

Supplemental Information

INVA PROTEIN: A NUDIX HYDROLASE REQUIRED FOR INFECTION BY PATHOGENIC *LEPTOSPIRA* IN CELL LINES AND ANIMALS*

Yihui Luo^{1,2,7}, Yan Liu^{3,7}, Dexter Sun⁴, David M. Ojcius⁵, Jinfang Zhao^{1,2}, Xuai Lin^{1,2}, Dong Wu³, Rongguang Zhang³, Ming Chen⁶, Lanjuan Li¹, Jie Yan^{1,2}

Supplemental Experimental Procedures

Leptospiral strains - Information about the leptospiral strains used in this study are presented in Table S1.

Primers- All the primers used in this study were synthesized by Invitrogen Co. Shanghai, China (Table S2).

Amplification and sequence analysis of invA genes in different leptospiral strains- Leptospiral DNA was extracted using a Bacterial Genomic DNA Extraction Kit (BioColor, Shanghai, China) and then dissolved in TE buffer. The density and purity of the extracted DNAs were measured by UV spectrophotometry. With a High Fidelity PCR Kit (TaKaRa, Dalian, China), in which the Taq-Pfu mixture was used as DNA polymerase, one pair of the primers N1F and N1R (Table S2) was used to amplify the entire *invA* gene. The product was detected in 1.5% ethidium bromide pre-stained agarose gels after electrophoresis. To obtain more accurate sequencing data, the PCR products were purified using a PCR Products Purification Kit (BioColor) and then ligated into plasmid pMD18-T using a T-A Cloning Kit (TaKaRa). The cloned *invA* genes from the leptospiral strains were sequenced by Invitrogen Co. (Shanghai, China). A phylogenetic tree from comparison among the actual InvA sequence of *L. interrogans* strain Lai and the sequences of other characterized microbial Nudix enzymes were analyzed using MEGA 4.0 software [1].

Supplemental References

1. Kumar, S., Nei, M., Dudley, J., and Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* **9**(4), 299-306.

FIGURE LEGENDS

Fig. S1. Strategy for generation of the *invA*⁻ mutant and *invA*^{com} revertant based on homologous recombination.

LA3978/P in the figure means that the DNA segment is composed of LA3978 gene and the promoter of *invA*-LA3976-LA3975 operon. See Methods for details.

Fig. S2. Distribution and sequencing of *invA* genes from *L. interrogans* species.

A: Detection of *invA* genes in different *Leptospira* species by PCR.

Lane M: DNA ladder (TaKaRa); Lane 1: blank control; Lanes 2 to 16: *invA* gene fragments amplified from the ten *L. interrogans* strains belonging to serovars Lai, Canicoda, Pyrogenes, Autumnalis, Australis, Pomona, Grippotyphosa, Hebdomadis, Paidjan and Wolffii, four *L. borgpetersenii* strains belonging to serovars Javanica, Ballum, Tarassovi and Mini, and one *L. weili* strain belonging to serovar Manhao 2. Lanes 17 and 18: negative PCR results for the *L. biflexa* strains belonging to serovars Pactoc and Andamana.

B: Sequences of *invA* genes from different *Leptospira* species.

(1): Nucleotide and amino acid sequences of *invA* genes from the ten *L. interrogans* strains belonging to serovars Lai, Canicoda, Pyrogenes, Autumnalis, Australis, Pomona, Grippotyphosa, Hebdomadis, Paidjan and Wolffii, and four *L. borgpetersenii* strains belonging to serovars Javanica, Ballum, Tarassovi and Mini.
(2): Nucleotide and amino acid sequences of the *invA* gene from the *L. weili* strain belonging to serovar Manhao 2. The underlined areas indicate the primer positions.

Fig. S3. Comparison of motif and phylogenesis among leptospiral InvA and other Nudix hydrolases.

A: Nudix motif and multiple alignments of leptospiral InvA and other bacterial Nudix hydrolases. The closest matches are to *R. prowazekii* InvA (No.Q9ZDT9 in NCBI), *B. bacilliformis* IalA (No. P35640 in NCBI), *H. pylori* NudA (No. O25826 in NCBI), *C. jejuni* dinucleoside oligophosphate hydrolase (No.Q9PHT5 in NCBI), *E. coli* YgdP (No. POA776 in NCBI), *P. aeruginosa* invasion protein homolog (No.Q9X4P2 in NCBI), *N. meningitidis* Ap₄A pyrophosphatase (No.Q9JT78 in NCBI), *H. influenzae* InvA (No.Q57045 in NCBI) and *S. typhimurium* hydrolase (No. P65555 in NCBI). The Nudix motif is indicated by “**”, and identical amino acids are marked with a black background.

