LUISA ZANOLLI MORENO

**Caracterização e análise comparativa de genomas de estirpes de *Leptospira*
isoladas no Brasil.**

São Paulo

2017

LUISA ZANOLLI MORENO

Caracterização e análise comparativa de genomas de estirpes de *Leptospira* isoladas no Brasil.

Tese apresentada ao Programa de Pós-Graduação em Epidemiologia Experimental Aplicada às Zoonoses da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo para obtenção do Título de Doutor em Ciências.

Departamento:

Medicina Veterinária Preventiva e Saúde Animal

Área de concentração:

Epidemiologia Experimental Aplicada às Zoonoses

Orientador:

Profa. Dra. Andrea Micke Moreno

São Paulo

2017

Autorizo a reprodução parcial ou total desta obra, para fins acadêmicos, desde que citada a fonte.

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO

(Biblioteca Virginie Buff D'Ápice da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo)

T. 3476
FMVZ

Moreno, Luisa Zanolli
Caracterização e análise comparativa de genomas de estirpes de *Leptospira* isoladas no Brasil / Luisa Zanolli Moreno. -- 2017.
98 f. : il.

Tese (Doutorado) - Universidade de São Paulo. Faculdade de Medicina Veterinária e Zootecnia. Departamento de Medicina Veterinária Preventiva e Saúde Animal, São Paulo, 2017.

Programa de Pós-Graduação: Epidemiologia Experimental Aplicada às Zoonoses.

Área de concentração: Epidemiologia Experimental Aplicada às Zoonoses.

Orientador: Profa. Dra. Andrea Micke Moreno.

1. *Leptospira*. 2. Espécies. 3. Sequenciamento. 4. Análise genômica comparativa. I.
Título.

CERTIFICADO

Certificamos que o Projeto intitulado “Caracterização de estirpes de *Leptospira* isoladas no Brasil através de tipificação por absorção cruzada de aglutininas (CAAT), eletroforese em campo pulsado (PFGE) e sequenciamento de multilocus (MLST)”, protocolado sob o nº 3093/2013, utilizando 56 (cinquenta e seis) estirpes de *Leptospira*, sob a responsabilidade do(a) Profa. Dra. Andrea Micke Moreno, foi aprovado em reunião de 14/8/2013 e está de acordo com os princípios éticos de experimentação animal da Comissão de Ética no Uso de Animais da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo.

We certify that the Research “Characterization of *Leptospira* strains isolated in Brazil through cross agglutinin absorption (CAAT), pulsed field gel electrophoresis (PFGE) and multilocus sequencing typing (MLST)”, protocol number 3093/2013, utilizing 56 (fifty-six) *Leptospira* strains, under the responsibility Profa. Dra. Andrea Micke Moreno, was approved in the meeting of day 8/14/2013 and agree with Ethical Principles in Animal Research adopted by Ethic Committee in the Use of Animals of the School of Veterinary Medicine and Animal Science of University of São Paulo.

São Paulo, 14 de agosto de 2013.

Denise Tabacchi Fantoni
Presidente

FOLHA DE AVALIAÇÃO

Autor: MORENO, Luisa Zanolli

Título: Caracterização e análise comparativa de genomas de estirpes de *Leptospira* isoladas no Brasil

Tese apresentada ao Programa de Pós-Graduação em Epidemiologia Experimental Aplicada às Zoonoses da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo para obtenção do título de Doutor em Ciências

Data: ____/____/____

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Julgamento: _____

Prof. Dr. _____

Instituição: _____ Julgamento: _____

Prof. Dr. _____

Instituição: _____ Julgamento: _____

Prof. Dr. _____

Instituição: _____ Julgamento: _____

Prof. Dr. _____

Instituição: _____ Julgamento: _____

AGRADECIMENTOS

Aos meus pais pelo amor, dedicação e apoio incondicional durante todos esses anos de estudo.

À minha irmã pela companhia, compreensão e muitas risadas.

Ao Roberto por todo amor, carinho, companheirismo e paciência nesta vida que escolhemos compartilhar.

À Profa. Andrea Micke Moreno pela orientação, paciência e companheirismo durante todos esses anos de trabalho. A senhora me deu todas as possibilidades, sem as quais eu não chegaria onde estou.

Aos colegas do Laboratório de Sanidade Suína e agregados (Paula, Rose, Carlos, Thais, Gabi, Vasco, Bárbara, Ketrin, Fabiana, dona Rose, Alê) pelo companheirismo, ajuda e paciência.

À Zenáide e Gisele do Laboratório de Zoonoses Bacterianas por toda ajuda e paciência, e especialmente ao Prof. Silvio pela disponibilização da coleção de cultura e parceria.

Aos colegas do Laboratório de Bacteriologia Veterinária – UFF (Ana Paula Loureiro, Prof. Walter) e do Centro de Desenvolvimento Tecnológico – UFPel (Frederico, Marcus, Prof. Odir) pela parceria e colaboração.

À Miriam, Mari, Patty, Camila, Beth e Lícia pelo companheirismo e amizade durante todos esses anos.

À FAPESP pelo auxílio financeiro neste projeto de pesquisa.

RESUMO

MORENO LZ. **Caracterização e análise comparativa de genomas de estirpes de *Leptospira* isoladas no Brasil.** [Characterization and comparative genomic analysis of *Leptospira* strains isolated in Brazil]. 2017. 98 f. Tese (Doutorado em Ciências) – Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2017.

O presente estudo teve como objetivo caracterizar o genoma de estirpes de *Leptospira* isoladas no Brasil e realizar a análise comparativa destes com os genomas disponíveis no banco de dados *GenBank*. Foram caracterizadas 17 estirpes isoladas de distintas espécies animais, em diferentes regiões do Brasil, no período de 1998 a 2012. Estas foram previamente tipificadas por sequenciamento do gene 16S rRNA e soroglutinação microscópica em seis espécies (*L. interrogans*, *L. santarosai*, *L. inadai*, *L. kirschneri*, *L. borgpetersenii* e *L. noguchii*) e mais de oito sorogrupos. Foi realizado o sequenciamento em plataforma Illumina^{TM} MiSeq e montagem dos genomas com algoritmo *ab initio*. Para ordenação e anotação foram utilizados genomas de referência das respectivas espécies estudadas. Foi realizada a análise *in silico* da Tipagem por Sequenciamento de Multilocus (MLST) para os três protocolos vigentes de *Leptospira*. A análise comparativa dos genomas, incluindo wgSNP, foi realizada intra-espécie avaliando as variações existentes entre os sorogrupos das espécies de *Leptospira* estudadas. As estirpes de *L. interrogans* apresentaram resultados na MLST congruentes com a sua identificação prévia. No caso de *L. kirschneri*, apenas uma estirpe apresentou novos alelos nos três protocolos de MLST e se distancia das demais estirpes brasileiras de *L. kirschneri*. As estirpes de *L. santarosai*, assim como as de *L. borgpetersenii* e *L. noguchii*, possuem novos alelos e/ou perfis alélicos para pelo menos dois dos protocolos vigentes de MLST, sendo que ainda se destacam em um agrupamento próprio de origem brasileira. Os genomas de *L. interrogans* apesar de apresentarem alta identidade e sintenia com a referência sorovar Copenhageni, também apresentaram regiões de diferença entre os respectivos sorogrupos. Os genomas dos sorogrupos Australis e Serjoe se destacaram por apresentarem inserções e deleções, respectivamente, principalmente no cromossomo 2. O genoma de *L. borgpetersenii* também

apresentou grande variação de composição, como esperado para espécie, sendo esta proporcionada por sequências de inserção e transposição de elementos móveis. Os sorogrupos Canicola e Pomona apresentaram maior proximidade entre si na análise wgSNP. Também foram identificados dois plasmídeos nos genomas do sorogrupo Canicola com alta identidade aos plasmídeos descritos na estirpe chinesa do mesmo sorovar. Na espécie *L. kirschneri*, a estirpe 47 (M36/05) apresentou alta identidade e sintonia com os genomas do sorovar Mozdok, como esperado, incluindo a estirpe brasileira de origem humana. Já a estirpe 55 (M110/06) se diferenciou dos demais genomas de *L. kirschneri* tanto no MLST quanto no wgSNP. O genoma brasileiro de *L. inadai* apresentou alta identidade à referência americana de origem humana, incluindo a presença de um bacteriófago próprio da espécie. A distinção das estirpes brasileiras de *L. santarosai* na MLST, também foi evidenciada na análise comparativa e no wgSNP, sendo que a estirpe 68 (M52/8-19), que não apresentou reatividade aos sorogrupos testados, ainda se diferencia das demais reafirmando a possibilidade de novo sorogrupo/sorovar. Dessa forma, o estudo genômico possibilitou a identificação de particularidades das estirpes brasileiras de *Leptospira*, incluindo a existência de elementos extra-cromossomais, proximidade com estirpes de origem humana indicando maior risco para saúde pública, além da possibilidade de novo sorogrupo de *L. santarosai*.

Palavras-chave: *Leptospira*. Espécies. Sequenciamento. Análise genômica comparativa.

ABSTRACT

MORENO LZ. **Characterization and comparative genomic analysis of *Leptospira* strains isolated in Brazil.** [Caracterização e análise comparativa de genomas de estirpes de *Leptospira* isoladas no Brasil]. 2017. 98 f. Tese (Doutorado em Ciências) – Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2017.

The present study aimed to characterize the genome of *Leptospira* strains isolated in Brazil and to perform their comparative analysis with GenBank available genomes. 17 strains isolated from distinct species, in different regions of Brazil, from 1998 to 2012 were characterized. These were previously typified through 16S rRNA sequencing and microscopic agglutination into six species (*L. interrogans*, *L. santarosai*, *L. inadai*, *L. kirschneri*, *L. borgpetersenii* and *L. noguchii*) and over eight serogroups. Illumina™ MiSeq sequencing and genome assembly with *ab initio* algorithm were performed. For ordering and annotation, reference genomes of the respective species were used. The *in silico* analysis of Multilocus Sequencing Typing (MLST) was performed for the three current *Leptospira* protocols. The comparative genomic analysis, including wgSNP, was performed intra-species evaluating the existing variations between the serogroups of the studied *Leptospira* species. The *L. interrogans* strains presented MLST results congruent with their previous identification. In the case of *L. kirschneri*, only one strain presented new alleles in the three MLST protocols and distanced itself from the other Brazilian *L. kirschneri* strains. The *L. santarosai* strains, as well as *L. borgpetersenii* and *L. noguchii*, presented new alleles and/or allelic profiles for at least two of the current MLST protocols, and still stand out in a separate group of Brazilian origin. Even though the *L. interrogans* genomes presented high identity and synteny with serovar Copenhageni reference, they also presented regions of difference between the respective serogroups. Serogroups Australis and Serjoe genomes stood out for having insertions and deletions, respectively, mainly in chromosome 2. The *L. borgpetersenii* genome also presented great variation of composition, as expected for the species, which is provided by insertion sequences and transposition of mobile elements. The serogroups Canicola and Pomona presented higher proximity in the wgSNP analysis. Two plasmids were also identified in the serogroup

Canicola genomes with high identity to the plasmids described in the Chinese strain of the same serovar. In the *L. kirschneri* species, the strain 47 (M36/05) presented high identity and synteny with the serovar Mozdok genomes, as expected, including the Brazilian strain of human origin. The strain 55 (M110/06) differed from other *L. kirschneri* genomes in both MLST and wgSNP. The Brazilian *L. inadai* genome presented high identity to the American reference of human origin including the presence of bacteriophage specific for the species. The distinction of the Brazilian *L. santarosai* strains in the MLST was also evidenced in the comparative analysis and in the wgSNP, and the strain 68 (M52 / 8-19), which showed no reactivity to the tested serogroups, also differs from the others reaffirming the possibility of a new serogroup/serovar. Therefore, the genomic study allowed the identification of particularities of Brazilian *Leptospira* strains, including the existence of extrachromosomal elements, proximity to strains of human origin indicating a greater risk for public health, in addition to the possibility of a new *L. santarosai* serogroup.

Keywords: *Leptospira*. Species. Sequencing. Comparative genomic analysis.

LISTA DE FIGURAS

Figura 1 -	Árvore de máxima verossimilhança baseada nas sequências concatenadas do protocolo de MLST de Boonsilp et al. (2013).	41
Figura 2 -	Árvore de máxima verossimilhança baseada nas sequências concatenadas do protocolo de MLST de Varni et al. (2014).	42
Figura 3-	Árvore de máxima verossimilhança baseada nas sequências concatenadas do protocolo de MLST de Ahmed et al. (2006).	43
Figura 4 -	MSTs dos três protocolos vigentes de MLST (A - Boonsilp et al. (2013); B - Varni et al. (2014); C - Ahmed et al. (2006)) específicas para as estirpes de <i>L. interrogans</i> com destaque para os sorogrupos Canicola, Pomona, Australis e Sejroe.	44
Figura 5 -	Representação esquemática do alinhamento dos cromossomos 1 (A) e 2 (B) das estirpes de <i>L. interrogans</i> sorovar Canicola utilizando como referência a estirpe patogênica <i>L. interrogans</i> sorovar Copenhageni L1.130.	47
Figura 6 -	A - Representação esquemática do alinhamento das sequências semelhantes ao plasmídeo pGui1. B – Gráfico de sintenia das sequências semelhantes ao plasmídeo pGui2 (barras vermelhas correspondem a >99% identidade).	48
Figura 7 -	Representação esquemática do alinhamento dos cromossomos 1 (A) e 2 (B) das estirpes de <i>L. interrogans</i> sorovar Pomona utilizando como referência a estirpe patogênica <i>L. interrogans</i> sorovar Copenhageni L1.130.	50
Figura 8 -	Representação esquemática do alinhamento do cromossomo 1 das estirpes 86 (SU5), PigK151 e L1.130.	52
Figura 9 -	Representação esquemática do alinhamento do cromossomo 2 das estirpes 86 (SU5), PigK151 e L1.130.	53
Figura 10 -	Representação esquemática do alinhamento dos cromossomos 1 (A) e 2 (B) da estirpe 76 (OV5) e das referências <i>L. interrogans</i> sorovar Copenhageni e <i>L. interrogans</i> sorovar Lai.	55

Figura 11 -	Representação esquemática do alinhamento das estirpes de <i>L. noguchii</i> utilizando como referência a estirpe CZ214 (Panama) (linha preta interna).	57
Figura 12 -	Representação esquemática do alinhamento do cromossomo 1 das estirpes de <i>L. borgpetersenii</i> .	58
Figura 13 -	Gráfico de sintenia das sequências do cromossomo 2 das estirpes de <i>L. borgpetersenii</i> (barras vermelhas correspondem a >99% identidade).	59
Figura 14 -	Representação esquemática do alinhamento das estirpes de <i>L. kirschenri</i> .	61
Figura 15 -	Comparação circular dos genomas das estirpes de <i>L. inadai</i> , <i>L. broomii</i> e <i>L. fainei</i> .	63
Figura 16 -	Alinhamento múltiplo dos genomas das estirpes de <i>L. inadai</i> , <i>L. broomii</i> e <i>L. fainei</i> com definição de blocos de sequências homólogas.	63
Figura 17 -	Gráfico de sintenia da sequência semelhante ao bacteriófago LinZ_10 (barras azuis correspondem a sequências invertidas > 99 % identidade).	64
Figura 18 -	Comparação circular do cromossomo 1 das estirpes de <i>L. santarosai</i> .	65
Figura 19 -	Comparação do cromossomo 2 das estirpes de <i>L. santarosai</i> .	66
Figura 20 -	Representação esquemática da análise de wgSNP das estirpes de <i>L. interrogans</i> .	68
Figura 21 -	Representação esquemática da análise de wgSNP das estirpes de <i>L. santarosai</i> .	69
Figura 22 -	Representação esquemática da análise de wgSNP das estirpes de <i>L. kirschneri</i> .	71

LISTA DE TABELAS

Tabela 1 -	Informações de origem e identificação prévia das estirpes de <i>Leptospira</i> selecionadas para sequenciamento do genoma.	31
Tabela 2 -	Distribuição dos alelos e sequências tipo (STs) identificados nos genomas estudados, segundo o protocolo de MLST.	39

LISTA DE QUADROS

Quadro 1 -	Informações das montagens dos genomas dos isolados de <i>Leptospira</i> spp. (N50 e tamanhos médios e total em pares de bases - pb).	38
Quadro 2 -	Total de posições (SNPs) selecionadas para análise de wgSNP de <i>L. interrogans</i> .	67
Quadro 3 -	Total de posições selecionadas para análise de wgSNP de <i>L. santarosai</i> .	70
Quadro 4 -	Total de posições selecionadas para análise de wgSNP de <i>L. kirschneri</i> .	70

SUMÁRIO

1. INTRODUÇÃO	16
2. REVISÃO DE LITERATURA	
2.1 Histórico	18
2.2 Etiologia	19
2.3 Classificação das leptospiras	20
2.3.1 Classificação sorológica	20
2.3.2 Classificação genotípica	21
2.4 Epidemiologia	22
2.5 Análise genômica das leptospiras	25
3. OBJETIVOS	
3.1 Objetivo geral	29
3.2 Objetivos específicos	29
4. MATERIAL E MÉTODOS	
4.1 Estirpes de leptospiras	30
4.2 Extração do DNA genômico	32
4.3 Sequenciamento de nova geração	32
4.4 Avaliação da qualidade do sequenciamento	32
4.5 Montagem, ordenação e anotação dos genomas	33
4.6 Análise <i>in silico</i> da Tipagem por Sequenciamento de Multilocus (MLST)	34
4.7 Análise comparativa de genomas	35
4.8 Análise de wgSNP	35

5. RESULTADOS	
5.1 Análise inicial dos genomas	37
5.2 Análise <i>in silico</i> da MLST	37
5.3 Análise comparativa de genomas	45
5.3.1 <i>L. interrogans</i> sorovar Canicola	45
5.3.2 <i>L. interrogans</i> sorovar Pomona	49
5.3.3 <i>L. interrogans</i> sorogrupo Australis	51
5.3.4 <i>L. interrogans</i> sorovar Hardjo tipo Hardjoprajitno	54
5.3.5 <i>L. noguchii</i> sorogrupo Panama	56
5.3.6 <i>L. borgpetersenii</i> sorovar Ballum	57
5.3.7 <i>L. kirschneri</i>	60
5.3.8 <i>L. inadai</i> sorovar Lyme	62
5.3.9 <i>L. santarosai</i>	64
5.4 Análise de wgSNP	67
6. DISCUSSÃO	72
7. CONCLUSÕES	84
8. REFERÊNCIAS	86
9. ANEXOS	
Anexo 1	96
Anexo 2	97
Anexo 3	98

1. INTRODUÇÃO

A caracterização genética das leptospiiras é essencial para o melhor entendimento da epidemiologia e da patogênese da doença em humanos e nos animais. A partir do desenvolvimento dos métodos moleculares de tipificação, os sistemas de classificação baseados nas similaridades genéticas começaram a ser utilizados em conjunto com a classificação antigênica clássica. Os estudos genômicos aplicados às espécies de *Leptospira* visavam, inicialmente, a melhor caracterização das espécies patogênicas e sua diferenciação das saprofíticas, além da compreensão dos mecanismos moleculares da patogênese da leptospirose. Com o incremento da bioinformática, foi possível identificar novos genes de interesse para desenvolvimento de novos testes diagnósticos, além de novos candidatos a vacinas de amplo espectro.

Além disso, a análise genômica comparativa apresenta grande valor para evidenciar o perfil evolutivo e as inter-relações filogenéticas das leptospiiras isoladas nas diferentes regiões do país, bem como, favorecer o estudo epidemiológico da doença nos diferentes hospedeiros e no ambiente. A aplicação da epidemiologia genômica representa uma grande evolução para o estudo e compreensão desta zoonose.

