembled using the Staden Software Package (Staden, Beal et al. 2000). Edited sequences were aligned using ClustalW v. 1.83 (Thompson, Higgins et al. 1994) for Windows and a phylogram generated using MrBayes v3.1.2 (Rodriguez, Oliver et al. 1990; Ronquist and Huelsenbeck 2003) for Windows with 2 simultaneous runs for 3,000,000 generations. The Tamura-Nei (1993) model of nucleotide substitution with gamma distributed rates and invariant sites (TrN+I+G) (Rodriguez, Oliver et al. 1990).

**DNA-DNA Hybridization Analysis**

Subcultures in liquid PLM-5 medium were incubated at 30°C for 7 days. DNA was extracted and purified from strains VAR010T, CEH010, CEH011, CEH033, CEH044, CEH162, MMD735, L. interrogans RGA, L. breamii 5399T, L. fainei BUT6T and L. inadai LymeT as described previously (Rodriguez, Oliver et al. 1990)DNA from strain VAR 010T was labeled with [32P]dCTP (Brenner, McWhorter et al. 1982) and DNA relatedness and percentage divergence between the strains were determined by the hydroxyapatite method, with 55°C used for optimal reassociation.
The G + C content (mol%) was determined for strain VAR 010T by the thermal denaturation method using a Beckman DU Series spectrophotometer (Beckman Coulter, Fullerton, CA) (Mandel, Igambi et al. 1970). All samples were run at least three times, using DNA from *Escherichia coli* K-12 as a control.

**Determination of leptospiral serogroup**

Leptospiral isolates at a density of $2 \times 10^8$ cells/mL were used in microscopic agglutination reactions with reference rabbit anti-sera raised against the comprehensive panel of leptospiral serogroups shown in Table 2.1 (Faine 1982). Individual titers higher than 1:100 were considered significant and reported.

**Biological Characterization**

**Growth Characteristics**

Growth of the unknown leptospiral isolate was determined in the presence of 225 μg/mL 8-azaguanine (8-AZA) at 30°C (Perolat, Chappel et al. 1998). *L. interrogans* serovar Icterohaemorrhagiae strain HAI188 (Segura, Ganoza et al. 2005), *L. fainei* serovar Hurstbridge strain BUT6T, and *L. biflexa* serovar Patoc strain Patoc I were included as representative pathogenic, intermediate and saprophytic strains, respectively. Growth in liquid EMJH, without 8-AZA, was used as a positive control.
Table 2.1. Comprehensive Panel of Leptospiral Serogroup Antisera Used to Characterize "*Leptospira licerasiae*" strain VAR10

<table>
<thead>
<tr>
<th>Serogroup</th>
<th>Serovar</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australis</td>
<td>Australis</td>
<td>Ballico</td>
</tr>
<tr>
<td></td>
<td>Bratislava</td>
<td>Jež Bratislava 941</td>
</tr>
<tr>
<td></td>
<td>Peruviana</td>
<td></td>
</tr>
<tr>
<td>Autumnalis</td>
<td>Autumnalis</td>
<td>Akiyami A</td>
</tr>
<tr>
<td>Ballum</td>
<td>Ballum</td>
<td>Mus 127</td>
</tr>
<tr>
<td>Bataviae</td>
<td>Bataviae</td>
<td>Van Tienen</td>
</tr>
<tr>
<td></td>
<td>Rioja</td>
<td>MR 12</td>
</tr>
<tr>
<td>Canicola</td>
<td>Canicola</td>
<td>Hond Utrecht IV</td>
</tr>
<tr>
<td>Celledoni</td>
<td>Celledoni</td>
<td>Celledoni</td>
</tr>
<tr>
<td>Cynopteri</td>
<td>Cynopteri</td>
<td>3522 C</td>
</tr>
<tr>
<td></td>
<td>Tingomaria</td>
<td>M-13</td>
</tr>
<tr>
<td>Djasiman</td>
<td>Djasiman</td>
<td>Moskva V</td>
</tr>
<tr>
<td>Grippotyphosa</td>
<td>Grippotyphosa</td>
<td></td>
</tr>
<tr>
<td>Hebdomadis</td>
<td>Borincana</td>
<td>HS 622</td>
</tr>
<tr>
<td>Hurstbridge</td>
<td>Hurstbridge</td>
<td>BUT 6</td>
</tr>
<tr>
<td>Icterohaemorrhagiae</td>
<td>Copenhageni</td>
<td>M 20</td>
</tr>
<tr>
<td></td>
<td>Icterohaemorrhagiae</td>
<td>RGA</td>
</tr>
<tr>
<td></td>
<td>Lai</td>
<td>Lai</td>
</tr>
<tr>
<td></td>
<td>Mankarso</td>
<td>Mankarso</td>
</tr>
<tr>
<td>Javanica</td>
<td>Javanica</td>
<td>Veldrat Batavia 46 24</td>
</tr>
<tr>
<td></td>
<td>Vargonicas</td>
<td></td>
</tr>
<tr>
<td>Louisiana</td>
<td>Orleans</td>
<td>LSU 2580</td>
</tr>
<tr>
<td>Manhao</td>
<td>Manhao 4</td>
<td>Li 130</td>
</tr>
<tr>
<td>Mini</td>
<td>Georgia</td>
<td>LT 117</td>
</tr>
<tr>
<td></td>
<td>Ruparupae</td>
<td>M 3</td>
</tr>
<tr>
<td>Panama</td>
<td>Panama</td>
<td>CZ 214</td>
</tr>
<tr>
<td>Pomona</td>
<td>Pomona</td>
<td>Pomona</td>
</tr>
<tr>
<td>Pyrogenes</td>
<td>Bagua</td>
<td>MW 12</td>
</tr>
<tr>
<td></td>
<td>Cenepa</td>
<td>MW 2</td>
</tr>
<tr>
<td></td>
<td>Pyrogenes</td>
<td>Salinem</td>
</tr>
<tr>
<td>Ranarum</td>
<td>Evansi</td>
<td>267-1348</td>
</tr>
<tr>
<td>Sarmin</td>
<td>Machiguenga</td>
<td>MMD 3</td>
</tr>
<tr>
<td></td>
<td>Sarmin</td>
<td>Sarmin</td>
</tr>
<tr>
<td>Sejroe</td>
<td>Wolffii</td>
<td>3705</td>
</tr>
<tr>
<td>Shermanni</td>
<td>Babudieri</td>
<td>CI 40</td>
</tr>
<tr>
<td>Tarassovi</td>
<td>Tarassovi</td>
<td>Perepelitsin</td>
</tr>
<tr>
<td>Semaranga</td>
<td>Patoc</td>
<td>Patoc 1</td>
</tr>
</tbody>
</table>
**LipL32 PCR**

To assess the presence of a PCR-amplifiable LipL32 gene in “L. licerasiae”, we used a modified PCR procedure (Levett 2005), and used DNA from *L. interrogans* serovars Copenhageni strain L1-130 (Nascimento, Ko et al. 2004) and HAI188 (Segura, Ganoza et al. 2005), respectively, as positive controls. All amplifications were performed on the PTC-200 system (MJ Research, Bio-Rad, Hercules, CA). Five µL of genomic DNA was amplified using the following protocol: 95°C for 15 s for enzyme activation, followed by 40 cycles of 95°C for 45 s, 55°C for 30 s and 72°C for 60 s. The annealing temperature was decreased by 1°C per cycle for the first 5 cycles.

**LipL32 Western Blot**

Western blot analysis to detect LipL32 in leptospiral strains was performed using 2 x 10⁷ leptospires/well separated by SDS-PAGE and transferred to a nitrocellulose membrane. The strains studied included “L. licerasiae” serovar Varillal strain VAR 010ᵀ, *L. interrogans* serovar Icterohaemorrhagiae strain HAI188 (Segura, Ganoza et al. 2005), *L. fainei* serovar Hurstbridge strain BUT6ᵀ (ATCC BAA-1109ᵀ), as well as *L. broomii* (ATCC BAA-1107ᵀ and BAA-1108), *L. weilii* serovar Celledoni (ATCC 43285ᵀ), *L. wolbachii* serovar Codice (ATCC 43284ᵀ), *L. biflexa* serovar Patoc (ATCC 23482ᵀ), and *Turneriella parva* (ATCC BAA-1111ᵀ) were obtained from the American Type Culture Collection, Virginia, USA. 100 μl of cultures were taken directly from the glycerol stock and centrifuged at 14,000 rpm for 30 minutes. The pellets were washed three times with
PBS, suspended in 100 µl of 1X SDS sample buffer and incubated in a boiling water bath for 10 min. Ten µl of the SDS-solubilized whole bacterial cell lysate were loaded onto 4-12% Bis-Tris SDS polyacrylamide gels (Invitrogen, Carlsbad, USA) and transferred to nitrocellulose membrane. The blot was blocked in PBS containing 5% BSA, incubated in anti-LipL32 rabbit polyclonal antiserum (diluted 1:2000, 2 hr at 21°C; kindly provided by Dr. David Haake, University of California, Los Angeles) followed by 1 hr incubation with 1:3000 dilution of phosphatase-labeled anti-rabbit IgG (Kierkegaard and Perry Laboratories, Gaithersberg, Maryland), and development with BCIP/NBT (Kierkegaard and Perry Laboratories).

**Experimental challenge infections by Leptospira**

Outbred female Syrian Golden hamsters were obtained from Charles River Laboratories (Wilmington, MA). Animal experiments were approved by the University of California, San Diego Institutional Animal Care and Use Committee and were performed in Biosafety Level 2 animal facilities approved by the Association for Assessment and Accreditation of Laboratory Animal Care under approved biological safety procedures. “L. licerasiae” strain VAR 010T and 2 other isolates from Iquitos, Peru (L. interrogans serovar Icterohaemorrhagiae, strains HAI188 and HAI156) that were isolated from leptospirosis patients were used to infect hamsters (N=2 in each group). Organisms fixed with 10% formalin were counted using a Petroff-Hauser counting chamber using dark-field microscopy. Groups of hamsters were inoculated intraperitoneally with 10⁸ organisms for each *Leptospira* strain; one animal was injected with EMJH leptospiral culture medium alone as a negative control. The animals were observed twice daily for
clinical signs of disease (hunching, decrease in oral intake, diarrhea, lethargy). On day 3 following infection, one member of each group was sacrificed and the organs (lung, liver, and kidney) were removed aseptically to determine the bacterial load by real-time quantitative PCR (Segura, Ganoza et al. 2005). The remaining animal in each group was sacrificed if moribund, and the organs were harvested and processed similarly. Samples for PCR were stored in 70% ethanol at −80°C until needed.

**DNA Preparation and Real-Time qPCR**

Total DNA for qPCR was prepared from three different pieces of weighed tissue samples using the DNeasy tissue kit (QIAGEN, USA) according to the manufacturer’s directions. Real-time qPCR was performed using a previously described primer pair and probe (Smythe, Smith et al. 2002) labeled with the fluorescent reporter dye FAM (6-carboxyfluorescein) at the 5’ end, and the fluorescent quencher TAMRA (6-carboxytetramethylrhodamine) at the 3’ end. The PCR primers, Lepto F (5’-CCC GCG CCC GCG TCC GAT TAG-3’) and Lepto R (5’ TCC ATT GTG GCC GRA/G ACA C-3’), allow amplification of the region between the positions 171 and 258 of the rrs (16S) gene, with an expected product size of 87 bp. The FAM-TAMRA labeled probe [5’-CTCA CCA AGC TCA CCA AGG CGA CGA TCG GTA GC-3’] spans the region from position 205 to 228. Reaction mixes were prepared using the Platinum Quantitative PCR supermix-UDG kit (Invitrogen, Carlsbad, CA, USA) with final primer and probe concentrations of 600 nM and 100 nM, respectively, and 5 μl DNA extract. Reactions were performed in triplicate. Amplification and fluorescent monitoring were performed using a DNA Engine Opticon® 2 thermal cycler (MJ Research) using the following
protocol: Incubate 2 min at 50.0°C; incubate 2 min at 95.0°C; incubate 30 s at 94.0°C; incubate 1 min at 50.0°C; plate read; repeat steps 2–5 for 44 more cycles.

To generate a standard curve, 13 mg of uninfected hamster kidney was spiked with $10^8$ leptospires, extracted as described above, and used to prepare a 10-fold dilution series for real-time qPCR. The tissue burden of *Leptospira* for each sample was quantified by interpolating threshold cycle (Ct) values against the standard curve. Samples with a C<sub>t</sub> value >40 were considered negative.

