OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

Pathogenomic Inference of Virulence-Associated Genes in *Leptospira interrogans*

Jason S. Lehmann[1], Derrick E. Fouts[2], Daniel H. Haft[2], Anthony P. Cannella[1], Jessica N. Ricaldi[1¤a], Lauren Brinkac[2], Derek Harkins[2], Scott Durkin[2], Ravi Sanka[2], Granger Sutton[2], Angelo Moreno[1¤b], Joseph M. Vinetz[1]*, Michael A. Matthias[1]*

1 Division of Infectious Diseases, Department of Medicine, University of California San Diego School of Medicine, La Jolla, California, United States of America, **2** J. Craig Venter Institute, Rockville, Maryland, United States of America

Abstract

Leptospirosis is a globally important, neglected zoonotic infection caused by spirochetes of the genus *Leptospira*. Since genetic transformation remains technically limited for pathogenic *Leptospira*, a systems biology pathogenomic approach was used to infer leptospiral virulence genes by whole genome comparison of culture-attenuated *Leptospira interrogans* serovar Lai with its virulent, isogenic parent. Among the 11 pathogen-specific protein-coding genes in which non-synonymous mutations were found, a putative soluble adenylate cyclase with host cell cAMP-elevating activity, and two members of a previously unstudied ~15 member paralogous gene family of unknown function were identified. This gene family was also uniquely found in the alpha-proteobacteria *Bartonella bacilliformis* and *Bartonella australis* that are geographically restricted to the Andes and Australia, respectively. How the pathogenic *Leptospira* and these two *Bartonella* species came to share this expanded gene family remains an evolutionary mystery. *In vivo* expression analyses demonstrated up-regulation of 10/11 *Leptospira* genes identified in the attenuation screen, and profound *in vivo*, tissue-specific up-regulation by members of the paralogous gene family, suggesting a direct role in virulence and host-pathogen interactions. The pathogenomic experimental design here is generalizable as a functional systems biology approach to studying bacterial pathogenesis and virulence and should encourage similar experimental studies of other pathogens.

Citation: Lehmann JS, Fouts DE, Haft DH, Cannella AP, Ricaldi JN, et al. (2013) Pathogenomic Inference of Virulence-Associated Genes in *Leptospira interrogans*. PLoS Negl Trop Dis 7(10): e2468. doi:10.1371/journal.pntd.0002468

Editor: Pamela L. C. Small, University of Tennessee, United States of America

Received April 5, 2013; **Accepted** August 23, 2013; **Published** October 3, 2013

Funding: This work was supported by NIH grants T32 GM008666 (NIH predoctoral training grant), R25GM083275 (“Mentoring Young Minds to Increase Diversity in the Biomedical Research”), R21AI064466 (“Microarray Analysis of Leptospiral Genomes”), RO1TW05860 (“Leptospirosis Transmission in the Peruvian Amazon”), and 1D43TW007120 (Fogarty Global Infectious Diseases Training Grant, “Endemic Infectious Diseases of the Peruvian Amazon”). This project was also funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under contract number HHSN272200900007C. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jvinetz@ucsd.edu (JMV); mmatthias@ucsd.edu (MAM)

¤a Current address: Instituto de Medicina Tropical “Alexander von Humboldt”, Universidad Peruana Cayetano Heredia, Lima, Peru.
¤b Current address: Department of Molecular Genetics and Microbiology, Duke University, Durham, North Carolina, United States of America.

Introduction

Leptospirosis, caused by spirochete bacteria of the genus *Leptospira*, is a zoonotic disease of high public health impact [1]. Globally, nearly 900,000 people are infected annually through contact with contaminated water, infected tissue or urine of mammalian reservoir hosts [2]. Phylogenetic analyses have resolved the genus into 3 distinct lineages, which are the focus of a pan-*Leptospira* genome project supported by the NIAID Genome Sequencing Center: nine pathogenic species; five intermediate species (eg. *L. fainei, L. licerasiae*); and six non-infectious saprophytic species (i.e. *L. biflexa*) (Fig. 1A) [3–6]. The greatest burden of disease is caused by the pathogenic species, mainly affecting people living in poverty and with poor sanitation [1,2,7]. Epidemics of leptospirosis associated with floods, monsoons, or hurricanes have a high morbidity and mortality with case fatality rates ranging as high as 20–25% in hospitalized patients leading to refractory shock, jaundice, renal failure, and pulmonary hemorrhage [1].

Despite its severity and global importance, the molecular pathogenesis of leptospirosis remains poorly understood [8]. *Leptospira* penetrate mucosal epithelium and damaged integument then hematogenously disseminate to localize within multiple organs, including the liver and kidney, within 72 hours. Leptospiremia may continue for up to two weeks after onset of symptoms with blood bacterial concentrations reaching as high as 10^6–10^7 organisms/mL in infected patients [9,10]. The only virulence factor genetically defined to date is the surface lipoprotein Loa22 [11], but mechanisms by which it contributes to disease pathogenesis remain unknown. Other virulence-associated genes include heme oxygenase [12], LPS [13], clpB [14], and flagellar components [15,16]. Although random transposon mutagenesis has been used to identify a few putative leptospiral virulence-related genes [17] [18], further progress has been hindered by the lack of efficient gene-targeted mutagenesis techniques in pathogenic *Leptospira* [8].

We used a functional systems biology (pathogenomic) approach to identify candidate virulence genes, by genomic comparison of a

Author Summary

Leptospirosis is one of the most common diseases transmitted by animals worldwide. It is important because it causes an often lethal febrile illnesses in tropical and subtropical areas associated with poor sanitation and agriculture. Leptospirosis may be epidemic, associated with natural disasters and flooding, or endemic in tropical regions. It is unknown how *Leptospira* cause disease and why different strains cause different severity of illness. In this study we attenuated (weakened) a highly virulent strain of *L. interrogans* by culturing it *in vitro* over several months. Comparison of the whole genome sequence before and after the attenuation process revealed a small set of genes that were mutated, and therefore associated with virulence. We discovered a putative soluble adenylate cyclase with host cell cAMP elevating activity, with implications for immune evasion and a new gene family that is upregulated *in vivo* during acute hamster infection. Interestingly, both *Bartonella bacilliformis* and *Bartonella australis* also have this unique gene family we describe in pathogenic *Leptospira*. This information aids in our understanding of *Leptospira* evolution and pathogenesis.

culture-attenuated *Leptospira interrogans* serovar Lai strain 56601 ($LD_{50}>10^9$)(Fig. 1B) with its virulent, isogenic parent ($LD_{50}<100$) [19]. *In vivo* relevance of identified candidate genes was determined by quantification of expression of candidate genes on day 4 after hamster infection in blood, liver, and kidney compared to *in vitro* culture.

Materials and Methods

Ethics statement

This study was carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health in AAALAC-approved facilities. The experimental animal work was approved by the Institutional Animal Care and Use Committee of the University of California San Diego under protocol S03128H.

Bacterial strain maintenance and attenuation of *L. interrogans* serovar Lai strain 56601

All strains were maintained *in vitro* in Ellinghausen-McCullough-Johnson -Harris (EMJH) media using standard protocols and are available from BEI Resources. *L. interrogans* serovar Lai strain 56601 was obtained from Dr. David Haake (UCLA). Virulence was selected for by serial passage through hamsters so that the P1 strain used in the present study had an LD_{50} of ~10 organisms. *L. interrogans* serovar Lai strain 56601_P1 was attenuated by 18 bi-weekly subcultures *in vitro*. Virulence was assessed every five to ten subcultures using three-week-old male Golden Syrian Hamsters. Following a final subculture, genomic DNA was prepared from this attenuated strain (designated P19) on which next generation sequencing was carried out.

Genome assembly of virulent P1 and attenuated P19 *L. interrogans* serovar Lai strain 56601 and non-synonymous SNV (nsSNV) detection

We generated 4,379,515 and 5,340,095 unpaired shotgun reads from *L. interrogans* serovar Lai 56601_P1 and 56601_P19, respectively using next generation sequencing technology. All reads were 36 bases long. Both genomes were assembled using the comparative assembler AMOScmp. The AMOSCmp-shortReads-alignmentTrimmed pipeline that runs within AMOScmp, was used to look for exact matches of each read to the published *L. interrogans* serovar Lai 56601 genome of at least 20 bp, permitting a maximum consensus error rate of 0.06% (i.e. at most two mismatches in any read). This script runs a reference-based trimming of the 3′-end of the reads prior to assembly. We found that trimming of at most 4 bases from the 3′-end of the reads based on their matches to the reference produced better assemblies than un-trimmed reads. The P1 assembly used 3,919,609 reads, leaving 459,906 unassembled singletons, while the P19 assembly used 4,915,295 leaving 424,800 singleton reads. The 56601_P1 genome was assembled into 167 contigs with an average length of 28,124 kb and an N50 length of 105,604 kb and the P19 genome into 97 contigs, average length 48,417 and N50 of 190,406. We checked the quality of both assemblies using the amosvalidate pipeline, which runs within AMOScmp. This pipeline identifies misassembly features such as increased read depth and correlated SNVs (i.e. one or more reads with the same SNV, which is unlikely to be due to sequencing error), both indicative of collapsed repeats. We found that both assemblies were high quality with at most 5 potential misassembly features in longer contigs. These potential misassemblies were inspected manually using the Hawkeye viewer and reassembled if necessary using minimus, which employs a stricter assembly algorithm. The unfinished 56601_P1a and 56601_P19 genomes were aligned and SNVs identified using the MUMmer v3.22 software package.