A coleção de estirpes de *Leptospira* mantida pelo Laboratório de Zoonoses Bacterianas da Faculdade de Medicina Veterinária e Zootecnia, da Universidade de São Paulo, é uma das mais completas e antigas do país, sendo fonte de material para diferentes grupos de pesquisa do Brasil e de outros países. Dessa forma, a completa caracterização e preservação desta coleção são de extrema importância para o estudo sobre a leptospirose humana e animal. Além disso, foram agregadas a esta coleção estirpes isoladas pelos grupos de pesquisa em leptospirose da Universidade Federal Fluminense e do Centro de Desenvolvimento Tecnológico da Universidade Federal de Pelotas.

A proposta do presente estudo foi caracterizar pela análise genômica comparativa estirpes de *Leptospira* isoladas de distintas espécies animais, em diferentes regiões geográficas do Brasil. Estas estirpes foram previamente tipificadas pela técnica de aglutinação microscópica com anticorpos poli e monoclonais e pelo sequenciamento do gene 16S rRNA para determinação de espécie e sorovar. Dentre estas, destaca-se a identificação de dois sorovares autóctones do país, além de uma estirpe que apresentou resultados inconclusivos sugerindo a existência de novo sorogrupo e/ou sorovar. Tratam-se, portanto, de estirpes de *Leptospira* de grande interesse para o estudo da epidemiologia da leptospirose no país.

2. REVISÃO DE LITERATURA

2.1 Histórico

A primeira descrição do quadro clínico de icterícia com insuficiência renal foi feita por Adolf Weil na Alemanha, em 1886 (Weil, 1886), e a sua denominação de Doença de Weil foi proposta por Goldschmidt, em 1887 (Sanfort, 1984). No entanto, já existiam descrições de quadros semelhantes no início do século XIX, na China e no Japão, relacionados com risco ocupacional da colheita do arroz (Inada et al., 1916; Faine, 1994). Em 1907, foi demonstrada a presença de espiroquetas no túbulo renal de um paciente que foram denominadas *Spirochaeta interrogans* (Stimson, 1907). A etiologia da leptospirose foi corroborada, apenas em 1915, com a detecção de microrganismos com forma espiralada no sangue de mineradores japoneses com “icterícia infecciosa” que foram denominados de *Spirochaeta icterohaemorrhagiae* (Inada et al., 1916).

A forma saprofítica desse microrganismo foi descrita por Wolbach & Binger (1914) a partir de filtrado de água doce das margens de uma lagoa nas proximidades de Boston; este microrganismo foi denominado de *Spirocheta biflexa*. A partir de 1916, essas espiroquetas também passaram a ser isoladas de roedores silvestres na Europa e nos Estados Unidos da América (EUA) (Noguchi, 1918); no entanto, foi Ido et al. (1917) que comprovaram o papel dos roedores na transmissão da doença para humanos e que determinou o estabelecimento da leptospirose como uma doença zoonótica. O gênero *Leptospira* foi proposto por Noguchi (1918) agrupando os microrganismos descritos acima e diferenciando-os dos demais gêneros de espiroquetas descritos até o momento.

No Brasil, os primeiros relatos de leptospirose foram publicados por Beaurepaire de Aragão (1917) e Carini (1918) que demonstraram a presença do microrganismo em ratos naturalmente infectados. No entanto, o quadro clínico da Doença de Weil foi primeiramente reconhecido por McDowell, em 1911, no estado do Pará (Alexander, 1960). A partir das

publicações Beaurepaire de Aragão e Carini e do aprimoramento das pesquisas sobre a relação dos quadros clínicos da leptospirose e da febre amarela, os relatos de animais portadores de *Leptospira* se disseminaram pelo país (Enrietti, 2001).

2.2 Etiologia

Leptospira spp. pertence ao filo *Spirochaetes*, família *Leptospiraceae* e ao gênero *Leptospira* (Johnson & Faine, 1984). Trata-se de bactérias Gram-negativas, aeróbias obrigatórias e com crescimento ótimo em temperatura entre 28-30°C. Apresentam estrutura flexível, com forma helicoidal e móvel, medindo aproximadamente 0,1 µm de diâmetro e 6 a 20 µm de comprimento, sendo visualizadas por microscopia de campo escuro (Adler & Faine, 1983; Perolat et al., 1990).

Bioquimicamente são caracterizadas por serem catalase positivas, oxidase negativas e capazes de utilizar ácidos graxos de cadeia longa como a única fonte de carbono e energia. Tradicionalmente, as leptospiros são cultivadas em meio líquido Ellinghausen & McCulloch modificado por Johnson & Harris (EMJH) acrescido de albumina bovina e no meio de cultura semissólido Fletcher suplementado com soro de coelho. No cultivo semissólido observa-se uma concentração bacteriana abaixo da superfície, relacionada à tensão de oxigênio ótima para o crescimento, denominada de anel de Dinger. O crescimento é lento sendo que as culturas são observadas por 7 a 21 dias e, no caso de isolamento primário, os cultivos devem ser mantidos por até 13 semanas. Devido à resistência ao antimicrobiano 5'-fluorouracil, este pode ser adicionado ao meio seletivo para o cultivo de leptospiros a partir de fontes contaminadas (Faine et al., 1999; Adler & de La Peña Moctezuma, 2010).

2.3 Classificação das leptospiros

Nas últimas décadas, foram definidos dois tipos de classificação para o gênero *Leptospira*, a sorológica e a genotípica, sendo estas fundamentadas em conceitos distintos e que têm sido utilizadas conjuntamente.

2.3.1 Classificação sorológica

Tradicionalmente, os membros do gênero *Leptospira* são divididos em espécies patogênicas, *Leptospira interrogans* sensu lato, e não patogênicas, *Leptospira biflexa* sensu lato. Estas são classificadas em mais de 300 sorovares sendo estes definidos com base na heterogeneidade dos carboidratos do lipopolissacarídeo da membrana externa (LPS) (Brenner et al., 1999; Levett, 2001). Os sorovares antigenicamente relacionados podem ser agrupados ainda em sorogrupos, sendo atualmente descritos 29 sorogrupos que, apesar de não apresentarem legitimidade taxonômica, são considerados importantes para compreensão da epidemiologia da leptospirose (Levett, 2001; Adler & de La Peña Moctezuma, 2010).

A identificação em nível de sorovar pode ser realizada pela caracterização com anticorpos monoclonais (mAbs) ou pelo teste de absorção cruzada de aglutininas (CAAT). Ambos se baseiam no reconhecimento de padrões de antígenos característicos dos sorovares. Os mAbs permitem a identificação precisa de antígenos e a tipificação de um alto número de estirpes em pequeno período de tempo. Além disso, diferenças nos perfis de aglutinação podem ser indicativos da existência de novos sorovares (Terpstra et al., 1985; 1986).

A técnica aceita mundialmente para identificação e determinação de novos sorogrupos e sorovares é a CAAT. No entanto, esta não é realizada em rotina, pois exige a manutenção

de amplos painéis de antissoros e antígenos vivos, além de ser uma técnica que requer tempo e especialização para sua realização (Faine et al., 1999; Galloway & Levett, 2010).

2.3.2 Classificação genotípica

Na classificação genotípica, todos os sorovares acima descritos estão incluídos, no entanto, as espécies não correspondem apenas à leptospiros patogênicas e não patogênicas (*Leptospiras* sensu lato). Yasuda et al. (1987) iniciaram a diferenciação dos sorovares em espécies pelo método de hibridização DNA-DNA, sendo definidas 10 espécies de *Leptospira* a partir do agrupamento dos sorovares que apresentavam mais de 70% de similaridade genética.

Atualmente, a classificação genotípica é baseada, principalmente, na sequência dos genes que codificam o RNA ribossomal (rRNA) e possibilita a distinção de, até o momento, 22 espécies dentro do gênero *Leptospira* (Bourhy et al., 2014). Recentemente, o *International Committee on Systematics of Prokaryotes, Subcommittee on the Taxonomy of Leptospiraceae*, aprovou a substituição da hibridização por métodos baseados em sequenciamento, incluindo a filogenia de um gene ou combinação de genes e a avaliação *in silico* da similaridade genômica (Levett & Smythe, 2014). Apesar de ser uma tecnologia amplamente difundida, o sequenciamento ainda pode ser um desafio para muitos laboratórios brasileiros, não apenas tecnológico, mas também financeiro. Dessa forma, a classificação genotípica das leptospiros também não é realizada de rotina na maioria dos laboratórios do país.

A partir da análise filogenética, associada às respectivas características bioquímicas, as espécies de *Leptospira* também passaram a ser classificadas em patogênicas, intermediárias e saprófitas (Levett, 2001; Ko et al., 2009). Dessa forma, são reconhecidas 10 espécies

patogênicas no gênero (*L. interrogans*, *L. kirschneri*, *L. borgpetersenii*, *L. santarosai*, *L. noguchii*, *L. weilii*, *L. alexanderi*, *L. kmetyi*, *L. alstonii* e *L. mayottensis*) capazes de infectar e causar doença em humanos e animais; cinco espécies intermediárias (*L. inadai*, *L. broomii*, *L. fainei*, *L. wolffii* e *L. licerasiae*) que foram isoladas de infecção em humanos e animais; e sete espécies saprófitas (*L. biflexa*, *L. wolbachii*, *L. meyeri*, *L. vanthielii*, *L. terpstrae*, *L. idonii* e *L. yanagawae*) que foram isoladas de diversos ambientes e não estão associadas à doença em humanos ou animais.

2.4 Epidemiologia

A leptospirose é uma zoonose de distribuição mundial, de maior incidência nas regiões de clima tropical, que acomete grande variedade de animais silvestres e domésticos, além do homem (Faine et al., 1999). Os animais infectados são classificados como hospedeiros de manutenção (reservatórios) ou hospedeiros acidentais. Os hospedeiros de manutenção são caracterizados como espécies nas quais a infecção é endêmica e a transmissão, em geral, ocorre por contato direto. Estes são acometidos pela infecção crônica nos túbulos renais que determina a eliminação de leptospiros na urina, de forma contínua ou intermitente, e que garante a persistência da doença na natureza. Os hospedeiros acidentais, incluindo o homem, se infectam pelo contato indireto com os hospedeiros de manutenção (Levett, 2001; Haake & Levett, 2015).

A principal porta de entrada é a pele com abrasões ou cortes e a conjuntiva; no entanto, a infecção também pode ocorrer via pele intacta com imersão ou contato prolongado com água (Levett, 2001). Considerando a importância da contaminação ambiental com a urina dos hospedeiros de manutenção e o fator facilitador da água na disseminação da doença, a maior incidência da leptospirose passou a ser associada a países em desenvolvimento com grandes áreas urbanas caracterizadas por pauperismo e agravamento das condições de vida

(existência de favelas, alta proliferação de ratos, acúmulo de lixo) e, principalmente, pela ocorrência de enchentes. A variação sazonal, com as chuvas de verão ou monções, também apresenta grande importância para a manutenção e disseminação da leptospirose (Levett, 2001; Haake & Levett, 2015).

Apesar disso, a leptospirose é tradicionalmente considerada como uma doença profissional, sendo que veterinários, fazendeiros, trabalhadores de abatedouros, inspetores e trabalhadores em desratização apresentam maior risco de se infectar por contato direto com animais. Considerando a contaminação ambiental e a água como fatores de risco, a doença também pode apresentar maior risco para os trabalhadores em esgotos, mineiros, soldados, rizicultores, cortadores de cana, dentre outros, além de pessoas em atividades recreacionais, como canoagem, *rafting* e pesca em água doce (Levett, 2001; Haake & Levett, 2015).

Diversos animais podem servir como hospedeiros de manutenção; no entanto, os roedores apresentam maior importância na manutenção e disseminação da leptospirose. O rato de esgoto ou ratazana (*Rattus norvegicus*) é considerado o principal reservatório de *Leptospira* em áreas urbanas no mundo (Levett, 2001; Haake & Levett, 2015). No Brasil, também se destaca o papel da capivara como hospedeiro de manutenção de diferentes espécies e sorovares de *Leptospira* em áreas rurais e suburbanas (Marvullo et al., 2009; Moreno et al., 2016a).

Algumas espécies animais podem servir de hospedeiro de manutenção para determinados sorovares de *Leptospira*, mas como hospedeiros acidentais de outros. Por exemplo, tradicionalmente, os ratos são considerados como reservatórios dos sorovares dos sorogrupos Icterohaemorrhagiae e Ballum; os bovinos são mais associados ao sorovar Hardjo, enquanto que os sorovares Pomona e Bratislava estão mais relacionados aos suínos, e o sorovar Canicola é mais descrito em cães (Bolin, 2000). No entanto, Hartskeerl & Terpstra (1996) já discutiam a existência de variação de hospedeiros de manutenção e

respectivos sorovares pelo mundo. A variação de espécies e sorovares de *Leptospira* nos animais domésticos continua a ser descrita (Lilenbaum et al., 2007; Cunha et al., 2016; Miotto et al., 2016) e demanda atenção para o melhor entendimento da epidemiologia da doença e seu controle e prevenção.

Também é observada a variação da incidência das espécies de *Leptospira* pelo mundo. *L. interrogans* e *L. borgpetersenii* são as espécies mais frequentes a nível mundial, sendo que *L. interrogans* sorogrupo Icterohaemorrhagiae é considerado de maior importância para saúde pública (Wang et al., 2015). Na África e no sudeste Asiático é descrito predomínio de *L. borgpetersenii*, *L. interrogans* e *L. kirschneri*, tanto em animais quanto em humanos (Desvars et al., 2012; Cosson et al., 2014; Allan et al., 2015). *L. santarosai* também apresenta importância como patógeno humano e animal na América Latina (Carmona-Gasca et al. 2011; Valverde et al. 2013) e, principalmente, no território asiático (Hsieh & Pan, 2004; Chou et al. 2012).

No Brasil, *L. interrogans* sorogrupo Icterohaemorrhagiae sorovares Icterohaemorrhagiae e Copenhageni são considerados os principais causadores de leptospirose urbana (McBride et al., 2005). O sorogrupo Sejroe, que inclui as espécies *L. borgpetersenii*, *L. interrogans* e *L. santarosai*, é descrito como mais frequente em bovinos, bubalinos e pequenos ruminantes (Vasconcellos et al., 2001; Sarmiento et al., 2012; Monte et al., 2015), enquanto que em suínos predomina *L. interrogans* sorovares Pomona, Icterohaemorrhagiae e Tarassovi (Bolin, 1994; Ferreira Neto et al., 1997; Ramos et al., 2006). No entanto, as espécies *L. noguchii* e *L. kirschneri* também já foram descritas no país ocasionando doença em animais e humanos (Silva et al., 2009; Martins et al., 2015; Cunha et al. 2016).

Apesar de a leptospirose ser uma doença de notificação compulsória no Brasil desde 1985, a incidência notificada da doença depende da suspeição e diagnóstico clínico e,

principalmente, do diagnóstico laboratorial. Considerando que a doença é frequentemente tratada como “febre de origem desconhecida”, responsiva à antibióticoterapia, e as dificuldades de acesso e acurácia do diagnóstico laboratorial, a incidência da leptospirose é frequentemente subestimada no país e no mundo (Levett, 2001; McBride et al., 2005; Haake & Levett, 2015).

2.5 Estudo genômico das leptospiras

A caracterização genômica e sua aplicação na epidemiologia genômica objetivam ampliar o horizonte do estudo da patogênese, evolução e epidemiologia da leptospirose. O sequenciamento e publicação de genomas de *Leptospira* foram iniciados recentemente. A dificuldade de cultivo e manutenção das estirpes, e consequente dificuldade de concentração de DNA, são fatores limitantes e desafios para o sequenciamento de genomas de *Leptospira* (Adler & de La Peña Moctezuma, 2010).

O gênero *Leptospira* é caracterizado por dois cromossomos circulares com 35–41% de conteúdo GC e tamanho de 3,9–4,6 Mb, dependendo da espécie. Até o momento, apenas as espécies *L. biflexa*, *L. borgpetersenii*, *L. interrogans* e *L. santarosai* possuem genomas de referência publicamente disponíveis. O estudo genômico de *Leptospira* se iniciou com Ren et al. (2003) e Nascimento et al. (2004) que publicaram os genomas de *L. interrogans* sorovar Lai e sorovar Copenhageni, respectivamente, enquanto Bulach et al. (2006) publicaram os genomas de *L. borgpetersenii* sorovar Hardjo-bovis estirpes L550 e JB197.

Além da dificuldade para obter material suficiente para preparo do DNA genômico, os autores ainda tinham o desafio da tecnologia disponível para o sequenciamento. Os genomas de *L. interrogans*, *L. borgpetersenii*, e até mesmo de *L. biflexa*, foram obtidos a partir do sequenciamento por *shotgun* utilizando biblioteca de clonagem em plasmídeos ou

cosmídeos. Além de ser uma técnica de processamento complexo, o sequenciamento por *shotgun* apresenta problemas de redundância e cobertura (Bouck et al., 1998), o que justifica a baixa cobertura (<10 X) e as dificuldades de montagem observadas nesses estudos. As diferentes metodologias para anotação dos respectivos genomas também foi um fator limitante e que pode ter contribuído para as variações observadas, principalmente em relação ao número de genes codificadores de proteínas hipotéticas (função desconhecida) (Nascimento et al., 2004; Adler et al., 2011).

No entanto, apesar da problemática de cobertura, montagem e anotação, esses estudos já possibilitaram a observação de importantes variações de tamanho e estrutura dos genomas das estirpes de *L. borgpetersenii* e *L. interrogans*. Bulach et al. (2006) caracterizaram a redução genômica de *L. borgpetersenii* que parece estar relacionada a sua adaptação a ambientes e hospedeiros diversos. Nascimento et al. (2004) fizeram a primeira análise genômica comparativa entre sorovares de uma mesma espécie, e descreveram uma inversão cromossômica e grande variação em genes estruturais hipotéticos e sequências de inserção.

Até então, esses estudos focavam, principalmente, nas espécies patogênicas para análise comparativa e compreensão dos mecanismos moleculares da patogênese da leptospirose. A partir da publicação dos genomas de estirpes saprófitas e intermediárias, *L. biflexa* sorovar Patoc (Picardeau et al., 2008) e *L. licerasiae* sorovar Varillal (Ricaldi et al., 2012), foram disponibilizadas mais informações para a análise da evolução das espécies de *Leptospira* e das diferenças entre as espécies e os sorovares, sejam estes patogênicos, saprófitas ou intermediários (Nascimento et al., 2004; Ricaldi et al., 2012).

Assim, foi observado que o genoma da *L. biflexa* é mais estável que o das espécies patogênicas que apresentam grande plasticidade genômica, mesmo entre sorovares da mesma espécie (Nascimento et al., 2004; Bulach et al., 2006; Picardeau et al., 2008).

Também foi comprovado que a espécie intermediária *L. licerasiae* é geneticamente mais próxima das espécies patogênicas do que da *L. biflexa* (Ricaldi et al., 2012).

Mais recentemente, com a disponibilização de *drafts*, genomas na forma de *contigs* e não em uma única sequência de cromossomo, de diversas estirpes de *Leptospira* foi possível ampliar a análise comparativa entre as espécies. Dessa forma, Xu et al. (2016) e Fouts et al. (2016) conseguiram descrever a evolução das linhagens de *Leptospira* (saprófita, intermediária e patogênica). O aparecimento das espécies intermediárias e patogênicas foi associado a eventos de aquisição de genes, caracterizando uma expansão genômica comparada às espécies saprófitas, e a plasticidade ou variação gênica parece ainda estar relacionada à diversidade dentre as espécies patogênicas estudadas.

No entanto, os desafios para utilização da análise genômica comparativa na caracterização do gênero *Leptospira* ainda persistem. A falta de genomas de referência para todas as espécies e sorovares e a problemática da qualidade dos *drafts* disponíveis comprometem etapas essenciais para finalização e comparação dos genomas. O alto número de genes hipotéticos nas anotações dos genomas de *Leptospira* também pode afetar a análise comparativa e evolutiva das estirpes.