*LigA Southern Blot*

A dioxigenin (DIG)-labeled 508 bp probe for detection of the pathogenic marker ligA gene was synthesized by PCR, using the following forward primer, 5´ - CAA AGT TGT ATG TCT TGG CCA C 3´ and reverse primer, 5´ - GGA AGA CCA AAC GAT CAG TGG - 3´. DNA from *L. interrogans* serovar Icterohaemorrhagiae strain HAI0188 was used as template. The PCR cycling profile consisted of 40 cycles of 95°C for 30s; 49°C for 30s; 72°C for 40s; and a final extension of 72°C for 7 min. A 16S rRNA gene probe to be used as a control was generated using primers lepto16S1006f, 5´ - TCA GCT CGT CGT GTC GTG AGA TG - 3´, designed from aligned leptospiral 16S sequences retrieved from GenBank, and rD1, 5’- AAG GAG GTG ATC CAG CC - 3´ (Weisburg, Barns et al. 1991). Genomic DNA was extracted from strain HAI188, “*L. licerasiae*” strain VAR 010<sup>T</sup>, *L. fainei* serovar Hurstbridge strain BUT6<sup>T</sup>, and *L. biflexa* serovar Patoc I<sup>T</sup> using the DNeasy Tissue Kit (Qiagen, USA), and was then digested with *Bam*HI (New England Biolabs, USA) according to manufacturer’s directions. Hybridization was carried out at 42°C. The membrane was washed with 2x SSC at room temperature and
0.1x SSC at 42°C. Bands were detected using anti-DIG-alkaline phosphatase Fab fragments (Roche, USA) and CDP-Star chemiluminescence substrate (Roche, USA).

2.3 Results

Patient Isolates of Leptospira

Of 881 patients presenting with a history of undifferentiated fever to a study site, 45 patients’ blood cultures yielded leptospires with typical morphology and motility as visualized under darkfield microscopy. Two of these leptospiral isolates from humans, obtained from blood cultures and identified as novel based on results presented below, were studied further. None of the 881 patients had urine cultures positive for this novel leptospire.

Case Descriptions of Patients with Novel Leptospires

Patient VAR10

A 31 year old female food vendor presented at the Varillal health post complaining of 2 days of fever, malaise, chills, headache and dizziness. She denied having gastrointestinal or urinary symptoms. The physical exam was unremarkable. The blood smear was negative for malaria. A blood sample for serology and culture was taken and the patient was sent home with antipyretics. Illness resolved after 5 days without any further treatment or complications.
**Patient HAI029**

A 19 year-old female student/domestic worker presented at the Hospital de Apoyo in Iquitos with a 5-day history of fever, malaise, headache, dizziness, chills, leg pain and weakness, abdominal pain, anorexia, nausea, and vomiting. A blood sample for serology and culture was taken during this hospital visit. Her illness spontaneously resolved with no complications.

**Follow Up**

Neither patients VAR10 nor HAI029 received antibiotic treatment. At 5 week follow up, all signs and symptoms of infection had completely resolved in both patients. Blood cultures from both patients were positive for Leptospira at 3 and 2 weeks after inoculation, respectively. The characterization of these leptospiral isolates as a new species and unique antigenic type is described below. The isolate from patient VAR10 (strain VAR 010T) has been deposited at the American Type Culture Collection as ATCC BAA-1110T, at the U.S. National Veterinary Services Laboratory, Ames, Iowa, USA, and at the WHO/FAO/OIE Collaborating Centre for Reference and Research on Leptospirosis, Australia and Western Pacific Region, Brisbane, Australia.

Serological results for patient VAR10 showed negative ELISA IgM serology on both acute and convalescent samples. MAT using the standard live *Leptospira* panel was negative. When the isolate from patient VAR10 was used as MAT antigen, the acute sample was negative but the convalescent sample (taken 42 days after the acute sample) was positive at a titer of 1:400.
The acute serum sample of patient HAI029 was negative for anti-leptospiral antibodies but was IgM positive by ELISA on a convalescent sample taken 31 days after the acute sample. The acute serum sample of patient HAI029 was negative by MAT by both the standard leptospiral panel plus “L. licerasiae” strain VAR 010T but convalescent samples reacted to serogroups Canicola, Icterohaemorrhagiae, Australis, and Sejroe at a titer of 1:1600, to serogroup Mini at a titer of 1:3200, and had a titer of 1:6400 against her own isolate subsequently identified as “L. licerasiae” serovar Varillal, and identical to strain VAR 010T as determined by PFGE and 16S rRNA gene sequencing (see below).

Characterization of Isolates

PFGE analysis was performed on the VAR 010T and HAI029 human isolates and the 8 non-serologically-typeable isolates from rats. PFGE fingerprints were compared to PFGE fingerprints from 206 other pathogenic, intermediate (all those included in this paper) and saprophytic serovars (Galloway and Levett, data not published). These rat and human isolates shared a previously undescribed fingerprint pattern (Figure 1.1).

16S rRNA Gene Sequencing

16S rRNA gene fragments of ~1.5 kb were amplified from genomic DNA extracted from all isolates in the study using the universal eubacterial primers fD1/rD1 (Weisburg, Barns et al. 1991). The internal primers lepto16S505f and lepto16S1006f were designed from consensus regions of published leptospiral 16S rRNA gene sequences and used to sequence an internal ~1.3 kb portion of the fD1/rD1 fragment.
The sequences of the 16S rRNA fragment from all strains with the new PFGE pattern were identical. Phylogenetic analysis revealed that these strains were more closely related to the intermediate leptospiral species *L. fainei* and *L. inadai* (Figure 2.2) than to the more pathogenic *Leptospira interrogans* group (Ganoza, Matthias et al. 2006).
Figure 2.1. Pulsed field gel electrophoresis analysis of leptospiral isolates obtained from rats and humans in the region of Iquitos, Peru. Indicated in parentheses is animal source of leptospiral isolate followed by location of trapping (see methods).

Figure 2.2. Phylogram of leptospiral 16S rRNA gene sequences generated by Bayesian phylogenetic analysis with simultaneous runs of 3,000,000 generations.
Serotyping

The comprehensive panel of reference rabbit anti-serogroup polyclonal antisera used in this study failed to agglutinate the leptospiral strains isolated from patients VAR10 and HAI029 and the 8 rat isolates in the MAT. “L. licerasiae” serovar Varillal strain \( V010^T \) was agglutinated by antisera to serovar Hurstbridge at a titer of 1:100 but by no other anti-serogroup antisera. Conversely, no other serogroup was agglutinated by the reference rabbit anti-serum raised against “L. licerasiae” serovar Varillal strain \( V010^T \). Because of the lack of significant seroreactivity of reference serogroup antisera against “L. licerasiae” serovar Varillal strain \( V010^T \), the cross-agglutination absorption test (CAAT) was not carried out. A similar approach was used to designate the Hurstbridge serovar of L. fainei serovar (Perolat, Chappel et al. 1998). These serotyping results were independently validated at the WHO/FAO/OIE Collaborating Centre for Reference and Research on Leptospirosis, Australia and Western Pacific Region (Dr. Lee Smythe, Supplementary Table 2). Rabbit anti-serum to “L. licerasiae” serovar Varillal strain \( V010^T \) (available from the National Veterinary Services Laboratory, Ames, Iowa) agglutinated the leptospires from patients VAR10 and HAI029, and the eight rat isolates, at a titer of 1:51,200. These serological results conclusively demonstrate that the two human and eight rat leptospires represent a new serogroup and a new serovar.
**DNA-DNA Hybridization**

Because leptospiral strains VAR 010\(^T\), CEHO10, CEHO11, CEHO33, CEHO44, and CEH162 grouped with the intermediate leptospires by 16S rRNA phylogenetic analysis, DNA-DNA hybridization was only carried out on the other intermediates *L. broomii* 5399\(^T\), *L. fainei* BUT6\(^T\) and *L. inadai* Lyme\(^T\) as well as *L. interrogans* RGA\(^T\) as an outgroup. As shown by DNA-DNA hybridization analysis (Table 2.2), leptospiral strains VAR 010\(^T\), CEHO10, CEHO11, CEHO33, CEHO44, and CEH162 showed no significant relatedness to *L. interrogans* RGA\(^T\), *L. broomii* 5399\(^T\), *L. fainei* BUT6\(^T\) or *L. inadai* Lyme\(^T\). However, there was strong relatedness between the strains VAR 010\(^T\), CEHO10, CEHO11, CEHO33, CEHO44, CEH162 and MMD735. These strains meet the criteria for the molecular definition of a species (Brenner, Kaufmann et al. 1999). The G + C content of *L. licerasiae* strain VAR 010\(^T\) was 43.9 mol%, within the range reported for other *Leptospira* species (Yasuda, Steigerwalt et al. 1987).

**Biological characterization**

To determine whether “*L. licerasiae*” serovar Varillal strain VAR 010\(^T\) had growth characteristics more typical of pathogenic or saprophytic *Leptospira*, growth in the presence of 8-azaguanine, a classic test to differentiate pathogenic from saprophytic leptospires (Perolat, Chappel et al. 1998), was performed. “*L. licerasiae*” serovar Varillal strain VAR 010\(^T\), *L. interrogans* serovar Icterohaemorrhagiae strain HAI188, and *L. fainei* serovar Hurstbridge strain BUT6\(^T\) failed to grow in the presence of 8-AZA after 4 weeks incubation at 30\(^\circ\)C; as a positive control, the saprophytic strain, *L. biflexa* strain Patoc I\(^T\), grew well in the presence of 8-azaguanine.
Table 2.2. DNA relatedness of “Leptospira licerasiae” strain VAR010\textsuperscript{T} to *Leptospira* species, *L. broomii* 5399\textsuperscript{T}, *L. fainei* BUT6\textsuperscript{T}, *L. inadai* Lyme\textsuperscript{T} and *L. interrogans* RGA\textsuperscript{T}.

<table>
<thead>
<tr>
<th>Source of unlabelled DNA</th>
<th>Results of reaction with labeled DNA from strain VAR010\textsuperscript{T}</th>
<th>RBR\textsuperscript{a}</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR010\textsuperscript{T}</td>
<td>100</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>CEH010</td>
<td>95</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>CEH011</td>
<td>92</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>CEH033</td>
<td>98</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>CEH044</td>
<td>91</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>CEH162</td>
<td>100</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>MMD735</td>
<td>96</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td><em>L. broomii</em> 5399\textsuperscript{T}</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. fainei</em> BUT6\textsuperscript{T}</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. inadai</em> Lyme\textsuperscript{T}</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. interrogans</em> RGA\textsuperscript{T}</td>
<td>34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}RBR, relative binding ratio; D, percent divergence. Reactions were performed at 55°C.
Southern blotting and PCR were performed to determine whether the LigA gene, encoding the putative virulence factor Lig A found in L. interrogans, might be present in “L. licerasiae” serovar Varillal strain VAR 010.T. PCR using 4 pairs of primers derived from L. interrogans serovar Copenhageni failed to produce a LigA band in “L. licerasiae” serovar Varillal strain VAR 010.T. Southern blot analysis for Lig A showed the expected bands in a strain of L. interrogans serovar Icterohaemorrhagiae strain HAI188, as expected, but failed to detect Lig A in “L. licerasiae” serovar Varillal strain VAR 010.T or in L. fainei serovar Hurstbridge strain BUT6T and L. biflexa serovar Patoc strain Patoc I.T (Figure 2.3).

Demonstration of a LipL32-related protein in “L. licerasiae”, L. fainei and L. biflexa

The presence of the lipoprotein lipL32 gene has been considered characteristic of pathogenic leptospires (Haake, Chao et al. 2000). PCR, using published primers (Haake, Chao et al. 2000; Levett 2005) to amplify LipL32, detected the expected product only in a pathogenic L. interrogans serovar Icterohaemorrhagiae strain HAI188, but not in “L. licerasiae” serovar Varillal strain VAR 010.T, HAI029, the rat-derived “L. licerasiae” strains, L. fainei or L. biflexa. However, Southern blotting revealed that in addition to strain HAI188, both L. fainei serovar Hurstbridge strain BUT6T and “L. licerasiae” serovar Varillal strain VAR 010.T, but not L. biflexa, had bands that hybridized to the L. kirschneri-derived LipL32 probe (data not shown).

Because of the potential for the phylogenetically distant Leptospira to have sufficiently diverged so that PCR amplification may have failed because of primer mismatch, we determined whether a LipL32 cross-reactive protein might be present in
Figure 2.3. Southern blot to determine the presence of LigA in “L. licerasiae” serovar Varillal. Lane 1, DIG-labeled marker; lane 2, *L. interrogans* serovar Icterohaemorrhagiae strain HAI188 (positive control); lane 3, “L. licerasiae” serovar Varillal; lane 4, *L. biflexa* serovar Patoc I^T^.
“L. licerasiae” serovar Varillal strain VAR 010ᵀ by Western immunoblot using a rabbit anti-L. kirschneri serovar Grippotyphosa LipL32 polyclonal antiserum. As expected, a protein of ~32 kDa was seen in L. interrogans serovar Icterohaemorrhagiae strain HAI188 (Figure 2.4). Surprisingly, a single, well-defined protein of ~27 kDa, less than the expected molecular mass of this protein, was detected in “L. licerasiae” strain VAR 010ᵀ. Western blot analysis showed a 32 kDa band that co-migrated with the L. interrogans LipL32 in L. broomii, L. weili, and L. fainei but failed to demonstrate any band in L. wolbachii, L. biflexa, and Turneriella parva (data not shown but provided for review). The unique size of the LipL32-cross reactive protein in “L. licerasiae” serovar Varillal strain VAR 010ᵀ supports lack of contamination of this culture by another LipL32-containing leptospire. To further rule out the possibility that the cultures may have been contaminated by a known pathogenic Leptospira known to express LipL32, serological typing and 16S rRNA gene sequencing were repeated on all cultures, which confirmed their expected identities (data not shown).