RT-qPCR in vivo gene expression analysis

Three wk old Golden Syrian Hamsters were infected via intraperitoneal injection with 10^8 low passage *L. interrogans* serovar Lai strain 56601. 96 hours post infection total RNA was collected using TRIzol (Invitrogen) from blood, liver, and kidney tissue, as well as from a 96-hour EMJH culture of *L. interrogans* grown at 30°C. Total RNA was reverse transcribed using a QuantiTect reverse transcription kit (Qiagen). cDNA was amplified using a CFX96 thermal cycler (Bio-Rad) using PerfeCta SYBR Green FastMix (Quanta Biosciences). PCR was carried out at 95°C for 3 min, a touchdown gradient of 14 cycles of (94°C 10 s, 80°C 45 s) decreasing 1°C/cycle, followed by 40 cycles of (94°C 30 s, 65°C 45 s). Ct values were normalized to the leptospiral 16S rRNA gene and expression fold change calculated using the Pfaffl method [20]. Primer sequences are listed in Table S3 in Text S1.

Domain architecture analysis of LA_4008 and other related AGC proteins

Domain architecture comparison of LA_4008 with orthologs of *Myxococcus xanthus*, *Corallococcus coralloides*, *Stigmatella aurantiaca*, and *Mycobacterium tuberculosis* using NCBI CD Search, SMART, and TPRPred. Protein homology analysis was carried out using BLAST using the following reference sequences: LA_4008 (NP_714188.1), MXAN_4545 (YP_632713.1), COCOR_04748 (YP_005370712.1), STAUR_4866 (YP_003954471.1), Rv0386 (CCP43116). The coverage for the query sequence, statistical significance (E-value), and maximum amino acid identify (“Max Ident”) are indicated at right for each predicted primary sequence. Identified domains were then graphically represented using the DOG 1.0 program (http://dog.biocuckoo.org)

Leptospira Concentrated Culture Supernatant (CCS)

L. interrogans Lai 56601 or *L. licerasiae* Varillal were grown in EMJH media +10% heat inactivated rabbit serum at 37°C on a rotating shaker for 96 hr. Culture was centrifuged for 30 min at

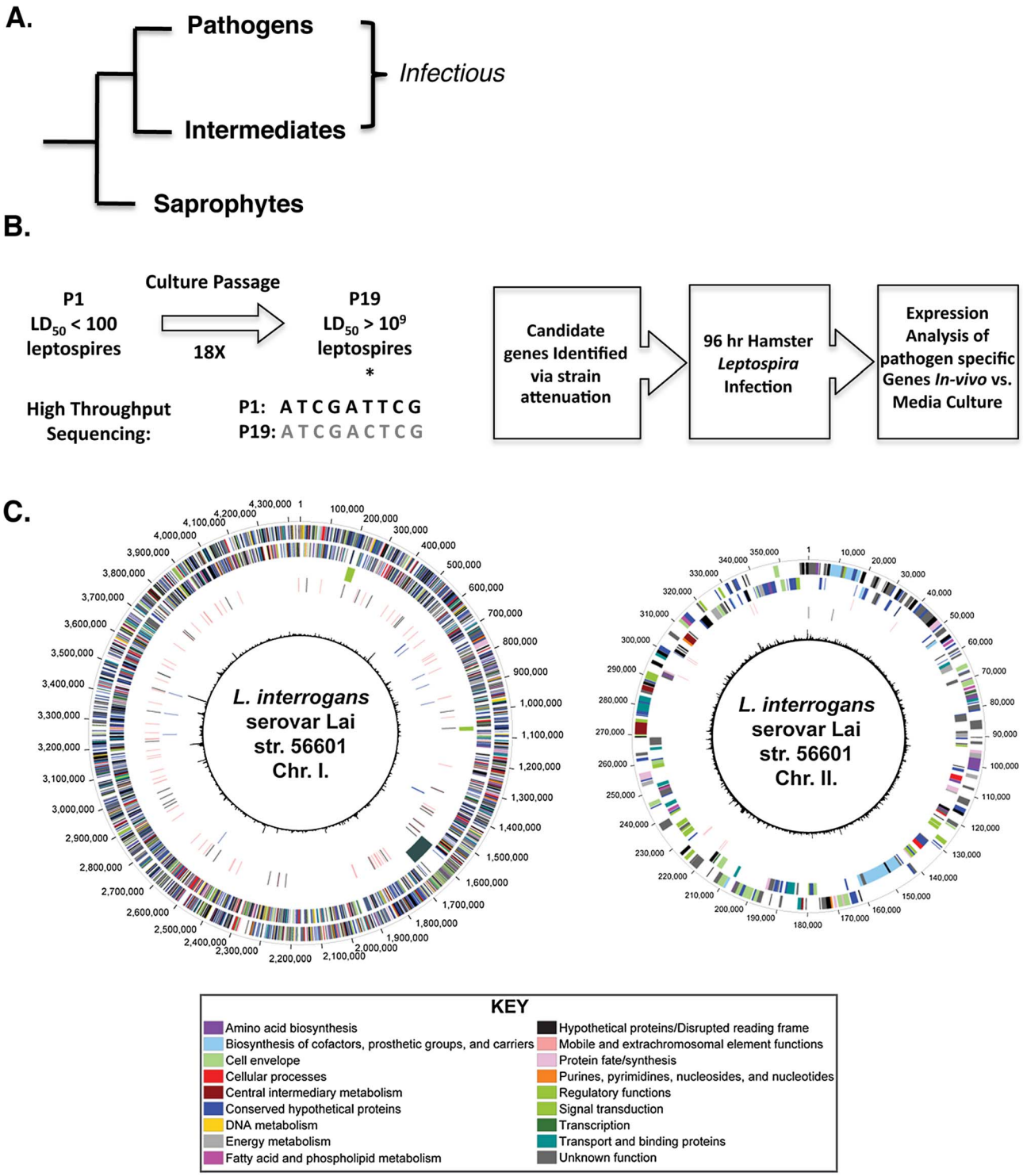


Figure 1. Pathogenomic analysis of *Leptospira interrogans* serovar Lai strain 556021 to identify virulence related genes. (**A**) Schematic of phylogenetic relatedness of "Pathogenic" (P), "Intermediate" (I) and "Saprophytic" (S) members of the genus *Leptospira*. (**B**) Workflow to identify putative virulence-associated genes. Asterisk denotes a hypothetical position in which a SNV has been identified (**C**) Genomic Locations of SNPs and PF07598 paralogs in the reference genome of *L. interrogans* serovar Lai strain 56601. Each concentric circle represents genomic data and is numbered from the outermost to the innermost circle. The outermost circles represent the predicted CDS on the + and − strands, respectively, colored by functional role categories (see key). The following circle descriptions apply to chromosome I. The third circle notes the location of predicted prophage regions (olive) and the LPS region (slate). The fourth circle indicates those CDS found to have non-synonymous amino acid substitutions (black) as well as the location of CDS annotated as "transposase" in Genbank (salmon). The fifth circle represents the location of the 12 PF07598 family members (blue). The innermost circle denotes atypical regions (χ^2 value). For chromosome II, the outermost and innermost circles are the same as for chromosome I; however, the third circle notes the location of transposases (salmon), while the fourth circle indicates the location of the CDS found to have non-synonymous amino acid substitutions (black).
doi:10.1371/journal.pntd.0002468.g001

10,000× g. Supernatant was decanted, filtered through a .22 µm syringe filter unit (Millipore), and concentrated 10× in an Amicon Ultra 10K MWCO centrifugal filter unit (Millipore). No leptospires were observed in the CCS after concentration using darkfield microscopy.

CCS cAMP elevating activity

CCS was incubated with monolayers of THP-1,a human monocyte/macrophage cell line. At 4, 6, and 20 hours cells were rinsed 3× in PBS and analyzed for cAMP (Direct cAMP EIA kit, Enzo Life Sciences). Secondly, CCS from *L. interrogans* Lai and *L. licerasiae* were incubated with THP-1 monolayers for 4 hours, and assayed for cAMP.

CCS immunodepletion studies

Rabbits were used to generate anti-peptide antisera against LA_4008 using a protein specific, sixteen amino acid fragment (SVEEDPLTREIDRKQK) conjugated to keyhole limpet hemocyanin as a carrier protein (Pacific Immunology, Ramona, CA). The IgG fractions from pre-immunization and production bleeds were purified using a Melon Gel IgG Purification kit (Thermo Scientific) and covalently linked to magnetic beads using a NanoLink BeadLink Kit (Solulink). Antibody linked beads were incubated with CCS overnight at 4°C on a rotating shaker. Beads were separated on a QuadroMACS separation unit (Miltenyi Biotec). Depleted CCS was applied to THP-1 monolayers and incubated for 4 hours. Cells were rinsed 3× in PBS and analyzed for total cAMP using the Direct cAMP EIA kit.