Além disso, a maior parte dos estudos descritos acima enfoca na caracterização das diferenças e da evolução das espécies de *Leptospira*, e não entre os respectivos sorogrupos/sorovares. Esses resultados também foram pouco aproveitados para o estudo da epidemiologia genômica do agente que possibilitaria a melhor caracterização da diversidade e dispersão das populações (sorovares/sorogrupos/ espécies) do gênero *Leptospira*.

No Brasil, o estudo genômico de *Leptospira* se iniciou com Nascimento et al. (2004) a partir da caracterização de *L. interrogans* sorovar Copenhageni estirpe L1.130, que é utilizada até hoje em ensaios de virulência e desafio vacinal no país. Os estudos seguintes

foram desenvolvidos por grupos de pesquisa do Instituto Butantan e da Fundação Oswaldo Cruz, sendo que estes apresentavam com maior foco na caracterização proteômica e aplicação para vacinologia reversa (Vieira et al., 2009; Santo, 2015; Fernandes et al. 2016). Mais recentemente, a Universidade Federal de Pelotas também aderiu ao estudo genômico de *Leptospira* com o sequenciamento e caracterização da estirpe brasileira de *L. borgpetersenii* sorogrupo Ballum (Eslabão, 2012). No entanto, esses estudos se limitaram a avaliar um pequeno número de espécies de *Leptospira* e não incluíam os sorovares autóctones do país. Além disso, esses dados também não foram aplicados no contexto epidemiológico da leptospirose no Brasil.

3. OBJETIVOS

3.1 Objetivo geral

Caracterizar estirpes de leptospiros isoladas de distintas espécies animais, em diferentes regiões geográficas do Brasil, no período de 1998 a 2012, pelo sequenciamento de nova geração e análise comparativa dos genomas.

3.2 Objetivos específicos

Os objetivos específicos são:

- Sequenciar o genoma de grupo selecionado de estirpes;
- Realizar a montagem e anotação dos genomas obtidos;
- Realizar a análise *in silico* da Tipagem por Sequenciamento de Multilocus (MLST);
- Realizar a análise comparativa dos genomas das estirpes estudadas com os genomas previamente publicados no banco de dados *GenBank*.

4. MATERIAL E MÉTODOS

4.1 Estirpes

Foram caracterizadas 17 estirpes de *Leptospira*, pertencentes à coleção de cultura de *Leptospira* que integra estirpes da Universidade de São Paulo, Universidade Federal Fluminense e Universidade Federal de Pelotas. As estirpes (Tabela 1) foram previamente tipificadas pela técnica de aglutinação microscópica com anticorpos monoclonais (mAbs), no Centro de Referência de Leptospirose da Organização Mundial da Saúde – *Royal Tropical Institute*, Amsterdã/Holanda, e anticorpos policlonais no Laboratório de Bacteriologia Veterinária da Universidade Federal Fluminense. Estas também foram previamente caracterizadas pela análise do número variável de sequências repetitivas em tandem (VNTR) (Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas) e/ou pelo sequenciamento do gene 16S rRNA (Universidade de São Paulo) para identificação das respectivas espécies. Apenas uma estirpe apresentou resultados inconclusivos sugerindo a existência de novo sorogrupo e/ou sorovar.

As estirpes encontravam-se em meio de cultura semissólido em estoque a temperatura ambiente - meio Fletcher (Difco) adicionado de 10% de soro de coelho. Dessa forma, as estirpes foram reativadas em meio Ellinghausen-McCullough-Johnson-Harris (EMJH - Difco), sendo mantidas a 30°C por sete dias. As estirpes foram reavaliadas quanto a sua pureza por observação em microscópio de campo escuro.

Foram sequenciados representantes das espécies e sorogrupos da coleção de cultura de *Leptospira*, incluindo as estirpes com perfil genotípico e fenotípico atípicos, além das estirpes que apresentaram resultados inconclusivos para identificação de espécie e/ou sorovar.

Tabela 1: Informações de origem e identificação prévia das estirpes de *Leptospira* selecionadas para sequenciamento do genoma.

Estirpe	Espécie de origem	Estado de origem	Ano de isolamento	Espécie/ Sorogrupo ¹
10 (M4/98)	Bubalino - <i>Bubalus bubalis</i>	SP	1998	<i>L. santarosai</i> / Sejroe ²
22 (LO-3)	Suíno – <i>Sus scrofa domestica</i>	PR	2000	<i>L. interrogans</i> / Canicola
23 (LO-4)	Suíno – <i>Sus scrofa domestica</i>	PR	2000	<i>L. interrogans</i> / Canicola
24 (LO-9)	Bovino - <i>Bos Taurus</i>	PR	2001	<i>L. santarosai</i> / Grippotyphosa
27 (2ACAP)	Capivara – <i>Hydrochaeris hydrochaeris</i>	SP	2001	<i>L. santarosai</i> / Grippotyphosa
34 (M34/99)	Roedor- <i>Rattus norvegicus</i>	SP	1999	<i>L. inadai</i> / Lyme
37 (GR5)	Suíno – <i>Sus scrofa domestica</i>	SP	2004	<i>L. interrogans</i> / Pomona
39 (LPF ³)	-	SP	-	<i>L. interrogans</i> / Pomona
47 (M36/05)	Roedor- <i>Rattus norvegicus</i>	SP	2005	<i>L. kirschneri</i> / Pomona
52 (M72/6-6)	Caprino - <i>Capra hircus</i>	RJ	2006	<i>L. santarosai</i> / Grippotyphosa
55 (M110/06)	Cachorro do Mato	MS	2006	<i>L. kirschneri</i> / Pomona
62 (LO-24)	Roedor- <i>Rattus norvegicus</i>	PR	2011	<i>L. borgpetersenii</i> / Ballum
68 (M52/8-19)	Capivara – <i>Hydrochaeris hydrochaeris</i>	SP	2008	<i>L. santarosai</i> / *
71 (UCR5RB1)	Equino - <i>Equus caballus</i>	RJ	2012	<i>L. kirschneri</i> / Grippotyphosa
72 (U73)	Bovino - <i>Bos taurus</i>	RJ	2012	<i>L. noguchii</i> / Panama
76 (OV5)	Ovino - <i>Ovis aries</i>	RJ	2012	<i>L. interrogans</i> / Serjoe
86 (SU5)	Suíno – <i>Sus scrofa domestica</i>	RJ	-	<i>L. interrogans</i> / Australis

¹ Espécie e sorogrupo determinados em tipagem prévia. ² Reagente igualmente para os sorovares Hardjoprajitno e Guaricura. ³ Estirpe de referência de *L. interrogans* sorogrupo Pomona sorovar Kennewicki estirpe Fromm. * Não apresentou reatividade a nenhum dos sorogrupos testados.

4.2 Extração do DNA genômico

O DNA genômico foi extraído utilizando-se o *illustra™ bacteria genomicPrep Mini Spin Kit* (GE Healthcare) seguindo as instruções do fabricante. A concentração e a qualidade do DNA foram avaliadas com NanoDrop™ 2000c (Thermo Scientific) por leitura em espectrofotômetro a 260 nm e cálculo da razão 260/280 (Sambrook & Russel, 2002), respectivamente.

4.3 Sequenciamento de nova geração

O sequenciamento dos genomas foi realizado no Laboratório Multiusuários Centralizado de Genômica Funcional Aplicada a Agropecuária e Agroenergia (ESALQ-USP) sob coordenação do professor Luiz Lehmann Coutinho. As bibliotecas foram preparadas com o *Nextera™ DNA Sample Prep Kit* (Illumina™) e o sequenciamento *paired-end* de fragmentos de 300 pares de bases (pb) foi realizado utilizando a plataforma Illumina™ MiSeq.

4.4 Avaliação da qualidade do sequenciamento

Para a avaliação da qualidade do sequenciamento foi utilizado o programa FastQC 0.10.1 (Babraham Bioinformatics). A partir deste foi realizada a avaliação da confiabilidade da identificação de cada base sequenciada (*base-calling*), pelo valor de qualidade Phred, para determinar a qualidade das sequências (*reads*) obtidas ao longo de sua extensão. Também foi avaliada a distribuição de frequências por tamanho das *reads*, o conteúdo GC e os níveis de duplicação e de sequências redundantes para determinar a necessidade de

“trimar” (cortar) às *reads*, conforme seu tamanho e a qualidade das respectivas bases. Foi adotado o critério de seleção de sequências maiores que 100 pb e qualidade de base acima de 20 (valor de Phred) para a continuidade da análise e montagem dos genomas. A cobertura estimada do sequenciamento foi obtida a partir do cálculo:

$$C = \frac{N^{\circ} \text{ reads } \times \text{ tamanho fragmento}}{\text{tamanho do genoma}}$$

considerando o total de *reads* após seleção por qualidade e tamanho, o comprimento médio das *reads* e o tamanho esperado do genoma.

4.5 Montagem, ordenação e anotação dos genomas

Para a montagem *ab initio* dos genomas, foram utilizados os programas CLC Genomics Workbench 7.5.1 (CLC Bio) e Geneious R10 (Biomatters Ltd). Foram utilizados os seguintes parâmetros em ambos os algoritmos: tamanho mínimo de *contig* de 1000 pb, cobertura mínima de cinco sequências para fusão de variantes, detecção automática da distância de pares, além de permitir *gaps* (lacunas). A avaliação da qualidade das respectivas montagens foi feita a partir dos valores de número de *contigs* obtidos, N50, tamanho médio dos *contigs* e tamanho total da montagem.

A ordenação dos *contigs* foi realizada com os programas *Mauve Contig Mover* (Rissman et al. 2009) e *CONTIGuator web server* (Galardini et al., 2011) utilizando como referência os genomas disponíveis de *L. interrogans* sorovar Copenhageni (NC_005823.1, NC_005824.1), *L. borgpetersenii* sorovar Hardjo-bovis (NC_008508.1, NC_008509.1), *L. santarosai* serovar Shermani (CP006694, CP006695), *L. interrogans* sorovar Bratislava (CP011410 e CP011411), *L. noguchii* sorovar Panama estirpe CZ214

(NZ_AKWY000000000), *L. inadai* sorovar Lyme estirpe 10 (NZ_AHMM000000000.2) e *L. kirschneri* sorovar Valbuzzi estirpe 200702274 (NZ_AHOC000000000). Os *contigs* que não alinharam com os cromossomos foram comparados com o banco de dados não redundantes do *GenBank* (NCBI) em busca de possíveis elementos extra-cromossomais.

A anotação automática dos genomas foi realizada pelo servidor BASys - *Bacterial Annotation System* (van Domselaar et al., 2005) e pela ferramenta *Prokaryotic Genome Annotation Pipeline* do NCBI (Tatusova et al., 2016). A curadoria das respectivas anotações foi feita manualmente com o programa Artemis (Carver et al., 2012) e a ferramenta BLASTp (*Basic Local Alignment Search Tool* para comparação de sequências de aminoácidos) do banco de dados *UniProt* (*UniProt Consortium*) para confirmação das sequências codificadoras (CDSs) identificadas e dos respectivos genes.

4.6 Análise *in silico* da Tipagem por Sequenciamento de Multilocus (MLST)

A predição dos alelos e sequências tipo (ST) da Tipagem por Sequenciamento de Multilocus (MLST) foi realizada pela ferramenta *MLST 1.8 server* (Larsen et al., 2012). Esta possibilita a busca dos genes dos três protocolos vigentes de MLST para *Leptospira* (Ahmed et al., 2006; Boonsilp et al., 2013; Varni et al., 2014) disponíveis no PubMLST (*Public databases for molecular typing and microbial genome diversity* - <http://pubmlst.org/>).

Para os novos alelos identificados, foi realizada a verificação manual e confirmação da cobertura e qualidade das mutações identificadas para determinação final dos alelos e submissão destes e dos novos perfis alélicos para os respectivos bancos de dados para a identificação das novas STs. Também foi realizada a análise filogenética das sequências concatenadas dos respectivos genes pelo método de máxima verossimilhança utilizando o

programa Bionumerics 7.6 (Applied Maths) para construção das respectivas árvores filogenéticas.

4.7 Análise comparativa de genomas

Para a análise comparativa dos genomas foram feitos alinhamentos múltiplos, gráficos de sintenia e comparação do conteúdo gênico e proteico. Para tanto, foram utilizados os programas Mauve (Darling et al., 2010) e ACT - *Artemis Comparison Tool* (Carver et al., 2005), além dos algoritmos BLASTn e BLASTp. O programa BRIG - *BLAST Ring Image Generator* (Alikhan et al., 2011) foi utilizado para gerar as imagens de comparação circular dos genomas e plasmídeos estudados.

Devido à grande diferença entre os genomas das distintas espécies do gênero *Leptospira*, optou-se por fazer a análise comparativa intra-espécie avaliando as variações existentes entre os sorogrupos e sorovares das espécies de *Leptospira* estudadas. A comparação foi feita com os genomas disponíveis no banco de dados de *Whole Genome Shotgun* (WGS) do NCBI - *National Center for Biotechnology Information*.

4.8 Análise de wgSNP

A análise de *whole genome single-nucleotide polymorphism* (wgSNP) foi realizada a partir do programa Bionumerics 7.6 (Applied Maths) apenas para as espécies *L. interrogans*, *L. kirschneri* e *L. santarosai*, por apresentarem maior número de estirpes estudadas. Inicialmente foi realizado o mapeamento das *reads* das respectivas estirpes com os genomas de referência: *L. interrogans* sorovar Copenhageni estirpe L1.130 (NC_005823.1, NC_005824.1), *L. santarosai* sorovar Shermani estirpe LT821

(NZ_CP006694.1, NZ_CP006695.1), *L. kirschneri* sorovar Valbuzzi estirpe 200702274 (NZ_AHOC000000000) e *L. kirschneri* sorovar Mozdok estirpe Brem 166 (NZ_AHQI000000000.1). Para tanto foram utilizados os parâmetros padrões do algoritmo de mapeamento do programa Bionumerics 7.6 (Applied Maths).

Para a identificação e seleção dos SNPs foram estabelecidos os seguintes filtros: distância mínima inter-SNP de 12 pb, cobertura mínima de 5 (mínimo de 1 Fw e 1 Rv), remoção de posições com pelo menos uma base ambígua ou não confiável e remoção de posições com pelo menos um *gap*. A partir dos SNPs selecionados foi montada uma matriz de SNPs que foi utilizada para a construção de um dendrograma. Para tanto, os SNPs foram considerados como dados categóricos e um coeficiente de similaridade “multi-estado” (*categorical - differences*) foi utilizado para a construção do dendrograma.

5. RESULTADOS

5.1 Análise inicial dos genomas

Após o recebimento dos arquivos do sequenciamento, as *reads* foram filtradas seguindo os padrões mínimos estabelecidos para tamanho e qualidade das sequências. A partir destas, foi realizada montagem *ab initio* dos genomas e a ordenação dos *contigs* obtidos. As informações de número de *contigs* obtidos, N50, tamanho médio dos *contigs* e tamanho total das respectivas montagens estão apresentadas no Quadro 1. Considerando a variação observada de 1,1 a 1,5 milhão de *reads* por estirpe, a cobertura estimada foi de aproximadamente 50 a 70x, considerando a espécie da estirpe sequenciada.

5.2 Análise *in silico* da MLST

Os perfis alélicos obtidos pela análise *in silico* da MLST são apresentados na Tabela 2. Os novos alelos e perfis alélicos identificados foram submetidos para os respectivos bancos de dados para a determinação das novas STs. Os novos números de alelos e STs que já foram atualizados nos respectivos bancos encontram-se destacados em vermelho.

No caso da estirpe 34 (M34/99), como os protocolos vigentes de MLST para *Leptospira* são restritos às espécies patogênicas (*L. interrogans*, *L. kirschneri*, *L. borgpetersenii*, *L. noguchii*, *L. santarosai*, *L. weilii* e *L. alexanderi*) estes, portanto, não se aplicam à espécie *L. inadai*.

Quadro 1: Informações das montagens dos genomas dos isolados de *Leptospira* spp. (N50 e tamanhos médios e total em pares de bases - pb).

Estirpe	Espécie	Sorogrupo/sorovar	Nº de <i>contigs</i>	N50	Tamanho médio	Tamanho total
22 (LO-3)	<i>L. interrogans</i>	Canicola / Canicola	319	24.742	14.749	4.704.942
23 (LO-4)	<i>L. interrogans</i>	Canicola / Canicola	49	200.400	96.916	4.748.911
37 (GR5)	<i>L. interrogans</i>	Pomona / Pomona	73	121.968	62.934	4.594.183
39 (LPF)	<i>L. interrogans</i>	Pomona / Pomona	406	19.349	11.270	4.575.676
76 (OV5)	<i>L. interrogans</i>	Sejroe / Hardjo*	46	202.280	103.828	4.776.130
86 (SU5)	<i>L. interrogans</i>	Australis /	133	72.764	36.906	4.908.622
72 (U73)	<i>L. noguchii</i>	Panama /	279	24.125	17.093	4.769.122
62 (LO-24)	<i>L. borgpetersenii</i>	Ballum / Ballum	37	175.012	107.482	3.964.077
71 (UCR5RB1)	<i>L. kirschneri</i>	Grippotyphosa /	89	95.650	48.875	4.341.139
47 (M36/05)	<i>L. kirschneri</i>	Pomona / Mozdok	44	241.247	80.262	4.460.519
55 (M110/06)	<i>L. kirschneri</i>	Pomona /	78	101.783	59.324	4.616.844
10 (M4/98)	<i>L. santarosai</i>	Sejroe / ¹	68	122.500	59.900	4.073.226
24 (LO-9)	<i>L. santarosai</i>	Grippotyphosa/Grippotyphosa	96	104.964	43.164	4.143.777
27 (2ACAP)	<i>L. santarosai</i>	Grippotyphosa/ Bananal	44	204.237	93.180	4.099.932
52 (M72/6-6)	<i>L. santarosai</i>	Grippotyphosa/ Grippotyphosa	59	200.288	69.688	4.111.639
68 (M52/8-19)	<i>L. santarosai</i>	-	44	163.662	95.948	4.221.754
34 (M34/99)	<i>L. inadai</i>	Lyme / Lyme	62	163.633	73.827	4.577.281

¹ Sorovar Hardjoprajitno / Guaricura. * *L. interrogans* sorogrupo Sejroe sorovar Hardjo tipo Hardjoprajitno.

Tabela 2: Distribuição dos alelos e sequências tipo (STs) identificados nos genomas estudados, segundo o protocolo de MLST.

