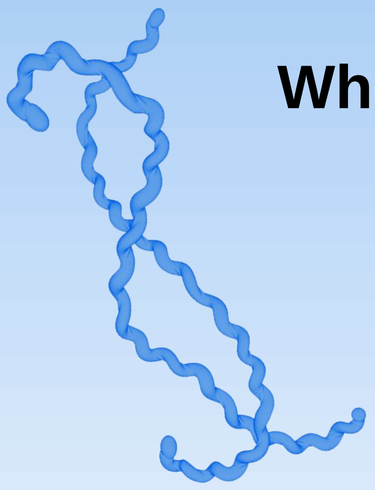


Whole genome sequencing and preliminary analysis of six *Leptospira* strains belonging to the serovar Hardjo

Alejandro Llanes[1], Carlos Restrepo[1] and Sree Rajeev[2*]

[1] INDICASAT AIP, Building 219, City of Knowledge, Panama City, Panama

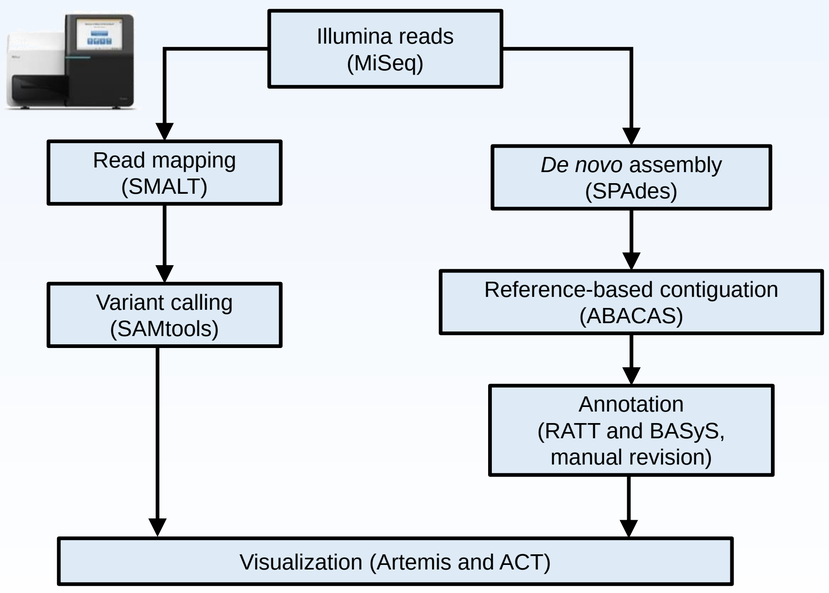
[2] Ross University School of Veterinary Medicine, Saint Kitts, West Indies

* E-mail: srajeev@rossvet.edu.kn

Leptospira borgpetersenii serovar Hardjo is a serovar maintained in the renal tubules of cattle and is implicated as a cause of reproductive loss in cattle. Two types of *L. borgpetersenii* serovar Hardjo (type A and type B) are reported with differences in genomic sequences and in clinical disease in the hamster model. In this study whole genome sequencing and assembly and a preliminary sequence analysis of three *Leptospira* field isolates obtained from cattle and three laboratory adapted strains (*L. borgpetersenii* serovar Hardjo type A, *L. borgpetersenii* serovar Hardjo type B, and *L.interrogans* serovar Hardjo) were performed. The genomes of *L. borgpetersenii* strains **L550** and **JB197** (Bulach et al. 2006) were used as reference for type A and type B Hardjo isolates, respectively for comparisons. The genome of *L. interrogans* serovar **Lai** (Ren et al. 2003) *L. interrogans* serovar **Copenhageni** (abbreviated **Cop**) (Nascimento et al. 2004) were used as reference for *L. interrogans* serovar Hardjo genome.

Leptospira isolates sequenced in this study

Label	Type	Species	Serovar	Origin
BK-6	Field isolate	*L. borgspetersenii*	Hardjo	Georgia, USA
BK-9	Field isolate	*L. borgspetersenii*	Hardjo	Georgia, USA
BK-30	Field isolate	*L. borgspetersenii*	Hardjo	Georgia, USA
LBH-A	Lab strain	*L. borgspetersenii*	Hardjo (type A)	Texas, USA
LBH-B	Lab strain	*L. borgspetersenii*	Hardjo (type B)	Texas, USA
LIH	Lab strain	*L. interrogans*	Hardjo	Indonesia