Phylogenetic analysis of PF07598 paralogous protein family

L. interrogans, *L. borgpetersenii* and *B. bacilliformis* full-length sequences were downloaded from the Uniprot databa se (http://www.uniprot.com) and aligned using MAFFT v7 (http://mafft.cbrc.jp/alignment/software) with default parameters. The evolutionary history was inferred by using the Maximum Likelihood method based on the Whelan and Goldman frequency model [21]. Statistical support of the tree topology was obtained from 500 bootstrap replicates. A discrete Gamma distribution was used to model evolutionary rate differences among sites. The rate variation model allowed for some sites to be evolutionarily invariable. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions containing gaps and missing data were eliminated. There were a total of 271 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 [22].

Pan-genomic analysis of attenuated genes and PF07598 orthologs

The genomic data analyzed here are publically available and are from newly generated, unpublished *Leptospira* whole genome sequence data produced by the JCVI as part of the white paper "*Leptospira* Genomics and Human Health," sponsored by the NIAID-funded Genome Sequencing Centers. PanOCT [23] was run using default settings with the exception that a dynamically determined pairwise cutoff was implemented, not available in the current release, but available upon request. The following genomes, representing all 20 *Leptospira* spp. were used: *L. alexanderi* sv. Manha 3 str. L 60^T (Genbank:AHMT00000000), *L. alstoni* sv. Pingchang str. 80-412 (Genbank:AOHD00000000), *L. biflexa* sv. Patoc str. Patoc I Paris (Genbank:CP000786), *L. borgpetersenii* sv. Javanica str. UI 09931 (Genbank:AHNP00000000), *L. broomii* sv. undetermined str. 5399^T (Genbank:AHMO00000000), *L. fainei* sv. Hurstbridge str. BUT 6^T (Genbank:AKWZ00000000), *L. inadai* sv. Lyme str. 10^T (Genbank:AHMM00000000), *L. interrogans* sv. Copenhageni str. Fiocruz L1-130 (Genbank:AE016823), *L. interrogans* sv. Copenhageni str. M20 (Genbank:AOGV00000000), *L. interrogans* sv. Lai str. 56601 (Genbank:AE010300), *L. kirschneri* sv. Cynopteri str. 3522 C^T (Genbank:AHMN00000000), *L. kmetyi* sv. undetermined str. Bejo-Iso9^T (Genbank:AHMP00000000), *L. licerasiae* sv. Varillal str. VAR 010^T (Genbank:AHOO00000000), *L. meyeri* sv. Hardjo str. Went 5 (Genbank:AKXE00000000), *L. noguchii* sv. Panama str. CZ 214^T (Genbank: AKWY00000000), *L. santarosai* sv. Shermani str. $1342K^T$ (AOHB00000000), *L. terpstrae* sv. Hualin str. LT $11\text{-}33^T$ (Genbank: AOGW00000000), *L. vanthielii* sv. Holland str. WaZ Holland (Genbank:AOGY 00000000), *L. weilii* sv. Ranarum str. ICF^T (Genbank: AOHC00000000), *L. wolbachii* sv. Codice str. CDC (Genbank: AOGZ00000000), *L. wolffii* sv. undetermined str. Khorat-$H2^T$ (Genbank:AKWX00000000), *L. yanagawae* sv. Saopaulo str. Sao $Paulo^T$ (Genbank:AOGX00000000).

Statistics

Data were analyzed using GraphPad Prism 5.0. Significance was assessed using one-way ANOVA followed by Tukey's HSD post hoc testing. P-values are reported as *** = $p<0.001$, ** = $p<0.01$, * = $p<0.05$.

Results

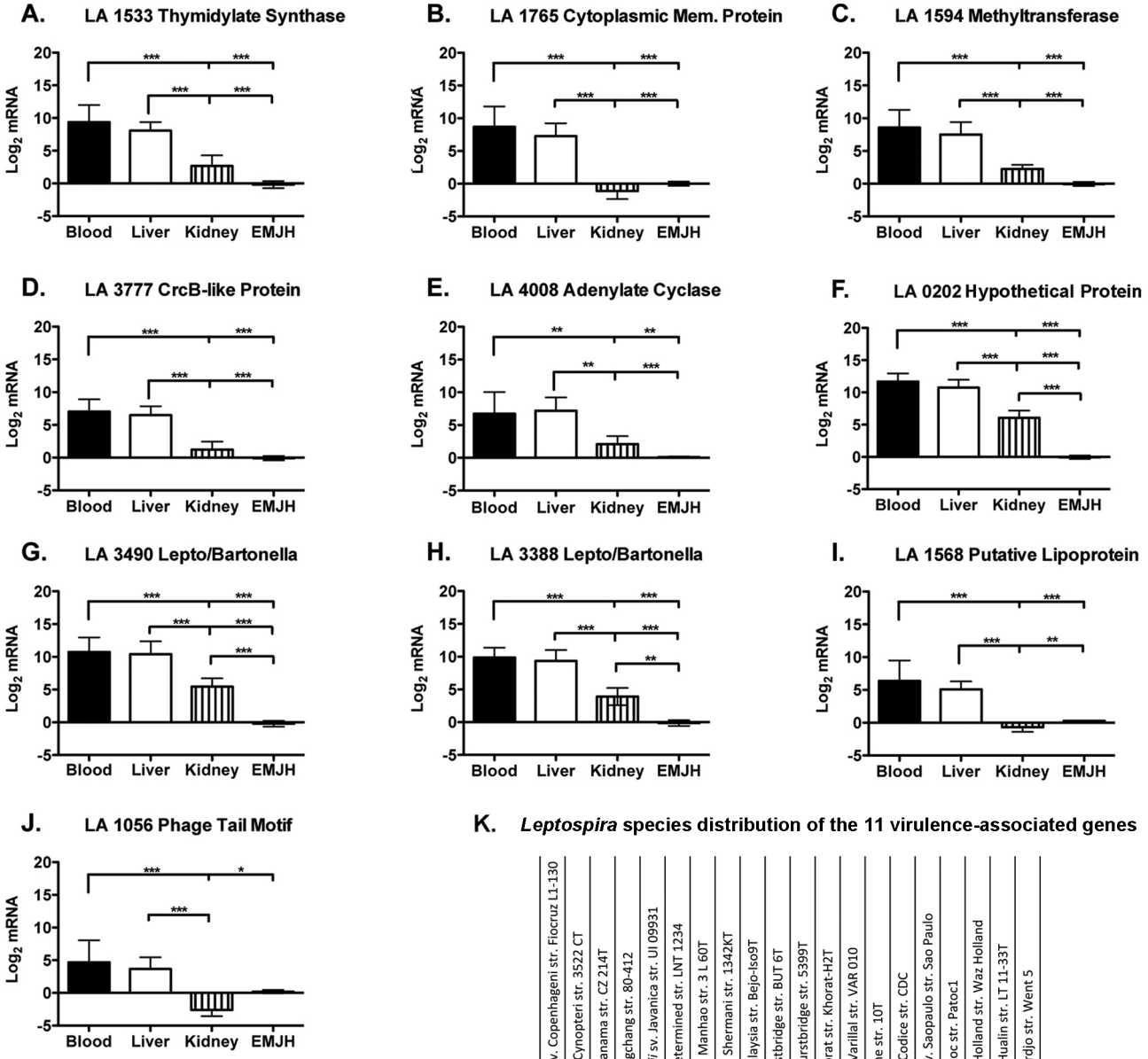
Pathogenomic identification of protein coding genes in *Leptospira interrogans* serovar Lai and patterns of tissue-specific up-regulation in vivo

Comparison of the wild type and attenuated *L. interrogans* Lai 55601 genomes identified 41 non-synonymous single nucleotide variants (nsSNVs) in a total of 35 protein-coding genes (CDS; Table S1 in in Text S1). P19 sequence analysis revealed that all SNVs were homogeneous within the culture population; minority populations were not detected at the limit of detection of the Illumina sequencing platform (<4%). For the purposes of this study, therefore, the bacterial populations were considered clonal.

Filtering to include CDS restricted to pathogenic *Leptospira* species identified 11 genes (Fig. 2K). These CDS are highly conserved among pathogenic *Leptospira* species (Fig. 2K). *In vivo* transcriptional analysis identified that of these 11 pathogen-specific genes, 10 were up-regulated *in vivo* during acute hamster infection (Fig. 2, normalized to the 16S rDNA gene, Fig. S1 in Text S1). Transcriptional up-regulation of CDS was as high as several thousand-fold, with a much higher dynamic range than found with *in vitro* conditions used in previously reported systems biology analyses (summarized in Table S2 in Text S1).