Estirpe	Ahmed et al., 2006 ¹							Boonsilp et al. (2013) ²								Varni et al. (2014) ³							
	<i>adk</i>	<i>icdA</i>	<i>lipL32</i>	<i>lipL41</i>	<i>rrs2</i>	<i>secY</i>	ST	<i>caiB</i>	<i>glmU</i>	<i>mreA</i>	<i>pfkB</i>	<i>pntA</i>	<i>sucA</i>	<i>tpiA</i>	ST	<i>adk</i>	<i>glmU</i>	<i>icdA</i>	<i>lipL32</i>	<i>lipL41</i>	<i>mreA</i>	<i>pntA</i>	ST
22 (LO-3)	2	1	1	1	2	6	34 [▲]	5	3	5	4	3	3	3	37 [▲]	3	2	3	3	3	1	3	3 [▲]
23 (LO-4)	2	1	1	1	2	6	34 [▲]	5	3	5	4	3	3	3	37 [▲]	3	2	3	3	3	1	3	3 [▲]
37 (GR5)	2	2	3	4	2	6	58 [▲]	16	3	5	4	3	3	3	140 [▲]	3	2	2	10	4	1	3	52 [▲]
39 (LPF)	2	2	3	4	2	6	58 [▲]	16	3	5	4	3	3	3	140 [▲]	3	2	2	10	4	1	3	52 [▲]
76 (OV5)	5	2	2	3	2	3	74	3	1	4	7	1	2	1	20 [▲]	4	1	2	2	2	4	2	6
86 (SU5)	2	2	2	2	1	5	49	6	6	2	13	13	2	2	51	3	7	2	2	5	9	7	98
72 (U73)	62	72	7	43	24	52	137	34	59	63	82	70	65	44	231	55	32	45	20	38	39	36	97
62 (LO-24)	NA	61	10	15	20	49	Nv ST	12	24	26	67	32	30	36	149 [▲]	56	26	46	34	50	36	37	99
71 (UCR5RB1)	13	24	11	16	12	22	94 [▲]	23	19	18	31	20	13	22	110 [▲]	28	25	21	8	7	7	11	100
47 (M36/05)	15	25	11	16	12	21	98 [▲]	23	13	18	33	25	15	22	117	7	5	22	8	7	7	5	101
55 (M110/06)	16	NA	12	19	NA	20	Nv ST	24	60	17	30	19	22	22	232	57	33	47	35	49	17	16	102
10 (M4/98)	51	46	22	23	10	40	140	43	58	62	54	48	49	48	225	58	34	28	36	51	40	38	103
24 (LO-9)	36	39	15	32	10	40	141	43	41	44	55	50	49	48	226	59	16	26	37	52	21	20	104
27 (2ACAP)	36	39	15	32	10	34	120	43	41	44	55	50	49	48	226	59	16	26	37	52	21	20	104
52 (M72/6-6)	NA	39	NA	23	10	NA	Nv ST	44	40	44	54	48	64	48	227	60	35	26	38	51	21	38	105
68 (M52/8-19)	NA	NA	22	32	10	NA	Nv ST	55	41	44	55	50	49	48	228	61	16	48	36	52	21	20	106

▲STs que incluem mais de um sorogrupo/sorovar. NA – Novos alelos. Nv ST – Novo ST. Em **vermelho** são destacados novos alelos e STs já atualizados nos respectivos bancos

1 – Identificação das STs: 34 - *L. interrogans* sorogrupo Canicola/Djasiman; 58 - *L. interrogans* sorogrupo Pomona sorovar Pomona/Kennewicki; 74 - *L. interrogans* sorogrupo Serjoe sorovar Hardjo tipo Prajitno; 94 - *L. kirschneri* sorogrupo Grippytyphosa/Canicola; 98 - *L. kirschneri* sorogrupo Pomona sorovar Mozdok/Tsaratsovo; 120 – *L. santarosai* sorogrupo Grippytyphosa sorovar Bananal.

2 – Identificação das STs: 20 - *L. interrogans* sorogrupo Serjoe sorovar Hardjo/Haemolytica; 37 - *L. interrogans* sorogrupo Canicola/Pyrogenes; 110- *L. kirschneri* sorogrupo Grippytyphosa sorovar Grippytyphosa/Valbuzzi/Vanderhoedeni; 117 - *L. kirschneri* sorogrupo Pomona sorovar Mozdok; 140 - *L. interrogans* sorogrupo Pomona/Grippytyphosa/Pyrogenes; 149 - *L. borgpetersenii* sorogrupo Ballum sorovar Ballum/Castellonis.

3 – Identificação das STs: 3 - *L. interrogans* sorogrupo Canicola/Pomona; 6 - *L. interrogans* sorogrupo Serjoe sorovar Hardjoprajitno; 52- *L. interrogans* sorogrupo Pomona/Sejroe.

Pode-se observar que as estirpes de *L. interrogans* apresentaram resultados na MLST congruentes com a sua identificação prévia. Isto também é evidenciado na análise filogenética das sequências concatenadas dos respectivos genes na qual, além de se agruparem com as demais STs da espécie *L. interrogans* (Figuras 1 a 3), as estirpes dos sorogrupos Canicola, Pomona, Australis e Sejroe se agrupam com as demais representantes dos respectivos sorogrupos (Figura 4); a única exceção é a estirpe 86 (SU5), do sorogrupo Australis, que apresenta nova combinação de alelos para o protocolo de Varni et al. (2014) (Figura 4-B).

No caso de *L. kirschneri*, as estirpes 71 (UCR5RB1) e 47 (M36/05) apresentam STs congruentes com a sua identificação prévia nos protocolos de Boonsilp et al. (2013) e Ahmed et al. (2006). No entanto, a estirpe 55 (M110/06), apesar de ser do mesmo sorogrupo que a 47 (M36/05) (sorogrupo Pomona), apresenta novos alelos nos três protocolos de MLST e se distancia das demais estirpes brasileiras de *L. kirschneri* em mais de três genes.

As estirpes de *L. santarosai*, assim como as de *L. borgpetersenii* e *L. noguchii*, possuem novos alelos e/ou perfis alélicos para pelo menos dois dos protocolos vigentes de MLST. No entanto, as estirpes brasileiras de *L. santarosai* são mais próximas entre si apresentando, em sua maioria, variação em um a três genes. Essa proximidade também é evidenciada na análise filogenética (Figuras 1 a 3). Pode-se observar que para os protocolos de Boonsilp et al. (2013) e Varni et al. (2014) (Figuras 1 e 2), a distinção das espécies é clara entre as STs e, apesar das estirpes estudadas se agruparem com as demais *L. santarosai*, nota-se um agrupamento próprio das STs de origem brasileira. Já o protocolo de Ahmed et al. (2006) apresenta maior dispersão das STs.

Figura 1: Árvore de máxima verossimilhança baseada nas sequências concatenadas do protocolo de MLST de Boonsilp et al. (2013).

Figura 2: Árvore de máxima verossimilhança baseada nas sequências concatenadas do protocolo de MLST de Varni et al. (2014).

Figura 3: Árvore de máxima verossimilhança baseada nas sequências concatenadas do protocolo de MLST de Ahmed et al. (2006).

Figura 4: MSTs dos três protocolos vigentes de MLST (A - Boonsilp et al. (2013); B - Varni et al. (2014); C - Ahmed et al. (2006)) específicas para as estirpes de *L. interrogans* com destaque para os sorogrupos Canicola, Pomona, Australis e Sejroe.

5.3 Análise comparativa de genomas

5.3.1 *L. interrogans* sorovar Canicola

A partir da montagem do genoma das estirpes 22 (LO-3) e 23 (LO-4) foi realizada a ordenação dos *contigs* pela referência *L. interrogans* sorovar Copenhageni estirpe L1.130 e a comparação com os *contigs* disponíveis das estirpes de *L. interrogans* sorovar Canicola P2655 (AOWJ000000000.1) e LV133 (AKWU000000000.2). A estirpe LV133 foi isolada de humano no Brasil, enquanto que a estirpe P2655 foi isolada de um rato em Portugal.

Na montagem inicial (*draft*), as estirpes 22 (LO-3) e 23 (LO-4) apresentaram, respectivamente, aproximadamente 4,2 e 4,3 Mb para o cromossomo 1 e 400 e 422 kb para o cromossomo 2. Os *contigs* que não alinharam com os cromossomos apresentaram alta identidade (>99%) com os plasmídeos da estirpe chinesa de *L. interrogans* sorovar Canicola pGui1 (NC_025136) e pGui2 (NC_025197) (Zhu et al., 2014). O mesmo processo de ordenação dos *contigs* e BLASTn com os plasmídeos foi realizado para as estirpes P2655 e LV133. Apenas a estirpe LV133 apresentou sequências semelhantes à pGui1 e pGui2, enquanto que a estirpe P2655 possui apenas pGui1.

Em relação ao conteúdo cromossomal, todas as estirpes apresentaram alta sintonia e 99% de identidade, ao nível de DNA (Figura 5). Estas possuem os principais fatores de virulência de *Leptospira*: lipoproteínas LipL32 e LipL41; fatores “immunoglobulin-like” LigA, LigB e LigC; flagelina; proteína “OmpA-like” Loa22. Os respectivos genes desses fatores apresentam 100% de identidade com os da estirpe patogênica *L. interrogans* sorovar Copenhageni L1.130. As estirpes 22 (LO-3) e 23 (LO-4) apresentam ainda genes codificadores de acetiltransferases AACs da superfamília GNAT (*Gcn5-related N-acetyltransferases*) que estão relacionadas à resistência a neomicina e gentamicina, além de

transportadores AcrB e MatE que estão relacionadas à resistência a fluoroquinolonas e a aminoglicosídeos.

Destaca-se ainda no cromossomo 1 (Figura 5-A), regiões de deleção (setas vermelhas) nas estirpes do sorovar Canicola comparadas à referência *L. interrogans* sorovar Copenhageni estirpe L1.130. Estas regiões variantes correspondem à transposases, proteínas hipotéticas e proteínas de membrana. Situação semelhante é também observada no cromossomo 2 (Figura 5-B). No entanto, pode-se observar que a estirpe 22 (LO-3) ainda se sobressai com maior variação de sua sequência mesmo em relação às demais estirpes do sorovar Canicola.

Em relação ao conteúdo plasmidial, todas as sequências semelhantes aos plasmídeos pGui apresentam conteúdo GC menor do 34% diferindo do conteúdo cromossomal. A estirpe P2655 possui uma sequência com identidade acima de 99% com o plasmídeo pGui1, enquanto que as sequências homólogas de pGui1 das estirpes brasileiras LV133, 22 (LO-3) e 23 (LO-4) apresentam variações de tamanhos (Figura 6-A). A principal diferença entre as estirpes brasileiras é observada na sequência similar ao plasmídeo pGui2. As estirpes de suínos apresentam a sequência pGui2 maior, aproximadamente 11,2 kb, do que a sequência chinesa original e o plasmídeo da estirpe humana LV133 (Figura 6-B). Esta sequência de inserção das estirpes de suínos corresponde, em sua maioria, a transposases e proteínas de domínio AAA ATPase relacionadas aos processos de replicação e recombinação.

Figura 5: Representação esquemática do alinhamento dos cromossomos 1 (A) e 2 (B) das estirpes de *L. interrogans* sorovar Canicola utilizando como referência a estirpe patogênica *L. interrogans* sorovar Copenhageni L1.130. Setas vermelhas indicam áreas de diferenças entre as sequências dos sorovares Canicola e Copenhageni.

B

Figura 6: A - Representação esquemática do alinhamento das sequências semelhantes ao plasmídeo pGui1. B – Gráfico de sintenia das sequências semelhantes ao plasmídeo pGui2 (barras vermelhas correspondem a >99% identidade).

5.3.2 *L. interrogans* sorovar Pomona

Na análise comparativa dos genomas das estirpes do sorovar Pomona, os *contigs* obtidos foram ordenados pela referência *L. interrogans* sorovar Copenhageni estirpe L1.130. Para a análise comparativa foram utilizadas as estirpes de *L. interrogans* sorovar Pomona UT364 (AHNX02000001), Fox32256 (AOHG01000001), CSL4002 (ANMZ01000001), CSL10083 (AOHJ01000001), Kennewicki LC82-25 (AHMK02000002) e Pomona (AFLT02000001). As estirpes UT364, Pomona e Kennewicki são originárias de humanos da Tailândia, Austrália e dos Estados Unidos, respectivamente, enquanto que as estirpes CSL10083 e CSL4002 foram isoladas de mamíferos marinhos da costa californiana e a estirpe Fox32256 foi isolada de uma raposa-das-ilhas também na Califórnia.

A partir da montagem e ordenação, as estirpes 37 (GR5) e 39 (LPF) apresentaram, aproximadamente, 4,18 e 4,19 Mb para o cromossomo 1 e 380 e 450 kb para o cromossomo 2, respectivamente. Não foram identificados elementos extra-cromossomais nas estirpes do sorovar Pomona quando comparadas com as sequências disponíveis no *GenBank* (incluindo os plasmídeos publicados dos sorovares Lai, Linhai e Canicola e de *L. biflexa*).

Em relação ao conteúdo cromossomal, as estirpes do sorovar Pomona apresentaram alta sintonia e identidade, ao nível de DNA. As estirpes possuem os principais fatores de virulência de *Leptospira* (lipoproteínas, fatores “*immunoglobulin-like*”, flagelina). Estas também possuem regiões, em ambos os cromossomos, distintas da referência *L. interrogans* sorovar Copenhageni estirpe L1.130, sendo estas regiões ausentes em todas as estirpes do sorovar Pomona (Figura 7).

Figura 7: Representação esquemática do alinhamento dos cromossomos 1 (A) e 2 (B) das estirpes de *L. interrogans* sorovar Pomona utilizando como referência a estirpe patogênica *L. interrogans* sorovar Copenhageni L1.130. Setas vermelhas indicam áreas de diferenças entre as sequências dos sorovares Pomona e Copenhageni.

5.3.3 *L. interrogans* sorogrupo Australis

Para ordenação e análise da estirpe 86 (SU5) foi utilizado o genoma da estirpe PigK151, isolada de um suíno nos EUA e recentemente publicada como exemplar de *L. interrogans* sorovar Bratislava (CP011410 e CP011411), com o qual os *contigs* da estirpe 86 (SU5) apresentaram alta identidade. Dessa forma, foi possível montar para a estirpe 86 (SU5) duas estruturas cromossomais com, aproximadamente, 4,52 Mb e 375 kb correspondendo aos cromossomos 1 e 2, respectivamente.

Destaca-se a variação de tamanho dos genomas das estirpes PigK151 e 86 (SU5) quando comparadas à referência disponível de *L. interrogans* sorovar Copenhageni estirpe L1.130. Apesar de apresentar alta sintonia e identidade com as estirpes L1.130 e PigK151, o genoma da estirpe 86 (SU5) apresenta áreas de inserções que não existem nas sequências das anteriores. Essas regiões podem ser evidenciadas nas figuras 8 e 9. No cromossomo 1 (Figura 8), essas áreas correspondem, em sua maioria, a proteínas de fagos e profagos, transposases e proteínas hipotéticas. Destacam-se também no cromossomo 1, algumas áreas de inversão dos genomas das estirpes PigK151 e 86 (SU5) em relação a referência L1.130 (Figura 8-B); essas áreas encontram-se flanqueadas por transposases, em sua maioria.

Já no cromossomo 2 (Figura 9), chama atenção uma região de inserção da estirpe 86 (SU5) (área destacada por colchete e asterisco na Figura 9) de, aproximadamente, 24 kb composta por proteínas hipotéticas e proteínas de domínio de função desconhecida (*DUFs proteins*), além de proteínas do sistema de fosfotransferases (PTS) e transferases que também estão presentes no genoma da estirpe PigK151. As demais áreas de diferença entre as estirpes 86 (SU5), PigK151 e L1.130 correspondem à transposases.

A

B

Figura 8: Representação esquemática do alinhamento do cromossomo 1 das estirpes 86 (SU5), PigK151 e L1.130. A - Comparação circular do cromossomo 1 com referência interna a estirpe 86 (SU5) (linha preta interna). Setas vermelhas indicam áreas de diferenças entre as sequências (inserções em 86 (SU5)). B – Alinhamento múltiplo com definição de blocos de sequências homólogas; as setas vermelhas indicam áreas de diferenças entre as sequências e as linhas pretas indicam exemplos de blocos de sequências invertidas.

Figura 9: Representação esquemática do alinhamento do cromossomo 2 das estirpes 86 (SU5), PigK151 e L1.130. A - Comparação circular do cromossomo 2 tendo como referência interna a estirpe 86 (SU5) (linha preta interna). Setas e colchete vermelhos indicam áreas de diferenças entre as sequências (inserções da 86 (SU5)). B – Gráfico de sintenia das sequências do cromossomo 2 (barras vermelhas correspondem a >99% identidade).

5.3.4 *L. interrogans* sorovar Hardjo tipo Hardjoprajitno

Os *contigs* da estirpe 76 (OV5) foram ordenados com a referência *L. interrogans* sorovar Copenhageni estirpe L1.130. Devido a inexistência de sequências disponíveis de *L. interrogans* sorogrupo Sejroe sorovar Hardjo, até o momento, a análise comparativa da estirpe 76 (OV5) foi limitada às referências *L. interrogans* sorovar Copenhageni estirpe L1.130 e *L. interrogans* sorovar Lai estirpe 56601.

Dessa forma, foi possível montar para a estirpe 76 (OV5) os cromossomos 1 e 2 com, aproximadamente, 4,39 Mb e 290 kb, respectivamente. No cromossomo 1 (Figura 10-A), pode-se observar áreas de inserções na estirpe 76 (OV5) que correspondem a proteínas hipotéticas, transposases e proteínas de membrana, em sua maioria, que podem estar associadas às diferenças entre os sorogrupos/sorovares. Em relação ao conteúdo gênico, não se observa muita variação entre as sequências sendo identificados os principais genes de virulência de *Leptospira*, além de genes relacionados a mecanismos de resistência aos antimicrobianos e a metais pesados, como *czcA* e *czcD*.

No cromossomo 2, observa-se uma diminuição (deleção) da sequência da estirpe 76 (OV5) quando comparada às referências (área indicada pelo colchete na Figura 10-B). Essa região de, aproximadamente, 60 kb é composta por proteínas hipotéticas, fosfatases, oxidoredutases, metiltransferases, proteína de choque térmico e proteínas de membranas, além de sequências de repetição e transposases distribuídas por toda sua extensão. Existe uma segunda área de deleção (~5 kb) no cromossomo 2, aproximadamente na posição 105 kb do genoma da estirpe L1.130, que também é flanqueada por sequências de repetição e transposases e é composta por proteínas hipotéticas, peptidase e adenilato ciclase.

A

B

Figura 10: Representação esquemática do alinhamento dos cromossomos 1 (A) e 2 (B) da estirpe 76 (OV5) e das referências *L. interrogans* sorovar Copenhageni e *L. interrogans* sorovar Lai. Setas e colchete vermelhos indicam áreas de diferenças entre as sequências.

5.3.5 *L. noguchii* sorogrupo Panama

No caso da *L. noguchii*, não existe disponível ainda uma referência de genoma “fechado”, ou seja, uma sequência única para os cromossomos 1 e 2 com curadoria apropriada de um dos bancos de dados do *International Nucleotide Sequence Database Collaboration* (*GenBank* /*DDBJ* /*EMBL*). Para *L. noguchii* existem apenas depósitos de *contigs* anotados disponíveis no banco de *Whole Genome Shotgun* (WGS), sendo indicado como genoma representativo *L. noguchii* sorovar Panama estirpe CZ214 (NZ_AKWY000000000).

Dessa forma, os *contigs* obtidos na montagem da estirpe 72 (U73) foram ordenados pelas sequências disponíveis da estirpe CZ214. As sequências de *L. noguchii* foram analisadas na forma de WGS, ou seja, sem a definição dos cromossomos. Em relação ao conteúdo gênico, foram identificados os principais genes de virulência que codificam lipoproteínas e fatores “*immunoglobulin-like*”, como *lipL41*, *lipL36*, *lolC/D* e *ligA*. Também foram identificados diversos genes relacionados a mecanismos de resistência aos antimicrobianos, incluindo bombas de efluxo (*mdtA*, *mdtB*, *norM*), além de genes de resistência a metais pesados como *sugE*, *czcA* e *czcD*.

A análise comparativa foi realizada com a estirpe CZ214 (Panama) e os demais *contigs* disponíveis de *L. noguchii* estirpes: ZUN142 (sorovar Autumnalis) (AHOP000000000.2), 1993005606 (AHMF000000000.2), 2001034031 (AKXB000000000.2), 2006001870 (AFLY000000000.2), 2007001578 (AHMH000000000.2) e Bonito (AOHH000000000.1). Pode-se observar na Figura 11 que, apesar de existir alta identidade entre as sequências, existem também áreas de diferença entre as estirpes 72 (U73), Autumnalis, 1993005606, 2001034031, 2006001870, 2007001578 e Bonito quando comparadas com a CZ214 (Panama). Essas regiões são compostas por proteínas hipotéticas e também possuem diversas proteínas relacionadas às sequências de inserção (IS) IS1500, IS1501, IS1502 e ISLin2.

Figura 11: Representação esquemática do alinhamento das estirpes de *L. noguchii* utilizando como referência a estirpe CZ214 (Panama) (linha preta interna). As setas vermelhas indicam áreas de diferenças entre as sequências.