**Experimental animal infections**

Both L. interrogans serovar Canicola strain HAI156 and L. interrogans serovar Icterohaemorrhagiae strain HAI188 caused severe disease in hamsters infected intraperitoneally with 10⁸ leptospires. HAI156- and HAI188- infected hamsters were sick on day 3 following challenge and moribund by day 5. In contrast, hamsters infected with “L. licerasiae” serovar Varillal strain VAR 010ᵀ showed no sign of illness (data not shown).
Figure 2.4. Western immunoblot of *Leptospira interrogans* serovar Copenhageni strain L1-130, “*L. licerasiae*” serovar Varillal, and *L. fainei* serovar Hurstbridge using rabbit polyclonal antisera to recombinant LipL32 of *L. kirschneri* serovar Grippotyphosa.
Quantitative real time PCR detected high levels of leptospires in organs of hamsters infected with HAI156- and HAI188, but leptospires were nearly completely eliminated by day 3 after infection in liver, lungs and kidneys of hamsters infected with “L. licerasiae” serovar Varillal strain VAR 010T (Fig 1.6), showing a major difference in virulence between these leptospiral species. A lack of symptomatic infection was found with experimental infection of more than 50 additional hamsters, as well as guinea pigs and SCID mice, with “L. licerasiae” serovar Varillal strain VAR 010T (Figure 2.5).

Prevalence of Leptospira licerasiae serovar Varillal seropositivity in the Iquitos region

During the study period, 1831 consecutive febrile patients were enrolled. Within these 1831 febrile patients on the data (means, including those with > 2 weeks of febrile illness and one sample only), 881 had a second serum sample available between 10 and 70 days after the first sample. Of these, 516 (58.6%) met criteria for acute leptospirosis. Of these, 367 (41%) reacted to “L. licerasiae” serovar Varillal strain VAR 010T only or had mixed reactions with “L. licerasiae” serovar Varillal strain VAR 010T and other serovars (155, 18%) with diagnostic titers highest against “L. licerasiae” serovar Varillal strain VAR 010T (Fig 1.7). The median percentage of febrile patients seropositive for “L. licerasiae” serovar Varillal strain VAR 010T was 29% and the interquartile range was 23-36%. A single high MAT titer against “L. licerasiae” serovar Varillal strain VAR 010T (≥1:800) was found in 40 patients in the acute sample, 57 in the second sample, and 16 patients had a titer of ≥1:800 in both (Figure 2.6).
Figure 2.5. Real time quantitative PCR analysis of experimental leptospiral infections of hamsters. HAI188 and HAI156, strains *L. interrogans* serogroups Icterohaemorrhagiae and Canicola isolated from patients in Iquitos, Peru. VAR10, “*L. licerasiae*” serovar Varillal strain VAR 010\(^T\). HAI188 and HAI156 caused a severe moribund state at days 4-5; none of the animals with VAR 010\(^T\) exhibited any signs of illness. Three 25 mg samples of each tissue were analyzed and error bars indicate the standard deviations of these three samples per tissue.
Figure 2.6 Seroprevalence of “Leptospira licerasiae” serovar Varillal in acute febrile patients in Iquitos, Peruvian Amazon (n = 881). Var 10 = “Leptospira licerasiae” serovar Varillal strain VAR 010r.
Apart from the high rate of “L. licerasiae” serovar Varillal seroreactivity in the acute febrile population in Iquitos, we found serological evidence of seroreactivity in sera from 11 distinct geographic locations in Peru (Table 3). Though seroreactivity was not as common (22/344; 6.7 % of seropositives) as in Iquitos, this finding does illustrate the widespread distribution of seroreactivity in Peru.

**Specificity of MAT for anti-Leptospira licerasiae serovar Varillal antibodies**

Due to the high frequency and titer of antibodies to “L. licerasiae” serovar Varillal strain VAR 010T, there was concern about the possibility that this leptospire might be cross-reactive with other organisms or that humans might have natural antibodies to this leptospire, so that seropositivity would be spurious and falsely positive. Fifty randomly collected, de-identified sera collected from inpatients at UCSD Medical Center were tested for antibodies against “L. licerasiae” serovar Varillal strain VAR 010T. There was no agglutination observed. We tested 180 sera collected from a serosurvey of healthy subjects in the north Lima town of Puente Piedra; of these, 2 had titers of 1:50, the rest being negative (Table 2.3).
Table 2.3. *Leptospira licerasiae* serovar Varillal Seroreactivity in Acute Leptospirosis Patients from Different Regions of Peru

<table>
<thead>
<tr>
<th>Region of Peru</th>
<th>Number of Febrile Patients Studied</th>
<th>Number Diagnosed with Acute Leptospirosis</th>
<th>Number Seropositive against <em>L. licerasiae</em> serovar Varillal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cajamarca-Jaen</td>
<td>180</td>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>Amazonas- Bagua</td>
<td>262</td>
<td>103</td>
<td>0</td>
</tr>
<tr>
<td>Huanuco</td>
<td>15</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Piura</td>
<td>16</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Huanuco-Tingo Maria</td>
<td>32</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Cusco</td>
<td>10</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>San Martin</td>
<td>21</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Ayacucho-San Francisco</td>
<td>162</td>
<td>97</td>
<td>15</td>
</tr>
<tr>
<td>Junin</td>
<td>7</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Ucayali</td>
<td>128</td>
<td>40</td>
<td>6</td>
</tr>
<tr>
<td>Lima Norte</td>
<td>41</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>Totals</td>
<td>874</td>
<td>344</td>
<td>23</td>
</tr>
</tbody>
</table>
2.4 Discussion

Here we report isolation of a new species of *Leptospira* with novel biological characteristics that causes in humans a non-specific syndrome of undifferentiated fever. We showed definitively through serological and molecular analysis using 16S rRNA gene sequencing and pulsed field gel electrophoresis that this new leptospire, provisionally named “*Leptospira licerasiae*” serovar Varillal strain VAR 010\(^T\), is antigenically unique, is a significant cause of acute leptospirosis in the Peruvian Amazon region of Iquitos, and has a *Rattus* reservoir. Recognition of “*Leptospira licerasiae*” serovar Varillal strain VAR 010\(^T\) as a new serovar is supported by the lack of agglutination of this strain by any serogroup reference serum and the lack of reactivity of anti- VAR 010\(^T\) serum raised in rabbits against the serovars of *Leptospira* strains representing the comprehensive panel of leptospiral serogroups. A similar situation was found with *L. fainei* serovar Hurstbridge, where the following evidence was adduced in support of this novel serovar: lack of significant cross-agglutination was observed with reference antisera representing the 24 pathogenic serogroups and the main saprophytic ones; lack of agglutination by antiserum raised against one of the strains against any serogroup (Perolat, Chappel et al. 1998). The serological characterization of the new serovar, Varillal, was conducted in two laboratories, one of which was the WHO/FAO/OIE Collaborating Centre for Reference and Research on Leptospirosis, Australia and Western Pacific Region, fulfilling the requirements for recognition of new serovars by the International Committee on Systematics of Prokaryotes, Subcommittee on the Taxonomy of *Leptospiraceae* (Stallman 1982).
The genus *Leptospira* presently consists of 13 named species and 4 un-named genomospecies (Perolat, Chappel et al. 1998),(Brenner, Kaufmann et al. 1999; Levett, Morey et al. 2005; Levett, Morey et al. 2006). Phylogenetic analysis reveals three clades, representing species that contain pathogenic serovars, non-pathogenic serovars and an intermediate group (Perolat, Chappel et al. 1998). The latter clade comprises three species, *Leptospira broomii*, *Leptospira inadai* and *Leptospira fainei* (Yasuda, Steigerwalt et al. 1987; Perolat, Chappel et al. 1998; Levett, Morey et al. 2006). Based on phylogenetic analysis, *L. licerasiae* is classified as an intermediate leptospiral species. Nonetheless, “*L. licerasiae*” serovar Varillal strain VAR 010<sup>T</sup> shares properties with pathogenic *Leptospira* such as sensitivity to 8-azaguanine, has a LipL32-related protein as revealed by Western and Southern blots, but does not appear to contain a *LigA*-related gene as determined by Southern blot. In contrast to *L. interrogans*, “*L. licerasiae*” serovar Varillal strain VAR 010<sup>T</sup> grew rapidly (similar to *L. biflexa*), did not grow significantly in vivo, and did not cause observable disease in experimentally infected animal. These observations suggest important biological and virulence differences between pathogenic and intermediate *Leptospira*.

Leptospirosis is typically thought of as an occupational disease originating from contact with water, soil or vegetation contaminated with the infected urine of carrier animals. The literature in recent years has shown that in under-developed areas of the world it is associated with environmental exposure during activities of daily living (Bharti, Nally et al. 2003). Neither patient had an occupation that would be considered a risk factor for leptospirosis. Patient A in this study was a food vendor and had contact with obvious risk factors having frequented a market area (Belen) with poor sanitation
and a high density of rats, and bathed in a natural pool, as there is no running water in her village. Patient B was a female student/domestic worker who lived in the city, did not frequent Belen (the urban slum area of Iquitos) and did not engage in other behavior that would place her at particular risk for contracting leptospirosis. However, both patients recalled seeing rats in and around their homes. Patient B did raise dogs and chickens, and the dogs urinated within the house. There were no established social or professional links between the patients, and their infections occurred in different places and times. Both cases presented with a mild, self-resolving febrile illness without secondary complications and show the ubiquity of exposure as part of the activities of daily living in this region.

Both patients initially had negative MAT and IgM results in their acute serum sample. While MAT of convalescent serum from Patient B was initially positive to a variety of serogroups, both acute and convalescent sera from Patient A were negative. However, when the test was repeated with the patient’s own isolate, Patient A was found to have circulating leptospiral antibodies. This pattern of leptospiral seroreactivity, known to be a common problem in the diagnosis of leptospirosis,\(^2\) underscores the importance of including region-specific leptospiral isolates in the panel of strains used in MAT for diagnosing leptospirosis and determining its true burden.

A curious finding in this study is that isolation of “\(L.\ licerasiae\)” serovar Varillal from humans was rare, only being obtained from 2 of 881 febrile patients, despite the far higher seroprevalence rate of antibodies to this serovar. Some might raise the concern that this rare isolation rate could reflect a laboratory contamination with this leptospiral species. We believe this possibility is unlikely for two reasons. First, the patients from
whom these isolates were obtained seroconverted to this novel serovar: patient 1 seroconverted to *L. licerasiae* serovar VAR 010\textsuperscript{T} but to no other leptospiral antigen by MAT, while patient 2 seroconverted with the highest titer against her own isolate of *L. licerasiae* at a titer higher than other leptospires. Second, we never obtained an isolate of *L. licerasiae* in any other culture of human or animal specimens other than those reported here, making the possibility of contaminated culture medium negligible. While the biological basis for the rare isolation of *L. licerasiae* remains speculative, we propose two hypotheses. First, it is possible that the two patients from whom this leptospire was isolated had an undefined, undetermined genetic predisposition that led to higher leptospiremia or failure to clear this relatively less virulent leptospire after exposure. Second, it is possible that varying degrees of heterologous, cross-reacting, anti-leptospiral immunity exist in the study population. This latter hypothesis is supported by the very high prevalence of anti-leptospiral antibodies in the Iquitos region, likely due to ubiquity of leptosporal exposure. It may be that these two patients, for some reason, never had been exposed to *Leptospira*, and thus were immunologically naïve and thus predisposed to a higher level of leptospiral bacteremia. Further prospective, population-based studies are needed to address these important questions.

This prospective study of acute febrile illness in Peru has shown that “*L. licerasiae*” is an important cause of fever in the Iquitos area and its surroundings, as evidenced by the number of patient sera that reacted predominantly or solely with serovar Varillal (298/425; 70\%). Isolation of “*L. licerasiae*” from rats suggests that this leptospiral species has at least *Rattus* spp. as a major reservoir host; we have not found this leptospiral species in other rodent, bat and marsupial species in the Peruvian Amazon.
Domestic rats are common in Iquitos: *R. norvegicus* and *R. rattus* are closely associated with human settlements in the area. *R. norvegicus* is more often encountered in urbanized areas, while *R. rattus* is the predominant rural species (data not shown). Six isolates identical to those isolated from both patients were recovered from rats caught in Belen, a city slum where sanitation is poor, rats are common and the risk of transmission to man is high. A further two VAR 010\(^\top\)-related isolates were recovered from rural rats. The isolation of a strain common to rats and found in 2 clinical cases, the ubiquity of the rat in Iquitos and the poor sanitation in most areas make the rat the likely source of leptospires in Iquitos. In other studies, we have succeeded in isolating *L. interrogans* serovar Icterohaemorrhagiae from 22 to 48% of peri-domestic *Rattus* species in villages near to and within urban areas of Iquitos (unpublished observations).