Identification of a putative leptospiral protein with host cAMP elevating activity

Of particular interest is LA_4008, a putative adenylate/guanylate cyclase (AGC) that lacks transmembrane helices typical of integral membrane cyclases involved in signal transduction, suggesting that this protein may be soluble. While another adenylate/guanylate cyclase was found in our screen in pathogens and intermediates (Table S1 in Text S1; LA_0027), this protein is predicted to be a housekeeping gene, a membrane- bound and intracellular, and not likely to be found in the extracellular milieu. Orthologs of LA_4008 are found only in pathogenic *Leptospira* and the intermediately pathogenic strain *L. fainei*, Fig. 2K. Other bacterial adenylate cyclases lacking transmembrane domains include the soluble cyclase class of toxins of the pathogens



K. ***Leptospira*** **species distribution of the 11 virulence-associated genes**

Protein	Accession	Function	P19 Mutation	*L. interrogans* sv. Copenhageni str. Fiocruz L1-130	*L. kirschneri* sv. Cynopteri str. 3522 CT	*L. noguchii* sv. Panama str. CZ 214T	*L. alstoni* sv. Pingchang str. 80-412	*L. borgpetersenii* sv. Javanica str. UI 09931	*L. weilii* sv. undetermined str. LNT 1234	*L. alexanderi* sv. Manhao str. 3 L 60T	*L. santarosai* sv. Shermani str. 1342KT	*L. kmetyi* sv. Malaysia str. Bejo-Iso9T	*L. fainei* sv. Hurstbridge str. BUT 6T	*L. broomii* sv. Hurstbridge str. 5399T	*L. wolffii* sv. Khorat str. Khorat-H2T	*L. licerasiae* sv. Varillal str. VAR 010	*L. inadai* sv. Lyme str. 10T	*L. wolbachii* sv. Codice str. CDC	*L. yanagawae* sv. Saopaulo str. Sao Paulo	*L. biflexa* sv. Patoc str. Patoc1	*L. vanthielii* sv. Holland str. Waz Holland	*L. terpstrae* sv. Hualin str. LT 11-33T	*L. meyeri* sv. Hardjo str. Went 5
				P	P	P	P	P	P	P	P	P	P	I	I	I	I	I	S	S	S	S	S	S
LA_4008	NP_714188.1	adenylate/guanylate cyclase	L46F	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-
LA_3777	NP_713957.2	CrcB-like protein	C53G	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
LA_0202	NP_710383.2	hypothetical protein	I167M	+	+	+	+	-	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-
LA_0979	NP_711160.1	hypothetical protein	Y222S	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_1056	NP_711237.2	hypothetical protein	G401A	+	+	+	+	-	-	-	-	+	+	+	-	-	-	+	-	-	-	-	-	-
LA_1533	NP_711714.2	thymidylate synthase	FRAMESHIFT	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-
LA_1568	NP_711749.1	beta-propeller repeat protein	N16T	+	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
LA_1594	NP_711775.1	methyltransferase	F382L	+	+	+	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_1765	NP_711946.2	cytoplasmic membrane proein	K2039Q; V2040F	+	+	+	+	-	+	+	+	+	-	+	+	+	+	-	-	-	-	-	-	-
LA_3388	NP_713568.1	hypothetical protein	G72E	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_3490	NP_713670.1	hypothetical protein	Y434D	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 2. ***In vivo*** **transcriptional analysis of putative virulence-associated genes.** *In vivo* relevance of the identified virulence-related genes, mRNA transcript levels of the genes identified by the pathogenomics approach was assessed by real time, reverse transcriptase quantitative PCR of blood, liver and kidney 4 d after hamster infection, compared to log phase *in vitro* cultured *Leptospira*. Leptospiral gene expression levels in infected tissue vs. EMJH were expressed logarithmically as the $\log_2$ of the fold change between the two conditions (**A–J**). 16S rRNA transcript levels (previously validated [61]) were used to normalize gene expression in tissues and under the different conditions (Fig. S1 in Text S1). Expression of 10/11 identified genes was detectable *in vivo* in all three tissues assayed; the exception was the hypothetical protein LA_0979. The remaining 10 genes were detected in all three tissues assayed. Expression varied between groups of animals, and interestingly, the highest levels of up-regulation were found in leptospires isolated from the blood of infected animals, with transcript levels also being up in bacteria from the liver. Virulence-associated genes were variably up-regulated in kidney. The data represented are the mean ± SEM of 3 independent experiments (n = 7 animals). (**K**) *Leptospira*

species distribution of the 11 virulence-associated genes identified and associated single nucleotide variants found in coding sequences of the avirulent passage (P19) strain. Protein code is according to the annotated protein database; Accession is the GenBank code for the protein.
doi:10.1371/journal.pntd.0002468.g002

Mycobacterium tuberculosis, Bordetella pertussis, Bacillus anthracis, Yersinia pestis, and Pseudomonas aeruginosa which modulate host cellular responses to infection [24]. Sequence analysis by SMART, TPRPred, and NCBI conserved domain (CD) search revealed a unique domain architecture for LA_4008 consisting of two tandem N-terminal class III cyclase homology domains followed immediately by an AAA-ATPase domain, and finally a series of C-terminal tetratricopeptide (TPR) domains (Fig. 3), that are known to mediate protein-protein interactions and have recently been recognized as components of bacterial virulence mechanisms [25,26]. LA_4008 also shares striking homology to a toxin (NCBI protein cluster PCLA_814229) shared by predatory species of the δ-proteobacterial order Myxococcales (Fig. 3). The domain structure shared by this protein cluster is reminiscent of another pathogenesis-related adenylate cyclase, Rv0386 of *Mycobacterium tuberculosis*. This domain structure, with precedent in both pathogenic and environmental bacteria, has been experimentally shown to increase cyclic AMP levels in host macrophages and impair the innate immune response to infection (Fig. 3), [27,28]. To test whether LA_4008 has the potential to elevate cyclic AMP in host cells, concentrated *L. interrogans* serovar Lai strain 56601 EMJH culture supernatant (CCS) was added to *in vitro* monolayer cultures of macrophage-like THP-1 cells and the cells were washed and lysed at various times over 20 hr, and intracellular cyclic AMP levels were assayed. All CCS preparations were microscopically confirmed to be absent of visible leptospires prior to all experiments. We observed that a *L. interrogans*-derived soluble factor from culture supernatant stimulated a transient rise in intracellular macrophage cAMP levels, peaking at four hours (Fig. 4A). Next, the cAMP elevating activity of CCS was compared between *L. interrogans* serovar Lai (which has LA_4008), and the intermediate *L. licerasiae* serovar Varillal (Fig. 4B) (which does not have an ortholog of LA_4008). The results of this experiment was consistent with the hypothesis that cAMP elevating activity may be related to pathogenic *Leptospira* species containing the LA_4008 AGC but not by intermediate *Leptospira* (Fig. 1A), and therefore is not a general feature of all infectious leptospires. To further test if LA_4008 is responsible for the elevated target cell cAMP, CCS was digested with proteinase K prior to addition to THP-1 cells (important because some bacterial LPS can also elevate cAMP levels in host cells [29]) and, more critically, immune-depleted with a specific anti-LA_4008 antibody before adding the CCS to THP-1 cells. As a control, CCS was also immune-depleted with pre-immune serum of the host animal in which the anti-LA_4008 antiserum was generated. Both protease treatment and specific immunodepletion, but not non-specific depletion blocked CCS-mediated increases in intracellular cAMP levels in THP-1 cells (Fig. 4C), consistent with the hypothesis that LA_4008 from *L. interrogans* Lai is a cAMP-elevating factor in host cells.

Identification of a paralogous gene family shared by pathogenic *Leptospira*, *Bartonella bacilliformis*, and *Bartonella australis* with profound, tissue-specific up-regulation in vivo in an acute leptospirosis infection model in hamsters

During our analysis of attenuation mutations we identified two members (LA_3490, LA_3388) of a newly discovered paralogous gene family that is shared between pathogenic *Leptospira* but conspicuously absent in the intermediate and saprophytic species. All full-length members of this family (PF07598/DUF1561) are predicted to have secretory signal peptides, although degenerate forms do occur. Past the signal peptide, Cys residues are invariant at twelve positions, and occur nowhere else, suggesting a conserved pattern of disulfide bond formation and implying extracellular function (Fig. S2 in Text S1). In a given genome, the most closely related paralogs are often tandem. Otherwise, gene neighborhood analysis provided no clue to protein function. Paralog counts in pathogenic *Leptospira* range from two in the leptospire *L. santarosai* to 12 in *L. kirschneri* serovar Cynopteri and *L. interrogans* (Fig. 5A). Interestingly the PF07598 gene family has also been recently described in the unrelated α-proteobacteria species *Bartonella bacilliformis* and *Bartonella australis*. *B. bacilliformis* has 15 paralogs in its genome with *B. australis* having nearly the same (Fig. 5B) [30]. In addition single gene copies were found in three animal-infecting ε-proteobacteria, *Helicobacter hepaticus*, *H. mustelae*, and *H. cetorum*.

There are great phylogenetic distances separating the genera that contain this gene family, but paralogs are restricted to select animal-infecting species within each lineage; suggesting that these proteins may be uniquely related to host adaptation. All 12 members of the leptospiral PF07598 gene family were analyzed for *in vivo* expression in hamsters acutely infected with virulent, wild type *L. interrogans* Lai 55601. All members of this gene family were up-regulated in blood and liver to varying degrees, with LA_3490 and LA_3388, both containing secretory signal peptide sequences, being most highly up-regulated (more than ~1000-fold); all members of this gene family were up-regulated in the circulation and liver to varying degrees. In contrast, up-regulation of other members this gene family significantly varied among experimental animals in kidney (Fig. 5C–E).

Other pathogenomically-identified putative virulence genes in *Leptospira* spp.

Other pathogenomically-associated virulence genes include the following:

LA_1056: This gene has two predicted transmembrane helices and shares a conserved PHA00965 domain with tail proteins found in Gram-positive bacteriophages. This protein shows similarity to phage tape measure proteins after repeated rounds PSI-BLAST. Recent studies involving the phage-encoded pblA in *Streptococcus mitis* have identified a sequence weakly reminiscent of a tape measure motif protein by PSI-BLAST as an adhesin-type molecule used for bacterial attachment to platelets [31–34].