5.3.6 *L. borgpetersenii* sorovar Ballum

Para a estirpe 62 (LO-24), foi utilizada a referência de *L. borgpetersenii* sorovar Hardjo-bovis estirpe L550 para ordenação dos *contigs*. Esta também foi utilizada na análise comparativa conjuntamente com a estirpe JB197 (sorovar Hardjo-bovis) (NC_008510.1; NC_008511.1). Apesar de apresentarem alta similaridade, as sequências da estirpe 62 (LO-24) também possuem muita variação na sua disposição (blocos de sequências invertidas nos genomas) (Figura 12-A). Essas regiões estão, em sua maioria, flanqueadas por sequências de inserção, como por exemplo, IS533, IS1501, IS1502 e IS1500A.

A

B

Figura 12: Representação esquemática do alinhamento do cromossomo 1 das estirpes de *L. borgpetersenii*. A – Alinhamento múltiplo com definição de blocos de seqüências homólogas; as linhas indicam os blocos de seqüências invertidas. B - Comparação circular dos cromossomos 1 tendo como referência interna a estirpe 62 (LO-24) (linha preta interna).

Em relação ao conteúdo gênico, foram identificados os principais genes de virulência, como *lolC/D*, *lipA/B*, *fliG/F* e *ligA*, além de genes de resistência aos antimicrobianos (*tetA*, *ermA* e genes codificadores de bombas de efluxo - *norM* e *ykkD*) e genes codificadores de proteínas de resistência a fagos. Destacam-se as diferenças nas sequências do cromossomo 2 (Figura 13); a estirpe 62 (LO-24) apresenta uma diminuição de, aproximadamente, 96 kb comparada às estirpes L550 e JB197. Foram identificadas duas áreas de deleção: a primeira (~57 kb – região A na Figura 13) é composta principalmente por proteínas hipotéticas, transposases, integrases, fosfolipases e proteínas de membrana, e é flanqueada por IS1533; a segunda área (~39 kb – região B na Figura 13) é composta por proteínas hipotéticas, permeases, proteínas da biossíntese de polissacarídeos, ferredoxina, transferases do sistema precorrina-cobalamina e IS1533, sendo que não são observadas outras ISs flanqueando a região.

Figura 13: Gráfico de sintenia das sequências do cromossomo 2 das estirpes de *L. borgpetersenii* (barras vermelhas correspondem a >99% identidade). Em destaque as áreas de deleção na estirpe 62 (LO-24) de ~57 kb (região A) e ~39 kb (região B).

5.3.7 *L. kirschneri*

Para análise das estirpes de *L. kirschneri*, como também não existe disponível ainda uma referência de genoma “fechado”, foi utilizado para ordenação dos *contigs* o genoma representativo *L. kirschneri* sorovar Valbuzzi estirpe 200702274 (NZ_AHOC000000000). A análise comparativa das estirpes 71 (UCR5RB1), 47 (M36/05) e 55 (M110/06) foi realizada em conjunto utilizando também os genomas representativos do sorovares Valbuzzi estirpe 200702274 e Mozdok estirpe Brem 166 (NZ_AHQI000000000.1).

Em relação ao conteúdo gênico, não se observa grande variação entre as estirpes 71 (UCR5RB1), 47 (M36/05) e 55 (M110/06). Foram identificados os principais genes de virulência como *ligA/T*, *lolC/D*, *lipA*, *fliG/F* e genes codificadores de proteínas flagelares, além de genes de resistência aos antimicrobianos (*tetA*, *ermA*) e codificadores de bombas de efluxo (*norM*, *mdtA* e *qacA*). Pode-se observar na Figura 14-A que a estirpe 71 (UCR5RB1) apresenta maior identidade com o sorovar Valbuzzi enquanto que a estirpe 47 (M36/05) é mais próxima do sorovar Mozdok. As regiões de diferença entre essas estirpes são compostas, em sua maioria, por proteínas hipotéticas, transposases e proteínas de membrana.

A estirpe 55 (M110/06) se diferencia das demais *L. kirschneri* estudadas. Pode-se observar na Figura 14-B que, apesar de existir similaridade entre as sequências, existem também várias áreas de diferença entre as estirpes 71 (UCR5RB1), Valbuzzi, 47 (M36/05) e Mozdok quando comparadas com a 55 (M110/06). Estas regiões variantes entre as estirpes de *L. kirschneri* correspondem a proteínas hipotéticas, transposases e proteínas de membrana.

Figura 14: Representação esquemática do alinhamento das estirpes de *L. kirschneri*. A - Comparação circular tendo como referência interna a estirpe 47 (M36/05) (linha preta interna). B - Comparação circular tendo como referência interna a estirpe 55 (M110/06) (linha preta interna).

5.3.8 *L. inadai* sorovar Lyme

De forma semelhante às estirpes de *L. noguchii* e *L. kirschneri*, ainda não existe disponível uma referência de genoma “fechado” de *L. inadai*, de tal forma que foi utilizado para ordenação dos *contigs* o genoma representativo *L. inadai* sorovar Lyme estirpe 10 (NZ_AHMM000000000.2) sem a definição dos cromossomos. Em relação ao conteúdo gênico, foram identificados alguns genes de virulência, como *ligA*, *lipA* e *lolC/D*, e genes codificadores de proteínas flagelares como *fliG/F*. Também foram identificados genes relacionados a mecanismos de resistência aos antimicrobianos (*tetA*), incluindo bombas de efluxo (*mdtC*, *norM*), além de genes de resistência a metais pesados e antissépticos (*merR*, *sugE* *qacA*).

Para a análise comparativa também foram utilizados os *contigs* disponíveis de *L. broomii* sorovar Hurstbridge estirpe 5399 (NZ_AHMO000000000) e *L. fainei* sorovar Hurstbridge estirpe BUT6 (NZ_AKWZ000000000), que correspondem às espécies mais próximas de *L. inadai*. Pode-se observar na Figura 15 que a estirpe 34 (M34/99) apresenta maior identidade com a sequência de *L. inadai* sorovar Lyme estirpe 10 e que existem áreas de diferença entre os genomas de *L. inadai* e *L. broomii* e *L. fainei*. Apesar das espécies apresentarem similaridade de conteúdo (>80% similaridade genética), as sequências também apresentam variação na sua disposição (blocos de sequências invertidas nos genomas) entre as espécies (Figura 16).

Foi identificado na estirpe 34 (M34/99) um fragmento extra-cromossomal que apresenta alta identidade (> 99%) com o bacteriófago LinZ_10 (KF114880) descrito em *L. inadai* sorovar Lyme estirpe 10. Existe apenas uma variação de ~2,3 kb entre as sequências devido a ausência das proteínas da família da IS66 na estirpe 34 (M34/99). Este fragmento, assim como o bacteriófago LinZ_10, é composto em sua maioria por proteínas hipotéticas. Adicionalmente, existem blocos de sequências invertidas (Figura 17). Dessa forma, a estirpe brasileira de *L. inadai* sorogrupo Lyme foi caracterizada com conteúdo cromossomal de 4.464.770 pb e o fragmento *phage-like* Lin_34 de 87.308 pb.

Figura 15: Comparação circular dos genomas das estirpes de *L. inadai*, *L. broomii* e *L. fainei*.

Figura 16: Alinhamento múltiplo dos genomas das estirpes de *L. inadai*, *L. broomii* e *L. fainei* com definição de blocos de sequências homólogas.

Figura 17: Gráfico de sintenia da sequência semelhante ao bacteriófago LinZ_10 (barras azuis correspondem a sequências invertidas > 99 % identidade).

5.3.9 *L. santarosai*

A partir da montagem do genoma das estirpes de *L. santarosai* foi realizada a ordenação dos *contigs* pela referência *L. santarosai* sorovar Shermani estirpe LT821 (NZ_CP006694.1, NZ_CP006695.1). As estirpes apresentaram, aproximadamente, 3,75 a 3,94 Mb para o cromossomo 1 e 218 a 354 kb para o cromossomo 2; não foram identificados elementos extra-cromossomais nas montagens realizadas.

Em relação ao conteúdo cromossomal, as estirpes apresentaram alta sintenia e identidade, ao nível de DNA. No cromossomo 1, são observadas regiões de diferença entre a referência e as estirpes brasileiras (Figura 18). Em relação ao conteúdo gênico, não se observa muita variação entre as sequências sendo identificados os principais genes de

virulência de *Leptospira*, além de genes relacionados a mecanismos de resistência aos antimicrobianos, principalmente à tetraciclina, fluoroquinolonas e β -lactâmicos. Também foram identificados genes de resistência a metais pesados, principalmente relacionados ao *operon* de resistência a cádmio, zinco e cobalto (complexo *czc*), a homeostase de cobre, além de resistência a amônio quaternário.

Figura 18: Comparação circular do cromossomo 1 das estirpes de *L. santarosai*.

No cromossomo 2, foram identificadas duas áreas de deleção nas estirpes 10 (M4/98) e 68 (M52/8-19): a primeira (~72 kb – região A na Figura 19-B) é composta, principalmente, por proteínas hipotéticas, transposases, fosfolipases, fosfatases e lipoproteínas, além de componentes das sequências de inserção IS1533 e IS407; a segunda área (~37 kb – região B na Figura 19-B) é composta por proteínas hipotéticas, permeases, proteínas da

biossíntese de polissacarídeos e transferases do sistema precorrina-cobalamina. No entanto, ambas as regiões não são propriamente flanqueadas por ISs. Em relação à estirpe 27 (2ACAP), esta apresenta maior similaridade à referência, enquanto que a estirpe 52 (M72/6-6) possui áreas de inserções (~40 kb) (Figura 19-A) compostas, principalmente, por proteínas hipotéticas.

Figura 19: Comparação do cromossomo 2 das estirpes de *L. santarosai*. A – Alinhamento múltiplo com definição de blocos de sequências homólogas. B – Gráfico de sintenia (barras vermelhas correspondem a sequências com >99% identidade).

5.4 Análise de wgSNP

Na análise de wgSNP das estirpes de *L. interrogans*, pode-se observar na Figura 20 que, para ambos os cromossomos, os sorogrupos de *L. interrogans* se distinguem em dois grupos, sendo um formado pelos sorogrupos Pomona e Canicola e o outro pelas demais estirpes estudadas. Para os sorogrupos Canicola e Pomona foi identificado, aproximadamente, o dobro de SNPs do que nos sorogrupos Sejroe e Australis o que justifica o distanciamento destes sorogrupos do genoma de referência e a maior proximidade entre eles (Quadro 2). Apesar dos sorogrupos Sejroe e Australis apresentarem maior proximidade com a referência Copenhageni, a similaridade deles entre si varia entre os cromossomos.

Quadro 2: Total de posições (SNPs) selecionadas para análise de wgSNP de *L. interrogans*.

Estirpe	Sorogrupo	Cromossomo 1	Cromossomo 2
22 (LO-3)	Canicola	11415	1153
23 (LO-4)	Canicola	11521	1175
37 (GR5)	Pomona	11628	1163
39 (LPF)	Pomona	11533	1196
75 (OV5)	Sejroe	6462	640
85 (SU5)	Australis	7288	688

Em relação à análise das estirpes de *L. santarosai*, pode-se observar na Figura 21 que, assim como na análise da MSLT, as estirpes brasileiras de *L. santarosai* se agrupam entre si e se distanciam do genoma de referência. Para o cromossomo 1, as estirpes do sorogrupo Grippytyphosa apresentam maior proximidade do que no cromossomo 2 (Quadro 3), sendo que as estirpes 24 (LO-9) e 27 (2ACAP) apresentam perfil praticamente “clonal” em ambos os cromossomos. Já a estirpe 68 (M52/8-19) se distancia das demais estirpes brasileiras conforme observado na análise comparativa e na MLST.

Figura 20: Representação esquemática da análise de wgSNP das estirpes de *L. interrogans*. Representação dos SNPs identificados e respectiva análise de agrupamento segundo cromossomo (A – cromossomo 1 e B – cromossomo 2).

Figura 21: Representação esquemática da análise de wgSNP das estirpes de *L. santarosai*. Representação dos SNPs identificados e respectiva análise de agrupamento segundo cromossomo (A – cromossomo 1 e B – cromossomo 2).

Quadro 3: Total de posições selecionadas para análise de wgSNP de *L. santarosai*.

Estirpe	Sorogrupo	Cromossomo 1	Cromossomo 2
24 (LO-9)	Grippotyphosa	3703	789
27 (2ACAP)	Grippotyphosa	4021	837
52 (M72/6-6)	Grippotyphosa	3969	820
10 (M4/98)	Sejroe	4003	837
68 (M52/8-19)	-	3799	786

No caso da análise das estirpes de *L. kirschenri*, foram utilizadas como referência *L. kirschneri* sorovar Valbuzzi estirpe 200702274 (NZ_AHOC000000000) e *L. kirschneri* sorovar Mozdok estirpe Brem 166 (NZ_AHQI000000000.1), sem a diferenciação dos cromossomos. Também foi utilizado o genoma da estirpe brasileira de *L. kirschneri* sorovar Mozdok de origem humana (parceria com Universidade Federal de Pelotas). Dessa forma, pode-se observar a proximidade dos genomas do sorovar Mozdok e o padrão de “clonalidade” das estirpes brasileiras (Figura 22-A). De forma semelhante, é evidente a similaridade da estirpe 71 (UCR5RB1) com o sorovar Valbuzzi (Figura 22-B) conforme observado anteriormente na análise genômica comparativa, além do distanciamento da estirpe 55 (M110/06) dos demais genomas (Quadro 4).

Quadro 4: Total de posições selecionadas para análise de wgSNP de *L. kirschneri*.

Estirpe	Sorogrupo	Sorovar Mozdok	Sorovar Valbuzzi
47 (M36/05)	Pomona	74	8142
61H*	Pomona	73	8279
55 (M110/06)	Pomona	19562	19879
71 (UCR5RB1)	Grippotyphosa	8312	520

* Estirpe brasileira de *L. kirschneri* sorovar Mozdok de origem humana (parceria com Universidade Federal de Pelotas).

Figura 22: Representação esquemática da análise de wgSNP das estirpes de *L. kirschneri*. Representação dos SNPs identificados e respectiva análise de agrupamento segundo referência utilizada (A – sorovar Mozdok e B – sorovar Valbuzzi).

6. DISCUSSÃO

A análise genômica já ampliou o horizonte do estudo da patogênese e evolução das espécies de *Leptospira*. A variação e plasticidade de conteúdo e estrutura genômica das espécies e sorovares já estudados (Nascimento et al., 2004; Picardeau et al., 2008; Ricaldi et al., 2012; Fouts et al., 2016) parecem ser um dos principais desafios para o estudo genômico de *Leptospira*. No entanto, é devido a falta de informação - de origem e identificação - dos genomas de *Leptospira* disponíveis no *GenBank* que a análise genômica comparativa é comprometida, pois impede a utilização de sequências que poderiam não apenas melhorar a análise e montagem dos genomas, mas também poderiam originar novas descobertas sobre o gênero.

A existência de genomas de referência, com montagem finalizada, representando apenas as espécies *L. biflexa*, *L. borgpetersenii*, *L. interrogans* e *L. santarosai*, também é outro obstáculo para o estudo genômico de *Leptospira*. Apesar de não interferir na montagem com algoritmo *ab initio*, a falta de referência compromete a ordenação, anotação e comparação de genomas, pois são etapas que necessitam de modelo (Mardis et al, 2002; Ekblom & Wolf, 2014). A utilização de *drafts*, genomas na forma de *contigs* ou *scaffolds* e não em uma única sequência de cromossomo, pode ser uma alternativa já que estes contêm o conteúdo gênico total apesar de não ordenado ou unido. No entanto, deve-se atentar para a qualidade e cobertura dessas sequências para não estar sujeito a erros decorrentes de sequenciamento e montagem (Mardis et al, 2002).

Dessa forma, no presente estudo, os genomas de referência de *L. borgpetersenii*, *L. interrogans* e *L. santarosai* foram utilizados para ordenação dos *contigs* obtidos, quando possível. Para as demais espécies, e também para a análise genômica comparativa, optou-se pela utilização dos *drafts* disponíveis no *GenBank* que apresentavam as melhores

informações de identificação das estirpes e de montagem e cobertura, de forma a minimizar as possibilidades de erro das análises decorrentes.

Neste estudo, optou-se também pela montagem com algoritmo *ab initio*, pois esta não exige genoma de referência, já que a montagem se baseia exclusivamente nas *reads* obtidas, e é a única abordagem que permite a identificação de grandes alterações de conteúdo cromossomal, principalmente inserções, e a presença de elementos extra-cromossomais (Nagarajan & Pop, 2013). Dessa forma, foi possível detectar, nas estirpes estudadas, as variações de composição cromossomal (inserções e deleções) nos genomas dos sorogrupos Australis e Serjoe e de *L. borgpetersenii*, além dos plasmídeos do sorovar Canicola e do bacteriófago de *L. inadai*.

Os resultados obtidos pela análise *in silico* da MLST demonstram como a epidemiologia molecular de *Leptospira* é complexa e ainda existem muitos dilemas e falhas a serem superados. Dos 17 genomas estudados, a MLST mostrou-se mais útil para a tipificação das estirpes de *L. interrogans* apresentando resultados congruentes com a sorotipagem. Para a espécie *L. santarosai*, assim como *L. borgpetersenii* e *L. noguchii*, a técnica apresentou maior heterogeneidade com a identificação de novos alelos e/ou perfis alélicos para pelo menos dois dos protocolos vigentes. No entanto, essa variabilidade pode ser decorrente, em parte, por se tratarem de espécies não tão estudadas quanto a *L. interrogans*.

Apesar de a MLST apresentar níveis mais elevados de discriminação dentro das espécies de *Leptospira*, comparada às demais técnicas moleculares, e ser indicada como método complementar para a classificação das estirpes (Thaipadungpanit et al., 2007; Ahmed et al., 2011; Boonsilp et al., 2013; Varni et al., 2014), ainda há grande conflito entre os protocolos vigentes. Embora os três protocolos tenham possibilitado a diferenciação das espécies de *Leptospira*, pode-se observar que o protocolo de Ahmed et al. (2006) apresentou maior dispersão das STs quando comparado ao de Boonsilp et al. (2013) e Varni

et al. (2014). Também deve-se considerar que, até hoje, ainda não foram caracterizadas todas as estirpes do banco do PubMLST para os três protocolos, de tal forma que o banco permanece incompleto e não é possível a comparação direta entre os protocolos. Além disso, embora os protocolos de Boonsilp et al. (2013) e Varni et al. (2014) sejam os mais utilizados e, conseqüentemente, os mais atualizados, estes ainda são restritos às sete espécies patogênicas do gênero (*L. interrogans*, *L. kirschneri*, *L. borgpetersenii*, *L. noguchii*, *L. santarosai*, *L. weilii* e *L. alexanderi*).

Varni et al. (2014) descreveram que a análise evolutiva dos perfis alélicos, do seu protocolo, demonstrava clara relação entre os genótipos e a origem geográfica e o hospedeiro das respectivas estirpes. No entanto, avaliando as informações do banco de dados do PubMLST e os resultados do presente estudo, pode-se observar que essa relação não é tão evidente para todas as espécies e os perfis alélicos parecem estar mais relacionados aos sorogrupos/sorovares das estirpes do que à sua origem.

Os genomas do sorogrupo Canicola, por exemplo, apresentaram a mesma ST identificada em estirpes do mesmo sorogrupo originárias de mais de cinco países e isoladas de cão, humano, bovino e suíno, considerando tanto o protocolo de Varni et al. (2014) quanto o de Boonsilp et al. (2013). Apesar de não evidenciar a relação entre “tipo” e origem das estirpes, esse resultado ainda demanda atenção, pois demonstra a infecção pelo sorogrupo Canicola em hospedeiros que não são classicamente associados a este (bovino, suíno e homem) elucidando a complexidade da epidemiologia da leptospirose.

