PLOS PATHOGENS

Evolutionary insights into the emergence of virulent *Leptospira* spirochetes

Alexandre Giraud-Gatineau[1], Cecilia Nieves[2], Luke B. Harrison[2], Nadia Benaroudj[1], Frédéric J. Veyrier[2], Mathieu Picardeau[1] *

1 Institut Pasteur, Université Paris Cité, Biology of Spirochetes Unit, Paris, France, 2 Bacterial Symbionts Evolution, Centre Armand-Frappier Santé Biotechnologie, Institut National de la Recherche Scientifique, Université du Québec, Laval, Canada

* mathieu.picardeau@pasteur.fr

Abstract

Pathogenic *Leptospira* are spirochete bacteria which cause leptospirosis, a re-emerging zoonotic disease of global importance. Here, we use a recently described lineage of environmental-adapted leptospires, which are evolutionarily the closest relatives of the highly virulent *Leptospira* species, to explore the key phenotypic traits and genetic determinants of *Leptospira* virulence. Through a comprehensive approach integrating phylogenomic comparisons with *in vitro* and *in vivo* phenotyping studies, we show that the evolution towards pathogenicity is associated with both a decrease of the ability to survive in the environment and the acquisition of strategies that enable successful host colonization. This includes the evasion of the mammalian complement system and the adaptations to avoid activation of the innate immune cells by the highly-virulent *Leptospira* species (also called P1+ species), unlike other species belonging to the phylogenetically related P1- and P2 groups, as well as saprophytes. Moreover, our analysis reveals specific genetic determinants that have undergone positive selection during the course of evolution in *Leptospira*, contributing directly to virulence and host adaptation as demonstrated by gain-of-function and knock-down studies. Taken together, our findings define a new vision on *Leptospira* pathogenicity, identifying virulence attributes associated with clinically relevant species, and provide insights into the evolution and emergence of these life-threatening pathogens.

Author summary

Leptospira is a highly heterogeneous bacterial genus and leptospires are ubiquitous bacteria found as free-living saprophytes or as pathogens that can cause disseminated infections, from asymptomatic carriage in rats to lethal acute infection in both humans and animals. Leptospirosis causes over one million cases and nearly 60,000 deaths annually. Despite leptospirosis being a re-emerging zoonosis, little is known about the ability of the etiologic agent to adapt to different hosts and cause disease. Here, combining genome analysis and phenotyping studies of representative species and mutant strains, we show that only a small group of species have the ability to evade the host immune system and

cause disease. In addition, our findings provide key insight into the emergence of pathogens from a saprophytic ancestor through events of gene gain and genome reduction.

Introduction

Spirochetes, which include the causative agents of Lyme disease, syphilis and leptospirosis, form an evolutionarily and morphologically unique phylum of bacteria. Despite their public health significance, spirochetes remain fastidious and challenging bacteria upon which to perform molecular genetic studies. As a result of their elusive nature, the underlying mechanisms for the emergence of these pathogens remain poorly understood.

Leptospirosis is a re-emergent zoonotic disease caused by pathogenic *Leptospira* species and accounts for approximately 1 million severe cases and 60,000 deaths every year [1]. The worldwide leptospirosis burden is expected to rise as climate and demographic changes fuel ideal conditions, including rising inequality which may contribute to rat-borne transmission which dominates human infection [1,2]. It is likely that *Leptospira* infections are underdiagnosed due to the non-specific clinical presentations, poor performance of diagnostics tests and lack of notification systems and diagnostic laboratory capacity in most highly endemic countries [2]. Pathogenic leptospires typically infect humans via contact of abraded skin or mucous membranes with water contaminated by the urine of animal reservoirs, leading to significant morbidity in tropical and subtropical countries during rainy seasons and heavy rainfalls [3], as well as significant economic losses in the livestock industry [2]. Leptospirosis is associated with a wide spectrum of clinical manifestations, ranging from asymptomatic infection to a syndrome of multi-organ failure and death, and very little is known about the bacterial determinants implicated in severe infections [4].

The knowledge obtained from the study of model bacteria does not always translate to spirochetes. This is particularly true of pathogenic *Leptospira* spp. which are extracellular pathogens that lack many canonical virulence factors, including type 3 to type 10 secretion systems and associated effectors, pathogenicity islands, and virulence plasmids. Pathogenic *Leptospira* also have unique characteristics such as atypical lipopolysaccharides (LPS), peculiar defense mechanisms to oxidative stress, an endoflagellar system, and a large fraction of genes of unknown function [4–6]. Characterizing the mechanisms by which *Leptospira* evolved to be pathogenic should uncover novel mechanisms of bacterial virulence.