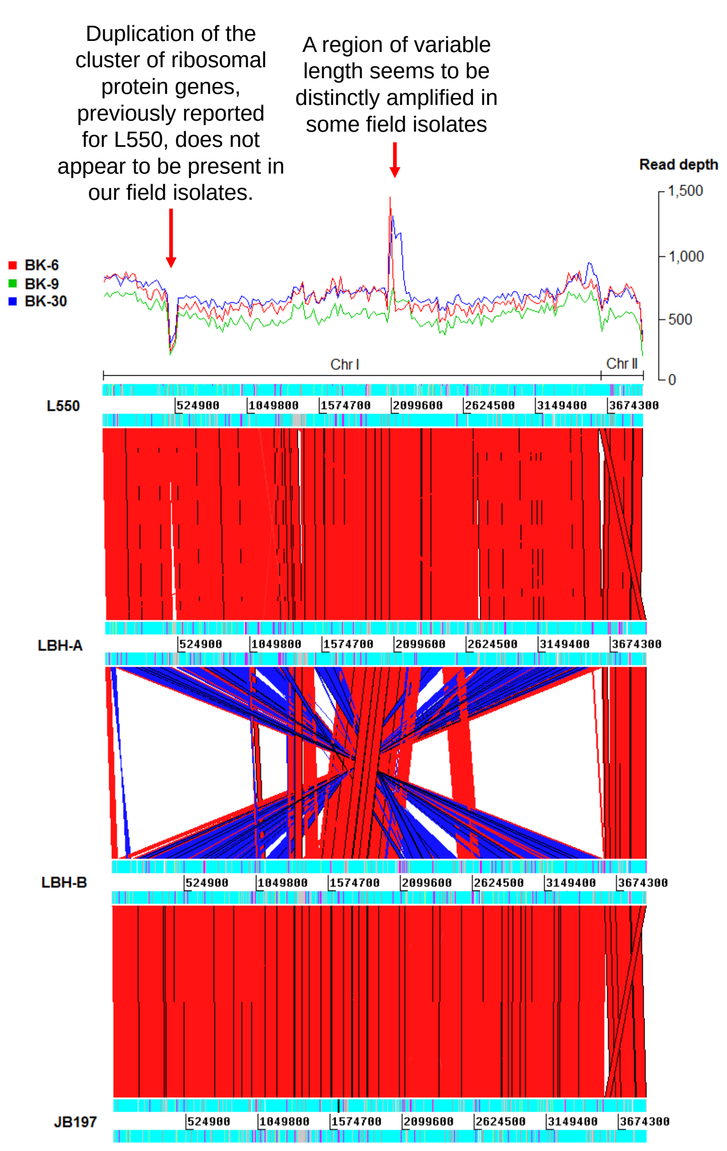
Read mapping and variant calling

Label	Total reads	Reference	Reads mapped	Reads unmapped	Total variants	SNP	INS	DEL
BK-6	9,467,714	L550	99.8%	0.2%	52	51	0	1
BK-9	7,716,262	L550	99.8%	0.2%	45	42	0	3
BK-30	9,903,708	L550	99.8%	0.2%	42	39	0	3
LBH-A	8,224,194	L550	99.9%	0.1%	43	40	0	3
LBH-B	7,467,366	JB197	99.6%	0.4%	3,848	3,777	58	13
LIH	7,929,346	Lai	97.4%	2.6%	28,828	28,561	192	75
		Cop	97.5%	2.5%	28,402	28,223	114	65

Read mapping allowed as to identify all the field isolates as type A. The number of variants is higher between type B isolate and its reference, than between the type A isolates and their reference.