Já os resultados observados para os genomas de *L. santarosai*, evidenciando o agrupamento das STs de origem brasileira, corroboram a afirmação de Varni et al. (2014) para esta espécie. A maior proximidade das STs brasileiras de *L. santarosai*, para o protocolo de Boonsilp et al. (2013), também já foi descrita por Kremer et al. (2015) e Miotto et al. (2016). Além disso, Miotto et al. (2016) sugerem que, apesar da proximidade entre as estirpes brasileiras de *L. santarosai*, as sequências originárias de hospedeiros

distintos também apresentam certa heterogeneidade. Considerando que os resultados do presente estudo não evidenciam a distinção de hospedeiros de *L. santarosai* pela MLST, já que estirpes de bovino e capivara apresentaram a mesma ST tanto no protocolo de Varni et al. (2014) quanto no de Boonsilp et al. (2013), mais estirpes devem ser avaliadas para confirmar se a técnica realmente possibilita a diferenciação das estirpes de *L. santarosai* por local e hospedeiro de origem.

Em relação à análise comparativa dos genomas de *L. interrogans* sorovar Canicola, é notável a similaridade entre as estirpes brasileiras (22 (LO-3), 23 (LO-4) e LV133) e até mesmo com a estirpe portuguesa P2655, apesar das diferenças de origem e hospedeiros. Pode-se observar, na Figura 5, que as quatro estirpes apresentam regiões, em ambos os cromossomos, distintas da referência *L. interrogans* sorovar Copenhageni possivelmente associadas a diferenças entre os sorovares. Ademais, as pequenas variações existentes entre essas estirpes parecem estar associadas à profundidade do sequenciamento e à montagem. Até o momento, ainda não há informação suficiente para confirmar se essas variações podem estar associadas à adaptação aos hospedeiros de origem; a análise da conservação de genes ortólogos e de mais estirpes seria uma abordagem para estudar essa relação.

Apesar disso, destaca-se o isolamento do sorovar Canicola em suínos, enquanto que o sorovar Pomona ainda não foi descrito em cães no país. Embora os sorovares não sejam totalmente exclusivos às espécies de hospedeiros, esses geralmente apresentam certa predileção (Schuller et al., 2015). Dessa forma, chama atenção as implicações epidemiológicas do isolamento do sorovar Canicola em suínos, com a possibilidade de disseminação e manutenção de *Leptospira* nas granjas pelos cães, assim como a alta similaridade com o genoma de estirpe isolada de humano também indica que essas estirpes de suínos podem apresentar risco para saúde pública.

A identificação de plasmídeos nas estirpes 22 (LO-3), 23 (LO-4), P2655 e LV133, confirmadas como sorogrupo/sorovar Canicola, assim como na descrição original dos

plasmídeos na estirpe chinesa Gui44 (Zhu et al., 2014), sugere que elementos extra-cromossomais sejam mais uma característica deste sorovar. No entanto, mais estirpes devem ser avaliadas para confirmar este fato. Ambos plasmídeos apresentam conteúdo GC menor que 34%, diferindo do conteúdo cromossomal, além de elementos relacionados à virulência (domínios “*immunoglobulin-like*”) e à replicação e recombinação, conforme previamente descrito por Zhu et al. (2014).

Em relação às variações de tamanho identificadas nas sequências semelhantes aos plasmídeos pGui1 e pGui2, estas podem estar relacionadas com a origem e/ou hospedeiros das estirpes estudadas. Considerando a sequência semelhante ao pGui1, esta foi identificada em todos os genomas estudados sendo observadas distintas áreas de deleção nas sequências de acordo com o hospedeiro de origem. Já em pGui2, as inserções detectadas nas estirpes isoladas de suínos não foram observadas nas estirpes de origem humana, tanto da China quanto do Brasil.

No entanto, deve-se lembrar também que as estirpes 22 (LO-3) e 23 (LO-4) foram mantidas em passagem serial em EMJH nos últimos dez anos, o que pode ter alterado o seu potencial patogênico e conteúdo genético. Considerando que Lehmann et al. (2016) já descreveram o aparecimento de SNPs em genes de virulência em decorrência da atenuação por cultura *in vitro*, avaliando apenas sete subculturas do sorovar Lai, a manutenção continuada das estirpes estudadas em meio EMJH poderia ter alterado significativamente não apenas o seu conteúdo cromossomal, mas também o plasmidial.

Nos genomas das estirpes de *L. interrogans* sorovar Pomona não foram identificados elementos extra-cromossomais. Nestes foram identificados os principais fatores de virulência de *Leptospira* e também apresentaram alta sintenia com os genomas de estirpes isoladas de humanos, confirmando seu potencial patogênico. As estirpes do sorovar Pomona também apresentam regiões, em ambos os cromossomos, distintas da referência *L. interrogans* sorovar Copenhageni que estão possivelmente associadas a diferenças entre os

sorovares. Embora o sorovar Pomona seja mais comumente associado aos suínos, a alta similaridade entre os genomas originários de suínos, animais marinhos e humanos possibilita o questionamento de se realmente existe uma predileção dos sorovares por espécies de hospedeiros.

Em relação a *L. interrogans* sorogrupo Australis estirpe 86 (SU5) destacam-se suas áreas de inserção quando comparadas às referências dos sorovares Copenhageni e Bratislava. Compostas, em sua maioria, por proteínas hipotéticas, proteínas de fagos e transposases, essas áreas podem estar associadas à inserção de elementos móveis entre os sorogrupos/sorovares. A proposição de que a variação dos genomas de diferentes sorogrupos/sorovares dentro de uma mesma espécie seja decorrente da movimentação de elementos genéticos móveis também é sustentada pelos resultados da análise comparativa, principalmente do cromossomo 2, da estirpe *L. interrogans* sorogrupo Sejroe 76 (OV5) na qual se observam áreas de deleção flanqueadas por sequências de repetição e transposases.

De forma semelhante, Nascimento et al. (2004) também relatavam que, apesar da alta similaridade genética, os genomas de *L. interrogans* sorovar Copenhageni e sorovar Lai apresentavam importantes diferenças estruturais, incluindo uma inversão cromossômica e grande variação no número e distribuição de elementos de sequências de inserção. No entanto, até o momento, a análise genômica comparativa foi mais utilizada para o estudo das diferenças e da evolução das espécies do gênero *Leptospira*, e não entre os respectivos sorogrupos/sorovares.

Xu et al. (2016) e Fouts et al. (2016) descreveram que a aquisição de genes, especialmente de virulência, determinou o aparecimento da linhagem patogênica das leptospiros e caracteriza a plasticidade genômica dessas espécies, contribuindo para sua adaptação à diversos ambientes e permitindo sua disseminação a novos hospedeiros. Os autores também confirmaram que essa expansão genômica está relacionada à transferência

horizontal de genes, justificando a existência de bacteriófagos, elementos de sequências de inserção e plasmídeos nessas espécies.

Já dentre as espécies patogênicas, Bulach et al. (2006) relataram a ocorrência de processos de redução genômica relacionados à evolução da espécie *L. borgpetersenii* diferenciando-a das demais espécies do gênero. Os autores descreveram que as deleções observadas no genoma de *L. borgpetersenii*, comparado ao de *L. interrogans*, poderiam prejudicar a adaptação do agente a ambientes diversos de forma a restringir sua transmissão à forma direta. A variação no tamanho e estrutura do genoma de *L. borgpetersenii* corrobora os resultados obtidos para a estirpe 62 (LO-24) (*L. borgpetersenii* sorogrupo Ballum), cujo genoma apresenta áreas de deleções, assim como blocos de sequências invertidas e alto número de réplicas de sequências de inserção.

Na análise comparativa com as estirpes de *L. borgpetersenii* sorovar Hardjo-bovis (estirpes L550 e JB197), também foi observado que, apesar de apresentarem alta similaridade, os genomas possuem muita variação na sua disposição (blocos de sequências invertidas), em sua maioria, associada a sequências de inserção, principalmente no cromossomo 1. Essas variações podem estar associadas às diferenças entre os sorogrupos/sorovares, pois Bulach et al. (2006) já haviam descrito a existência de áreas de inversões e alto número de sequências de inserção entre genomas do mesmo sorovar (*L. borgpetersenii* sorovar Hardjo-bovis estirpes L550 e JB197). Da mesma forma, as deleções identificadas no cromossomo 2 da estirpe 62 (LO-24) também podem estar relacionadas às diferenças entre os sorogrupos/sorovares de *L. borgpetersenii*.

Deleções semelhantes também foram identificadas no cromossomo 2 das estirpes 10 (M4/98) e 68 (M52/8-19) de *L. santarosai*; apesar de apresentarem menor tamanho na estirpe de *L. borgpetersenii*, as deleções se apresentam nas mesmas regiões do cromossomo 2 de ambas as espécies. A similaridade observada entre as estirpes estudadas de *L. santarosai* e *L. borgpetersenii* corrobora os relatos prévios de que ambas as espécies são

intimamente relacionadas (Chou et al., 2012; 2014). Chou et al. (2012) fizeram o primeiro relato do genoma da referência *L. santarosai* sorovar Shermani estirpe LT821, no qual observaram uma cobertura maior que 80% do genoma de *L. borgpetersenii* pelas *reads* da estirpe LT821. Apesar de exibirem variação do conteúdo gênico, *L. santarosai* e *L. borgpetersenii* apresentam genomas com tamanhos semelhantes e ambas são caracterizadas por alto número de elementos móveis (Chou et al., 2014).

Dessa forma, considerando os eventos de redução genômica em *L. borgpetersenii*, descritos por Bulach et al. (2006), e de transferência horizontal de genes entre as espécies de *Leptospira* (Xu et al., 2016), pode-se sugerir que estes processos também ocorram em *L. santarosai*. Também é notável que, para o cromossomo 1, as estirpes brasileiras apresentam regiões distintas da referência *L. santarosai* sorovar Shermani. Pode-se observar, na Figura 18, que algumas dessas regiões são comuns a todas as estirpes brasileiras de *L. santarosai*, enquanto que outras parecem ser específicas de determinados sorogrupos. Conforme discutido para a MLST, a diversidade das estirpes brasileiras de *L. santarosai* já era esperada e corrobora os relatos prévios (Kremer et al., 2015; Miotto et al., 2016). No entanto, mais estudos comparativos são necessários para avaliar a variabilidade dos sorogrupos e das estirpes brasileiras de *L. santarosai*, e também para confirmar a suposição de redução genômica nesta espécie.

Em relação aos genomas de *L. noguchii* e *L. kirschneri*, o maior desafio para finalização e análise comparativa dos genomas foi a falta de genomas representativos e de referência. Apesar disso, foi possível aperfeiçoar a montagem, considerando a profundidade obtida no sequenciamento, e fazer a comparação dos *contigs* obtidos com as sequências disponíveis no *GenBank* sem a separação dos cromossomos. Conforme evidenciado para os genomas de *L. interrogans*, pode-se observar alta similaridade intra-espécie e algumas regiões de diferença compostas por proteínas hipotéticas, transposases e proteínas de membrana, em sua maioria, possivelmente relacionadas às variações entre os sorogrupos/sorovares.

No caso de *L. kirschneri*, as estirpes 47 (M36/05) e 71 (UCR5RB1) apresentaram alta identidade com os genomas de mesma identificação (sorovares Mozdok e Valbuzzi, respectivamente – Figura 14) tanto pela análise comparativa quanto pelo wgSNP. Ambas as análises também evidenciaram o distanciamento da estirpe 55 (M110/06) dos demais genomas de *L. kirschneri*. Apesar da alta similaridade das estirpes brasileiras, foram identificadas inserções no genoma da estirpe 55 (M110/06) e regiões de variação compostas por proteínas hipotéticas, transposases e proteínas de membrana dentre as estirpes 47 (M36/05) e 71 (UCR5RB1), que podem estar relacionadas às variações entre os sorogrupos/sorovares.

Destaca-se também que a estirpe 47 (M36/05), identificada como *L. kirschneri* sorogrupo Pomona sorovar Mozdok, também apresenta alta identidade com a estirpe brasileira 61H isolada de humano. Considerando a distância geográfica e cronológica entre as estirpes, uma vez que a estirpe 61H foi isolada de uma amostra de sangue humano de Pelotas, na região metropolitana do Rio Grande do Sul, em 2013 (Cunha et al. 2016), é possível inferir que *L. kirschneri* sorogrupo Pomona sorovar Mozdok já circula nas regiões sudeste e sul do país pelas últimas duas décadas. A estirpe aparentemente se adaptou ao roedor como reservatório e apresenta alto potencial patogênico a partir dos resultados de ensaio de virulência em hamsters (dados não apresentados).

Cunha et al. (2016) também descreveram o isolamento de *L. kirschneri* sorovar Mozdok de um cão assintomático, em Pelotas, sugerindo que o sorovar já se adaptou a diferentes hospedeiros reservatórios em áreas urbanas apresentando risco para saúde pública. Dessa forma, a partir dos resultados da análise genômica e dos dados epidemiológicos, pode-se sugerir a existência de uma linhagem brasileira do sorovar Mozdok que apresenta maior risco para saúde pública, do que o esperado, e que já circula por grandes áreas urbanas do país há pelo menos duas décadas. A falha de identificação e caracterização da estirpe é

decorrente do fato do sorovar Mozdok não fazer parte da bateria de antígenos dos testes de rotina para diagnóstico de leptospirose.

De forma semelhante, a espécie *L. inadai* provavelmente era despercebida na realidade brasileira da leptospirose por não fazer parte da bateria de antígenos dos testes de rotina. O presente estudo apresentou o primeiro isolamento e caracterização de *L. inadai* no Brasil e o segundo relato desta espécie na América do Sul (Chiriboga et al., 2015). A primeira descrição da espécie ocorreu nos Estados Unidos da América de um paciente com doença de Lyme; apesar de a estirpe ser caracterizada como patogênica para animais de laboratório, esta não interferiu no quadro clínico do paciente. Esta estirpe é a referência de *L. inadai* sorovar Lyme (estirpe 10) que foi formalmente descrita por Yasuda et al. (1987).

Apesar de se tratar de uma espécie intermediária, já foram descritos casos fatais de infecção em humanos na Índia (Gangadhar et al., 2008), aonde os hospedeiros principais parecem ser roedores (Gangadhar et al., 2000). Nas Américas, Barragan et al. (2015) detectaram DNA de *L. inadai* em um paciente febril e também relataram a detecção da espécie em cães, suínos, bovinos e roedores da mesma província. A identificação de *L. inadai* em roedor no Brasil sugere que esta espécie também é circulante no país e passa despercebida por não fazer parte da bateria de antígenos dos testes de rotina. Além disso, a alta similaridade com o genoma da referência americana indica que a estirpe brasileira pode representar risco para saúde pública e até mesmo a possibilidade de uma linhagem do sorovar Lyme comum nas Américas.

Portanto, recomenda-se que as estirpes brasileiras de *L. kirschneri* sorovar Mozdok (47 (M36/05)) e *L. inadai* sorovar Lyme (34 (M34/99)) sejam incluídas na bateria de antígenos para diagnóstico laboratorial de leptospirose, de tal forma que se amplie o espectro de identificação e simultaneamente o torne mais sensível para as estirpes circulantes no país. De forma semelhante, a inclusão das estirpes autóctones dos sorovares Guaricura (10 (M4/98)) e Bananal (27 (2ACAP)) também contribuiria

positivamente para o diagnóstico da doença no país, conforme previamente observado por Sarmiento et al. (2012) que detectaram um aumento significativo do número de animais reativos, assim como da frequência e distribuição dos sorovares, a partir da inclusão de estirpes autóctones na bateria de antígenos para diagnóstico da leptospirose.

Por fim, em relação à análise de wgSNP, esta foi utilizada com o propósito de “tipagem genômica epidemiológica” como mais uma técnica para caracterização da variação das estirpes de uma mesma espécie. Esta técnica se fundamenta na identificação e seleção de polimorfismos de base única (SNPs), a partir do mapeamento de genomas, que podem ser utilizados em uma análise de agrupamento ou filogenética para demonstrar similaridade/proximidade, e até mesmo evolução, das estirpes estudadas. A tipificação genômica por SNPs já é amplamente utilizada para epidemiologia genômica de bactérias patogênicas, com a caracterização da diversidade e dispersão das populações bacterianas (Holt et al., 2010; 2015; Wong et al., 2015), apesar de ainda não ter sido aplicada nos estudos de *Leptospira*.

Os resultados obtidos demonstram a associação entre os SNPs analisados e os sorogrupos/sorovares das três espécies estudadas. No caso de *L. interrogans*, foi observada a proximidade entre os sorogrupos Pomona e Canicola corroborando os resultados prévios de tipagem molecular (Moreno et al., 2016b). Destaca-se que ambos os sorogrupos ainda se distanciam do sorovar Copenhageni (referência) que apresentou, inesperadamente, maior proximidade dos sorogrupos Sejroe e Australis. Mais estudos incluindo maior número de estirpes dos respectivos sorogrupos são necessários para confirmar se esta poderia ser uma tendência evolutiva das linhagens destes sorogrupos dentro da espécie *L. interrogans*.

Devido à necessidade do mapeamento, o ponto crítico da análise de wgSNP é a seleção do genoma de referência. A análise dos genomas de *L. kirschneri* utilizando referências de sorovares distintos corrobora, não apenas a associação dos SNPs com os sorogrupos/sorovares das espécies de *Leptospira*, mas também a importância e influência

dos genomas de referência utilizados para o mapeamento. A seleção do genoma de referência pode limitar a análise e influenciar drasticamente os resultados. Dessa forma, os critérios de seleção devem ser padronizados buscando, além de similaridade com as estirpes estudadas, confiança na montagem do genoma que será utilizado como referência. No caso do gênero *Leptospira*, muitos dos genomas disponíveis no *GenBank* não são fechados (*drafts*), de tal forma que ainda deve-se considerar a possibilidade da referência selecionada estar incompleta.

Em relação à análise de wgSNP de *L. santarosai*, foi corroborada a tendência de agrupamento das estirpes brasileiras entre si e que ainda se distanciam da referência. Foi detectada também maior variação entre as estirpes do sorogrupo Grippytyphosa, do que o esperado, não sendo observada associação com a origem e/ou hospedeiros das respectivas estirpes. No entanto, considerando o pequeno número de estirpes avaliadas e sua variação de origem, seria necessário analisar mais estirpes do sorogrupo Grippytyphosa para confirmar se a variação observada realmente é uma característica própria deste e se ainda poderia estar relacionada à origem das estirpes. Destaca-se também, na análise de wgSNP, o distanciamento da estirpe 68 (M52/8-19) das demais estirpes brasileiras, corroborando os resultados para a MLST e análise genômica comparativa que, associados a ausência de reatividade da estirpe na sorogrupagem, sugerem a possibilidade de um novo sorogrupo e/ou sorovar de *L. santarosai* no país.

7. CONCLUSÕES

Com base nos resultados encontrados pode-se concluir que:

- A MLST mostrou-se útil para a tipificação das estirpes de *L. interrogans* apresentando resultados congruentes com a sorotipagem. No entanto, para as espécies *L. santarosai*, *L. borgpetersenii* e *L. noguchii*, a técnica apresentou maior heterogeneidade com a identificação de novos alelos e/ou perfis alélicos para pelo menos dois dos protocolos vigentes. As estirpes de *L. santarosai* ainda se destacaram por formarem um agrupamento próprio de origem brasileira.

- Dentre os genomas de *L. interrogans*, destacaram-se os do sorovar Canicola nos quais foram identificados dois plasmídeos com alta identidade aos descritos na estirpe chinesa do mesmo sorovar. Já, nos genomas dos sorogrupos Australis e Serjoe foram observadas inserções e deleções, principalmente no cromossomo 2, que podem estar associadas à diferenciação dos sorogrupos.

- A variação de conteúdo e estrutura observada nos genomas de *L. borgpetersenii* parece ser mediada por sequências de inserção e pela transposição de elementos móveis.