The last few years have seen several comparative genomics studies, revealing insights into the epidemiology, evolution, and genetic contents of *Leptospira* [7–12]. Notably, we have shown that the genus *Leptospira* is composed of 2 subclades (S1, S2) of free-living non-pathogenic species and 2 subclades (P1, P2) composed of species with variable pathogenic potential (**Fig 1A**). The subclade P1 can be further divided in two phylogenetically related groups termed P1+ (high-virulence pathogens) and P1- (low-virulence pathogens) [9]. Members of the P1+ group are established pathogenic species that have been reported to cause infections in both human and animals and the vast majority of *Leptospira* strains isolated from mammals belong to the P1+ group (**Fig 1A**). Among the P1+, *L. interrogans* accounts for the majority of human infections and is, by far, the most studied of these organisms and consequently, most described virulence factors have been reported from this species. Interestingly, virulence-associated genes are over-represented in the entire P1 subclade with no clear distinction between the P1+ and the P1- groups (**Fig 1A and S2 Table**). Species in the P1- group and P2 subclade are mostly environmental isolates that have not been phenotypically well-characterized. Nevertheless, there is no evidence for an animal reservoir for most of the P1- and P2 species and
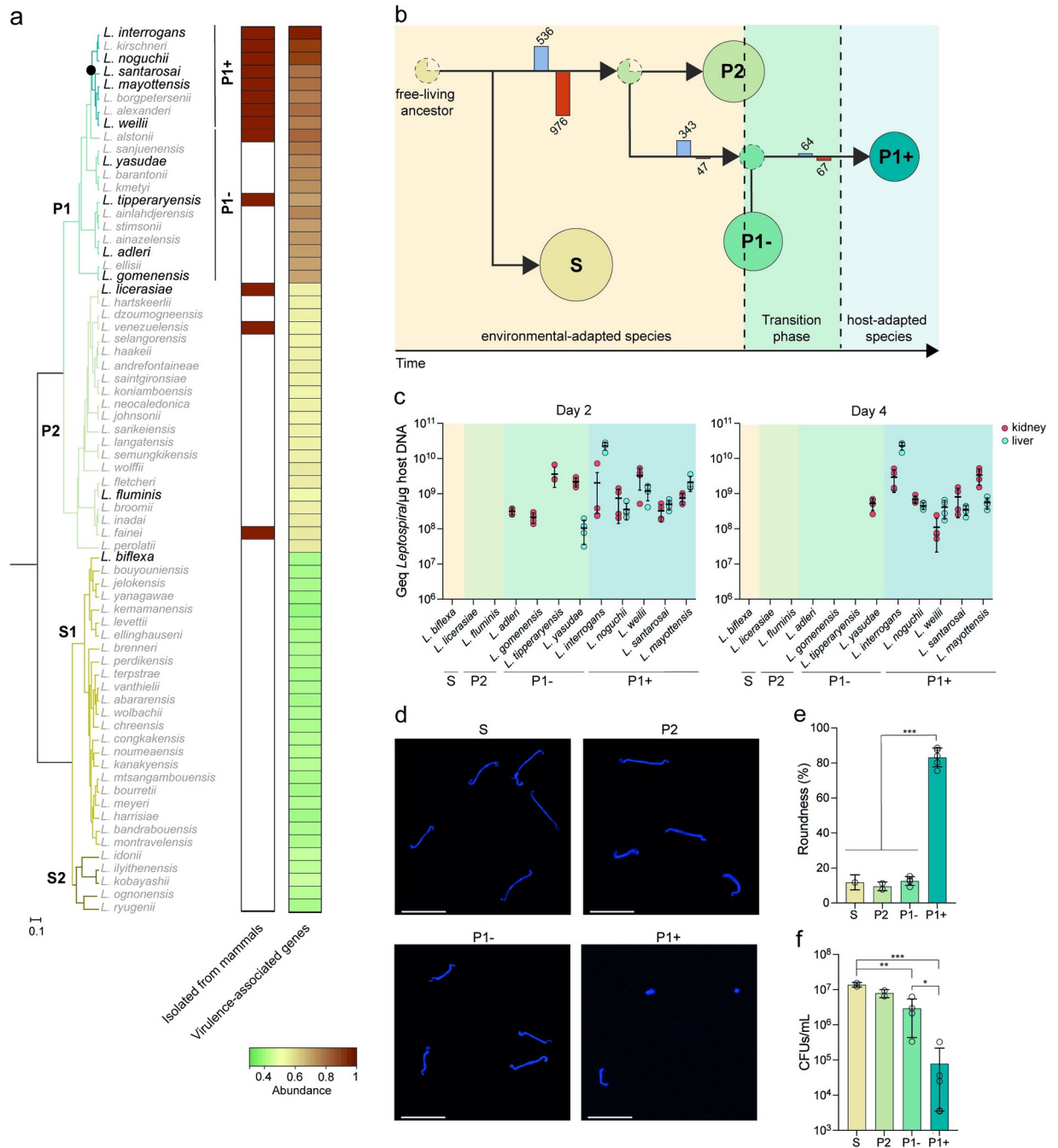
Fig 1. Phylogeny and evolution of host- and non-host adapted *Leptospira* lineages. (a) Distribution of virulence-associated genes and sources of isolation across *Leptospira* species. The phylogenetic tree of *Leptospira* is based on soft-core genes (present in at least 95% of the genomes). The subclade P1, formerly referred to as the "pathogens" lineage, can be separated into two distinct groups: P1+ and P1-. P1 + consists of species associated with severe infections and diverged after a specific node of evolution (filled circle), while P1- comprises species that have not been isolated from patients and are considered as "low-virulent pathogens". The species used in this study are indicated in the phylogenetic tree in bold typeface. Species in the P1- group and P2 subclade isolated from animals are indicated by a red rectangle according to previous studies (*L. alstonii* (frogs) [19], *L. tipperaryensis* (shrew) [20], *L. liceasiae* (humans and rats) [21–23], *L. venezuelensis* (rodents, cattle and humans) [15], and *L. fainei* (pigs and wild boars) [24,25]). The distribution of virulence-associated genes (**S2 Table**) within the genus *Leptospira* is also shown using a heat map representation. **(b)** Schematic representation of the *Leptospira* genealogy. Evolutionary model and reconstruction of ancestral phenotypes in the genus *Leptospira* by PastML analysis using all maximum likelihood methods [26]. Branches are annotated with bars representing the sum of gene gain (blue bar) and loss (red bar). S, P2, P1- and P1+ clades and groups are indicated by spheres (whose size corresponds to the number of species) while most-recent common ancestors are indicated by dashed spheres. The dotted circles represent the most-recent common ancestors of each *Leptospira* group (S, P2, P1- and P1+), and the color indicates the most likely

phenotype of that ancestor. **(c)** Virulence of representative *Leptospira* species in the hamster model. The virulence of *Leptospira* was assessed by infecting hamsters (n = 4) with 10^8 leptospires by the intraperitoneal route. After 2 and 4 days of infection, the burden was assessed in kidney (red symbols) and liver (blue symbols) by quantitative PCR. Data are means ± SD; the absence of values indicates that *Leptospira* DNA was not detected. **(d-f)** Survival of representative *Leptospira* species in water. Leptospires were incubated at RT in filter-sterilized spring water. At 21 days, leptospires were harvested, labelled with DAPI and analyzed by confocal microscopy **(d)** (scale bar: 10μm). The roundness of DAPI-positive leptospires was performed using Icy software **(e)** (n = 100 leptospires). The survival of *Leptospira* in filter-sterilized spring water after 21 days was determined by CFU **(f)**. S: *L. biflexa*; P2: *L. licerasiae, L. fluminis*; P1- group: *L. adleri, L. gomenensis, L. tipperyarensis, L. yasudae*; P1 + subgroup: *L. interrogans, L. noguchii, L. weilii, L. santarosai, L. mayottensis*.

only asymptomatic to mild infections have been sporadically reported [9,11,13–17] (**Fig 1A**). This defined phylogenetic structure, which mostly correlates with ecological niches and pathogenicity, provides a unique opportunity for the investigation of both the evolutionary events and the molecular mechanisms involved in the emergence of pathogenicity in *Leptospira*.