LA_1765: This protein has similarity to spvB, a protein from a group of plasmid-encoded virulence genes that mediate lethal infection in nontyphoid *Salmonella* strains [35].

LA_1533: a flavin-dependent thymidylate synthase. This unusual and newly described class of enzyme is expressed by many clinically relevant pathogens, including *Bacillus anthracis*, *Borellia burgdorferi*, *Campylobacter jejuni*, *Clostridium difficile*, *Helicobacter pylori*, *Mycobacterium tuberculosis*, and *Treponema pallidum* during infection as part of an alternative thymidine synthesis pathway [36–38].

LA_0202: a gene of unknown function previously reported to be transcriptionally up-regulated in virulent *L. interrogans* Lai 55601 when compared to another avirulent strain [39].

LA_1568: a putative lipoprotein with β-propeller repeats that has not been previously studied. Lipoproteins are important mediators of spirochete virulence, with the *L. interrogans* genome

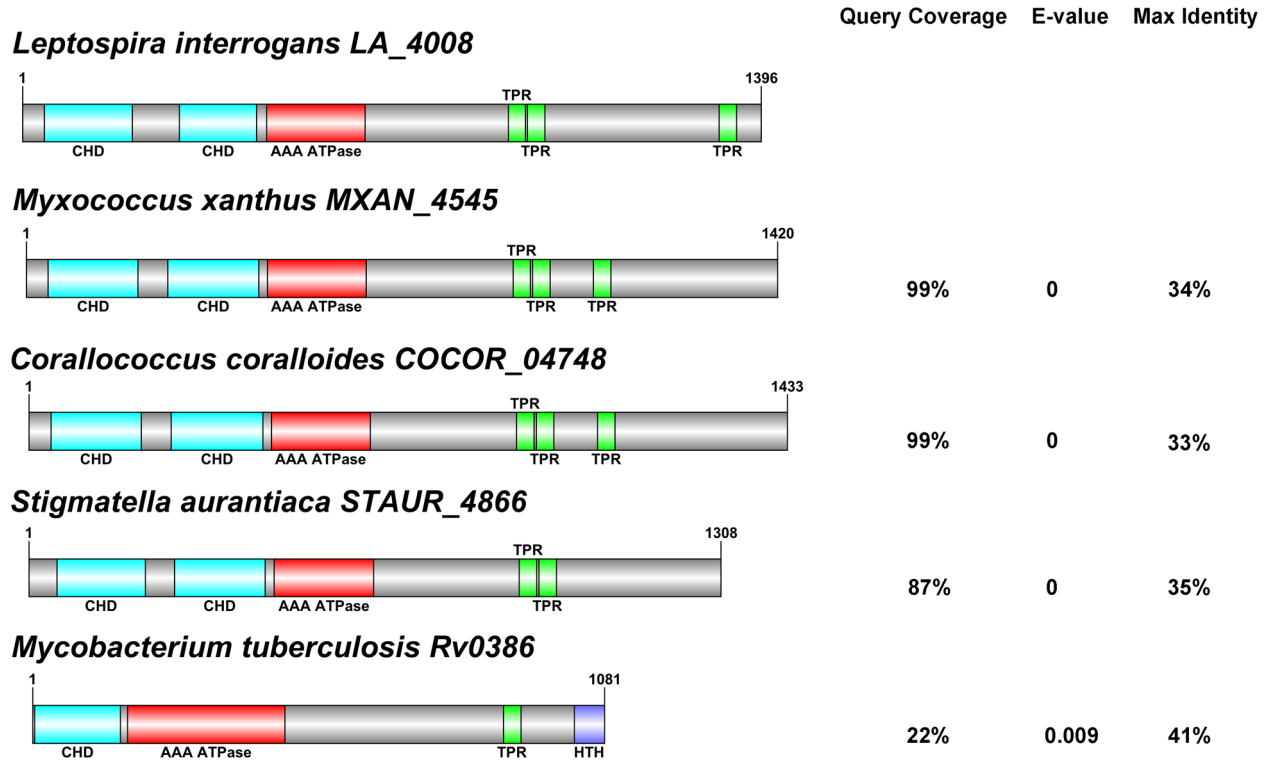


Figure 3. Ortholog sequence analysis of pathogenic *Leptospira* adenylate/guanylate cyclase compared to predatory environmental bacteria and the pathogen, *Mycobacterium tuberculosis*. Domain architecture comparison of LA_4008 with orthologs of *Myxococcus xanthus*, *Corallococcus coralloides*, *Stigmatella aurantiaca*, and *Mycobacterium tuberculosis* using NCBI CD Search, SMART, and TPRPred. Protein homology analysis was carried out using BLAST using the following reference sequences: LA_4008 (NP_714188.1), MXAN_4545 (YP_632713.1), COCOR_04748 (YP_005370712.1), STAUR_4866 (YP_003954471.1), Rv0386 (CCP43116). The coverage for the query sequence, statistical significance (E-value), and maximum amino acid identify ("Max Ident") are indicated at right for each predicted primary sequence.
doi:10.1371/journal.pntd.0002468.g003

encoding over one hundred lipoproteins [40], the function and localization of many remain unclear.

Discussion

Here we describe the use of pathogenomics to identify novel potential virulence genes in the pathogenic spirochete *Leptospira interrogans*. Previous work to identify mechanisms of pathogenesis by gene knockouts and transposon mutagenesis has not yet yielded detailed mechanistic insights into the role of individual genes play in the pathogenesis of leptospirosis. It has long been known in the leptospirosis field that serial *in vitro* passage of pathogenic *Leptospira* yielded attenuated organisms; the converse, serial passage of liver homogenates of infected animals selects for virulence. A previous study explored the genomic and proteomic differences between a pathogenic *L. interrogans* serovar Lai strain 56601, and an avirulent strain IPAV [41]. These data must be carefully considered because the analyzed strains are not isogenic (the IPAV strain is of unknown provenance since details of its original isolation are unavailable) nor do they provide any *in vivo* relevance for identified genes, focusing instead on proteomic differences between strains during *in vitro* EMJH culture. Our current study, which employed whole genome sequence comparison of an attenuated strain with its isogenic pathogenic parent, yielded a small set of protein coding genes (CDS) with point mutations. While most of the 11 specific mutations found here cannot be quantitatively attributed to specific aspects of virulence or pathogenicity, our pathogenomic approach yielded the identification of a novel leptospiral AGC with cAMP elevating activity in host cells and a hitherto unstudied large gene family that is broadly up-regulated, in a tissue-specific manner, *in vivo* during an animal model of acute leptospirosis.

The identification of a non-transmembrane bound AGC in pathogenic *Leptospira* is particularly important for two reasons. First, the primary structure implies a non-housekeeping function since the protein is not predicted to be membrane-associated, unlike the housekeeping AGC. Second, the host cell cAMP elevating activity of LA_4008 reported in this study is the first demonstrated evidence of a possible biological mechanism that could contribute to virulence for *Leptospira*. Although long established and accepted as a virulence mechanism in other pathogens, the evidence of elevation of host cAMP levels by *L. interrogans* suggests a previously unknown mechanism of pathogenesis and immune evasion for this bacterium, especially given recent evidence that pathogenic *Leptospira* may reside within macrophages *in vivo* [42,43]. Manipulation of intracellular cAMP levels in immune cells may be an important means of attenuating host responses to infection [44], an enticing hypothesis given the up-regulation of this gene upon leptospiral entry into the bloodstream observed in this study. Many human pathogens exploit host cell cAMP signaling during infection, for example, the pore-forming toxin CyaA of the respiratory pathogen *Bordetella pertussis* penetrates host cells where it catalyzes the unregulated conversion of cellular ATP to cAMP, thereby impairing superoxide production, chemotaxis, cytokine production, and phagocytosis [45–47]. Similar effects are caused by the edema factor (EF) of *Bacillus anthracis*, the ExoY toxin of *Pseudomonas aeruginosa*, and the AGC