Molecular and serological analysis of human- and rat-derived strains revealed that they comprise a single novel leptospiral species and serovar. The fact that the PFGE fingerprint patterns were found to be identical to each other, but did not match any of the patterns in the CDC database (P.N. Levett and R. Galloway, unpublished data) supports our contention that the strains are novel leptospires. The isolates are members of a new serovar and serogroup as none were agglutinated by any of the reference anti-sera in our panel, although they had trace reactions to serogroup Hurstbridge. Given the lack of reactivity to the leptospiral serogroups represented by the rabbit reference sera used in the present study, the reference serological test, the cross-absorption agglutination test, was not necessary to define Varillal as a new serovar or antigenic type, similar to what was found for *L. fainei* serovar Hurstbridge (Perolat, Chappel et al. 1998)]. Phylogenetic
analysis of 16S rRNA gene sequence demonstrated that the strains comprised a homogenous genetic group separate from all other described leptospiral species. These strains, much like the recently described *L. broomii* (Levett, Morey et al. 2006), *L. fainei* serovar Hurstbridge (Perolat, Chappel et al. 1998) and *L. inadai* serovar Lyme (Schmid, Steere et al. 1986), are intermediate between the two larger saprophytic and pathogenic groups of *Leptospira* and, as such, share characteristics similar to both pathogenic and saprophytic leptospires. DNA-DNA hybridization further confirmed that *L. licerasiae* is a new *Leptospira* species. Our logic was similar to that taken in using DNA-DNA hybridization to further confirm *L. broomii* as a new *Leptospira* species (Levett, Morey et al. 2006). Because 16S rRNA gene sequencing places *L. licerasiae* into the intermediate *Leptospira* group close to *L. fainei*, *L. broomii*, etc., the only relevant DNA-DNA hybridization analysis is to differentiate the closest known clade partners identified by 16S rRNA gene sequence, so as to be able to confirm the distinctness of these 16S-rRNA gene-defined intermediate *Leptospira* species. The DNA-DNA hybridization analysis reported here does indeed confirm the distinctness of *L. licerasiae* from the other known intermediate *Leptospira*.

We report the existence of a new species of leptospire, provisionally named “*Leptospira licerasiae*” serovar Varillal, which causes acute leptospirosis in the Peruvian Amazon. We have proposed this name to recognize the contribution of Professor Julia Liceras de Hidalgo who obtained the first leptospiral isolates in Peru (Liceras de Hidalgo and Hidalgo R 1968; Liceras de Hidalgo and Hidalgo 1970; Liceras de Hidalgo, Hidalgo et al. 1971; Liceras de Hidalgo 1975; Liceras de Hidalgo 1981; Liceras de Hidalgo, Hidalgo et al. 1981; Liceras de Hidalgo and Mejia 1981). We propose a new serogroup, Iquitos,
based on the lack of agglutination with a comprehensive panel of reference antisera comprised of all serogroups except for Lyme and Sehgali (which in the case of serovar Lyme had cross-reaction with serovar Celledoni at a titer of 1:400 (Schmid, Steere et al. 1986), and in the case of Sehgali had a broad level of cross-reactivity 25 serovars and 12 serogroups ranging from titers of 1:80 to 1:1280 (Vijayachari, Hartskeerl et al. 2004).

Based on serological data that take advantage of its antigenic uniqueness, “Leptospira licerasiae” serovar Varillal appears to be an important cause of leptospirosis in the Peruvian Amazon region, but is uncommon elsewhere in Peru. The peridomestic rat is likely the major reservoir of this new species. Elucidation of virulence differences between pathogenic and intermediate leptospires will provide insight into leptospiral evolution and disease mechanisms, and may contribute to the control and amelioration of leptospirosis in the developing world.

**Note on Taxonomy:**

To fulfill the rules of the International Code of Nomenclature of Bacteria (Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, Clark WA. International code of nomenclature of bacteria (1990 revision). Washington, DC: American Society for Microbiology, 1992) we provide the following description of the novel species identified in this report.

Description of *Leptospira licerasiae* sp. nov. *Leptospira licerasiae* (li.ce.ra' si.ae. N.L. fem gen. n. licerasiae of Liceras, to honor Dr. Julia Liceras de Hidalgo, who obtained the first leptospiral isolates in Peru). Isolated from the blood of human patients...
with febrile illness and from kidneys of rats in Peru. Morphology is as described previously for the genus (Yanagihara Y 1984; Brenner, Kaufmann et al. 1999). The G + C ratio is 43.9 mol%. The type strain is VAR 010\textsuperscript{T} (=ATCC BAA-1110\textsuperscript{T} = WPR VAR 010\textsuperscript{T}), and has been deposited at the American Type Culture Collection, Manassas, Virginia, the National Veterinary Services Laboratory, U.S. Department of Agriculture, Ames, Iowa, and the WHO/FAO/OIE Collaborating Centre for Reference and Research on Leptospirosis, Australia and Western Pacific Region.

Database deposition

The 16S rRNA sequences for the Leptospira licerasiae isolates reported in this paper have been deposited in GenBank with the following accession numbers (strain, accession number): CEH006, EF612278; CEH011, EF612279; CEH033, EF612280; CEH044, EF612281; CEH162, EF612282; MMD735, EF612283; VAR 010\textsuperscript{T}, EF612284; CEH010, EF612285; CEH174, EF612286; MMD835, EF612287; HAI029, EF612288.

Chapter 2, in full, is a reprint of the material as it appears in PLOS Neglected Diseases, 2008, Ricaldi, JN, Matthias MA, Cespedes M, Diaz, MM, Galloway RL, Saito M, Steigerwalt, AG, Patra, KP, Vidal Ore C, Gotuzzo E, Gilman RH, Levett PN, Vinetz JM. The dissertation author was the co-primary investigator and author of this paper.
Chapter 3
Genomic features of *L. licerasiae*

3.1 Introduction

The genetic basis of leptospiral pathogenicity remains unclear. The lack of genetic manipulation tools for pathogenic leptospires has made it difficult to confirm the role of potential virulence factors (Adler and Faine 2006). Previously suggested virulence factors for leptospiral infection include lipoproteins, adhesins and hemolysins (Merien, Truccolo et al. 2000; Lee SH 2002; Cullen 2004; Ristow, Bourhy et al. 2007). However, few experimental data are available to independently confirm their role.

Only recently has random mutagenesis been described in a saprophyte species of *Leptospira* and it has allowed experimental investigation into the role of one of the leptospiral surface lipoproteins in infection (Ristow, Bourhy et al. 2007). Still, the process of identifying individual potential virulence factors from a library is labor intensive, and not guaranteed to contain all potential genes involved in virulence.

An alternative approach has been used in other microorganisms. This approach uses the annotated genomes of different species, both pathogenic and non pathogenic for identifying potential virulence factors (Mokady, Gophna et al. 2005). Currently, six annotated leptospiral genomes are available.

Recently developed sequencing methods have decreased the time and cost necessary to obtain complete genomes. This technology (pyrosequencing) has allowed us
to construct a draft genome of a newly discovered intermediate species of *Leptospira: Leptospira licerasiae* (Ricaldi, Matthias et al. 2008). The sequencing of this fast growing leptospire will not only allow us to compare the genetic differences between pathogen, intermediate and saprophytic leptospires, but will allow us to determine pathways associated with virulence and survival in the environment using a systems biology approach. Until now the wide variability among the leptospiral genomes has made it difficult to differentiate potential virulence factors from factors needed to survive in different environments, the genome of this intermediate species will allow for clearer comparisons and understanding.

Understanding the interactions of how genes interact to product virulence at a complete genomic level is now possible with a whole genome analysis and allows us to use a system biology approach to understand these processes.

### 3.2 Methods

**Bacterial strains**

A new species of *Leptospira, Leptospira licerasiae* serovar Varillal strain V10 was sequenced. This new species was isolated from a patient in the Iquitos area of the Peruvian Amazon (Ricaldi, Matthias et al. 2008). The strain had been serially passaged in the laboratory. Its identity was confirmed previous to genome sequencing by Sanger sequencing of the 16s gene.

One liter of culture was grown to a density of approximately $10^8$ organisms/mL. DNA was extracted using Trizol (Invitrogen Life Technologies). Briefly, 1 L of log phase culture was pelleted by centrifugation. Cells were lysed using Trizol. After the
sample was incubated at room temperature for 5 minutes, 0.2 mL of chloroform /each mL of Trizol were added, followed by vigorous shaking and a 3 minute incubation at room temperature. The sample was then centrifuged at 12000 g x 15 minutes at 4 C. This allowed for phase separation and removal of the aqueous phase. The DNA was then precipitated using ethanol.

**Sequencing**

*L. licerasiae* genomic DNA was sequenced using 454 pyrosequencing technology. A Genome Sequencer FLX System and GS FLX Titanium® reagents were used. Whole genomic DNA was sheared to generate random libraries that underwent emulsion-based clonal amplification. Briefly, DNA fragments of about 300 bp are linked into microbeads using a 1:1 ratio. PCR microreactors are assembled from micelles forming an emulsion that captures the microbeads. A fiber optic slide (PicoTiterPlate, Roche) is coated with the microbeads (one bead per well) and then exposed to DNA nucleotides one at a time. A light signal is generated by the firefly enzyme luciferase when a nucleotide is incorporated into a DNA strand. The light signal intensity is proportional to the number of nucleotides incorporated. The light generated is detected by a CCD sensor and the images for all the beads are processed simultaneously and assembled by a series of modules: the Overlapper, the Unitigger, and the Multialigner. The first one finds overlaps between the reads, the second turns the overlapping reads into contigs, and the last generates scores and consensus calls. High quality reads were defined as those with less than 5% of indeterminate flow reads. (Margulies, Egholm et al. 2005;Urich, Lanzen et al. 2008;McClain, Shaffer et al. 2009).
Annotation

A pipeline was established based on the SOP for Bacterial Genome Annotation at BioHealthBase for bacterial pathogens (NIAID / http://www.biohealthbase.org).

First, Glimmer3 (Microbial Gene-Finding System, University of Maryland) and GenemarkS (Besemer, Lomsadze et al. 2001) were used for predicting complete open reading frames (ORFs) using the published leptospiral start codons (atg, ttg, gtg, ctg, and att), their frequencies (0.8001, 0.1331, 0.0608, 0.0054 and 0.0001, respectively) and a minimum gene length of 108 bp (based the length of known leptospiral CDSs) as parameters (Bulach, Seemann et al. 2006). All other parameters were left at default values. Glimmer3 uses an interpolated Markov model that was trained with the published Leptospira genomes. GeneMarkS uses the Gibbs sampling alignment program and a heuristic Markov model of coding and non-coding regions to predict ORFs.

Reverse-Position Specific BLAST (rpsblast) of the ORFs was done against the Clusters of Orthologous Groups of proteins (COGs) database and for assigning Pfam categories.

SpLip, a spirochetal lipoprotein prediction tool, was used to determine potential lipoproteins in Leptospira licerasiae. This program was trained using the two published pathogenic Leptospira genomes: L. interrogans serovar Lai and Copenhageni (Setubal, Reis et al. 2006).

RNA analysis for non-coding RNAs was done using the program INFERNAL. Finally, tRNAscanSE was used to locate the tRNAs.
Both Glimmer3 (version 3.02) and Genemark were used to determine possible ORFs and make a draft annotation. The Glimmer3 ORFs were used for reciprocal Blast analysis to infer protein orthology with an e-value of $10^{10}$ or less. The highest matching reciprocal hits were manually curated to separate orthologs from paralogs. Comparisons of the existing orthologs between species were done using spreadsheets.

Gene, protein and genome comparisons

Individual gene and protein comparisons were done using the BLAST server. The RAST server (Rapid Annotation using Subsystem Technology) and the SEED project website (subsystem annotation environment) were used to compare metabolic functions across genomes and to look at the localization of specific genes (Overbeek, Begley et al. 2005; Aziz, Bartels et al. 2008).
3.3 Results

*General genome features*

454 pyrosequencing of *L. licerasiae* yielded 192,506 reads, 76,771,856 bases, 191,814 aligned reads and 75,521,924 aligned bases (Table 3.1). The reads obtained were processed using GS FLX Titanium cluster. This software aligned 185,152 of the reads into 59 contigs with 4,154,666 (99.45%) Q40 plus bases (Table 3.2). Of these 59 contigs, 33 were considered large contigs with an average length of 126,601 and a median of 233,766 bases and 5 contigs were smaller than 500 bp and were not further analysed. The G/C content of these contigs was 41.6%.
Table 3.1. Sequencing output

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number</th>
<th>Length</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number Of Reads</td>
<td>192,506</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Number Of Bases</td>
<td>767,718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number Searches</td>
<td>16,881</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seed Hits Found</td>
<td>8,687,540</td>
<td>514.63</td>
<td>514.63%</td>
</tr>
<tr>
<td>Overlaps Found</td>
<td>638,474</td>
<td>37.82</td>
<td>7.35%</td>
</tr>
<tr>
<td>Overlaps Reported</td>
<td>611,238</td>
<td>36.21</td>
<td>7.04%</td>
</tr>
<tr>
<td>Overlaps Used</td>
<td>375,873</td>
<td>22.27</td>
<td>61.49%</td>
</tr>
</tbody>
</table>

Table 3.2. Alignment consensus results outcome parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number</th>
<th>Length</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of aligned reads</td>
<td>191,814</td>
<td></td>
<td>99.64%</td>
</tr>
<tr>
<td>Number of aligned bases</td>
<td>75,521,924</td>
<td></td>
<td>98.37%</td>
</tr>
<tr>
<td>Inferred read error</td>
<td>753,775</td>
<td></td>
<td>1.00%</td>
</tr>
<tr>
<td>Number assembled reads</td>
<td>185,152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number partial reads</td>
<td>4,498</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of singleton reads</td>
<td>420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of repeats</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of outliers</td>
<td>2,086</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of reads too short</td>
<td>272</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.3 Comparison with other genome data on. The *L. licerasiae* G/C% and CDS number approximate those of *L. borgpetersenii*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>L. interrogans copenhageni Fiocruz</th>
<th>L. borgpetersenii Hardjo</th>
<th>L. biflexa Patoc Ames</th>
<th>L. licerasiae Varillal Var10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome Size (bp)</td>
<td>4,627,366</td>
<td>3,956,088</td>
<td>~ 4,171,655</td>
<td></td>
</tr>
<tr>
<td>GC (mol %)</td>
<td>35.1</td>
<td>41.1%</td>
<td>38.6</td>
<td>41.6%</td>
</tr>
<tr>
<td>CDS</td>
<td>3659</td>
<td>2880</td>
<td>3598</td>
<td>&gt;2,990</td>
</tr>
<tr>
<td>Species-specific genes</td>
<td>684</td>
<td>150</td>
<td>874</td>
<td>464</td>
</tr>
</tbody>
</table>
The largest contig had 460,507 bases. The contigs were numbered from 001 to 054 for further analysis.