The objectives of this study were to better characterize the P1- species, in particular with regard to their potential virulence, and to identify the genetic and phenotypic changes that characterize the emergence of P1+ species and their close associations with different hosts. This work provides an evolutionary framework for understanding the emergence of pathogenic *Leptospira* lineages.

Results

Stepwise evolution of *Leptospira* from environmental saprophyte to life-threatening pathogen

To investigate the emergence of pathogens in *Leptospira*, we reconstructed the evolutionary history of the genus *Leptospira* and demonstrated that host-adapted pathogens evolved from environmental saprophytes (**Fig 1B and S3 Table**). Our analysis showed that P1+ species emerged stepwise from clades of leptospires with P2 and subsequently P1- -like characteristics (**Fig 1B**). Indeed, the most-recent common ancestor (MRCA) of the P (P1 and P2) clade was more likely to have P2-like phenotype than a P1- phenotype while the most recent common ancestor of the P1 clade was reconstructed to have a P1- phenotype.

To contrast their characteristics, we chose to study in more detail the phenotypes of 5 representative P1+ species (*L. interrogans, L. weilii, L. noguchii, L. mayottensis, and L. santarosai*) and 4 representative P1- species (*L. adleri, L. gomenensis, L. tipperaryensis, and L. yasudae*). Species sampled from P2 (*L. licerasiae and L. fluminis*) and S1 (*L. biflexa*) were also added in our analysis as additional reference strains (**Fig 1 and Table 1**). We first assessed the burden of infection for representative species in the kidney and liver using the hamster model of acute leptospirosis. Although the P1- species can be detected at day 2 post-infection (pi), only the P1 + species were detected in organs at day 4 pi, with the exception of *L. yasudae* which was detected in kidneys only (**Fig 1C**). S2 and P2 species were not detected at either time point. These results suggest that P1- species do not establish persistent infections in a susceptible host but do demonstrate a greater ability to survive in the host at very early times compared to P2 and S species.

We then examined whether P1+ and P1- have the same ability to persist in the environment. After 21 days in spring water, P1+ species exhibited a 100-fold decrease in survival compared to other species and adopted a compromised round shape, while other species retained their helical shapes (**Fig 1D–1F**). In addition, the P1+ species showed slower *in vitro* growth and reduced *in vitro* microbial and metabolic activity (redox activity, fluorescein diacetate hydrolysis and ATP production) in comparison to P1-, P2 and S species (**S1A–S1D Fig**). Interestingly, a gradual decrease in microbial activity and ATP production was observed across *Leptospira* groups on the evolutionary trajectory towards P1+. This is consistent with the already

Table 1. List of isolates used in this study.

Species	Serogroup	Strain	Subclade/ group	Host	Source	GenBank
L. interrogans	Pyrogenes	UP-MMC-NIID-LP	P1+	Human	Philippines	GCA_001047635.1
L. mayottensis	Mini	200901116	P1+	Human	Mayotte	GCA_000306675.3
L. noguchii	Panama	201102933	P1+	Human	Guadeloupe	GCA_000306255.2
L. santarosai	Shermani	56164	P1+	Human	USA	GCA_000313175.2
L. weilii	Celledoni	14535	P1+	Human	Laos	GCA_000243995.3
L. adleri	unknown	201601302	P1-	Soil	Mayotte	GCA_002811845.1
L. gomenensis	unknown	KG8-B22	P1-	Soil	New Caledonia	GCA_004770155.1
L. tipperaryensis	unknown	GWTS#1	P1-	Shrew	Ireland	GCA_001729245.1
L. yasudae	unknown	M12A	P1-	Water	Mayotte	GCA_004770515.1
L. fluminis	unknown	SCS5	P2	Soil	Malaysia	GCA_004771275.1
L. licerasiae	unknown	VAR010	P2	Human	Peru	GCA_000244755.3
L. biflexa	Semaranga	Patoc	S1	Water	Italy	GCA_000017685.1

https://doi.org/10.1371/journal.ppat.1012161.t001

described ongoing process of genome decay of P1+ [7–9,18] characterized by an overrepresentation of mobile elements and pseudogenes, with 33% of non-functional genes being linked to metabolic processes (**S1E–S1H and S4** Figs).

Only P1+ species have developed strategies to escape host immunity

Pathogenic *Leptospira*, like most other pathogenic spirochetes, are stealth pathogens that can escape the recognition by the host innate immune system [27]. We thus asked if P1- species had evolved similar strategies to escape or modulate host immunity during infection. First, we tested resistance to the complement system by assessing the survival of P1+ and non-infectious or low-virulent (P1-, P2 and S1) isolates in presence of normal human serum. Although, the majority of the genes encoding the resistance to the complement system described in *L. interrogans* are present in P1 and P2 but not in S (**S2 Fig**), only P1+ isolates resisted to the complement-mediated killing and had the highest survival (86.27%±23.2) in human serum in comparison to other isolates (0.16%±0.01, 0.40%±0.28 and 1.83%±2.75 for S, P2 and P1-, respectively) (**Fig 2A**). Analysis of Membrane Attack Complex (MAC) deposition on the cell membrane of *Leptospira* showed higher levels of deposition for S1, P2 and P1- compared to P1+ (83.00%±2.01, 81.25%±1.2, 65.64%±1.19 and 13.91%±1.21 for S, P2, P1- and P1+ respectively), which is associated with low survival (**Fig 2B**). Of note, the level of deposition for P1- was intermediate between P2 and P1+. Concordant with these results, maximal C3b deposition and bacterial opsonization in human macrophages was obtained in S1, P2 and P1- species (**S3A–S3B Fig**).