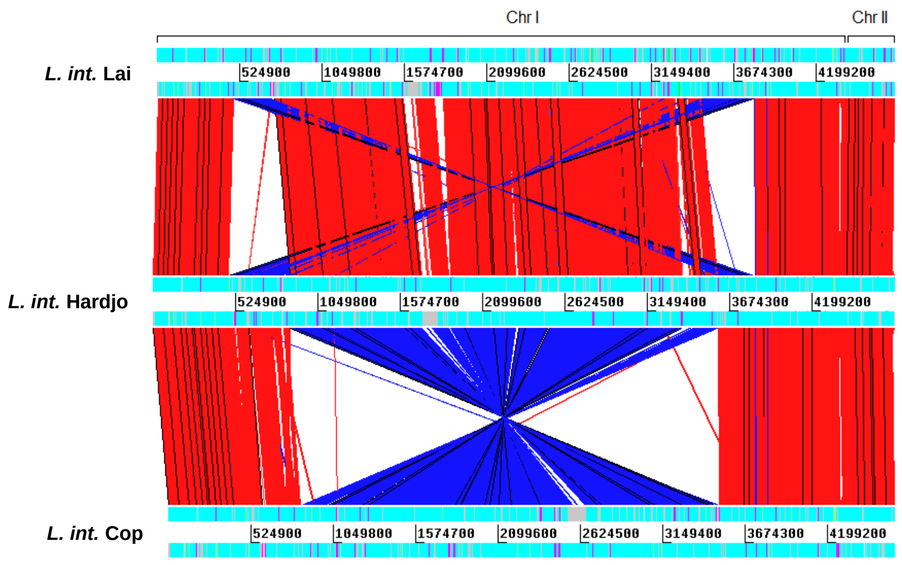
LIH showed a reasonably higher number of variants when compared to Lai and Copenhageni, however, the number seems to be slightly lower between LIH and the later.

***L. borgpetersenii* serovar Hardjo (Hardjobovis)**



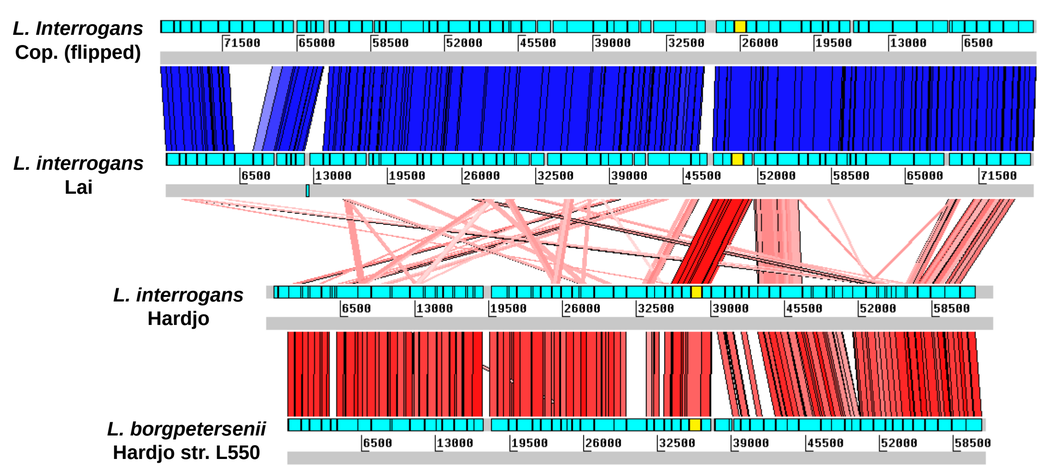
Although L550 and LB197 genomes where used as template for contiguation of the LBH-A and LBH-B genomes, fragments that could be assembled *de novo* were very similar in structure to those of the corresponding reference genomes.

***L. Interrogans* serovar Hardjo (Hardjoprajitno)**



Comparison of the ordered contigs resulting from *de novo* assembly revealed a lower number of predicted rearrangements between the genomes of serovar Lai and Hardjo, when compared to serovar Copenhageni. There are several regions with predicted protein-coding genes that are absent from both the Lai and Copenhageni genomes.

The *rfb* loci (LPS biosynthesis genes)



The part of the *rfb* locus that is divergent among species is compared at the amino acid level. As expected, the loci are very similar between LIH and L550, despite they are different species. The four serovars share high similarity in five genes located near the middle of the cluster, centered on the gene predicted to encode a 3-deoxy-manno-octulosonate cytidylyltransferase (yellow).

Legend

- Sequence match
- Inverted match

NOTE: Red/blue intensity is proportional to the percent identity of the match.

- Protein-coding genes
- Transfer RNA genes
- Transposase genes

References

Bulach DM et al. (2006) *PNAS* **103**(39):14560-5.

Nascimento AL et al. (2004) *J Bacteriol.* **186**(7):2164-72.

Ren SX et al. (2003) *Nature*. **422**(6934):888-93.

Information for all the sequencing experiments generated in this work can be found in NCBI BioProjects PRJNA296675, PRJNA296677, PRJNA296679, PRJNA296687, PRJNA296689 and PRJNA296694. All sequence data will be accessible upon publication.

ROSS UNIVERSITY
SCHOOL OF VETERINARY MEDICINE

INDICASAT AIP
INSTITUTO DE INVESTIGACIONES CIENTÍFICAS Y SERVICIOS DE ALTA TECNOLOGÍA