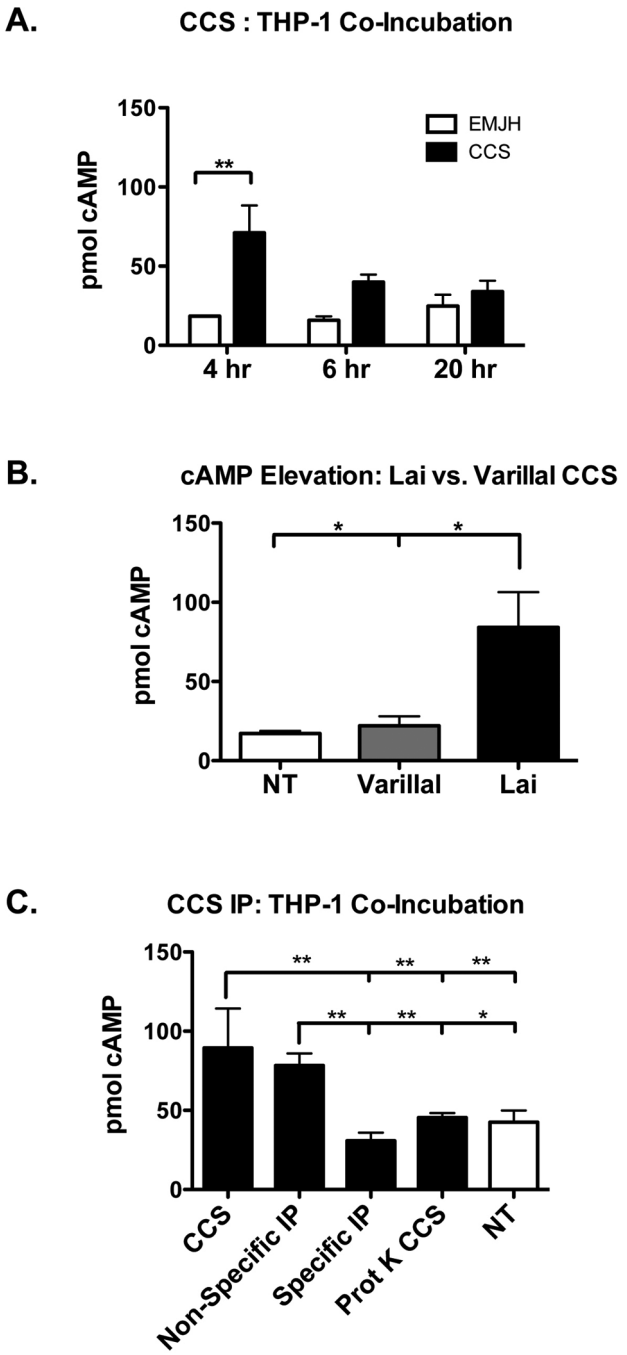


Figure 4. Confirmation of cAMP induction in target mammalian cells by LA_4008 activity in leptospiral culture supernatant. (**A**) THP-1 cell monolayers were treated with leptospire-free concentrated culture supernatant (CCS) from *L. interrogans* Lai or EMJH negative control. (**B**) THP-1 monolayers were treated with CCS from *L. interrogans* Lai or *L. licerasiae* Varillal, NT = not treated. (**C**) THP-1 cell monolayers were treated with CCS, CCS that was immunoprecipitated (IP) with specific anti-peptide antibody raised in rabbits and non-specific anti-LA 4008 antibody, and CCS that was digested with proteinase K. Values in all experiments are represented as the mean (n = 3) ± SD.
doi:10.1371/journal.pntd.0002468.g004

toxin of *Yersinia pestis* [24,48–50]. Due to an unexpected loss of the cryogenically preserved stock cultures, we were unable to assess the cAMP elevating activity of the attenuated P19 strain. However, we would hypothesize that the attenuated SNV-containing variant could have either absolute elimination or quantitative reduction in cAMP elevating activity; this possibility will be addressed directly in ongoing experiments by quantifying the effect of recombinantly producing wild type and mutant LA_4008 on THP-1 and other target cells. We also believe that any observed reduction in activity would have been a quantitative not qualitative difference. Regardless, our findings demonstrate that LA_4008 contributes to a transitory increase in cAMP levels in host cells, and that further experiments are certainly needed to assess the functional consequences of cAMP intoxication in host immune cells during leptospirosis. To formally determine the role of LA_4008 in *Leptopira* pathogenesis is the subject of ongoing experiments, including determining whether this protein modulates mechanisms of evading host defenses.

The identification of a paralogous protein family shared by pathogenic *Leptospira* spp., and two α-proteobacteria *B. bacilliformis*, and *B. australis* was particularly intriguing. The observation that this gene family expanded in pathogenic *Leptospira* and the two *Bartonella* spp. suggests that ancestors of these pathogens must have co-existed at some time and place in the past. Phylogenetic analysis suggests a common origin of this gene family, and revealed a greater divergence in the *Bartonella* members, indicated by greater branch length differences (Fig. 5B). Regardless of the source of the primary ortholog, the founding gene was presumably transferred after the branching of pathogenic *Leptospira* from the other clades of *Leptospira*, although it is also possible that gene loss occurred in intermediates or saprophytes evolved from pathogens. Although we cannot speculate on the molecular mechanism of gene transfer, it is interesting to consider the conditions that would have been conducive to such an event. *L. interrogans* is a globally distributed bacterium that can infect many vertebrate hosts as well as live in the environment; it is considered an extracellular parasite, although evidence is mounting that *Leptospira* [42,43,51] are able to persist within macrophages and transverse epithelial cells [52]. *B bacilliformis* and *B. australis* are facultative intracellular pathogens found only in a specific region of South America [53] and Australia respectively. The PF07598 family shared between pathogenic *Leptospira* might be shared by other *Bartonella* species that have yet to be sequenced or even identified, such as those recently found in Thailand [54,55]. The maintenance of multiple members of this paralogous gene family clearly must confer a selective advantage to these pathogens. We performed a meta-analysis of 6 previous studies [12,56–60] that explored transcriptional responses of *L. interrogans* during exposure to host-like physiological conditions (Table S2 in Text S1), and discovered that the expression levels of several of these genes occurs in response to multiple stimuli. This implies that *L interrogans* responds to signals from the host milieu that lead to the alteration of expression of these genes in a differential manner during its infection cycle.

The identification of leptospiral AGC and PF07598 gene family orthologs in specific species of evolutionary distant alpha and delta-proteobacteria was an unexpected and exciting discovery. Given the broad host range of *Leptospira* as well as their environmental persistence, the horizontal gene transfers our findings imply emphasize how the soil context within the unique transmission cycle of *Leptospira* has likely shaped the evolution of pathogenic mechanisms for these bacteria.

Our investigation was not without limitations. The attenuation experiment was done only once. While genes of pathogenetic interest were identified here, whether these mutations occurred stochastically or not remains to be determined. Accumulation of mutations during the attenuation process was not assessed so that step-wise accumulation of mutations could not be attributed to a

A.

Protein	Accession	Function	*L. interrogans* sv. Copenhageni str. Fiocruz L1-130	*L. kirschneri* sv. Cynopteri str. 3522 CT	*L. noguchii* sv. Panama str. CZ 214T	*L. alstoni* sv. Pingchang str. 80-412	*L. borgpetersenii* sv. Javanica str. UI 09931	*L. weilii* sv. undetermined str. LNT 1234	*L. alexanderi* sv. Manhao str. 3 L 60T	*L. santarosai* sv. Shermani str. 1342KT	*L. kmetyi* sv. Malaysia str. Bejo-Iso9T	*L. fainei* sv. Hurstbridge str. BUT 6T	*L. broomii* sv. Hurstbridge str. 5399T	*L. wolffii* sv. Khorat str. Khorat-H2T	*L. licerasiae* sv. Varillal str. VAR 010	*L. inadai* sv. Lyme str. 10T	*L. wolbachii* sv. Codice str. CDC	*L. yanagawae* sv. Saopaulo str. Sao Paulo	*L. biflexa* sv. Patoc str. Patoc1	*L. vanthielii* sv. Holland str. Waz Holland	*L. terpstrae* sv. Hualin str. LT 11-33T	*L. meyeri* sv. Hardjo str. Went 5
			P	P	P	P	P	P	P	P	P	I	I	I	I	I	S	S	S	S	S	S	S
LA_3388	NP_713568.1	hypothetical protein	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_3490	NP_713670.1	hypothetical protein	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_0589	NP_710770.1	hypothetical protein	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_0591	NP_710772.1	hypothetical protein	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_0620	NP_710801.1	hypothetical protein	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_0769	NP_710950.2	hypothetical protein	+	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_0835	NP_711016.1	hypothetical protein	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_0934	NP_711115.1	hypothetical protein	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_1400	NP_711581.2	hypothetical protein	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_1402	NP_711583.2	hypothetical protein	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_2628	NP_712809.1	hypothetical protein	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LA_3271	NP_713451.1	hypothetical protein	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