Next, we showed that P1+ strains were significantly less internalized and less adherent to THP-1 macrophages than other species, (**Fig 2C–2E**). Concordantly, levels of pro-inflammatory cytokines (IL-1β, TNF-α and IL-6) were significantly lower in P1+-infected macrophages in comparison to other isolates, including P1- (**Fig 2F**). To understand the molecular mechanisms underpinning the macrophage response, we analyzed the activation of the master transcriptional regulator of pro- and anti-inflammatory host response, NF-κB. P1+-infected macrophages showed a low level of NF-κB translocation into the nuclei (**Fig 2G–2H**). In contrast, NF-κB was mainly localized in the nucleus of S, P2 and P1- isolates. This was correlated with a higher induction of transcription of several inflammatory genes in non-P1+ isolates (**Fig 2I**).

Taken together, these data show that only P1+ isolates have the capacity to resist the complement system and to avoid internalization by macrophages, which is correlated with a less
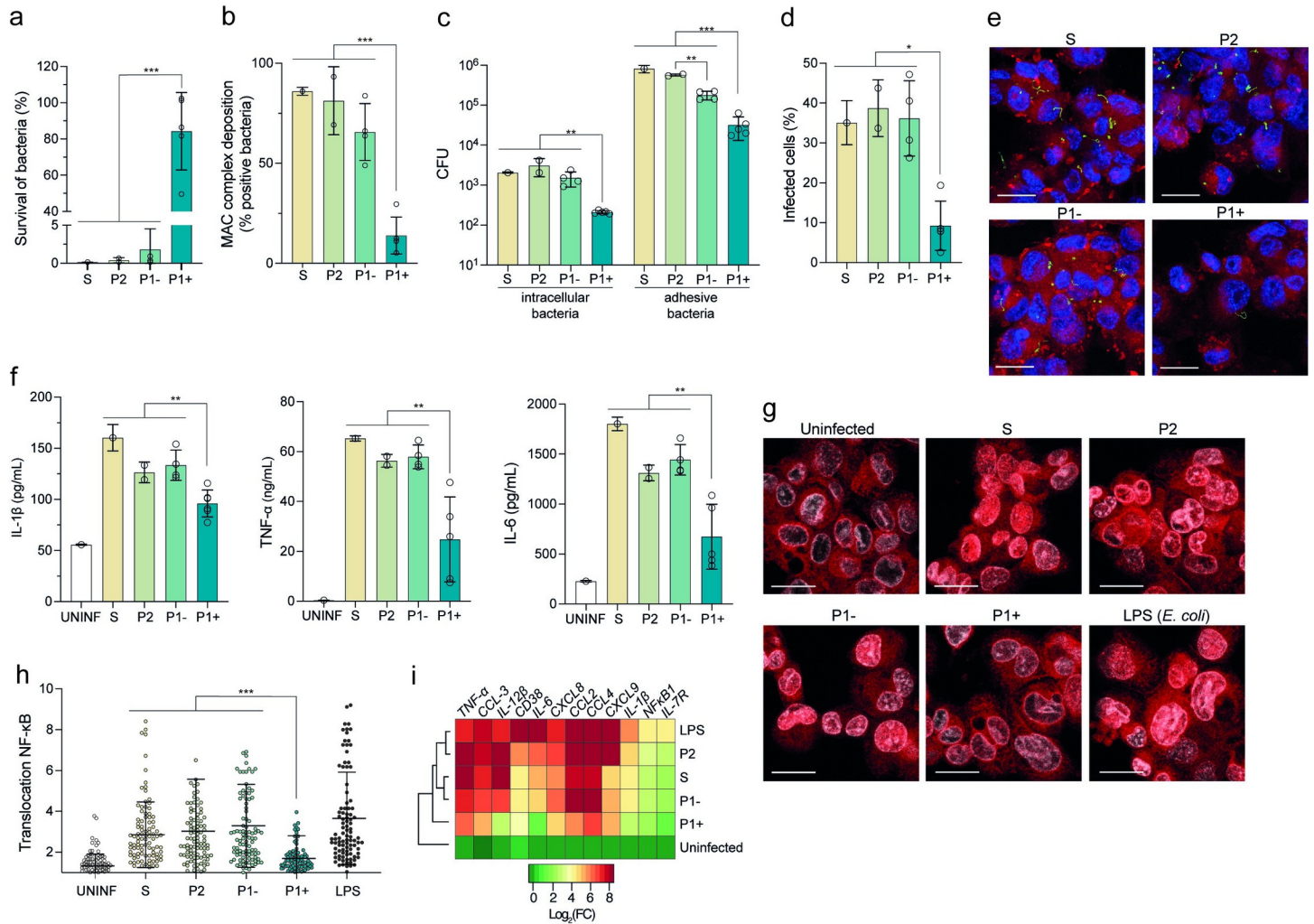
Fig 2. Only P1+ species escape the human complement system and have reduced interaction with human macrophages. (a) Survival of *Leptospira* upon exposure to human serum. Each *Leptospira* species was incubated in 20% of normal or heat-inactivated human serum for 2 hr. Living bacteria were enumerated by CFU (counted in triplicate). The survival was compared to the inactivated-human serum. Unpaired two-tailed Student's t test was used. ***p<0.0001. **(b)** MAC deposition in *Leptospira* was detected by indirect immunofluorescence. CFSE-stained *Leptospira* species were incubated with human serum for 30 min, fixed and then incubated with an anti-MAC antibody (C5b9). Indirect immunofluorescence was quantified by flow cytometry. The percentage was calculated by comparing the number of positive MAC-*Leptospira* to the number of negative MAC-*Leptospira*. Unpaired two-tailed Student's t test was used. ***p<0.0001. **(c)** Bacterial adhesion and entry into human macrophages. To assess bacterial internalization and adhesive bacteria, infected THP-1 macrophages at 2 hr post-infection (PI) were washed with PBS and lysed directly before (adhesive bacteria) and after gentamicin treatment (bacterial internalization). Bacteria were enumerated by CFU (counted in triplicate). **(d-e)** Infection of human macrophages with CFSE-stained *Leptospira*. After 2 hr pi, macrophages were labelled with LysoTracker (Red). The fluorescence was analyzed by confocal microscopy. DAPI (blue) was used to visualize nuclei, CFSE (green) was used to visualize leptospires (scale bar: 20 μm). Quantification of CFSE positive macrophages (infected macrophages) compared to CFSE negative macrophages (uninfected macrophages) was performed using Icy software. Unpaired two-tailed Student's t test was used. *p<0.01. **(f)** Human macrophage response to *Leptospira* infection. After 6 hr pi, IL-1β, TNF-α and IL-6 cytokines release from supernatant were measured by ELISA. Unpaired two-tailed Student's t test was used. **p<0.001. **(g)** Representative fluorescence microscopy images of macrophages uninfected or infected with leptospires for 6 hr. Macrophages were labeled with antibody against NF-κB (red). DAPI (white) was used to visualize nuclei, respectively (scale bar: 20 μm). The NF-κB translocation to the nucleus can be visualized by the increase of NF-κB fluorescence intensity (red here) into the nucleus (white). **(h)** Ratio between nuclear and cytosolic NF-κB fluorescence intensity (n > 100 cells per condition, two-way ANOVA test; ***p<0,01). LPS: *Escherichia coli* LPS. **(i)** Heatmap showing relative expression of several genes regulated by NF-κB after 6hr pi for *Leptospira* infected macrophages. Expression of genes was analyzed and normalized using *gapdh* gene. Hierarchical clustering was performed using Ward's method. Data are the mean ± SD (panels a-d, f, h, and i) or representative (panels e and g) of three independent biological replicates. S: *L. biflexa*; P2: *L. licerasiae, L. fluminis*; P1-subgroup: *L. adleri, L. gomenensis, L. tipperyarensis, L. yasudae*; P1+ subgroup: *L. interrogans, L. noguchii, L. weilii, L. santarosai, L. mayottensis*.