The contigs resulting from the previous analysis were processed using Glimmer3 and GeneMarkS for determination of open reading frames (ORF). Glimmer3 found 2990 complete ORFs that included both a start and an end codon. 1791 of these ORFs were able to be assigned to a protein family (Pfam) (Appendix). The largest protein families found were: ABC transporters (34 members), histidine kinase/ DNA gyrase B and HSP90-like ATPase (34 members), stage II sporulation protein E (33 members), AcrB/AcrD/AcrF family (23 members) and methyl accepting chemotaxis protein (MCP) signaling domain (21 members). All of the other families had less than 20 members. This distribution is similar to the one found in other leptospires.

2418 of the ORFs had a matching Clusters of Orthologous Groups of proteins (COG) function (Table 3.3). Among these the largest groups were General function (346) and Cell envelope biogenesis and outer membrane (206). 169 ORFs had no known function according to this analysis, hence termed hypothetical proteins.

Origin of replication

The origin of replication in *L. licerasiae* was determined based on the localization of the cluster of *dnaA, dnaN, gyrA* and *gyrB*. The arrangement of the chromosome initiation site is very similar to that of *L. interrogans* serovar Copenhageni and serovar Lai, which differ from the initiation site of other spirochetes
Figure 3.1. The chromosome initiation site in *L. licerasiae* is similar to that of both species of pathogenic leptospires and the saprophytic *L. biflexa*, presenting all four genes in the same orientation and order.
"T. pallidum" Nichols is missing gyrA and the orientation of the B. burgdorferi initiation site differs from Leptospira species. (Seshadri, Myers et al. 2004) (Figure 3.1).

**Transcription, translation and gene expression regulation**

Among the genes necessary for transcription and translation, *L. licerasiae* contains orthologs of the subunits α, β and β′ of RNA polymerase. *L. licerasiae* has 37 tRNA species, the same number as in *L. interrogans* and *L. borgpettersenii* (Picardeau, Bulach et al. 2008). These are distributed among several contigs (Table 3.4). All of the tRNA synthetase genes are present, ten of them seem to be single genes and the other 27 seem to be forming 8 clusters. This distribution will be confirmed once the genome gaps are closed. 2 copies of 5S and one of 16S and 23S are present. Each ribosomal RNA is found in a separate and small contig. Their distribution on the genome awaits genome closure.

Seventeen sigma factors including σ^32^ (which is involved in the transcription of heat shock proteins), σ^54^, σ^70^, heat induced transcriptional regulators and 3 extra cytoplasm function (ECF)-type σ factors are present in *L. licerasiae*. It also contains elongation and termination factors nusA, nusB, nusG, greA and rho.

The *L. licerasiae* draft genome contains 63 genes that encode sensor histidine kinases and response regulators. These numbers are similar to the ones found in other leptospiral species and could be associated with the ability of leptospires to survive under different conditions.

The genome also contains orthologs of heat-shock protein 90 and heat-shock protein HtpX. There is also a dnaK cluster similar in organization to the one in the other
known leptospiral genomes (Figure 3.2). Other heat shock response components found in *L. licerasiae* include heat-inducible transcriptional repressor *hrcA* and *groES*. The specific role of each of this transcription factors in *Leptospira* remains to be elucidated.

The GroEL ortholog in *L. licerasiae* seems to be disrupted by a frame shift in aa 157 due to the insertion of an adenine (Figure 3.3). Depending on which ORF is recognized, GroEL would be missing either the ring oligomerisation interface residues, or the residues involved with stacking interactions, both of which are part of the double stacking ring that GroEL forms and inside which proteins are folded.

The controller of bacterial stress response *rpoS* ($\sigma^{38}$) is missing, as has been reported for the other sequenced species of *Leptospira* and in *T. pallidum* (Fraser 1998; Matsunaga, Sanchez et al. 2005).

This intermediate *Leptospira* also contains an ortholog of $\sigma^{WhiG}$. Besides its function as a sigma factor this gene contains a domain that has been involved in sporulation and glycogen biosynthesis in other bacteria and is also present on other species of *Leptospira*. This gene appears to be missing the C-terminal sigma70-r4 domain and the DNA binding residues due to a stop codon caused by a guanine deletion after aa 216. This could lead to a loss of function.

Similarly to the other leptospiral species, this intermediate leptospiral species has 16 orthologs of anti-sigma factors including *rsbW* and *rsbU*.
Figure 3.2. Nucleotide sequence and translation of *L. licerasiae* orf00124contig001 in frame 2 (green). This amino acid sequence matches amino acids 165-483 of GroEL in *L. interrogans*. The amino acid sequence in frame 3 (blue) corresponds to amino acids 1-157. If the adenine highlighted in the red box is removed then orf in frame 3 (blue) continues with the amino acid sequence from orf00124 and forms a 483 amino acid protein matching GroEL.
Figure 3.3 *dnaK* cluster in *L. licerasiae* is similar in organization to those in *L. interrogans* Lai and L1130 and *L. biflexa*. It contains the *hrcA* gene that is absent in *Treponema* species.
Replication, repair and recombination

*L. licerasiae* has orthologs of sub-units α, β, γ/τ and δ of DNA polymerase III. It also contains the SOS-inducible *uvrABCD, recA* and DNA polymerase IV. DNA polymerase IV appears to be necessary for leptospires to survive the environment and exposure to UV light (Xue, Yan et al. 2008). Orthologs for subunits A and B of topoisomerase I and topoisomerase II are also present.

Three orthologs of recombination repair proteins and one of the recombination factor protein (RarA) can be found. These are also present in the other species of *Leptospira.*

Included in the genome there are *parA* and *parB*-like genes, their significance cannot be determined at this time. Five integrases/ recombinases and one ISLbp transposase are also found in this genome draft. At this time IS elements were not located, as they might be repeated in-between contigs. The number of them will be known once this genome is closed.

Metabolism

The main studies in the metabolism of leptospires were completed over 25 years ago (Baseman and Cox 1969). There are no studies investigating the metabolism of these spirochetes with current technology, so most metabolic pathways remain to be elucidated. Most of our knowledge is based on the known nutritional requirements of leptospires (Faine 1959; Johnson, Walby et al. 1973) and the very early work in the field (Staneck, Henneberry et al. 1973; Finn and Jones 1976).
In the *L. licerasiae* genome, classification based on the database of Clusters of Orthologous Group of proteins (COG) found components of the lipid metabolism and electron transfer systems, carbohydrate transport and metabolism including pentose pathways, as in the other species of *Leptospira*. Enzymes related to glycerol, amino acid and nucleotide metabolism were also found. The organization of the LOS cluster in *L. licerasiae* cannot be determined at this time.

Among the proteins that could be involved in the processing of lipoproteins were found orthologs of an outer membrane lipoprotein carrier protein, a lipoprotein releasing system LolE permease component and a prolipoprotein diacylglyceryltransferase.

**Comparative genomics**

Of the 2990 CDS found in *L. licerasiae* in this study, 684 (22.8%) do not have an ortholog in the other leptospiral species. Of the remaining CDS *L. licerasiae* shares the most with *L. interrogans* (1864, 62%) followed by *L. borgpetersenii* (1661, 55.5%) and *L. biflexa* (1372, 45.8%)(Figure 3.5). A total of 1253 CDS (41.9% of *L. licerasiae*) are common to all of the leptospiral genomes (Figure 3.4).

Genes shared by the four leptospiral species include the flagellar synthesis genes *fliDEGMPQRW*, flagellar motor protein encoding genes *motAB*, flagellar assembly protein gene *flhBF* and flagellar basal body genes *flgBCEGKL*. These genes and several TPR-repeat proteins are also present in other spirochetes and are thought to be involved in motility and host pathogen interactions.
It is also interesting to find a high number of paralogs of *L. biflexa* CDS in *L. licerasiae* (820) and viceversa (758), which could be suggestive of gene duplications. When comparing the to *L. interrogans* against the *L. licerasiae* genome and the other sequenced leptospiral genomes, 684 proteins can be identified as unique to this pathogenic leptospire. This group includes 565 hypothetical proteins, 25 putative lipoproteins, 41 transposases and 53 proteins with an annotated function. This last group is where potential virulence factors would be found (Figure 3.5).

**Potential virulence factors**

Several leptospiral proteins potentially implicated in virulence have been suggested in various reports. For this analysis several strategies were used to identify potential virulence factors. First, potential virulence factors were identified in the database originated from reciprocal blast. If no orthologs were found, individual sequences were blasted against the *L. licerasiae* contigs using blast (tblastn). Finally the SEED viewer was used to compare genome areas and identify specific ORFs encoding putative virulence factors. Among the potential virulence factors found are lipoproteins, hemolysins and post-translational modification pathways.
Figure 3.4. Two way comparisons between all genomes, showing number of unique genes, shared orthologs and paralogs.
Figure 3.5. (A) Comparisons between all genomes. Three way comparison data is not shown. (B) Distribution of proteins unique to *L. interrogans* L1130, characteristics of other proteins in Table 4.1.
**Lipoproteins**

Several lipoproteins have been suggested as potential virulence factors not only in *Leptospira*, but also in other spirochetes and gram-negative bacteria (Haake 2000; Rezwan, Grau et al. 2007). For *Leptospira* both surface and non-surface lipoproteins have been described. Surface lipoproteins seem to be important in interactions with the environment, potential hosts and possible virulence factors (Haake 2000). Few leptospiral lipoproteins have been demonstrated experimentally to be lipidated. The other proteins called lipoproteins are classified as such based on the sequence of the signal peptide. To be considered part of a lipoprotein the signal peptide must contain a positively charged N-terminal and a sequence know as a lipobox in the C-terminal. A lipobox consists of four or five amino acids that include a cystein (position +1) and must contain at least one leucine, phenylalanine, valine, isoleucine or tyrosine in position -3 or -4. SpLip, a prediction algorithm designed for identifying lipobox sequences in spirochetes was used to predict lipoproteins in the sequence of *L. licerasiae*.

Analysis of the *L. licerasiae* contigs using SpLip identified 119 potential (containing a known lipobox sequence) lipoproteins and 17 possible lipoproteins (contains a sequence that has criteria to be a lipobox but has not been previously described). The list of leptospiral lipoprotein orthologs obtained by reciprocal blast was compared to the list of lipoproteins predicted by SpLip. Orthologs of the surface lipoproteins LipL32, LipL41, LipL46, LigB, and Loa22 are present in *L. licerasiae*. In the LipL21 and the Loa22 orthologs the lipobox sequence has been replaced by other amino acids (that do not have the characteristics to be a lipobox (IFAAC to LLTYC in LipL21, SFLTC to...
SLALC in Loa22) (Figure 3.6). Orthologs of the non-surface proteins LipL31, LruA/LipL71 and LruB were also found, but the lipobox sequence had changed and lost lipobox characteristics or in the case of LruA/LipL71, are missing the entire signal peptide (Figure 3.7). There is an ortholog of LipL45 that is a possible lipoprotein, 2 out of 4 LipL45-like proteins are identified as lipoproteins, 2 are not. The \emph{L. licerasiae} genome does not appear to have a ortholog of LipL36.

The \emph{L. licerasiae} genome also contains orthologs of LipL48, RlpA-like lipoprotein and several hypothetical lipoproteins. Thirty-six of the lipoproteins predicted by SpLip in this species do not have a ortholog in the other species of \emph{Leptospira}. The potential role of lipoproteins in virulence will be further discussed in the next chapter.