B: Molecular phylogenetic relationship between leptospiral InvA and other bacterial Nudix hydrolases.

Table S1. Leptospiral strains used in this study

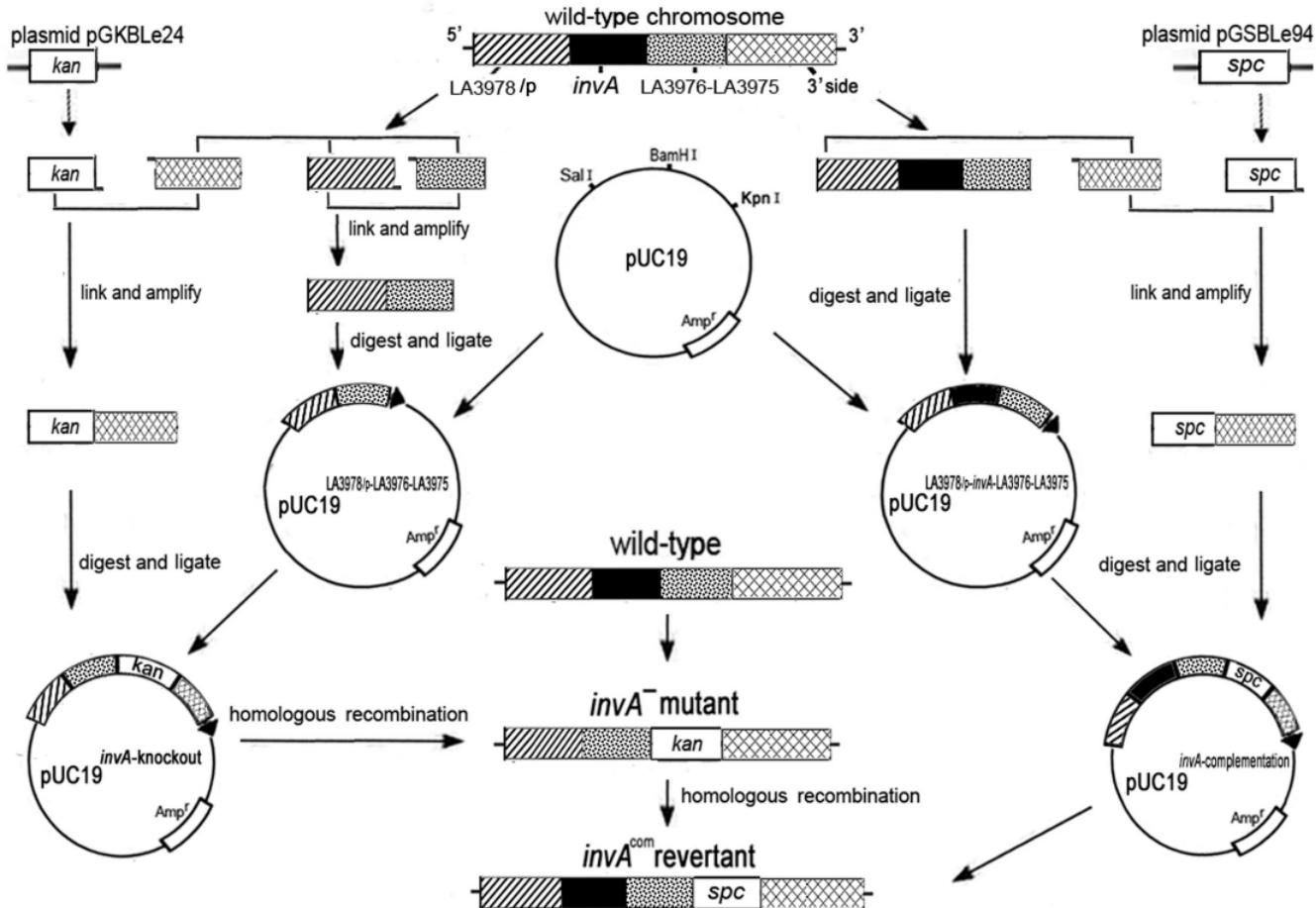
Strains	Serovars	Serogroups	Genospecies
Lai	Lai	Icterohaemorrhagiae	<i>L. interrogans</i>
Lin	Canicola	Canicola	<i>L. interrogans</i>
Tian	Pyrogenes	Pyrogenes	<i>L. interrogans</i>
Lin 4	Autumnalis	Autumnalis	<i>L. interrogans</i>
65-9	Australis	Australis	<i>L. interrogans</i>
Luo	Pomona	Pomona	<i>L. interrogans</i>
Lin 6	Grippotyphosa	Grippotyphosa	<i>L. interrogans</i>
P 7	Hebdomadis	Hebdomadis	<i>L. interrogans</i>
L 37	Paidjan	Bataviae	<i>L. interrogans</i>
L 183	Wolffi	Sejroe	<i>L. interrogans</i>
M 10	Javanica	Javanica	<i>L. borgpetersenii</i>
Pishu	Ballum	Ballum	<i>L. borgpetersenii</i>
55-52	Tarassovi	Tarassovi	<i>L. borgpetersenii</i>
Nan 10	Mini	Mini	<i>L. borgpetersenii</i>
L 105	Manhao 2	Manhao	<i>L. weilii</i>
Patoc I	Patoc	Semaraga	<i>L. biflexa</i>
CH11	Andamana	Andamana	<i>L. biflexa</i>