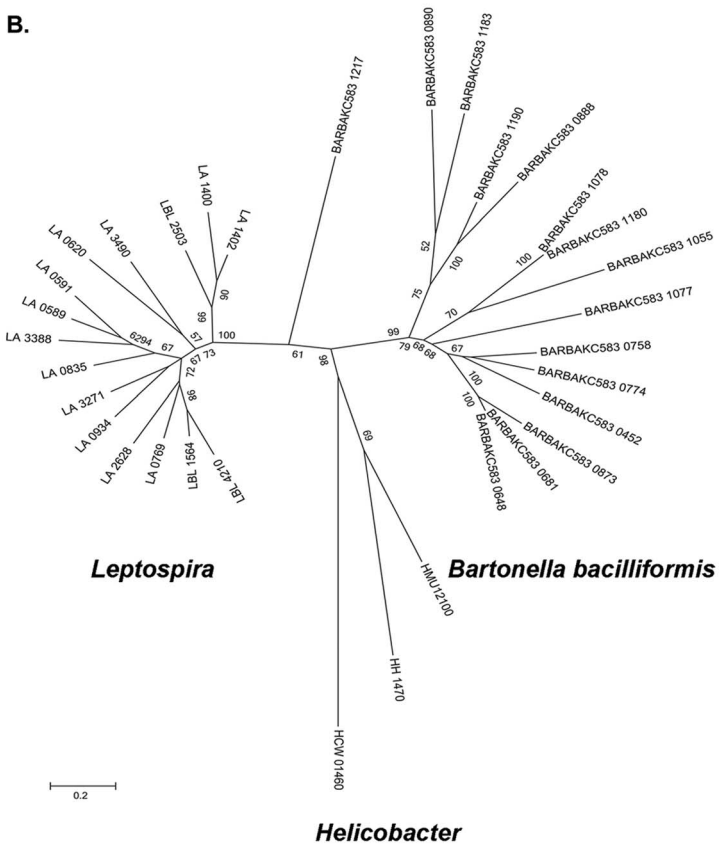


Figure 5. Phylogenetic and in vivo gene expression analysis of the PF07598 paralogous gene family shared by pathogenic *Leptospira* and *Bartonella bacilliformis*. (**A**) Distribution of the paralogous gene family shared by *Leptospira* and *Bartonella bacilliformis* in the genus *Leptospira*. P, pathogen; I, intermediate; S, saprophyte. (**B**) An unrooted phylogenetic tree was constructed of protein sequences from all identifiable homologs of the DUF1561 protein family found in GenBank and the PATRIC databases, which included predicted sequences from the following bacteria (*Helicobacter* spp. and *B. bacilliformis* genome locus tags and protein sequences used for constructing the tree are listed in Table S4 in Text S1): *L. interrogans* Lai, *L. borgpeterseni* Hardjo; *Helicobacter cetorum*, *H. hepaticus* and *H. mustelae*; and *B. bacilliformis* full-length sequences were aligned using MAFFT. Node labels represent support from 500 bootstrap replicates. Tree drawn to scale, with branch lengths measured in the number of substitutions per site. All positions containing gaps and missing data were eliminated. Analyses were conducted in MEGA5. (**C–E**) *In vivo* relevance of the leptospiral paralogous gene family was assessed in the acute hamster infection model as described in Fig. 1. Transcript levels of the genes were assessed by real time, reverse transcriptase quantitative PCR of blood, liver and kidney 4 days after hamster infection and compared to log phase *in vitro* cultured *Leptospira*. Leptospiral gene expression levels in infected tissue vs. EMJH medium alone were expressed logarithmically as the $\log_2$ of the fold change between the two conditions. Solid bars indicate proteins containing predicted signal peptides that suggest extracellular presence, i.e. secretion or cell-surface, of the protein, consistent with bacterial interaction with the host. Data represented are the mean ± SEM of 3 independent experiments (n = 7 animals).
doi:10.1371/journal.pntd.0002468.g005

level of virulence. Proteomic comparisons between *ex-vivo*-isolated and EMJH cultured leptospires were not performed, as our study only focused on gene transcriptional levels, which do not necessarily correlate with protein expression levels. It would be interesting to undertake such *ex-vivo* proteomic investigations in *Leptospira*; especially given the vast transcriptional up-regulation of identified genes upon entry into host tissues. Further investigation remains to define the precise mechanisms of how the identified genes in our study relate to the virulence and pathogenesis of leptospirosis, as a majority of these genes have undiscovered functions.

We show here that a systems biology-pathogenomic approach to infer virulence-related genes in *Leptospira interrogans* identified a notable set of hitherto unstudied genes with both pathogenetic and evolutionary significance, including a putative soluble adenylate/guanylate cyclase (AGC), and a paralogous gene family shared by pathogenic *Leptospira* and the distantly related pathogens *B. bacilliformis*, a human-specific pathogen geographically restricted to the Andes mountains of South America, and *B. australis*, a species currently known to only infect kangaroos. This pathogenomic approach is generalizable beyond prokaryotes and particularly relevant to novel virulence gene identification in any pathogen capable of *in vitro* attenuation. Given the recalcitrant nature of pathogenic leptospires to genetic manipulation, this approach represents an improved method to identify important virulence genes in pathogens whose pathogenesis remains poorly defined by current research strategies, and highlights the extraordinary insights into bacterial pathogenesis and evolutionary biology that large scale genomic sequencing can produce in the context of simple experimentation. These genes will hopefully spur much needed research into the pathogenesis of this neglected disease, but many may also represent rational choices for new vaccine studies.

Supporting Information

Text S1 Supporting information. Includes: **Figure S1.** Validation of 16S rDNA Gene to Normalize Leptospira *In Vivo* Gene Expression. **Figure S2.** Alignment of *Bartonella bacilliformis* and *Leptospira interrogans* serovar Lai anonymous paralog families. **Table S1.** *Leptospira* Species Distribution of Pathogenomically-Discovered Genes. **Table S2.** Differential Expression of Gene Family Members During Exposure of *L. interrogans* to Host-like Conditions. **Table S3.** Primers used for In-vivo RT-qPCR Analysis. **Table S4.** Genome locus tags and GenBank protein sequence accession numbers for *Bartonella bacilliformis* and *Helicobacter* spp. PF07598 family homologs used to construct Figure 5A. (DOCX)

Acknowledgments

We thank Dr. Douglas Berg (UC San Diego) for critical review of the manuscript and are grateful to Dr. David Haake (UCLA) for providing an isolate of *Leptospira interrogans* serovar Lai strain 55601. We are grateful to Professor Michael Minnick of the University of Montana for discussions about *Bartonella* and for his calling to our attention the recent availability of the complete genome sequence of *Bartonella australis* in GenBank. We thank Paula Maguina, Staff Research Associate, UC San Diego, for her important and key scientific and logistical contributions to the work reported here, and Jason Tanseco for contributions and early discussions of this project.

Author Contributions

Conceived and designed the experiments: JSL APC JNR JMV MAM. Performed the experiments: JSL APC JNR AM MAM. Analyzed the data: JSL DEF DHH APC JNR LB DH SD RS GS AM JMV MAM. Contributed reagents/materials/analysis tools: JSL DEF DHH APC JNR LB DH SD RS GS AM JMV MAM. Wrote the paper: JSL DEF DHH APC JNR JMV MAM.

References

1. Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, et al. (2003) Leptospirosis: A zoonotic disease of global importance. Lancet Infect Dis 3: 757–771.
2. LERG (2011) Report of the Second Meeting of the Leptospirosis Burden Epidemiology Reference Group. Geneva, Switzerland: World Health Organization.
3. Ricaldi JN, Fouts DE, Selengut JD, Harkins DM, Patra KP, et al. (2012) Whole Genome Analysis of *Leptospira licerasiae* Provides Insight into Leptospiral Evolution and Pathogenicity. PLoS Negl Trop Dis 6 (10): e1853; doi:1810.1371/journal.pntd.0001853.
4. Matthias MA, Diaz MM, Campos KJ, Calderon M, Willig MR, et al. (2005) Diversity of bat-associated Leptospira in the Peruvian Amazon inferred by bayesian phylogenetic analysis of 16S ribosomal DNA sequences. Am J Trop Med Hyg 73: 964–974.
5. Brenner DJ, Kaufmann AF, Sulzer KR, Steigerwalt AG, Rogers FC, et al. (1999) Further determination of DNA relatedness between serogroups and serovars in the family Leptospiraceae with a proposal for Leptospira alexanderi sp. nov. and four new Leptospira genomospecies. Int J Syst Bacteriol 49 Pt 2: 839–858.
6. Yasuda PH, Steigerwalt AG, Sulzer KR, Kaufmann AF, Rogers F, et al. (1987) Deoxyribonucleic acid relatedness between serogroups and serovars in the family *Leptospiraceae* with proposals for seven new *Leptospira* species. Int J Syst Bacteriol 37: 407–415.
7. Ko AI, Galvao Reis M, Ribeiro Dourado CM, Johnson WD, Jr., Riley LW (1999) Urban epidemic of severe leptospirosis in Brazil. Salvador Leptospirosis Study Group. Lancet 354: 820–825.
8. Ko AI, Goarant C, Picardeau M (2009) Leptospira: the dawn of the molecular genetics era for an emerging zoonotic pathogen. Nat Rev Microbiol 7: 736–747.
9. Lourdault K, Aviat F, Picardeau M (2009) Use of quantitative real-time PCR for studying the dissemination of Leptospira interrogans in the guinea pig infection model of leptospirosis. J Med Microbiol 58: 648–655.
10. Segura E, Ganoza C, Campos K, Ricaldi JN, Torres S, et al. (2005) Clinical spectrum of pulmonary involvement in leptospirosis in an endemic region, with quantification of leptospiral burden. Clin Infect Dis 40: 343–351.
11. Ristow P, Bourhy P, da Cruz McBride FW, Figueira CP, Huerre M, et al. (2007) The OmpA-like protein Loa22 is essential for leptospiral virulence. PLoS Pathog 3: e97.