**Iron metabolism**

Iron is an essential cofactor for the growth and development of multiple microorganisms. Iron scavenging is a major a virulence factor \textit{in vivo} for pathogens. Iron uptake in gram negative bacteria involves a complex network of proteins and regulators: outer membrane receptors capable of detecting iron sources (Ton-B receptors), protein-dependent ABC-transmembrane transporters, a complex that provides energy for the transport formed by biopolymer transporters, (TonB-ExbB-ExbD) and a ferric uptake protein regulator (FUR)(Louvel, Bommezzadri et al. 2006). Previous studies have shown that in \emph{Leptospira} the expression and processing of lipoproteins like LipL32 can be iron-dependent.(Cullen, Cordwell et al. 2002). It has also been described that pathogenic and saprophytic leptospires have different mechanisms of iron uptake (Louvel, Bommezzadri et al. 2006).
A cluster of a TonB dependent receptor, ExbB/ExbG/TonB/ and another that seems to have lost the Ton B receptor can be found in this the genome draft of *L. licerasiae* (Figure 3.7). How this loss affects iron metabolism in *L. licerasiae* remains to be investigated.
Figure 3.6. ExbB/ExbD/TonB cluster present in contig005 (top) and contig006 (bottom) in *L. licerasiae*. Contig006 lacks a TonB dependant receptor, present in both *L. interrogans* serovar Copenhageni L1130 and Lai but not in *L. biflexa*. In the bottom figure transposases are colored light blue or cyan.
3.4 Discussion

Sequencing an intermediate leptospire gives an unprecedented way to compare leptospiral genomes. This comparison could help to establish which genes contribute to the ability to infect the host and which genes allow dissemination in the host and influence the severity of disease.

The sequences obtained with 454 sequencing generated enough data to allow for a draft genome reconstruction with an accuracy of over 99%.

Based on this data, the *L. licerasiae* genome appears to be larger that the genomes of *L. biflexa* and *L. borgpetersenii*, but smaller that the *L. interrogans* genome.

It has been suggested that the genome reduction in *L. borgpetersenii* has resulted in reduction of the potential of environmental transmission (Bulach, Zuerner et al. 2006). It could be argued that the smaller size of the *L. licerasiae* compared to the genomes of pathogenic leptospires is a result of partial genome reduction.

It could also be argued that the genome reduction in *L. borgpetersenii* is associated with the longer life span of its hosts (cattle), therefore selecting for those leptospires more adapted to the host. It has been reported that *L. borgpetersenii* transmission seems to occur mostly from host to host, with no environmental exposure. In fact, DNA polymerase IV is missing from *L. borgpetersenii*, which could lead to its inability to survive exposure to UV light in the environment and, but would not be necessary inside the host (Xue, Yan et al. 2008).

In comparison, *L. licerasiae* has been mostly isolated from rodents, therefore not supporting the genome reduction based on host lifespan hypothesis (Ricaldi, Matthias et
al. 2008). Therefore, other factors must have lead to the loss of characteristics that would allow this leptospire to survive inside the host or of factors that confer to it the ability to cause disease.

The size of the *L. licerasiae* genome lies in between that of *L. biflexa* and *L. interrogans*, but its G/C content (41.6%) appears to be closer to *L. borgpetersenii* (41.1%). The number of CDS predicted in this analysis is slightly higher than that of *L. borgpetersenii* and we predict this number will only increase once the genome is closed. Still, a comparison of the set of individual genes from each genome against the other genomes showed that *L. licerasiae* has more genes in common with *L. interrogans* L1130 than with any of the other species (1864) (Figure 3.5).

Species-specific genes are present in all of the analyzed genomes. *L. biflexa* has the highest number of unique genes (874) and *L. borgpetersenii* the least (150). *L. licerasiae* has 464 unique genes in this genome draft. A considerable amount of predicted genes in all of the genomes is annotated as hypothetical proteins.

It has been reported that leptospiral genomes are propense to wide genomic rearrangement even in closely related strains. (Zuerner, 2006). Figure 3.4 shows syntheny plots comparing *L. licerasiae* with *L. interrogans* and *L. borgpetersenii*. Wide rearrangement occurs in comparison with the two other genomes. The presence of IS elements could account for this variability. These elements are present in all the other sequenced species of Leptospira. In our draft genome we have not been able to identify IS elements, one possible explanation is that their similarity and repetitive nature are responsible for the gaps in sequence assembly.
Even though the organization of most of the *L. licerasiae* clusters mentioned in this study is similar to that of the pathogens, the intergenic distances seem to vary slightly, either decreasing or increasing (1 – 2 Kbp), but it is similar to the one in *L. biflexa*. It has been described that a smaller intergenic distance is associated with genes that are transcribed as one operon. Whether the changes in intergenic distance in *L. licerasiae* disrupt the translation of the operons will have to be further investigated experimentally. This finding could be associated with differences in the induction of expression. *L. licerasiae* has 4 extracytoplasm function (ECF) type σ factors compared to 11 found in other leptospiral genomes.

No striking differences between the genomes were found in this analysis. A systems biology approach will be necessary to understand the metabolic pathways each species uses and therefore in which environments they are able to survive.

The reciprocal-blast best-hit approach used in this study finds orthology between genes, but the process to confirm whether domains or amino acids involved in specific interactions are conserved has to be done manually. At the same time the annotation of genes that appear to be orthologs differs frequently, making the process of comparison more difficult.

The characteristics of the *L. licerasiae* genome seem to correspond to the taxonomic classification of this species as intermediate. This species appears to have some genetic characteristics of the pathogens and some of the saprophytes.

There have been few experimental attempts to try to differentiate factors that are necessary to infect the host, from those necessary to survive and replicate inside the host.
and those that are necessary to cause active disease. The findings in this genome might allow us to do so in a more directed approach.

In the next chapter the existing experimental evidence for potential virulence factors will be discussed.

Chapter 3, in part is currently being prepared for submission for publication of the material. Ricaldi, JN, Matthias MA, Vinetz, JM. The dissertation author was the primary investigator and author of this material.
4.1 Introduction

Despite the worldwide prevalence of leptospirosis, few virulence determinants of pathogenic *Leptospira* have been thoroughly characterized. Most of the data obtained focuses on one particular protein or gene. The near lack of genetic tools makes the analysis of loss of function very difficult and the analysis of interactions even more so. The comparison of different leptospiral genomes that have become available will allow for a wider approach, ideally by reconstructing and comparing pathways using a “bottom-up” approach, that is reconstructing pathways based on genes present in each genome, inferring the chemical scale reactions they participate in and then using a directed experimental approach to validate the models produced *in silico*.

This analysis aimed to identify genes that are potential virulence factors among genes with assigned function that are unique to the pathogen *L. interrogans* and present evidence supporting their role.

Sialic acids are used by pathogens to evade the immune system. In early studies of *Treponema* pathogenesis, activation of the alternative complement pathway was enhanced by enzymatic removal of sialic acid, therefore suggesting a role for sialic acid in the evasion of complement (Fitzgerald 1987). Sialic acid mediates complement resistance by promoting the binding of complement factor H.
Factor H is composed of a linear array of 20 homologous short consensus repeat (SCR) domains with many functional sites. Three of these sites are involved in binding C3b and regulating complement activation; others, SCR7, SCR20, and possibly SCR12–14 bind to heparin or sialic acid, where they could have a role in discriminating activators from non activators and could be responsible for host recognition (Jokiranta 2000; Pangburn, Pangburn et al. 2000). In mammals, the affinity of fH is relatively high for C3b molecules deposited on self-structures whereas C3b on foreign targets has a relatively low affinity for fH, a fact that leads to opsonization and subsequent target damage. Because the non-activating surfaces usually contain negatively charged surface structures (sialic acids and glycosaminoglycans), interactions between fH and polyanions have been intensively characterized (Jokiranta, Hellwage et al. 2000). A factor H binding protein (LfhA) has been recently reported in Leptospira (Verma, Hellwage et al. 2006), still, more thorough functional studies remain to be published.

Sialic acids (Sias) are N- and O- derivatives of 5-amino-3,5-dideoxy-D-glycero-D-galacto-nonulosonic acid (neuraminic acid). They are the only nine carbon sugars present in bacteria and also include pseudaminic acid and legionaminic acid. Pseudaminic acid is necessary for flagella biogenesis in Campylobacter and Helicobacter (Goon, Kelly et al. 2003; Schoenhofen, McNally et al. 2006). Legionaminic acid is attached to Legionella LPS and may support adherence to the membrane of amoebae in the natural environment and to the membrane of alveolar macrophages in the human lung (Neumeister, Faigle et al. 1998). Neisseria meningitidis, has polysialic acids in the surface that are necessary for disseminated infection, and their expression constitute one of the steps in variation from encapsulated to unencapsulated state. Meningococcci
sialylate lipooligosaccharides which decreases the binding of C3b and induces resistance to complement fixation, binding of complement factor H (fH). Through this mechanism, it mediates serum resistance and inhibits activation of the alternative complement pathway (Ram, Sharma et al. 1998; Vogel and Frosch 1999). Capsular Sia is a critical virulence factor of Group B Streptococcus (GBS) functioning to inhibit complement activation and impairing opsonophagocytic clearance (Pluschke, Mayden et al. 1983; Marques, Kasper et al. 1992; Ram, Sharma et al. 1998).

Although the existence of sialic acid in the surface of Leptospira has not been reported, it has been recently described, using a bioinformatics approach, that Leptospira contains a Sia genes cluster containing orthologs of neuA, neuB, neuC and neuD (Lewis 2006). In this study a biochemical approach analysis was used to investigate the presence of sialic acid in L. interrogans L1130.

The structure and composition of the lipid modifications of leptospiral lipoproteins remain unknown. All known spirochetal lipoproteins contain a signal peptide at their N-terminal, which maintains structural characteristics in all lipoproteins. These characteristics include a lipoprotein signal peptidase (Lsp)/leader peptidase (Lep) cleavage site known as a “lipobox” in the C-terminal of the signal peptide, a hydrophobic region and an N-terminal region. In the signal peptide, the positively charged N-region is composed of 2 or more amino acids. The hydrophobic region is composed of 6 or more amino acids and should not have any charged amino acids such as Lys, Arg, Glu and His. Finally, the C-terminal of the signal peptide contains the lipobox (Figure 4.7A).
The lipobox region and its surrounding area are thought to determine whether the lipoprotein will be in the outer or in the inner membrane.

In this study, the sequences of known lipoproteins in *L. interrogans* L1130 have been compared with their orthologs in *L. licerasiae*. The analysis was focused on the lipoproteins signal peptide since changes in this sequence could determine changes in lipidation, surface localization and potentially, virulence.
4.2 Methods

Mild acid hydrolysis and DMB-derivatization

In order to determine whether *L. interrogans* expresses Sia, mild acid hydrolysis was performed on extensively washed pellets from a centrifuged culture strain L1130 to release covalently bound surface carbohydrates. The resulting soluble fraction was filtered to remove large molecular weight components and derivatized with DMB (1,2-diamino-4,5-methylene dioxybenzene), a reagent that reacts specifically with the α-keto acid portion of Sia molecules. Derivatized Sias were then resolved by reverse phase HPLC as previously described (Lewis, Nizet et al. 2004), using standard Sias in parallel injections to determine elution times. Leptospiral culture media (EMJH) was also analyzed and this profile was used as a control for the origin of the detected Sias.

Solid phase lectin binding

In order to probe the abundance and nature of the sialylated molecules on *L. interrogans*, whole cell lysates prepared from the *L. interrogans* L1130 genome strain were fractionated using a lectin-based solid phase assay (Q Proteome Sialic Acid kit - Quiagen).

Affinity purification of sialylated species was accomplished utilizing spin-columns containing three different immobilized Sia-binding lectins: wheat germ agglutinin (WGA), *Sambucus nigra* agglutinin (SNA), and *Maackia amurensis* lectin (MAL). Molecules captured by each of these lectins and then eluted were analyzed by SDS-PAGE followed by silver staining.
Mass spectrometry

To determine whether *L. interrogans* uses Sias for posttranslational modification of proteins, pooled affinity-purified material was subjected to denaturation, reduction, and alkylation, followed by trypsin digestion and MS/MS analysis using a nano-flow LC-tandem mass spectrometer. Peptide mass fingerprints identified in the affinity-purified material were used to identify *L. interrogans* proteins by searching against the NCBInr bacterial genome database.

Glycoprotein staining

A commercially available glycoprotein staining kit (Pierce) was used on a SDS-PAGE electrophoresis gel to identify glycoproteins in *L. interrogans* L1130 lysate.

Metabolic staining

Metabolic staining of live leptospires was performed using an azide-labeled sialic acid precursor (tetraacetylated N-azidoacetyl-d-mannosamine). The precursor was added to a culture of *L. interrogans* and incubated for 3 days at 30°C. Cell were washed and lysed, followed by chloroform:methanol protein precipitation. The precipitated product was resuspended, bound in a Cu mediated reaction to a biotin labeled-alkyne, subjected to SDS-PAGE and transferred to a nitrocellulose membrane. HRP-streptavidin and colorimetric subtrate were used to visualize labeled protein.
**LPS extraction**

A 1L culture of *L. interrogans* was grown in serum free-low endotoxin media to a density of $10^8$ cells/mL. The cells were washed in pyrogen free PBS and used for LPS extraction using the Tri-Reagent method. Briefly, 100 mg of lyophilized bacteria were resuspended in Tri-Reagent method and incubated at room temperature for 15 min. Chlorophrom was added, followed by vigorous vortexing and centrifugation for phase separation. The aqueous phase was removed and water added to the organic phase and the vortexing and centrifugation are repeated. Both aqueous phases are combined and precipitated using cold ethanol magnesium precipitation. The resulting material was subjected to SDS-PAGE, silver staining and DMB-derivatization.