- Para a espécie *L. kirschneri*, foi identificada uma linhagem brasileira do sorogrupo Pomona sorovar Mozdok, isolada de roedor e humano, que pode apresentar maior risco para saúde pública. Já a estirpe 55 (M110/06) se diferenciou dos demais genomas de *L. kirschneri*, tanto na MLST quanto na análise comparativa, e necessita de mais estudos para sua caracterização.

- Foi feita a primeira identificação e caracterização de *L. inadai* no Brasil; sugere-se que esta espécie é circulante no país e provavelmente passa despercebida por não fazer parte da bateria de antígenos dos testes de rotina. A alta similaridade com o genoma da referência

americana isolada de humano indica que a estirpe brasileira pode apresentar risco para saúde pública.

- As estirpes brasileiras de *L. santarosai* se distinguem, tanto na MLST quanto na análise comparativa, das demais estirpes da espécie apresentando agrupamento próprio independente dos sorogrupos. A diferenciação genômica da estirpe 68 (M52/8-19), que não apresentou reatividade aos sorogrupos testados, sugere a possibilidade de novo sorogrupo/sorovar de *L. santarosai*.

- A análise de wgSNP mostrou-se útil para a caracterização das espécies *L. interrogans*, *L. kirschneri* e *L. santarosai*, demonstrando associação entre os SNPs identificados e os sorogrupos/sorovares das espécies estudadas. A sua maior aplicação para o estudo de *Leptospira* pode confirmar a utilidade desta ferramenta para a epidemiologia genômica da leptospirose.

8. REFERÊNCIAS

- Adler B, Lo M, Seemann T, Murray GL. Pathogenesis of leptospirosis: The influence of genomics. *Vet Microbiol.* 2011; 153: 73–81.
- Adler B, de Pena-Moctezuma MA. *Leptospira* and leptospirosis. *Vet Microbiol.* 2010; 140: 287-296.
- Adler BA, Faine S. Species- and genus-specific antigens in *Leptospira*, revealed by monoclonal antibodies and enzyme immunoassay. *Zentbl Bakteriol.* 1983; 255: 317-322.
- Ahmed A, Thaipadungpanit J, Boonsilp S, Wuthiekanun V, Nalam K, Spratt BG, Aanensen DM, Smythe LD, Ahmed N, Feil EJ, et al. Comparison of two multilocus sequence based genotyping schemes for *Leptospira* species. *PLoS Negl Trop Dis.* 2011; 5: e1374.
- Ahmed N, Manjulata DS, Valverde M, Vijayachari P, Machangu RS, Ellis WA, Hartskeerl RA. Multilocus sequence typing method for identification and genetic classification of pathogenic *Leptospira* species. *Ann Clin Microbiol Antimicrob.* 2006; 5: 1-10.
- Alexander AD. The distribution of leptospirosis in Latin America. *Bull World Health Organ.* 1960; 23: 113-25.
- Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons, *BMC Genomics.* 2011; 12: 402.
- Allan KJ, Biggs HM, Halliday JEB, Kazwala RR, Maro VP, Cleaveland S, Crump JA. Epidemiology of Leptospirosis in Africa: A Systematic Review of a Neglected Zoonosis and a Paradigm for ‘One Health’ in Africa. *PLoS Negl Trop Dis.* 2015; 9: e0003899.
- Beaurepaire Aragão HB. Sobre a presença do *Spirochaeta icterohemorrhagiae* nos ratos no Rio de Janeiro. *Brasil Med.* 1917; 31: 329-330.
- Bolin C. Leptospirosis. In: Brown C, Bolin C, editors. *Emerging Diseases of Animals.* Washington, DC: ASM Press; 2000. p. 185–200.
- Bolin CA. Diagnosis of leptospirosis in swine. *Swine Health Prod.* 1994; 2: 23-24.
- Boonsilp S, Thaipadungpanit J, Amornchai P, Wuthiekanun V, Bailey MS, Holden MT, Zhang C, Jiang X, Koizumi N, Taylor K, et al. A Single multilocus sequence typing (MLST) scheme for seven pathogenic *leptospira* species. *PLoS Negl Trop Dis.* 2013; 7: e1954.

- Bouck J, Miller W, Gorrell JH, Muzny D, Gibbs RA. Analysis of the quality and utility of random shotgun sequencing at low redundancies. *Genome Res.* 1998; 8: 1074–1084.
- Bourhy P, Collet L, Brisse S, Picardeau M. *Leptospira mayottensis* sp. nov., a pathogenic species of the genus *Leptospira* isolated from humans. *Int J Syst Evol Microbiol.* 2014; 64: 4061–4067.
- Brenner DJ, Kaufmann AF, Sulzer KR, Steigerwalt AG, Rogers FC, Weyant RS. Further determination of DNA relatedness between serogroups and serovars in the family *Leptospiraceae* with a proposal for *Leptospira alexanderi* sp. nov. and four new *Leptospira* genomospecies. *Int J Syst Bacteriol.* 1999; 49: 839–858.
- Bulach DM, Zuerner RL, Wilson P, Seemann T, McGrath A, Cullen PA, Davis J, Johnson M, Kuczek E, Alt DP, et al. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. *Proc Natl Acad Sci U S A.* 2006; 103: 14560–14565.
- Carini A. Espirochaetose icterohemorrágica nos ratos em São Paulo. *An Paul Med Cir.* 1918; 9:70.
- Carmona–Gasca CA, Lara LL, Castillo–Sánchez LO, Ramírez–Ortega JM, Ko A, Palomera CL, de la Peña–Moctezuma A. Detection of *Leptospira santarosai* and *L. kirschneri* in cattle: new isolates with potential impact in bovine production and public health. *Vet Mex.* 2011; 42: 277-288.
- Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics.* 2012; 28: 464-469.
- Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. *Bioinformatics.* 2005; 16: 3422-3.
- Chiriboga J, Barragan V, Arroyo G, Sosa A, Birdsell DN, España K, Mora A, Espín E, Mejía ME, Morales M, et al. High prevalence of intermediate *Leptospira* spp. DNA in febrile humans from urban and rural Ecuador. *Emerg Infect Dis.* 2015; 21: 2141-2147.
- Chou L-F, Chen T-W, Ko Y-C, Pan M-J, Tian Y-C, Chiu C-H, Tang P, Hung C-C, Yang C-W. Potential impact on kidney infection: a whole-genome analysis of *Leptospira santarosai* serovar Shermani. *Emerg Microbes Infec.* 2014; 3:e82.

- Chou LF, Chen YT, Lu CW, Ko YC, Tang CY, Pan MJ, Tian YC, Chiu CH, Hung CC, Yang CW. Sequence of *Leptospira santarosai* serovar Shermani genome and prediction of virulence-associated genes. *Gene*. 2012; 511: 364–370.
- Cosson JF, Picardeau M, Mielcarek M, Tatard C, Chaval Y, Suputtamongkol Y, Buchy P, Jittapalapong S, Herbreteau V, Morand S. Epidemiology of *Leptospira* transmitted by rodents in Southeast Asia. *PLoS Medicine*. 2014; 8:1-10.
- Cunha CEP, Felix SR, Seixas Neto ACP, Campello-Felix A, Kremer FS, Monte LG, Amaral MG, de Oliveira Nobre M, da Silva ÉF, Hartleben CP, et al. Infection with *Leptospira kirschneri* serovar Mozdok: first report from the Southern Hemisphere. *Am J Trop Med Hyg*. 2016; 94: 519-521.
- Darling AE, Mau B, Perna NT. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*. 2010; 5: e11147.
- Desvars A, Naze F, Vourc'h G, Cardinale E, Picardeau M, Michault A, Bourhy P. Similarities in *Leptospira* serogroup and species distribution in animals and humans in the Indian ocean island of Mayotte. *Am J Trop Med Hyg*. 2012; 87: 134–140.
- Ekblom R, Wolf JBW. A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl*. 2014; 7: 1026-1042.
- Enrietti MA. Contribuição ao conhecimento da incidência de leptospirosas em murédeos, caninos e suínos no Paraná. *Braz Arch Biol Technol*. 2001; vol jubilee: 311-342.
- Eslabão MR. Sequencing, assembly and genome annotation of a new isolated of *Leptospira borgpetersenii* [dissertação]. Pelotas: Universidade Federal de Pelotas, 2012.
- Faine S, Adler B, Bolin C, Perolat P. *Leptospira* and leptospirosis. 3rd ed. Melbourne: MediSci; 1999.
- Faine S. *Leptospira* and leptospirosis. Boca Raton: CRC Press; 1994.
- Fernandes LG, Siqueira GH, Teixeira AR, Silva LP, Figueredo JM, Cosate MR, Vieira ML, Nascimento AL. *Leptospira* spp.: Novel insights into host-pathogen interactions. *Vet Immunol Immunopathol*. 2016; 176: 50-57.
- Ferreira Neto JS, Vasconcellos SA, Ito FH, Moretti AS, Camargo CA, Sakamoto SM, Marangon S, Turilli C, Martini M. *Leptospira interrogans* serovar Icterohaemorrhagiae seropositivity and the reproductive performance of sows. *Prev Vet Med*. 1997; 31: 87-93.

- Fouts DE, Matthias MA, Adhikarla H, Adler B, Amorim-Santos L, Berg DE, Bulach D, Buschiazzo A, Chang YF, Galloway RL, et al. What makes a bacterial species pathogenic?: Comparative genomic analysis of the genus *Leptospira*. PLoS Negl Trop Dis. 2016; 10: e0004403.
- Galardini M, Biondi EG, Bazzicalupo M., Mengoni A. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. Source Code Biol Med. 2011; 6: 11.
- Galloway RL, Levett PN. Application and validation of PFGE for serovar identification of *Leptospira* clinical isolates. PLoS Negl Trop Dis. 2010; 4: 1–7.
- Gangadhar NL, Prabhudas K, Bhushan S, Sulthana M, Barbuddhe SB, Rehaman H. *Leptospira* infection in animals and humans: a potential public health risk in India. Rev Sci Tech. 2008; 27: 885-892.
- Gangadhar NL, Rajasekhar M, Smythe LD, Norris MA, Symonds ML, Dohnt MF. Reservoir hosts of *Leptospira inadai* in India. Rev Sci Tech OIE. 2000; 19: 793–799.
- Haake DA, Levett PN. Leptospirosis in Humans. Curr Top Microbiol Immunol. 2015; 387: 65–97.
- Hartskeerl RA, Terpstra WJ. Leptospirosis in wild animals. Vet Quart. 1996; 18: S149–S150.
- Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, Jenney A, Connor TR, Hsu LY, Severin J, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. Proc Natl Acad Sci USA. 2015; 112: E3574-81.
- Holt KE, Baker S, Dongol S, Basnyat B, Adhikari N, Thorson S, Pulickal AS, Song Y, Parkhill J, Farrar JJ, et al. High-throughput bacterial SNP typing identifies distinct clusters of *Salmonella Typhi* causing typhoid in Nepalese children. BMC Infect Dis. 2010; 10: 144.
- Hsieh WJ, Pan MJ. Identification *Leptospira santarosai* serovar Shermani specific sequences by suppression subtractive hybridization. FEMS Microbiol Lett. 2004; 235: 117–124.
- Ido Y, Hoki R, Ito H, Wani H. The rat as a carrier of *Spirochaeta icterohaemorrhagiae*, the causative agent of Weil's disease (spirochaetosis icterohaemorrhagica). J Exp Med. 1917; 26: 341–353.

- Inada R, Ido Y, Hoki R, Kaneko R, Ito H. The etiology, mode of infection, and specific therapy of Weil's disease (spirochaetosis icterohaemorrhagica). *J Exp Med*. 1916; 23: 377–402.
- Johnson RC, Faine S. *Leptospira*. In: Krieg NR, Holt JG, editors. *Bergey's manual of systematic bacteriology*. Baltimore: Williams & Wilkins; 1984. p. 62–67.
- Ko AI, Goarant C, Picardeau M. *Leptospira*: the dawn of the molecular genetics era for an emerging zoonotic pathogen. *Nat Rev Microbiol*. 2009; 7: 736–747.
- Kremer FS, Eslabão MR, Provisor M, Woloski RD, Ramires OV, Moreno LZ, Moreno AM, Hamond C, Lilenbaum W, Dellagostin OA. Draft genome sequences of *Leptospira santarosai* strains U160, U164, and U233, isolated from asymptomatic cattle. *Genome Announc*. 2015; 3: e00910-15.
- Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O. Multilocus sequence typing of total genome sequenced bacteria. *J Clin Microbiol*. 2012; 50: 1355-1361.
- Lehmann JS, Corey VC, Ricaldi JN, Vinetz JM, Winzeler EA, Matthias MA. Whole Genome Shotgun sequencing shows selection on *Leptospira* regulatory proteins during *in vitro* culture attenuation. *American J Trop Med Hyg*. 2016; 94: 302-313.
- Levett PN, Smythe L. International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of *Leptospiraceae* - Minutes of the closed meeting, 9 October 2013, Fukuoka, Japan. *Int J Syst Evol Micr*. 2014; 64: 1071–1072.
- Levett PN. Leptospirosis. *Clin Microbiol Rev*. 2001; 14: 296–326.
- Lilenbaum W, Morais ZM, Gonçalves AP, Souza GO, Richtzenhain L, Vasconcellos SA. First isolation of leptospires from dairy goats in Brazil. *Braz J Microbiol*. 2007; 38: 507-510.
- Mardis E, McPherson J, Martienssen R, Wilson RK, McCombie WR. What is finished, and why does it matter. *Genome Res*. 2002; 12: 669-671.
- Martins G, Loureiro AP, Hamond C, Pinna MH, Bremont S, Bourhy P, Lilenbaum W. First isolation of *Leptospira noguchii* serogroups Panama and Autumnalis from cattle. *Epidemiol Infect*. 2015; 143: 1538–1541.

- Marvullo MFV, Silva JCR, Ferreira PM, Morais ZM, Moreno AM, Doto DS, Paixão R, Baccaro R, Vasconcellos SA, Ferreira Neto JS. Experimental leptospirosis in capybaras (*Hydrochaeris hydrochaeris*) infected with *Leptospira interrogans* serovar Pomona. J Zoo Wildl Med. 2009; 40: 726–730.
- McBride AJ, Athanazio DA, Reis MG, Ko AI. Leptospirosis. Curr Opin Infect Dis. 2005; 18: 376–386.
- Miotto BA, Moreno LZ, Guilloux AGA, Sousa GO, Loureiro AP, Moreno AM, Lilenbaum W, Vasconcellos SA, Heinemann MB, Hagiwara MK. Molecular and serological characterization of the first *Leptospira santarosai* strain isolated from a dog. Acta Tropica. 2016; 162: 1-4.
- Monte LG, Ridieri KF, Jorge S, Oliveira NR, Hartwig DD, Amaral MG, Hartleben CP, Dellagostin OA. Immunological and molecular characterization of *Leptospira interrogans* isolated from a bovine foetus. Comp Immunol Microbiol Infect Dis. 2015; 40: 41-45.
- Moreno LZ, Miraglia F, Marvulo MFV, Silva JCR, Paula CD, Costa BLP, Morais ZM, Ferreira F, Ferreira Neto JS, Dellagostin OA, et al. Characterization of *Leptospira santarosai* serogroup grippotyphosa serovar bananal isolated from Capybara (*Hydrochaeris hydrochaeris*) in Brazil. J Wildlife Dis. 2016a; 52: 688-693.
- Moreno LZ, Miraglia F, Lilenbaum W, Neto JSF, Freitas JC, Morais ZM, Hartskeerl RA, Costa BLP, Vasconcellos SA, Moreno AM. Profiling of *Leptospira interrogans*, *L. santarosai*, *L. meyeri* and *L. borgpetersenii* by SE-AFLP, PFGE and broth microdilution – a continuous attempt to species and serovars differentiation. Emerg Microbes Infect. 2016b; 5: e17.
- Nagarajan N, Pop M. Sequence assembly demystified. Nat Rev Gen. 2013; 14: 157–167.
- Nascimento AL, Ko AI, Martins EA, Monteiro-Vitorello CB, Ho PL, Haake DA, Verjovski-Almeida S, Hartskeerl RA, Marques MV, Oliveira MC, et al. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. J Bacteriol. 2004; 186: 2164–2172.
- Noguchi H. Morphological characteristics and nomenclature of *Leptospira* (*Spirochaeta*) *icterohaemorrhagiae* (Inada and Ido). J Exp Med. 1918; 27: 575-592.

- Perolat P, Grimont F, Regnault B, Grimont PAD, Fournie E, Thevenet H, Baranton G. rRNA gene restriction patterns of *Leptospira*: a molecular typing system. *Res Microbiol*. 1990; 141: 159–171.
- Picardeau M, Bulach DM, Bouchier C, Zuerner RL, Zidane N, Wilson PJ, Creno S, Kuczek ES, Bommezzadri S, Davis JC, et al. Genome sequence of the saprophyte *Leptospira biflexa* provides insights into the evolution of *Leptospira* and the pathogenesis of leptospirosis. *PLoS One*. 2008; 3: e1607.
- Ramos ACF, Souza GN, Lilenbaum W. Influence of leptospirosis on reproductive performance of sows in Brazil. *Theriogenology*. 2006; 66: 1021-1025.
- Ren SX, Fu G, Jiang XG, Zeng R, Miao YG, Xu H, Zhang YX, Xiong H, Lu G, Lu LF, et al. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature*. 2003; 422: 888–893.
- Ricaldi JN, Fouts DE, Selengut JD, Harkins DM, Patra KP, Moreno A, Lehmann JS, Purushe J, Sanka R, Torres M, et al. Whole genome analysis of *Leptospira licerasiae* provides insight into leptospiral evolution and pathogenicity. *PLoS Negl Trop Dis*. 2012; 6: e1853.
- Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics*. 2009; 25: 2071–2073.
- Sambrook JE, Russel DW. *Molecular cloning: a laboratory manual*. 3 ed. New York: Cold Spring Harbor Laboratory Press; 2002. p. 2700.
- Sanford JP. Leptospirosis - time for a booster. *New England J Med*. 1984; 310: 524-525.
- Santo LA. *Leptospira interrogans* sorovar Copenhageni e Icterohaemorrhagiae: relação evolutiva, diferenças genéticas e associação com desfecho clínico [tese]. Salvador: Fundação Oswaldo Cruz, Centro de Pesquisa Gonçalo Moniz, Salvador; 2015.
- Sarmiento AMC, Azevedo SS, Morais ZM, Souza GO, Oliveira FCS, Gonçalves AP, Miraglia F, Vasconcellos SA. Emprego de estirpes *Leptospira* spp. isoladas no Brasil na microtécnica de soroaglutinação microscópica aplicada ao diagnóstico da leptospirose em rebanhos bovinos de oito estados brasileiros. *Pes Vet Bras*. 2012; 32: 601-606.

- Schuller S, Francey T, Hartmann K, Hugonnard M, Kohn B, Nally JE, Sykes J. European consensus statement on leptospirosis in dogs and cats. *J Small Anim Pract.* 2015; 56: 159–179.
- Silva EF, Cerqueira GM, Seyffert N, Seixas FK, Hartwig DD, Athanazio DA, Pinto LS, Queiroz A, Ko AI, Brod CS, Dellagostin AO. *Leptospira noguchii* and human and animal leptospirosis, Southern Brazil. *Emerg Infect Dis.* 2009; 15: 621-623
- Stimson AM. Note on an organism found in yellow-fever tissue. *Public Health Rep.* 1907; 22: 541.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 2016; 44: 6614-6624.
- Terpstra WJ, Schoone GJ, Schegget JT. Detection of leptospiral DNA by nucleic acid hybridization with ³²P- and biotin-labelled probes. *J Med Microbiol.* 1986; 22: 23-28.
- Terpstra WJ, Ligthart GS, Schoone GJ. ELISA for the detection of specific IgM and IgG in human leptospirosis. *J Gen Microbiol.* 1985; 131: 377–385.
- Thaipadungpanit J, Wuthiekanun V, Chierakul W, Smythe LD, Petkanchanapong W, Limpaboon R, Apiwatanaporn A, Slack AT, Suputtamongkol Y, White NJ, et al. A dominant clone of *Leptospira interrogans* associated with an outbreak of human leptospirosis in Thailand. *PLoS Negl Trop Dis.* 2007; 1: e56.
- Valverde MA, Goris M, González V, Anchía ME, Díaz P, Ahmed A, Hartskeerl RA. New serovars of *Leptospira* isolated from patients in Costa Rica: implications for the public health. *J Med Microbiol.* 2013; 62: 1263–1271.
- van Domselaar GH, Stothard P, Shrivastava S, Cruz JA, Guo A, Dong X, Lu P, Szafron D, Greiner R, Wishart DS. BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res.* 2005; 33: W455-459.
- Varni V, Ruybal P, Lauthier JJ, Tomasini N, Brihuega B, Koval A, Caimi K. Reassessment of MLST schemes for *Leptospira* spp. typing worldwide. *Infect Genet Evol.* 2014; 22: 216-222.
- Vasconcellos SA, Oliveira JCF, Morais ZM, Baruselli PS, Amaral R, Pinheiro SR, Ferreira F, Ferreira Neto JS, Schouberg A, Hartskeerl RA. Isolation of *Leptospira santarosai*,

serovar Guaricura from buffaloes (*Bubalus bubalis*) in Vale do Ribeira, São Paulo, Brazil. Braz J Microbiol. 2001; 32: 298-300.