Table S2. Primers used in this study

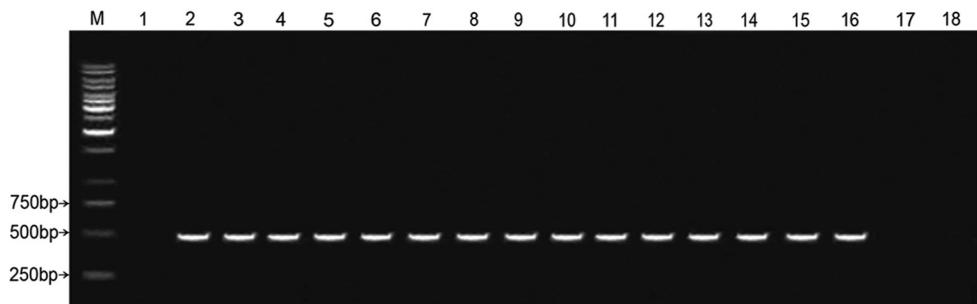
Primer	Sequence (5' to 3')	Target	Size (bp)
N1	F: CGCCATATG (Nde I) GACAAACCTACCGA R: CGCCTCGAG (Xho I) CGATCTATTCCGATGTC	<i>invA</i> gene detection and cloning	486
N2	F: AAGATTCACTTAAAATTAAA R: ATCATGAGAAATACGAAAACC	Identification of <i>invA</i> ⁻ mutant and <i>invA</i> ^{com} revertant	4080 4741
N3	F: ATG GAC AAA CCC TAC CGA R: TGG AAA TTG CCA AGA ACC	<i>invA</i> mRNA detection by qPCR	150
K	F: CGCGGATCC(BamH I)ATCGGCTCCGTCGATACTATG R: agaatcgatattatacgtaggCATCAGAGTATGGACAGTTGC	Amplification of <i>kan</i> cassette	1062
L1	F: CGCGTCGAC(Sal I) CTAATTCTTCGATTGGATC R: aggtggataaaaaaggacatACGAATAAGCTCCGTTAAC	5'-homologous arm for <i>invA</i> gene knock-out	977
L2	F: agttaacggagcttattcgATGTCTTTTATCCCACCT R: CGCGGATCC(BamH I)TCATTCTTAATTCAAAAGT	Amplification of LA3976-LA3975 genes	715
S	F:CGCGGATCC(BamH I)AACCGCTCCGAGCTTCAAGG R: agaatcgatattatacgtaggAACCGCTAACGACACTG	Amplification of <i>spc</i> cassette	1235
L3	F: gcaactgtccatactctgtatCCTACGTATAATCGATTCT R: CGCGGTACC(Kpn I)GAAATCATACGACCCCTGGAAG	3'-homologous arm for <i>invA</i> gene knock-out	1086
L4	F: cagggtcttaacttacgcgttCCTACGTATAATCGATTCT R: CGCGGTACC(Kpn I)GAAATCATACGACCCCTGGAAG	3'-homologous arm for <i>invA</i> gene reversion	1086
16S	F: CTTCCTGTGCCTCAGCGTCAGT R: CGCAGCCTGCACTTGAAC	16S rRNA for internal reference in qPCR	145
LA3976	F: CGGTCCACCTTGATTCTCAA R: AAACCTCCGCATCCCAATCGTC	Detection for mRNA level of LA3976 gene by qPCR	140
LA3975	F: GCTAAGCGGTTATACTCAA R: TCTTCCCAAGTAATCAGG	Detection for mRNA level of LA3975 gene by qPCR	147