**PCR of sialic acid cluster genes**

Primers based on the genome of *L. interrogans* L1130 were designed for the detection of genes in the sialic acid cluster as follows: sasfrontF (5’- TCC GGA AAT GCG AAT GAT G-3’), sasfrontR (5’- CAC CGG GCA AAA GAC TAA CCT-3’), sasendF (5’- CGG ATA TAG CGG ACG ATG TAA-3’), sasendR (5’- CGC CAA AAA GCC AAG GAA-3’), neuA2F (5’-TGA AGC GGC AAA AAG AGC-3’), neuA2R (5’-TGA AAT AAC ATG CCG ACA AAT A-3’), neuCfrontF (5’- CGC TAC GGG AAT GCA TCT GTC TC-3’), neuCfrontR (5’- CCC ATT CCC CCA ACC AAA AA-3’), neuCendF (5’- GGC GAG GAT CCT TCT AAT GTT TTT-3’) and neuCendR (5’- ACT CGC TCC GCC TTC ACC A-3’). PCR reactions were prepared using 0.2 µM of each primer in a 20µL reaction using DNA from *L. interrogans* Lai, *L. interrogans* L1130, the intermediates *L. licerasiae* and *L. fainei* and *L. biflexa* Patoc. neuA2 and neuBfront
reactions used an annealing temperature of 52°C. neuCfront, neuCend, sasfront and sasend were run using an annealing temperature of 58°C. A 16S gene PCR reaction using previously published primers fLIP and rLIP1 was used as a control for integrity of DNA.

**NeuA2 southern blot**

Genomic DNA samples of *Salmonella enterica*, *L. interrogans* Lai, *L. interrogans* L1130, *L. biflexa* Patoc, rat isolate CEH008 (*L. licerasiae*), wild rodent isolate MMD4847 (*L. licerasiae*), wild rodent isolate MMD3731 (*L. interrogans*) and *L. interrogans* Evansi were prepared into plugs using 1% agarose and 0.5x TBE. These were subjected to depurination and denaturing conditions. DNA was then transferred to a positively charged membrane via overnight capillary transfer with 20x SSC. Finally the DNA was cross-linked to the membrane using short wave DNA for 15 minutes.

10 mL of pre-hybridization solution (QuikHyb, Stratagene) were warmed to 40°C prior to hybridization. Hybridization was done overnight at 40-42°C using the same solution and adding 10µL of DIG-labelled PCR product of primer neuA2F (5’-TGA AGC GGC AAA AAG AGC-3’) and neuA2R (5’-TGA AAT AAC ATG CCG ACA AAT A-3’). 2x SSC at room temperature and with 1x SSC at 68°C were used for stringency washes. A chemiluminescense substrate and an alkaline phophatase bound anti-DIG antibody were used to demonstrate binding of the probe.

**Palmitate labeling**

An adaptation of the method described by Haake was used for palmitate labeling (Haake, Chao et al. 2000). 12 mL cultures of *L. interrogans* L1130 and *L. licerasiae* were
grown to a density of $10^8$ cells/mL in EMJH media and intrinsically labeled by adding [U-$^{14}$C] palmitate (50 μCi (1.85 MBq)). Incubation at 30 C in a shaking incubator continued for 48 h until the cells reached a density of $10^9$cells/mL. Cells were then pelleted and washed. Cell lysates were then separated using SPS-PAGE in 12% Tris-Temed gels. The gels were dried at exposed to film for 4 weeks.

**Western blots**

The gels used to observe palmitate labeling were rehydrated in 4% methanol and transferred to a nitrocellulose membrane. Western blots were performed using 1:1000 dilution of rabbit anti-LipL32 and anti-Loa22 antibodies (provided kindly by David Haake) in 5% non-fat dry milk in PBS-Tween 20. HRP-labeled goat anti-rabbit antibodies (KPL) were used as secondary antibodies in a 1:10 000 dilution. West Pico HRP substrate (Pierce Scientific) or a colorimetric subtrate (TMB) were used to visualize antibody binding.

**Live immunofluorescence assay**

To determine surface localization of lipoproteins in *L. interrogans* and *L. licerasiae* rabbit antibodies anti-LipL32, -Loa22 and –LipL71 (provided by David Haake) were used. A donkey anti-rabbit rhodamine labeled antibody was used as a secondary antibody. CFSE was used for live staining of leptospires. 2 mL of live stationary phase cultures grown in EMJH were washed with TBS twice and then incubated with a 1:250 dilution of the primary antibodies for 2 hours. Cultures were washed twice and then incubated with the secondary antibody in the dark at room
temperature for 2 hours. After washing twice a 1:250 dilution of CFSE was added followed by a 30-minute incubation.
4.3 Results

Genome analysis

When comparing two pathogen and a saprophyte genomes with the *L. licerasiae* genome, 684 genes unique to *L. interrogans* L1130 were identified (Table 4.1). 565 of these genes do not have an assigned function.

Two of the 3 enzymes involved in sialic acid synthesis are found in the list of proteins unique to *L. interrogans* L1130 (Table 4.1). Among the unique genes in *L. interrogans* L1130 identified in this analysis there are orthologs of *neuA* (YP_002102.1) and *neuC* (YP_002107.01). Both of these genes are necessary for sialic acid metabolism (Figure 4.1).

DMB-derivatization and HPLS analysis

A large early-eluting HPLC peak corresponding to KDO (2-keto-3-deoxy-D-manno-octulosonic acid) and an α-keto acid present in the core region of lipopolysaccharide were found (Fig. 4.3 B). As predicted by the presence of Sia biosynthetic gene homologues, a peak corresponding to the elution time of N-acetyleneuraminic acid (Neu5Ac) was also identified (Fig. 4.3 B). LCMS analysis showed that mass spectrum that matches the expected mass of DMB-derivatized Neu5Ac in (Fig 4.3 C) was present at the elution time of the HPLC peak in question (Fig 4.3B), demonstrating conclusively that *L. interrogans* Copenhageni L1130 expresses this Sia.
# Table 4.1. Unique proteins in *L. interrogans* serovar Copenhageni strain L1130.

**Potential virulence factors**

<table>
<thead>
<tr>
<th>Description</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Dehydropantoate 2-Reductase</td>
<td>gi</td>
</tr>
<tr>
<td>Acetyl Transferase</td>
<td>gi</td>
</tr>
<tr>
<td>Acetyl Transferase</td>
<td>gi</td>
</tr>
<tr>
<td>Acetyl Transferase</td>
<td>gi</td>
</tr>
<tr>
<td>Acetyl Transferase</td>
<td>gi</td>
</tr>
<tr>
<td>Alpha/Beta Fold Family Hydrolase</td>
<td>gi</td>
</tr>
<tr>
<td>Aminopeptidase</td>
<td>gi</td>
</tr>
<tr>
<td>Ankyrin G Domain-Containing Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Ankyrin-Like Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Capsule Biosynthesis Protein</td>
<td>gi</td>
</tr>
<tr>
<td>CMP-N-Acetyleneuraminic Acid Synthetase</td>
<td>gi</td>
</tr>
<tr>
<td>Cyclopropane-Fatty-Acyl-Phospholipid Synthase/Methyltransferase</td>
<td>gi</td>
</tr>
<tr>
<td>Cytoplasmic Membrane Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Cytoplasmic Membrane Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Cytoplasmic Membrane Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Cytoplasmic Membrane Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Deor Family Transcriptional Regulator</td>
<td>gi</td>
</tr>
<tr>
<td>Dolichyl-Phosphate Mannose Synthase</td>
<td>gi</td>
</tr>
<tr>
<td>Dtdp-4-Dehydrorhamnose Epimerase</td>
<td>gi</td>
</tr>
<tr>
<td>FAD-Dependent Oxidoreductase Family Protein</td>
<td>gi</td>
</tr>
<tr>
<td>FAD-Dependent Oxiredutase</td>
<td>gi</td>
</tr>
<tr>
<td>Fruiting Body Developmental Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Fruiting Body Developmental Protein</td>
<td>gi</td>
</tr>
<tr>
<td>Glyoxalase</td>
<td>gi</td>
</tr>
<tr>
<td>GNAT Family Acetyltransferase</td>
<td>gi</td>
</tr>
<tr>
<td>Hydrolase</td>
<td>gi</td>
</tr>
<tr>
<td>Table 4.1 continued</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Integral Membrane Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Lexa Family Transcriptional Repressor</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Methionyl-Trna(Fmet) N-Formyltransferase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Methyltransferase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Mrr</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Mutator Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Mutt-Like Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Nuclease-Like Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Nucleotide-Diphosphate-Sugar Epimerase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Nucleotidylytransferase Domain-Containing Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Phenazine Biosynthesis-Like Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Polysaccharide Deacetylase Family Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Potassium-Transporting Atpase C Chain</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Putative Glycosyltransferase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Putative Molybdenum Cofactor Biosynthesis Protein</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Putative O-Methyl Transferase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Putative Ring Hydroxylating Dioxygenase Alpha-Subunit</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Putative UDP-N-Acetylglucosamine-2-Epimerase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Response Regulator</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>SAM-Dependent Methyltransferase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Tautomerase</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Tetr Family Transcriptional Regulator</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Thermolysin</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Thermolysin Homolog Precursor</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Thermolysin Precursor</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Thermolysin Precursor</strong></td>
<td>gi</td>
</tr>
<tr>
<td><strong>Type I Restriction Enzyme</strong></td>
<td>gi</td>
</tr>
</tbody>
</table>
Figure 4.1. Potential sialic acid metabolic pathway in *Leptospira*. UDP-N-acetyl glucosamine is transformed into the activated form of sialic acid (CMP-Neu5Ac) by NeuC, NeuB and NeuA. Two homologs of each of these enzymes are present in the *L. interrogans* L1130 genome, the accession numbers are shown.
Figure 4.2. Sialic acid clusters in *L. interrogans* L1130 (A) and *L. interrogans* Lai (B). Both show identical organization, with gen sequences differing in only 2 nucleotides and 1 amino acid.
**Figure 4.3.** The genes comprising the sialic acid pathway are only present in *L. interrogans*. A genetically distant ortholog of *neuB* is present in all of the leptospiral species studied, its role is unknown.
Controls consisting of *Leptospira* culture media were also analysed, this profile showed peaks at the elution times of Neu5Ac and Neu5Gc, after washing only the Neu5Ac peak remains evidencing total removal of media components (Figure 4.3D). When leptospires grown in serum free media were analyzed, the Neu5Ac peak was also present, further confirming the leptospiral origin of this sialic acid (Figure 4.3D)

**Solid-phase binding assay**

Silver staining after SDS-Page gel electrophoresis evidences bands at ~21 kDa and ~25 kDa that are present at similar intensities in the MAL and SNA lanes, other bands appear to be enriched by affinity purification using one or the other lectin. A faint band at ~43 kDa is apparent in the material isolated by MAL, but not by SNA. Alternatively, bands at ~15, ~37, and ~41 kDa are much stronger in the SNA-purified sample (Fig. 4.4). The affinity-purified material was subjected to DMB-derivatization and HPLC analysis, which confirmed the presence of Neu5Ac in the samples by detecting a peak at the expected time (data not shown).

As a control, buffers used the solid phase assay were analyzed in parallel lanes of the gel, revealing that the faint bands present at ~60 kDa were not specific for sialylated *L. interrogans* molecules (data not shown).
Figure 4.4. A) The DMB- derivatization reaction consists of the binding of a fluorescent molecule (DMB, middle) to the sialic acid molecule to produce a fluorescent compound that can be identified by HPLS. B) HPLC profile shows a large peak for KDO and a smaller peak that corresponds to the elution time of Neu5Ac. C) Mass spectrum of the Neu5Ac confirms the identity of this molecule. D) Neu5Ac and Neu5Gc are found in EMJH alone, in washed leptospires only Neu5Ac is present. Neu5Ac is also present in leptospires grown in serum free media.
Figure 4.5. A) Silver stained gel of elution samples of the solid phase binding assay, showing different patterns of binding for each lectin, but a common band of about 20 kD (red box) B) Glycoprotein staining of cell lysate of *L. interrogans* L1130 shows a band of ~20 kD. C) Western blot of metabolically labeled glycoproteins showing a prominent band of ~180 kDa and a much lighter one at ~20 kD. D) Western blot using anti-Loa22 antibody showing a band of similar size as the ones seen in previous panels.
Mass spectrometry analysis

Three proteins were identified by mass spectrometry: Loa22, LipL32, and LipL41 by analysis, all of which have been described in previous publications as surface-exposed lipid-linked outer membrane proteins of *L. interrogans* (Shang, Summers et al. 1996; Koizumi and Watanabe 2003; Cullen, Xu et al. 2005; Nally, Whitelegge et al. 2005; Nally, Whitelegge et al. 2007). Previous characterizations of these proteins indicated that their actual masses are considerably larger than their predicted masses; a finding attributed to lipid modification.