severe early inflammatory response. In contrast, P1-, as well as P2 and S1 species, are sensitive to the complement system and are recognized by macrophages, triggering an inflammatory response.

## P1+-specific genetic determinants of *in vivo* virulence

Comparative genomics was then used to identify genes and/or pathways linked to the pathogenicity, revealing that the P1+ group is characterized by the acquisition of 64 genes and the loss of 67 genes (**Figs 1B and 3A and S4 and S5 Tables**). Protein-encoding genes exclusively present in P1+ species, and potentially involved in the adaptation to the host, include uncharacterized proteins (26%), lipoproteins (20%), transposases (8%), and the established virulence factors collagenase [28], sphingomyelinases [29], and virulence-modifying (VM) proteins (19%) [30] (**Fig 3A**). Among these 64 specific genes, we found seven genes under positive selection (d*N*/d*S* >1; p-value <0.05) including genes encoding the collagenase, lipoproteins, VM proteins and an uncharacterized protein (**Fig 3B and S6 Table**). To experimentally evaluate the role of these proteins in host adaptation, we selected three genes (*LIMLP_09380* [*hypothetical protein*], *LIMLP_03665* [*Collagenase A*], and *LIMLP_11655* [*virulence-modifying protein*]*)* that are induced during *in vivo*-like conditions (**S5 Fig**) as previously shown [31–33] (**S7 Table**). Each gene was heterologously expressed in the P1- species *L. adleri* and *L. yasudae*. Expression was confirmed by RT-qPCR (**S8 Table**) and *LIMLP_03665*-expressing strains exhibited collagenase activity (**S6 Fig**). Production of the three *L. interrogans* proteins led to increased burdens of the low-virulent strains in hamsters in at least one of the tested conditions (**Fig 3C–3D**). To further confirm the involvement of these genes in virulence, CRISPR-*dcas9*-based transcriptional silencing of *LIMLP_03665*, *LIMLP_11655* and *LIMLP_03665* in the pathogen *L. interrogans* resulted in an attenuation of virulence in the hamster model (**Figs S7 and 3E**). We also found that *LIMLP_09380* expression in P1- strains induced resistance to complement-mediated killing and lowered MAC deposition (**Figs 3F and S8**), while *LIMLP_11655* expression reduced the inflammatory response of P1- strain-infected macrophages (**Fig 3F and 3G**). The role of *LIMLP_09380* in inducing resistance to complement-mediated killing and the involvement of *LIMLP_11655* in reducing the inflammatory response of infected macrophages in the pathogen *L. interrogans* were validated with the respective knock-down mutants (**S9 Fig**).

Discussion

Pathogenic *Leptospira* (P1+) evolved from environmental bacteria in progressive trajectory of host-adaptation to animals through deep time, likely beginning with the appearance of the first mammals [3]. Other *Leptospira* species are mostly environmental isolates but some P2 and P1- species may be responsible for asymptomatic to mild infections in both humans and animals [13–17,22,23]. Our results, when summarized and subjected to a multivariate hierarchical clustering analysis, highlight a distinct separation of P1+ isolates from others *Leptospira* subgroups. Interestingly, the P1- isolates, which possess most of the known virulence factors and are phylogenetically closely related to P1+, were clustered with the non-infectious or low-virulent P2 and S isolates (**Fig 4**).

We found that only P1+ strains can establish persistent colonization in the acute animal model of infection showing that these bacteria possess mechanisms enabling the spirochete to survive longer in the blood and to proliferate in target organs, consistent with previous observations [11]. Along the same lines, we show that only the P1+ species can escape immune surveillance and complement-mediated killing. Structural differences within the LPS lipid A may contribute to differential recognition by host immune cells observed between P1+ and P1-/P2 species [34] (**Fig 5**). Adaptation of P1+ species to a wide range of hosts, unlike other species, also correlates with a more diverse and complex LPS O-antigen biosynthesis gene locus [7].