F: forward primer; R: reverse primer. Letters in lowercase indicate linking sequences.

Supplemental Figure S1



Supplemental Figure S2

A**B Nucleotide sequence**

(1)	1	ATGGACAAACCTACC <u>GAAAAA</u> TGTCGGATGGTGTATTAACTCTCGTGAGAGGTT
(2)	1
(1)	61	TTGGTTGGAGAAAGATTGAATTTCTAGGTTCTGGCAATTCCACAAGGTGGAATTGAC
(2)	61
(1)	121	GACGGATGAAGATCCGATCAAGGCAGCCATGAGAGAATTATATGAAGAAGTCGGAATCGAT
(2)	121
(1)	181	TCTGGAAAATCGTAGCTGAATATCCAGATTGGATTTCCTATGACTTCCC <u>AAAACCTT</u>
(2)	181 T
(1)	241	CCTCTAACCGTCATCTTCAAAAATATAGGGGCAACTTCAA <u>AAAGTGGTTCTTATCTAT</u>
(2)	241
(1)	301	TGGGACGGGGAA <u>AGTGGATCAATGTGATTGGATATT</u> CATGAAAGAGAATTGGAACCGTT
(2)	301 A
(1)	361	CGTTTATTCCATA <u>AAAAAACACGTTGAATACAGTCGTTCTT</u> AAAAAAGATGTATAT
(2)	361
(1)	421	TATAAAATTGTAATGACTTGGCCTAAGAT <u>CCAAA</u> ACTTTTGCAAGACATCGGAAAT
(2)	421 A
(1)	481	<u>AGATCGTAA</u>
(2)	481

Amino acid sequence

(1)	1	MDKPYRKNVGMVFNSRGEVLVGERLNFLGSWQFPQGGIDDDEDPIKAAMRELYEEVGID
(2)	1
(1)	61	SGKIVAEYPDWISYDFPENLPLNRLQKYRGQLQKWFLIYWDGEVDQCDLDIHEREFGTV
(2)	61 F
(1)	121	RFPPIKNTLNTVVPFKKD VYY KIVND F GP K IQNFLQD I GNRS
(2)	121 E

Supplemental Figure S3

A

<i>L. interrogans</i>	: M-----KPYRENVGMMVWPNNSRG----PVLVGEPRINFLGS-----	WQFPC	: 36
<i>R. prowazekii</i>	: MRNSSNKYLDLPLYREGVGMILNADN-----GIVFGCPRDTKIS-----	SWQFPC	: 45
<i>B. bacilliformis</i>	: M-----TMVDFDKTLEYRKVGIVVNREG-----QVNIGERLITSSHTYAEVSKI-----	WQFPC	: 51
<i>H. pylori</i>	: M-----YRENVAPIIMSPDYPNACAEVVFABREDIEG-----	WQFPC	: 41
<i>C. jejuni</i>	: M-----ENKEK-----YRPNVAPLVLSSSYVPFECMFIARHRSMDN-----	WQFPC	: 42
<i>E. coli</i>	: M-----DDG-----YRPNVGVIVCNRKQ-----QVMWARRFGQHS-----	WQFPC	: 37
<i>P. aeruginosa</i>	: M-----SDG-----FRPNVGIILIANERAG-----QVLWARRINQEA-----	WQFPC	: 37
<i>N. meningitidis</i>	: M-----REG-----YRPNVGIILINNRN-----EVWCGSRVREHS-----	WQFPC	: 37
<i>H. influenzae</i>	: M-----FDG-----YRPNVGVIVCNRKG-----QVLWARRGQNS-----	WQFPC	: 37
<i>S. typhimurium</i>	: M-----DDG-----YRPNVGVIVCNRKQ-----QVMWARRFGQHS-----	WQFPC	: 37