Vieira ML, Pimenta DC, de Moraes ZM, Vasconcellos SA, Nascimento ALT. Proteome analysis of *Leptospira interrogans* virulent strain. Open Microbiol J. 2009; 3: 69-74.

Wang Y, Zhuang X, Zhong Y, Zhang C, Zhang Y, Zeng L, Zhu Y, He P, Dong K, Pal U, Guo X, Qin J. Distribution of plasmids in distinct *Leptospira* pathogenic species. PLoS Negl Trop Dis. 2015; 9: e0004220.

Weil A. Ueber eine eigenthümliche, mit Milztumor, Icterus und Nephritis einhergehende, acute Infectionskrankheit. Dtsch Arch Klin Med. 1886; 39: 209.

Wolbach SB, Binger CAL. Notes on filterable spirochaete from fresh water. *Spirochaeta biflexa* (new species). J Med Res. 1914; 30: 23-26.

Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, Kingsley RA, Thomson NR, Keane JA, Weill FX, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella Typhi* identifies inter- and intracontinental transmission events. Nat Genet. 2015; 47: 632-639.

Xu Y, Zhu Y, Wang Y, Chang YF, Zhang Y, Jiang X, Zhuang X, Zhu Y, Zhang J, Zeng L, et al. Whole genome sequencing revealed host adaptation-focused genomic plasticity of pathogenic *Leptospira*. Sci Rep. 2016; 6: 200-20.

Yasuda PH, Steigerwalt AG, Sulzer KR, Kaufmann AF, Rogers F, Brenner DJ. Deoxyribonucleic acid relatedness between serogroups and serovars in the family *Leptospiraceae* with proposals for seven new *Leptospira* species. Int J Syst Bacteriol. 1987; 37: 407-415.

Zhu WN, Huang LL, Zeng LB, Zhuang XR, Chen CY, Wang YZ, Qin JH, Zhu YZ, Guo XK. Isolation and characterization of two novel plasmids from pathogenic *Leptospira interrogans* serogroup Canicola Serovar Canicola strain Gui44. PLoS Negl Trop Dis. 2014; 8: e3103.

9. ANEXOS

Anexo 1: Artigo publicado referente à primeira etapa do projeto – caracterização fenotípica e genotípica das estirpes da coleção de cultura para seleção do grupo de estirpes para sequenciamento de genoma.

OPEN

Emerging Microbes and Infections (2016) 5, e17; doi:10.1038/emi.2016.16
www.nature.com/emi

ORIGINAL ARTICLE

Profiling of *Leptospira interrogans*, *L. santarosai*, *L. meyeri* and *L. borgpetersenii* by SE-AFLP, PFGE and susceptibility testing—a continuous attempt at species and serovar differentiation

Luisa Z Moreno¹, Fabiana Miraglia^{1,2}, Walter Lilenbaum², José SF Neto¹, Julio C Freitas³, Zenaide M Morais¹, Rudy A Hartskeerl⁴, Barbara LP da Costa¹, Silvio A Vasconcellos¹ and Andrea M Moreno¹

Leptospirosis is a widespread systemic zoonosis, considered as reemerging in certain developing countries. Although the cross agglutinin absorption test is still considered the standard method for *Leptospira* identification, it presents several disadvantages. The aim of this study was to characterize *Leptospira* spp. isolated from various hosts by genotyping and broth microdilution susceptibility testing in an attempt to differentiate *Leptospira* species, serogroups and serovars. Forty-seven isolates were studied. They were previously serotyped, and species confirmation was performed by 16S rRNA sequencing. Single-enzyme amplified fragment length polymorphism (SE-AFLP) and pulsed-field gel electrophoresis (PFGE) analysis enabled the distinction of *L. interrogans* from *L. santarosai*, *L. meyeri* and *L. borgpetersenii* in two main clusters. Among *L. interrogans*, it was possible to differentiate into two new clusters the serogroup Icterohaemorrhagiae from the serogroups Canicola and Pomona. *L. santarosai* isolates presented higher genetic variation than the other species in both techniques. Interestingly, the minimum inhibitory concentration (MIC) cluster analysis also provided *Leptospira* serogroup differentiation. Further studies are necessary regarding serovar Bananal isolates, as they presented the highest MIC values for most of the antimicrobials tested. All studied techniques successfully distinguished *Leptospira* species and serogroups. Despite being library-dependent methods, these approaches are less labor intensive and more economically viable, particularly SE-AFLP, and can be implemented in most reference laboratories worldwide to enable faster *Leptospira* typing.

Emerging Microbes and Infections (2016) 5, e17; doi:10.1038/emi.2016.16; published online 9 March 2016

Keywords: leptospira; MIC; PFGE; SE-AFLP; sequencing; serotyping

INTRODUCTION

Leptospirosis is a worldwide systemic zoonosis, with higher incidence in tropical climates.¹ This disease is caused by bacteria of the *Leptospira* genus, which are classified into 21 species and nearly 300 serovars organized into 29 serogroups. Among these 21 established species, nine are characterized as pathogenic and frequently isolated from humans and animals; five are considered to be intermediately pathogenic with the ability to infect humans and animals, although less frequently and with variable clinical signs; and seven are considered saprophytic environmental non-pathogenic species.^{2–5}

Until now, the species identification level was defined by DNA-DNA hybridization data, whereas *Leptospira* serogroup and serovar classification is based on the expression of the surface antigens.² Further serological identification is complicated because various serovars can be distributed among different species.⁶ Although the cross-agglutinin absorption test (CAAT) is still considered the

standard method for *Leptospira* identification, it is highly laborious and expensive because it requires the maintenance of all reference strains and the production of respective antisera.

In this context, molecular methods, with higher discriminatory power and the ability to establish the molecular epidemiology of the isolates and intra-serovar variation, have been applied for *Leptospira* characterization.^{3,7–10} Despite the widespread use of sequence-based molecular methods, with low- to high-throughput scales, they require the use of expensive equipment, rigorously standardized sample preparation protocols and complex bioinformatics analysis. Some genotyping methods of restriction patterns can be quicker and easier to perform; digital analysis enables the standardization and more accurate interpretation of band patterns¹⁰ and is more economically viable for the vast majority of researchers from developing countries.

This study aimed to characterize *Leptospira* spp. isolated from various hosts in Brazil, at different time periods, by single-enzyme

¹Laboratory of Molecular Epidemiology and Antimicrobial Resistance/Laboratory of Bacterial Zoonosis, School of Veterinary Medicine and Animal Science, University of São Paulo, São Paulo/SP 05508-270, Brazil; ²Veterinary Bacteriology Laboratory, Department of Microbiology and Parasitology, Fluminense Federal University, Fluminense Mello, 101 Niterói/RJ 24210-130, Brazil; ³Laboratory of Animal Leptospirosis, Londrina State University, Rod Celso Garcia Cid, Londrina/PR 86057-970, Brazil and ⁴WHO/FAO/OIE and National Leptospirosis Reference Centre, KIT Biomedical Research, Royal Tropical Institute, Amsterdam 391105, Netherlands

Correspondence: AM Moreno

E-mail: morenam@usp.br

Received 23 July 2015; revised 25 September 2015; accepted 23 November 2015

Anexo 2: Artigo publicado referente à caracterização do genoma de *L. noguchii*.

Draft Genome Sequence of Brazilian *Leptospira noguchii* Serogroup Panama Strain U73, Isolated from Cattle

Luisa Z. Moreno,^a Ana P. Loureiro,^b Fabiana Miraglia,^b Carlos E. C. Matajira,^a Frederico S. Kremer,^c Marcos R. Eslabao,^c Odir A. Dellagostin,^c Walter Lilenbaum,^b Andrea M. Moreno^a

Laboratório de Epidemiologia Molecular e Resistência a Antimicrobianos, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, Brazil;^a Departamento de Microbiologia e Parasitologia, Laboratório de Bacteriologia Veterinária, Universidade Federal Fluminense, Rio de Janeiro, Brazil;^b Centro de Desenvolvimento Tecnológico, Núcleo de Biotecnologia, Universidade Federal de Pelotas, Pelotas, Brazil^c

***Leptospira noguchii* is a current zoonotic pathogen in Brazil. Here, we report the draft genome sequence of the Brazilian *L. noguchii* serogroup Panama strain U73, isolated from asymptomatic cattle urine.**

Received 25 August 2015 Accepted 3 September 2015 Published 15 October 2015

Citation Moreno LZ, Loureiro AP, Miraglia F, Matajira CC, Kremer FS, Eslabao MR, Dellagostin OA, Lilenbaum W, Moreno AM. 2015. Draft genome sequence of Brazilian *Leptospira noguchii* serogroup Panama strain U73, isolated from cattle. *Genome Announc* 3(5):e01179-15. doi:10.1128/genomeA.01179-15.

Copyright © 2015 Moreno et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Andrea M. Moreno, morenoam@usp.br.

Leptospira noguchii is one of the pathogenic species of the *Leptospira* genus. It is widely described as a zoonotic pathogen in the Americas (1–5). In Brazil, *L. noguchii* has already been isolated from diseased humans and a dog (3) and also from asymptomatic sheep and cattle (4, 5). Although the epidemiology of *L. noguchii* infection has not yet been elucidated, its isolation from humans and livestock reinforces the importance of this species in the Brazilian leptospirosis scenario.

Here, we present the draft genome of *L. noguchii* serogroup Panama strain U73, isolated from asymptomatic cattle urine in Rio de Janeiro, Brazil. Isolation, serogrouping, and analysis of partial *rrs* and *secY* sequences of this strain were previously described by Martins et al. (5). Genomic DNA was extracted and purified with the illustra bacteria genomicPrep mini spin kit (GE Healthcare) and used for paired-end library preparation with the Nextera DNA sample prep kit (Illumina) and sequencing through the Illumina MiSeq platform.

Sequencing resulted in 789,252 paired-end quality filtered reads with approximately 50-fold coverage. The *de novo* assembly was performed with CLC Main Workbench version 7.5.1 (CLC Bio) and Geneious version 8.0.5 (Biomatters Ltd.) and resulted in 279 scaffolds with an N_{50} of 24,125. The U73 draft genome comprises ~4.76 Mbp, with an overall G+C content of 35.63%. Automatic genome annotation was performed with BASys (Bacterial Annotation System) (6).

In total, 4,535 coding sequences, 37 tRNAs, and 4 rRNAs were identified. In addition, ISFinder (7) was used for the prediction of insertion sequence (IS) elements; IS1500, IS1501, IS1502, and IS-Lin2 were identified in the U73 draft genome. Regarding the virulence factors, strain U73 presents genes encoding lipoproteins and immunoglobulin-like proteins, such as *lipL41*, *lipL36*, *lolC/D*, and *ligA*, which correspond to the main virulence factors of *L. interrogans*. Interestingly, several antimicrobial resistance genes were identified, including efflux pumps (*mdtA*, *mdtB*, *norM*), as well as metal resistance genes (*sugE*, *czcA*, *czcD*).

A multilocus sequence typing (MLST) analysis was also performed using the MLST server (8). Novel sequence types (STs)

were identified for the U73 strain among the *Leptospira* MLST 2 and 3 schemes (ST97 and ST137, respectively). Despite the wide range of alleles and the strain's incongruity between the schemes, the U73 strain presented greater proximity with *L. noguchii* strains from serogroups Autumnalis and Louisiana at MLST schemes 2 and 3, respectively. There are only two more *L. noguchii* serogroup Panama in the *Leptospira* MLST database; however, they are more distant from the U73 strain than the aforementioned serogroups.

Although being pathogenic, *L. noguchii* has been scarcely studied. Sequencing of new strains and the increase of genetic and epidemiological data on this *Leptospira* species will enhance the knowledge about *L. noguchii* infection, transmission, and control.

Nucleotide sequence accession number. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number [LHQ700000000](https://www.ncbi.nlm.nih.gov/nuclseq/LHQ700000000).

ACKNOWLEDGMENTS

This study was supported by FAPESP (grants 2011/18290-0 and 2013/17136-2), CAPES, and CNPq. (A.M.M., O.A.D., and W.L. are CNPq fellows.)

REFERENCES

- Brenner DJ, Kaufmann AF, Sulzer KR, Stelgerwalt AG, Rogers FC, Weyant RS. 1999. Further determination of DNA relatedness between serogroups and serovars in the family *Leptospiraceae* with a proposal for *Leptospira alexanderi* sp. nov. and four new *Leptospira* genomospecies. *Int J Syst Bacteriol* 49:839–858. <http://dx.doi.org/10.1099/00207173-49-2-839>.
- Faine SB, Adler B, Bolin C, Perolat P. 1999. *Leptospira* and leptospirosis, 2nd ed. MediSci, Melbourne, Australia.
- Silva EF, Cerqueira GM, Seyffert N, Setas FK, Hartwig DD, Athanazio DA, Pinto LS, Queiroz A, Ko AI, Brod CS, Dellagostin OA. 2009. *Leptospira noguchii* and human and animal leptospirosis, Southern Brazil. *Emerg Infect Dis* 15:621–623. <http://dx.doi.org/10.3201/e1504.071669>.
- Silva EF, Brod CS, Cerqueira GM, Bourscheidt D, Seyffert N, Queiroz A, Santos CS, Ko AI, Dellagostin OA. 2007. Isolation of *Leptospira noguchii* from sheep. *Vet Microbiol* 121:144–149. <http://dx.doi.org/10.1016/j.vetmic.2006.11.010>.
- Martins G, Loureiro AP, Hamond C, Pinna MH, Bremont S, Bourhy P,

Anexo 3: Artigo publicado referente à caracterização e análise comparativa do genoma *L. kirschneri* sorovar Mozdok.

Mem Inst Oswaldo Cruz, Rio de Janeiro: 1-3, 2016 1

Comparative genomic analysis of Brazilian *Leptospira kirschneri* serogroup Pomona serovar Mozdok

Luisa Z Moreno¹, Frederico S Kremer², Fabiana Miraglia², Ana P Loureiro², Marcus R Eslabao³, Odir A Dellagostin³, Walter Lilenbaum², Andrea M Moreno¹†

¹Universidade de São Paulo, Faculdade de Medicina Veterinária e Zootecnia, Laboratório de Epidemiologia Molecular e Resistência a Antimicrobianos, São Paulo, SP, Brasil

²Universidade Federal Fluminense, Departamento de Microbiologia e Parasitologia, Laboratório de Bacteriologia Veterinária, Niterói, RJ, Brasil ³Universidade Federal de Pelotas, Centro de Desenvolvimento Tecnológico, Pelotas, RS, Brasil

Leptospira kirschneri is one of the pathogenic species of the *Leptospira* genus. Human and animal infection from *L. kirschneri* gained further attention over the last few decades. Here we present the isolation and characterisation of Brazilian *L. kirschneri* serogroup Pomona serovar Mozdok strain M36/05 and the comparative genomic analysis with Brazilian human strain 61H. The M36/05 strain caused pulmonary hemorrhagic lesions in the hamster model, showing high virulence. The studied genomes presented high symmetrical identity and the in silico multilocus sequence typing analysis resulted in a new allelic profile (ST101) that so far has only been associated with the Brazilian *L. kirschneri* serogroup Pomona serovar Mozdok strains. Considering the environmental conditions and high genomic similarity observed between strains, we suggest the existence of a Brazilian *L. kirschneri* serogroup Pomona serovar Mozdok lineage that could represent a high public health risk; further studies are necessary to confirm the lineage significance and distribution.

Key words: *L. kirschneri* - serovar Mozdok - genomics

Leptospirosis is an emerging worldwide zoonosis which is caused by spirochetes of the *Leptospira* genus (Levett 2001). To date, the genus comprises 11 pathogenic species (Bourhy et al. 2014) of which *Leptospira interrogans* is the most associated with human infection. Nevertheless, *L. kirschneri* infection has gained further attention over the last few decades.

Masuzawa et al. (2006) reported one of the first cases of *L. kirschneri* human infection by direct contact with southern flying squirrels imported from the United States; two Japanese workers were infected by *L. kirschneri* serovar Grippotyphosa from the imported American exotic pets. The *L. kirschneri* serogroup Pomona serovar Mozdok has also been related to human infection in Cuba (Obregón et al. 2007). In Europe, the serovar Mozdok is described as endemic in wild rodents (Majetic et al. 2014) and has also been associated with canine leptospirosis (Renaud et al. 2013).

In Brazil, *L. kirschneri* serogroup Pomona serovar Mozdok has recently been described as the causative agent of human and canine infection in different time points (Cunha et al. 2016). Here we present the isolation and characterisation of *L. kirschneri* serogroup Pomona serovar Mozdok strain M36/05 and the comparative genomic analysis with previously described Brazilian *L. kirschneri* serogroup Pomona serovar Mozdok human strain 61H.

The M36/05 strain was isolated from the kidney of a captured urban black rat (*Rattus rattus*) in Suzano, metropolitan region of São Paulo State, Brazil, in 2005. For isolation, 5 g of kidney sample was collected and homogenised in 30 mL of Sørensen saline, and 100 µL aliquots of 10⁻¹ to 10⁻³ dilutions were inoculated into duplicate tubes containing EMJH (DIFCO, USA) enriched with 15% rabbit serum, 5-fluorouracil and nalidixic acid. Once isolated, the strain was stored in Fletcher's medium (DIFCO/USA), enriched with 15% rabbit serum and maintained in EMJH broth (DIFCO/USA) at 30°C, as part of the *Leptospira* collection of the Laboratory of Bacterial Zoonosis - University of São Paulo.

Serogrouping was performed at the Laboratory of Veterinary Bacteriology - Fluminense Federal University. The isolate was subjected to microscopic agglutination test (MAT) using a panel of polyclonal rabbit antisera of 32 reference serovars representing the 24 known serogroups (provided by Royal Tropical Institute - KIT, Amsterdam). The M36/05 strain presented high agglutination rates with serogroup Pomona antisera which was then identified as the strain presumptive serogroup (Dikken & Kmety 1978).

To evaluate the strain virulence, five Golden Syrian hamsters (*Mesocricetus auratus*) were infected with M36/05 strain (10⁶ leptospores) through intraperitoneal route. The animal experiment was conducted with the approval of the Ethics Committee from the School of Veterinary Medicine and Animal Science - University of São Paulo (2244/2011). Clinical symptoms were checked daily. Within five days post-infection, the M36/05 strain caused lethal leptospirosis and the infected hamsters were euthanised. The inoculated animals developed acute infection characterised mainly by haemorrhagic pulmonary lesions (Fig. 1).

doi: 10.1590/0074-02760160174

Financial support: CNPq, CAPES, FAPESP (2011/18290-0, 2013/17136-2).

† Corresponding author: morenoam@usp.br

Received 27 April 2016

Accepted 13 June 2016