Although P1-, P2 and S species are cleared by host defenses at day 4 pi, P1- species can be differentiated from P2 and S1 species as they can be detected in the early stage of infection,
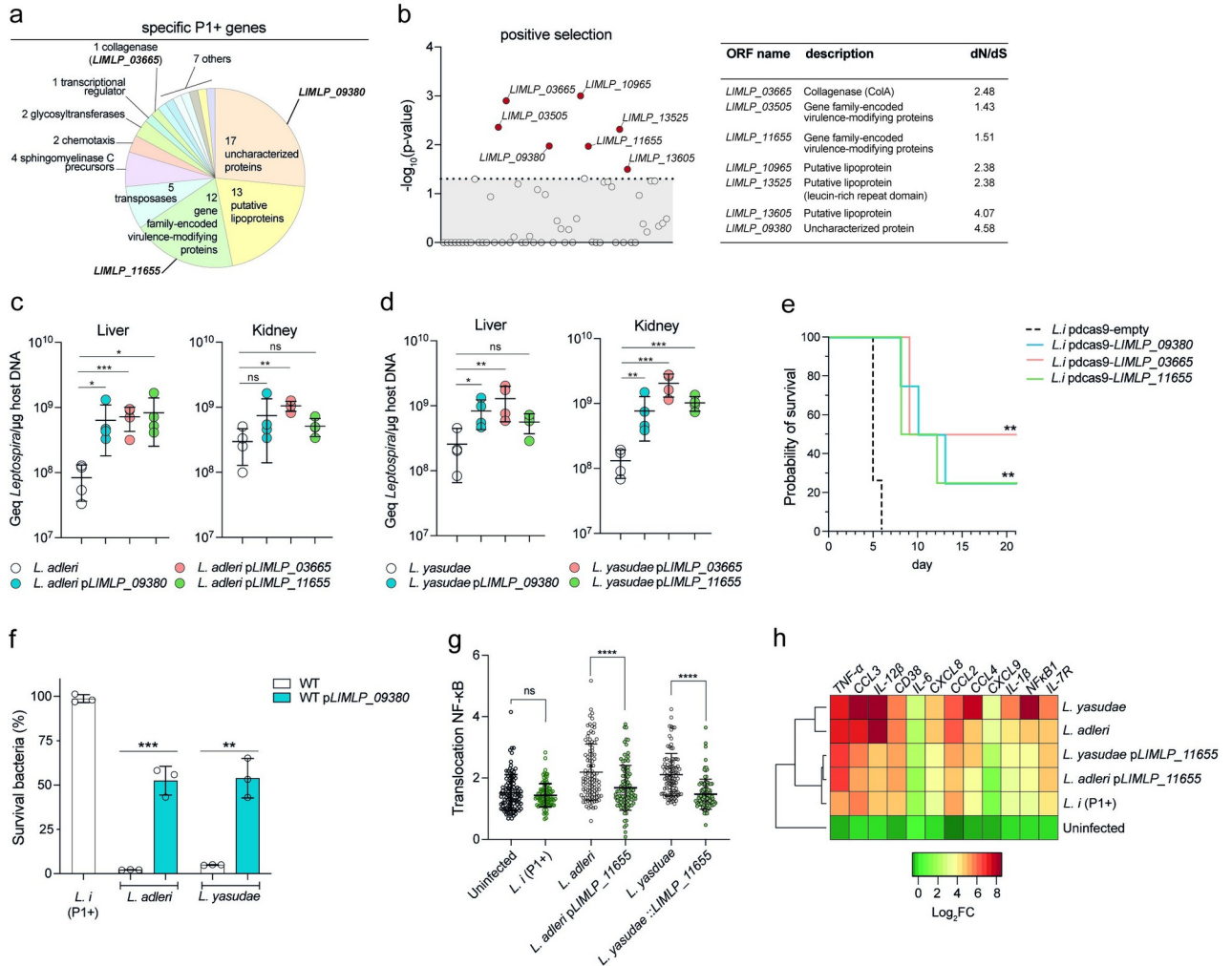
Fig 3. Acquisition of specific P1+ genes involved in host-associated lifestyle. (a) Genes gained in P1+ species. Annotation of the 64 P1+-specific genes is indicated. **(b)** Positive selection analysis of the P1+-specific genes. Significative positive selection was determined using the PoSeiDon pipeline. Significant genes (p-value <0.05, dashed line) and rate of non-synonymous (d*N*) and synonymous (d*S*) in the alignment of orthologous sequences are indicated. **(c-d)** Heterologous expression of putative virulence factors in P1- strains affects the virulence characteristics. The virulence of *L. adleri* (c) and *L. yasudae* (d) P1- strains was assessed by infecting hamsters (n = 4) by intraperitoneal route with 10^8 leptospires. After 1 day of infection, leptospiral load in kidney and liver was assessed by quantitative PCR. pLIMLP_09380, pLIMLP_11655 and pLIMLP_03665 correspond to constructs that enable the heterologous expression of *L. interrogans* genes in *L. adleri* and *L. yasudae*. Unpaired two-tailed Student's t test was used. *p< 0.01, **p<0.001, ***p<0.0001, ns: not significant. **(e)** Silencing of putative virulence factors in the pathogen *L. interrogans* affects the virulence characteristics. Survival of hamsters (n = 4) infected intraperitoneally with 10^6 *Leptospira* for each construct. *L.i.* pdcas9-empty, *L.i.* pdcas9-LIMLP_09380, *L.i.* pdcas9-LIMLP_11655 and *L.i.* pdcas9-LIMLP_03665 correspond to knock-down *L. interrogans* mutant strains of the corresponding genes using dcas9. Statistical significance in comparison with *L.i* pdcas9-empty was determined by a Log rank Mantel Cox test (**p<0.0021). **(f)** Heterologous expression of *LIMLP_09380* in P1- strains affects the survival in human serum. P1- species *(L. adleri* and *L. yasudae)* producing or not *LIMLP_09380* were incubated in 20% of human serum or inactivated-human serum for 2 hr; *L. interrogans* WT *(L.i)* is shown here as a reference for P1+ species. After incubation, the bacteria were enumerated by CFU (counted in triplicate). The percentage of surviving bacteria was calculated using the inactivated-human serum as normalization. Unpaired two-tailed Student's t test was used. **p<0.001, ***p<0.0001. **(g-h)** Heterologous expression of LIMLP_11655 in P1- strains affects the innate immune response of macrophages. Ratio between nuclear and cytosolic NF-κB fluorescence intensity (n > 100 cells per condition, two-way ANOVA test; ****p<0,001; ns: not significant) in the different *Leptospira* strains (g). Heatmap showing relative expression of several genes regulated by NF-κB after 6hr pi for *Leptospira* infected cells. Expression of genes were analyzed and normalized using *gapdh* gene. Hierarchical clustering procedure of *Leptospira* genus was performed using Ward's method (h).