***** motif (GX₅EX₇REUXEEEXGU, where U = I, L or V)

<i>L. interrogans</i>	:	GGIEIDDEDPIKRAMREIYEEVGI	-DSGKIVIAEYEDWISYDFFENI	PLNRLHILQKRYR	:	90							
<i>R. prowazekii</i>	:	GGIVPGEETPSIAPMREIYEEIGS	-NKGYIIAESRKWQSYSDVFSPE	PLKWLNGNMR	:	99							
<i>B. bacilliformis</i>	:	GGIEIDEGEETPLDARREIYEEITG	-RSVNLIAKEVQDWFCYDFDEQEL	IGHVLLNNQYR	:	105							
<i>H. pylori</i>	:	GGIEIDEGEETPLEALYREIYEEIGT	-NEIEIIAQYPFWIAYDFFPSNE	EHKFYSD-ED	:	93							
<i>C. jejuni</i>	:	GGIDKGESVKNAFLREIYEEIGT	-DEVELIAAEYPFWIYDFFPSKVI	VEKRMYP-MD	:	94							
<i>E. coli</i>	:	GGINPGESAEQAMYREIYEEVGL	LSRKDVRLIASTRNWIRYKIPK	KRLVRWDTKPVC	:	93							
<i>P. aeruginosa</i>	:	GGINIDRETPEEALYREIYEEVGL	EAAGDVRVIIACTRGWYRVLQF	PLVRTHSQPLCI	:	93							
<i>N. meningitidis</i>	:	GGIKFGESSETPAMREIYEEVGLL	FPLQHWRVIIGRTRWYF	DYDVEVNMMWRRERWGSYR	:	93							
<i>H. influenzae</i>	:	GGINDESASEQAMYREIYEEVGL	EQPKDVRVLLYWSKRWYKIPK	KRLVRWDTSKPCMI	:	93							
<i>S. typhimurium</i>	:	GGINPGESAEQAMYREIYEEVGL	LSRKDVRLIASTRNWIRYKIPK	EKRLVRWDTSKPVCI	:	93							
		GGI	E	A	SE6	PE	G	66	W	Y	P		

<i>L. interrogans</i>	: GQ	IQKWFELIYWYDGEVQCELDIHER---	EFGT	VRFPIPKNTIINIVVPEKKEVYVR	: 142
<i>R. prowazekii</i>	: GQ	KQRWFELIIRTFGNNNKIINHTESNP--	EFGC	QMRWESLDELDISIIPFKRFLYQA	: 151
<i>B. bacilliformis</i>	: GQ	MKQKWFELAFQSFQIGEETIVINSPENSNAK	EFGC	QMKWNLEVLPSLSIVSFKEKRVYMK	: 161
<i>H. pylori</i>	: GQ	IQKRYFLVRLKHANN-IDLNKHTP---	EFGA	QFIHLKDLIKKIIIPFKRQVYRQ	: 144
<i>C. jejuni</i>	: GQ	IQKRYFLVRLKHGAT-ININTKHP---	EFGD	QFVFSVKQIFEMINHFKRKNVYVR	: 145
<i>E. coli</i>	: GQ	OKQKWFELQEVSGDAEINMQTSSTP--	EFGC	WRWVSYWPFYVRQVSVSEKRVYRR	: 146
<i>P. aeruginosa</i>	: GQ	OKQKWFELIPLMSDEARVEMDTISKP--	EFGC	WRWVSYWPFYIGCQVCFKREVVYRR	: 146
<i>N. meningitidis</i>	: GQ	QCWYFLIIRITGRCDVNIRATRHP--	EFGC	WRWVSYWPFYDZVWIDFKRKEVYLG	: 146
<i>H. influenzae</i>	: GQ	KQRWFELIQLVSEDEKNINMQTTKSP--	EFGC	WRWVSYWPFYVRQVSVSEKRVYRR	: 146
<i>S. typhimurium</i>	: GQ	OKQKWFELIQLMSADAEINMQTSSTP--	EFGC	WRWVSYWPFYVRQVSVSEKRVYRR	: 146
		Q Q 5.51	6	EFG 5 6 EFK 4 6Y	

B