12. Lo M, Bulach DM, Powell DR, Haake DA, Matsunaga J, et al. (2006) Effects of temperature on gene expression patterns in Leptospira interrogans serovar Lai as assessed by whole-genome microarrays. Infect Immun 74: 5848–5859.
13. Murray GL, Srikram A, Henry R, Hartskeerl RA, Sermswan RW, et al. (2010) Mutations affecting Leptospira interrogans lipopolysaccharide attenuate virulence. Mol Microbiol 78: 701–709.
14. Lourdault K, Cerqueira GM, Wunder EA, Jr., Picardeau M (2011) Inactivation of clpB in the pathogen Leptospira interrogans reduces virulence and resistance to stress conditions. Infect Immun 79: 3711–3717.
15. Liao S, Sun A, Ojcius DM, Wu S, Zhao J, et al. (2009) Inactivation of the fliY gene encoding a flagellar motor switch protein attenuates mobility and virulence of Leptospira interrogans strain Lai. BMC Microbiol 9: 253.
16. Lambert A, Picardeau M, Haake DA, Sermswan RW, Srikram A, et al. (2012) FlaA proteins in Leptospira interrogans are essential for motility and virulence but are not required for formation of the flagellum sheath. Infect Immun 80: 2019–2025.
17. Bourhy P, Louvel H, Saint Girons I, Picardeau M (2005) Random insertional mutagenesis of Leptospira interrogans, the agent of leptospirosis, using a mariner transposon. J Bacteriol 187: 3255–3258.
18. Murray GL, Morel V, Cerqueira GM, Croda J, Srikram A, et al. (2009) Genome-wide transposon mutagenesis in pathogenic Leptospira species. Infect Immun 77: 810–816.
19. Ren SX, Fu G, Jiang XG, Zeng R, Miao YG, et al. (2003) Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing. Nature 422: 888–893.
20. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29: e45.
21. Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18: 691–699.
22. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.
23. Fouts DE, Brinkac L, Beck E, Inman J, Sutton G (2012) PanOCT: automated clustering of orthologs using conserved gene neighborhood for pan-genomic analysis of bacterial strains and closely related species. Nucleic Acids Res 40: e172.
24. Ahuja N, Kumar P, Bhatnagar R (2004) The adenylate cyclase toxins. Crit Rev Microbiol 30: 187–196.
25. Cerveny L, Straskova A, Dankova V, Hartlova A, Ceckova M, et al. (2013) Tetratricopeptide repeat motifs in the world of bacterial pathogens: role in virulence mechanisms. Infect Immun 81: 629–635.
26. D'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. Trends Biochem Sci 28: 655–662.
27. Castro LI, Hermsen C, Schultz JE, Linder JU (2005) Adenylyl cyclase Rv0386 from Mycobacterium tuberculosis H37Rv uses a novel mode for substrate selection. FEBS J 272: 3085–3092.
28. Agarwal N, Lamichhane G, Gupta R, Nolan S, Bishai WR (2009) Cyclic AMP intoxication of macrophages by a Mycobacterium tuberculosis adenylate cyclase. Nature 460: 98–102.
29. Chen CC, Chiu KT, Sun YT, Chen WC (1999) Role of the cyclic AMP-protein kinase A pathway in lipopolysaccharide-induced nitric oxide synthase expression in RAW 264.7 macrophages. Involvement of cyclooxygenase-2. J Biol Chem 274: 31559–31564.
30. Guy L, Nystedt B, Toft C, Zaremba-Niedzwiedzka K, Berglund EC, et al. (2013) A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen Bartonella. PLoS Genet 9: e1003393.
31. Bensing BA, Rubens CE, Sullam PM (2001) Genetic loci of Streptococcus mitis that mediate binding to human platelets. Infect Immun 69: 1373–1380.
32. Bensing BA, Siboo IR, Sullam PM (2001) Proteins PblA and PblB of Streptococcus mitis, which promote binding to human platelets, are encoded within a lysogenic bacteriophage. Infect Immun 69: 6186–6192.
33. Mitchell J, Siboo IR, Takamatsu D, Chambers HF, Sullam PM (2007) Mechanism of cell surface expression of the Streptococcus mitis platelet binding proteins PblA and PblB. Mol Microbiol 64: 844–857.
34. Mitchell J, Sullam PM (2009) Streptococcus mitis phage-encoded adhesins mediate attachment to {alpha}2-8-linked sialic acid residues on platelet membrane gangliosides. Infect Immun 77: 3485–3490.
35. El-Gedaily A, Paesold G, Krause M (1997) Expression profile and subcellular location of the plasmid-encoded virulence (Spv) proteins in wild-type Salmonella dublin. Infect Immun 65: 3406–3411.
36. Koehn EM, Fleischmann T, Conrad JA, Palfey BA, Lesley SA, et al. (2009) An unusual mechanism of thymidylate biosynthesis in organisms containing the thyX gene. Nature 458: 919–923.
37. Myllykallio H, Lipowski G, Leduc D, Filee J, Forterre P, et al. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. Science 297: 105–107.
38. Leduc D, Graziani S, Lipowski G, Marchand C, Le Marechal P, et al. (2004) Functional evidence for active site location of tetrameric thymidylate synthase X at the interphase of three monomers. Proc Natl Acad Sci U S A 101: 7252–7257.
39. Qin JH, Zhang Q, Zhang ZM, Zhong Y, Yang Y, et al. (2008) Identification of a novel prophage-like gene cluster actively expressed in both virulent and avirulent strains of Leptospira interrogans serovar Lai. Infect Immun 76: 2411–2419.
40. Setubal JC, Reis M, Matsunaga J, Haake DA (2006) Lipoprotein computational prediction in spirochaetal genomes. Microbiology 152: 113–121.
41. Zhong Y, Chang X, Cao XJ, Zhang Y, Zheng H, et al. (2011) Comparative proteogenomic analysis of the Leptospira interrogans virulence-attenuated strain IPAV against the pathogenic strain 56601. Cell Res 21: 1210–1229.
42. Toma C, Okura N, Takayama C, Suzuki T (2011) Characteristic features of intracellular pathogenic Leptospira in infected murine macrophages. Cell Microbiol 13: 1783–1792.
43. Li S, Ojcius DM, Liao S, Li L, Xue F, et al. (2010) Replication or death: distinct fates of pathogenic Leptospira strain Lai within macrophages of human or mouse origin. Innate Immun 16: 80–92.
44. Serezani CH, Ballinger MN, Aronoff DM, Peters-Golden M (2008) Cyclic AMP: master regulator of innate immune cell function. Am J Respir Cell Mol Biol 39: 127–132.
45. Confer DL, Eaton JW (1982) Phagocyte impotence caused by an invasive bacterial adenylate cyclase. Science 217: 948–950.
46. Pearson RD, Symes P, Conboy M, Weiss AA, Hewlett EL (1987) Inhibition of monocyte oxidative responses by Bordetella pertussis adenylate cyclase toxin. J Immunol 139: 2749–2754.
47. Vojtova J, Kamanova J, Sebo P (2006) Bordetella adenylate cyclase toxin: a swift saboteur of host defense. Curr Opin Microbiol 9: 69–75.
48. Leppla SH (1982) Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. Proc Natl Acad Sci U S A 79: 3162–3166.
49. Yahr TL, Vallis AJ, Hancock MK, Barbieri JT, Frank DW (1998) ExoY, an adenylate cyclase secreted by the Pseudomonas aeruginosa type III system. Proc Natl Acad Sci U S A 95: 13899–13904.
50. Shevchenko LA, Mishankin BN (1987) [Adenylate cyclase of the causative agent of plague: its purification and properties]. Zh Mikrobiol Epidemiol Immunobiol: 15–20.
51. Li L, Ojcius DM, Yan J (2007) Comparison of invasion of fibroblasts and macrophages by high- and low-virulence Leptospira strains: colonization of the host-cell nucleus and induction of necrosis by the virulent strain. Arch Microbiol 188: 591–598.
52. Barocchi MA, Ko AI, Reis MG, McDonald KL, Riley LW (2002) Rapid translocation of polarized MDCK cell monolayers by Leptospira interrogans, an invasive but nonintracellular pathogen. Infect Immun 70: 6926–6932.
53. Sanchez Clemente N, Ugarte-Gil CA, Solorzano N, Maguina C, Pachas P, et al. (2012) Bartonella bacilliformis: a systematic review of the literature to guide the research agenda for elimination. PLoS Negl Trop Dis 6: e1819.
54. Saisongkorh W, Rolain JM, Suputtamongkol Y, Raoult D (2009) Emerging Bartonella in humans and animals in Asia and Australia. J Med Assoc Thai 92: 707–731.
55. Kosoy M, Bai Y, Sheff K, Morway C, Baggett H, et al. (2010) Identification of Bartonella infections in febrile human patients from Thailand and their potential animal reservoirs. Am J Trop Med Hyg 82: 1140–1145.
56. Xue F, Dong H, Wu J, Wu Z, Hu W, et al. (2010) Transcriptional responses of Leptospira interrogans to host innate immunity: significant changes in metabolism, oxygen tolerance, and outer membrane. PLoS Negl Trop Dis 4: e857.
57. Lo M, Cordwell SJ, Bulach DM, Adler B (2009) Comparative transcriptional and translational analysis of leptospiral outer membrane protein expression in response to temperature. PLoS Negl Trop Dis 3: e560.
58. Patarakul K, Lo M, Adler B (2010) Global transcriptomic response of Leptospira interrogans serovar Copenhageni upon exposure to serum. BMC Microbiol 10: 31.
59. Lo M, Murray GL, Khoo CA, Haake DA, Zuerner RL, et al. (2010) Transcriptional response of Leptospira interrogans to iron limitation and characterization of a PerR homolog. Infect Immun 78: 4850–4859.
60. Matsunaga J, Lo M, Bulach DM, Zuerner RL, Adler B, et al. (2007) Response of Leptospira interrogans to physiologic osmolarity: relevance in signaling the environment-to-host transition. Infect Immun 75: 2864–2874.
61. Carrillo-Casas EM, Hernandez-Castro R, Suarez-Guemes F, de la Pena-Moctezuma A (2008) Selection of the internal control gene for real-time quantitative rt-PCR assays in temperature treated Leptospira. Curr Microbiol 56: 539–546.