suggesting a limited degree of host adaptation or pre-adaptation in response to other uncharacterized selective pressures. Thus, it has been shown that adaption of environmental microbes to a host-independent factor can incidentally increase their ability to cause an infection [35].
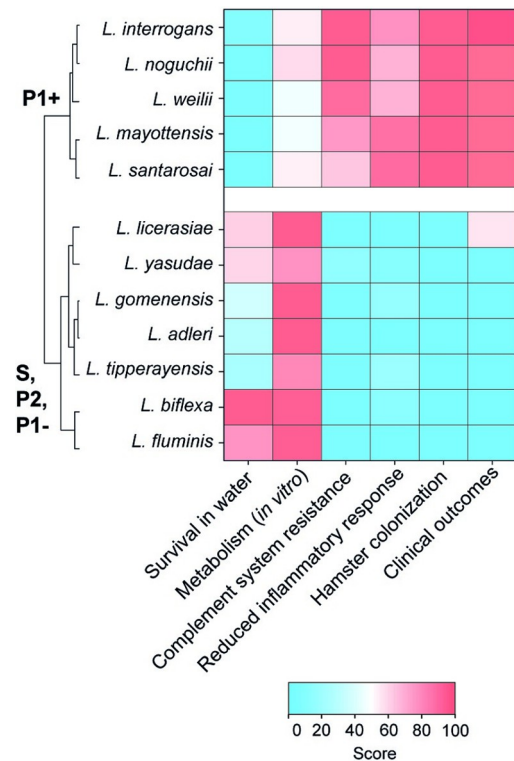
Fig 4. Heatmap representation of the main features of representative *Leptospira* species described in this study. With the exception of *L. licerasiae* [21–23], only P1+ species are responsible for infections in humans. Hierarchical clustering was performed using Ward's method.

https://doi.org/10.1371/journal.ppat.1012161.g004

Our findings suggest that the gradual evolution from P1-like species to P1+ species with the capacity for host colonization was based on only a few key genetic evolutionary innovations through the acquisition and/or the loss of genes, along with gene family expansion and pseudogenization. Acquisition events of virulence genes as well as gene loss during the process of adaptation to the host are characteristic features in the evolution of several specialized pathogens and has been previously demonstrated for the agents of plague, melioidosis and tuberculosis [36–38]. In leptospires, several potential mechanisms of gene acquisition could be hypothesized. Natural transformation has not been convincingly described in *Leptospira* spp. but the conjugation of plasmids can be achieved *in vitro* [39], which strongly suggests that this is a mechanism of gene acquisition *in vivo*. This hypothesis is supported by the significant plasmid diversity which has been observed in *Leptospira* spp., and there is evidence of inter-specific plasmid exchanges [40]. We hypothesize that the dynamics of plasmid exchange may have enriched leptospiral chromosomal DNA with plasmids-derived genes. In addition, putative prophages are found in *Leptospira* spp. Some prophages have even been observed to have been excised from the chromosome and to have formed circular replicons analogous to circular plasmids [41,42]. Thes phages appear to be present in species from the P clades, but are mostly absent in the S clades, suggesting that they may have had a major role in the emergence of the P1+ species [43] by sequentially enriching the P chromosomes with phage-derived genes. Putative virulence factors specific to P1+ isolates include families of sphingomyelinase-like proteins and leucine-rich repeat (LRR) proteins, both probably mediating host-pathogen interactions [29,44]. We also show experimentally that, among the specific genetic determinants of P1+, at least three genes (*LIMLP_09380* [*hypothetical protein*], *LIMLP_03665* [*Collagenase A*], and *LIMLP_11655*
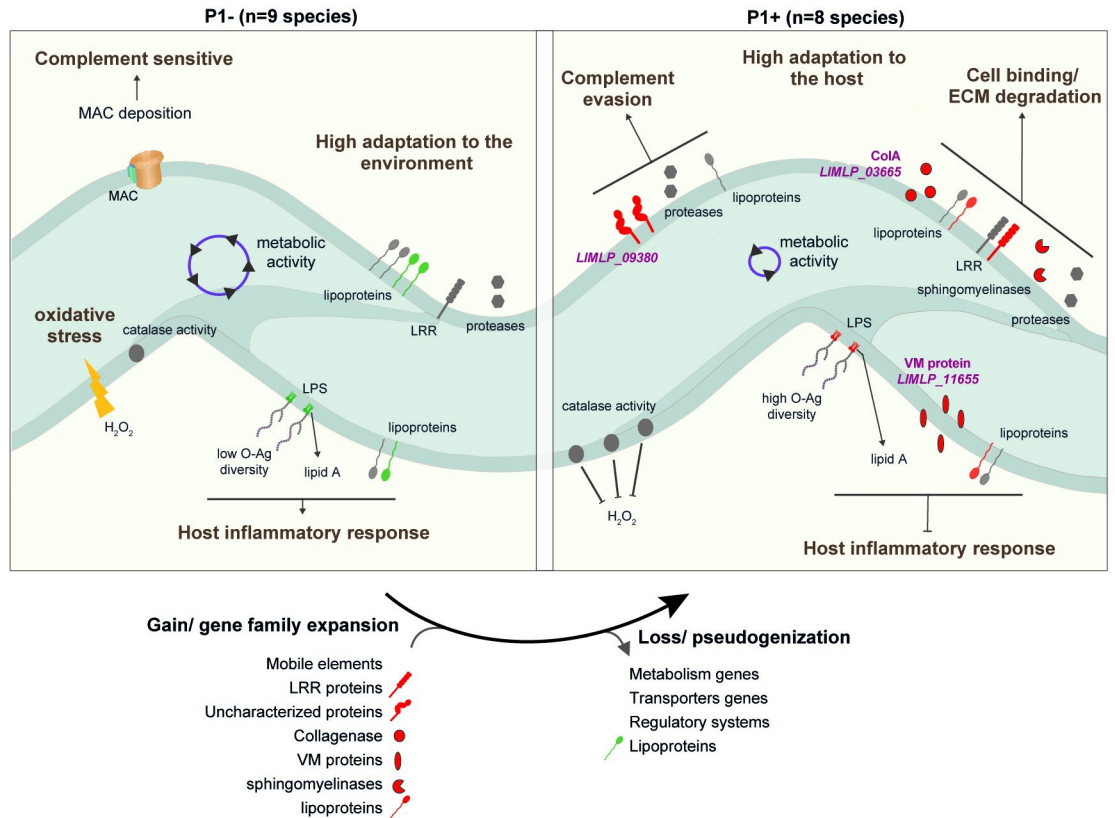
Fig 5. Schematic representation of the evolutionary transition from environmental-adapted *Leptospira* species (P1- group) to host-adapted *Leptospira* species (P1+ group). Host-associated factors found in both P1- and P1+ species (S2 Fig) are indicated in grey. Factors found exclusively in P1+ and P1- are indicated in red and green, respectively. Specific lipoproteins, Leucine-rich repeat (LRR)-encoding proteins, sphingomyelinase-like proteins, virulence-modifying (VM) proteins and uncharacterized proteins are prominent among P1+ isolates. The different factors contributing to host adaptation of P1+ species are represented, including a collagenase (encoded by LIMLP_03665), a hypothetical protein (encoded by LIMLP_09380) and a VM protein (encoded by LIMLP_11655). LIMLP_09380 participates in evasion from complement-mediated killing and the VM proteins is involved in prevention of host inflammatory response. In addition, several factors are contributing to cell binding and ECM (extracellular matrix) degradation in P1+ species. The lipopolysaccharides (LPS) of P1+ species, which have a higher complexity than those of other *Leptospira* species [7,34] may differentially interfere with the host and confer resistance to immune surveillance. The reduced *in vitro* microbial and metabolic activities of P1+ species in comparison to P1- species might be also important for adaptation to the host. The catalase activity of P1+ species is higher than in P1- species, allowing them to better tolerate H_2O_2 as encountered inside a host.