LPS extraction and analysis

Extracted LPS was subjected to SDS-PAGE and modified silver stain to evidence presence of LPS (Figure 4.5). This material was subjected to DMB-derivatization and HPLC. The analysis shows as expected a large peak at the KDO elution time and a much smaller peak, equivalent to 0.352 pmoles of Neu5Ac.
Figure 4.6. Top figure shows modified silver stain evidencing leptospiral LPS in three samples. Bottom figure show DMB-HPLC analysis of the same material showing a Neu5Ac peak.
**PCR of sialic acid cluster genes**

In order to confirm the presence of the sialic acid metabolism pathway genes PCR analysis of pathogen, intermediate and saprophytic strain was performed. The reaction amplified genes in *L. interrogans* L1130 and Lai but not in *L. licerasiae*, *L. fainei* or *L. biflexa* (Fig 4.6). The control reaction using 16s primers showed DNA amplification in all the samples.

**NeuA2 southern blot**

To further confirm the presence or absence of sialic acid genes in different species of *Leptospira*, a probe based on the *neuA2* gene sequenced was used for southern blotting. Figure 4.6 shows binding for *L. interrogans* strains but not for *L. licerasiae* or *L. biflexa*. These findings support the previous evidence for sialic acid genes only in pathogenic species.

**Palmitate labeling**

Palmitate labeling using a previously described method showed wide incorporation of palmitate both in *L. interrogans* and *L. licerasiae*, albeit with a different pattern. Prominent band at ~25, ~32, ~45 and ~70 kDa can be seem in *L. interrogans*. In *L. licerasiae* both the ~25 ~45 kDa bands are missing. Western blots done in the same gels used for detection of the radioactive labelling show the presence of LipL32 and Loa22 in both species.
Figure 4.7 A) PCR of sialic acid cluster genes shows DNS amplification in pathogenic species. Viability of DNA is confirmed by amplification of the 16S gene. B) Genomic DNA of 1) *S. enterica*, 2) *L. interrogans* Lai, 3) *L. interrogans* L1130, 4) *L. biflexa* Patoc, 5) Rat isolate CEH 008 (*L. licerasiae*), 6) Wild rodent isolate MMD 4847 (*L. licerasiae*), 7) Wild rodent isolate MMD 3731 identified as serovar Icterohaemorrhagiae, 8) *L. interrogans* Evansi, 9) *S. enterica* stained using ethidium bromide in a agarose gel (top) and probed with a DIG-labeled anti-NeuA2 oligonucleotide probe.
Figure 4.8. A) Configuration of lipobox signal peptide, showing N-, H- and lipobox regions (adapted from Setubal, 2006). B) Comparison of lipoproteins from *L. interrogans* L1130 (upper case) and *L. licerasiae* (lower case). The positively charged N-terminal (green), the H-region (black) and the lipobox (red) differ but maintain function in LipL32, LipL41, LipL46 and LigB. LipL45 has a potential Lipobox (light blue). LipL21, LruB and LipL21 have lost the lipobox sequence.
Figure 4.9. Palmitate labeling and LipL32 Western blots in *L. interrogans* L1130 (top), and *L. licerasiae* (bottom). M = Marker, WB1,WB2 and P1, P2 = Western blots and palmitate labeling, two separate experiments. C = Coomassie stain.
Figure 4.10. Palmitate labeling and Loa22 Western blots in *L. interrogans* L1130 (top), and *L. licerasiae* (bottom). M = Marker. WB1, WB2 and P1, P2 = palmitate labeling, two separate experiments. C = Coomassie stain. *Contaminant band
Figure 4.11. A and B) Western blots using anti-Loa22 (A) and anti-LipL32 (B) antibodies, showing the slightly smaller size of the proteins in *L. licerasiae* (M= marker, 1= *L. interrogans*, 2= *L. licerasiae*. C) Shows two separate experiments of palmitate labeling showing different patterns of lipidation in *L. licerasiae*. 
Live immunofluorescence

Live immunofluorescence assays using anti-LipL32, -Loa22 and –Lipl71 antibodies and a rhodamine labeled secondary antibody show surface localization of the three proteins in *L. interrogans* L1130, while only showing surface localization of LipL32 and LipL71 in *L. licerasiae*. CFSE staining confirms the co-localization in the surface of the leptospires (Figure 4.12).
Figure 4.12. Live immuno-fluorescence staining using anti-Loa22, -LipL32 and -71 antibodies showing surface expression of LipL32 and Lip71 in *L. licerasiae*, but no Loa22.
4.4 Discussion

Several pathogens use sialic acid as part of a mechanism of immune evasion. Some of them do not encode endogenous biosynthetic pathways for Sia biosynthesis, but rather scavenge Sias from the host milieu for expression on their surfaces (Vimr and Lichtensteiger 2002). The presence of the Sia cluster genes in the published genome of *Leptospira interrogans* serovar Copenhageni L1-130 (Nascimento, Verjovski-Almeida et al. 2004) suggests that pathogenic *Leptospira* species could produce sialic acid. The *L. interrogans* L1-130 genome contains in a cluster two homologues each for genes encoding the enzymes UDP-N-acetylglucosamine 2-epimerase (neuC), N-acetylneuraminate synthase (neuB), and N-acetylneuraminate cytidyl transferase (neuA) (Figure 4.2), which correspond respectively to the group B *Streptococcus* (GBS) and *E. coli* K1 genes neuC, neuB and neuA, each essential for the bacterium’s de novo biosynthesis and surface expression of Sia (Vimr, Kalivoda et al. 2004). A nearly identical cluster is present in *L. interrogans* Lai. The contiguity of this cluster to the LOS cluster could be suggestive of its surface localization. A third protein containing a neuB domain, in a location independent of the above mention cluster is found in *L. interrogans* L1130 and has homologs in all the other known genomes. Its role is unknown.

In this analysis in order to determine definitively whether the Sia expressed by *L. interrogans* is synthesized by the organism, the spirochetes were grown in a chemically defined medium (Bey and Johnson 1978) containing no Sias, rather than the routine medium (EMJH) used for leptospiral culture. Despite some differences in the overall appearance of the HPLC profile, Neu5Ac is still clearly present in *L. interrogans* grown
in the absence of an exogenous Sia source. This experiment confirms that, like GBS and
*E. coli* K1 possessing the *neuA-C* gene cluster, *Leptospira interrogans* uses an
endogenous biosynthetic pathway for Neu5A expression.

The presence of the sialic acid cluster in *L. interrogans* but not in *L. licerasiae* or
*L. biflexa* support the hypothesis that leptospires also use it as an immune evasion
mechanism. The lack of a Sia cluster in *L. borgpetersenii* could suggest that sialic acid is
necessary during the initial phase of invasion, when leptospires go from the environment
to the host. Since *L. borgpetersenii* in transmitted from host-to-host, it would not use this
mechanism (Bulach, Zuerner et al. 2006).

All of the leptospiral genomes analyzed in this study contain a protein with a
*neuB* domain located away from the Sia cluster (Figure 4.11). These *neuB* homologs are
more similar to the *Legionella* homolog of *neuB* than to the other homologs of *neuB* in
*Leptospira*.

The wide distribution of these genetically distant *neuB* proteins suggests a more
functional role possibly similar to the one in *Legionella* in the production of flagellum-
associated legionaminic acid. Its location close to flagellum assembly genes supports its
having a role similar to the legionaminic acid producing enzymes in *Legionella* and
*Clostridium*. At this point not enough evidence is available to demonstrate this
hypothesis.

This study tried to identify the location of sialic acid in the leptospiral surface,
initially by affinity purifying sialylated cell components. The specificity of the affinity
assay can be supported by the fact that despite WGA’s relaxed specificity for Sias with a
more effective recognition of terminal N-acetylglucosamine, no additional bands were
purified compared to when compared to MAL and SNA. This fact suggests that terminal N-acetylglcosamine is not found on *L. interrogans*. In contrast to WGA, MAL and SNA are quite specific for Sias, but favor different linkages of Sias to an underlying structure - - MAL prefers a2→3-linked Sias while SNA prefers a2→6-linked Sias. This finding suggests that *L. interrogans* may modify surface structures with both a2→3- and a2→6-linked Sias. Specifically, Alternatively, bands at ~15, ~37, and ~41 kDa are much stronger in the SNA-purified sample, suggesting that the predominant form of Sias on *L. interrogans* may be a2→6-linked, a rare occurrence among bacterial pathogens (Figure 4.4).

Affinity purification followed by MS peptide identification identified three known surface proteins as potentially sialylated compounds: Loa22, LipL32 and LipL41. Loa22 was identified with the highest number of peptide matches. Loa22 is an outer membrane protein in found the genome of *L. interrogans* serovar Lai, *L. interrogans* serovar Manilae, *L. interrogans* L1130 and *L. borgpetersenii* serovar Hardjo-bovis L550. Koizumi and colleagues have suggested that Loa22 is a lipoprotein, based on its hydrophobicity plot, the presence of a signal peptide and a lipobox in the N-terminal of the predicted protein and positive palmitate labeling of a Loa22/phoA fusion protein in *E. coli*. (Koizumi 2003). In addition, recent data also shows that Loa22 is the only leptospiral protein that is observed to be upregulated in-vivo (Nally, Whitelegge et al. 2007) and it is the only leptospiral protein that has been shown to be necessary for virulence (Ristow, Bourhy et al. 2007). Previous published data had shown that *L. licerasiae* contained a smaller form of LipL32 (Ricaldi, Matthies et al. 2008).
Experiments described in this manuscript shown the presence of Loa22 in *L. licerasiae* by western blot but in a smaller size that the *L. interrogans* Loa22 (Fig 4.12).

From the over 600 unique genes in *L. interrogans* 565 are putative lipoproteins. The composition of the spirochetal lipoboxes in the lipoprotein signal peptide seem to be more variable than those in *E. coli* (Setubal, Reis et al. 2006). When analyzing the sequences of the know surface lipoproteins present in *L. licerasiae*, clear differences were seen in the N-terminal signal peptide while the rest of the protein sequence seemed to be more conserved. Further analysis revealed that the lipobox sequence had changed in all of them. When these sequences were analyzed using SpLip LipL21, Loa22, LipL21 and LruB had lost the criteria to be considered a lipobox. LipL 45 now had a possible lipobox sequence. These changes in sequence could determine changes in the membrane export and localization of these proteins and affect post-translational modifications including glycosylation. The differences found in palmitation patterns between *L. licerasiae* and *L. interrogans* are evidence of an effect produced by the loss of the lipobox.

Previous published data had shown that *L. licerasiae* contained a smaller form of LipL32 (Ricaldi, Matthias et al. 2008). Experiments described in this manuscript shown the presence of Loa22 in *L. licerasiae* by western blot but in a smaller size that the *L. interrogans* Loa22 (Fig 4.11). Based on the data presented it could be suggested that the smaller size of the *L. licerasiae* proteins might suggest a loss or change of post-translational modifications. This assertion is supported by the loss of the lipobox sequence that determines the lipidation of the lipoproteins (Figure 4.12).

Data from the live immuno-fluorescence assay show that *L. licerasiae* does not seems to express Loa22 on its surface or expresses it in considerably lesser amounts that
*L. interrogans*. Taking into account that Loa22 is the only proven virulence factor, it could be argued that it is the loss of the lipobox in *L. licerasiae* determines its loss of virulence.

With the data shown in this manuscript it is not possible to doubtlessly assert that lipoproteins are the sialylated component in the surface of *Leptospira*. Considering the close interaction of LPS and surface lipoproteins, the finding of minute amount of Sias in LPS could suggest a either lipid modifications associating and therefore “contaminating” LPS or that sialylated portions of LPS become attached to lipid modifications of lipoprotein and make them susceptible of lectin affinity purification.

Regardless of the location of sialic acid in the surface, these data add *L. interrogans* to the list of important human pathogens with surface expression of Neu5Ac, and prove the utility of phylogenetic comparison to predict Sia biosynthesis by the presence of a conserved disease cluster. In human evolution, a mutation in the gene encoding the enzyme CMP-N-acetylneuraminic acid hydroxylase led to loss of expression the common mammalian Sia N-glycolylneuraminic acid (Neu5Gc) and compensatory overexpression of Neu5Ac on all cell surfaces (Muchmore, Diaz et al. 1998). In addition to demonstrated roles in complement evasion, intracellular survival (Monteiro, Hirata et al. 2004), and biofilm formation (Jurcisek, Greiner et al. 2005), evidence is emerging that human pathogens with Neu5Ac on their surface engage Sia receptors (Siglecs) the leukocyte cell surface to further influence the host-bacterial interaction (Jones, Virji et al. 2003; Carlin, Lewis et al. 2007). Curiously, *L. interrogans* shares with several of the well characterized Neu5Ac-expressing human bacterial pathogens such as GBS, *N. meningitides*, *E. coli* K1, and *Haemophilus influenzae* a
certain neurotropism, with meningitis an important component of their clinical disease spectrum.

As more genetic tools and new animal models for study of *L. interrogans* are further refined, the analysis of the contribution of sialic acid biosynthesis and surface expression to the virulence of this pathogen should be more thoroughly investigated.

Chapter 4, in part is currently being prepared for submission for publication of the Material. Ricaldi, JN, Borkowsky A, Matthias MA, Nizet V, Vinetz, JM, Lewis AL. The dissertation author was the primary investigator and author of this material.


Sandwith, F. M. (1905). "The Medical Diseases of Egypt." 100-105