https://doi.org/10.1371/journal.ppat.1012161.g005

[*virulence-modifying protein*]*)* actively contribute to *in vivo* virulence through evasion of host defenses. They include a protein belonging to a paralogous family of proteins called Virulence-Modifying proteins (VM proteins) which are not present in P1- and expanded in the most virulent species (P1+) whose genomes can encode 12 or more paralogs [30], as well as a collagenase which, when mutated, reduces virulence [28]. On the other hand, genes lost in P1+ isolates are predominantly associated with metabolism (31%, S4 Fig and S5 Table) and this is corroborated with reduced metabolic activity and ability to survive in water when compared to other *Leptospira* species. At one extreme of this evolutionary trajectory is the obligate pathogen *L. borgpetersenii*, which cannot persist in the environment and whose genome is characterized by gene decay through extensive IS-mediated genome rearrangement and pseudogenization [18]. Most P1+ species, however, can still persist in the environment and represent different intermediate stages of genome reduction relative to *L. borgpetersenii* on the evolutionary transition to an obligate parasitic lifestyle. Genome evolution is a key determinant in host adaptation, but other

mechanisms should also be considered. Previous studies had demonstrated that genes controlling adhesion, uptake and intracellular survival within macrophages and neutralization of the complement system [45,46] are involved in *Leptospira* pathogenicity. However, we have shown here that most of these genes involved in virulence and host adaptation are also present within the P1- group (**S2 Fig**), implying that differential gene expression of some virulence determinants might also be at the origin of the different pathogenicity of P1- and P1+ species. Supporting this hypothesis, we have recently demonstrated that P1+ isolates exhibit greater resistance to peroxide, a characteristic associated with a constitutive higher expression of *katE*-encoding catalase compared to P1- isolates, which display a lower tolerance to peroxide [47]. Therefore, a rewiring of transcriptional circuits could also contribute to the emergence of pathogenicity in *Leptospira* genus.

In summary, our study shows that the evolution of virulence in leptospires, defined here as the ability to adapt and persist in the host, occurred with the successive appearance of intermediate phenotypes (**Fig 5**). Firstly, a reduction in the metabolism-related genetic repertoire occurred in the most-recent common ancestor of P2 species persisting through the phylogeny of P1- species until the emergence of P1+ species. Secondly, mechanisms that support host colonization evolved in the most-recent common ancestor of the P1 clade, as demonstrated by the detection of P1- species in the organs of the acute animal model during the early stages of infection. Finally, an enrichment in mobile elements appears in P1+ species, which contributes to reductive genomic evolution. This phenomenon further reduces the content of metabolic genes and marks a crucial stage in the evolution of *Leptospira*, as these species adapt to different hosts by acquiring genes essential for evading the host immune response and experience a concomitant relaxation of purifying selection on genes important for survival in the environment. Of note, our results suggest that P1- species could, under certain conditions (e.g. immunocompromised individuals), be responsible for infections in humans. However, the diagnostic tools currently available are not capable of detecting P1-/P2 isolates, and therefore their relevance to human and animal health has yet to be determined. Similarly, the sporadic reports of P1-/P2 infection in animals need to be further investigated to determine whether some animals may serve as reservoirs for low-virulent species. In addition, all P1+ species are probably not equally virulent. Further studies should include large scale analyses integrating clinical data, phenotypic analysis and comparative genomics of P1+ isolates to provide a better understanding of bacterial factors associated with severe infections.

Our findings refine our understanding of virulence in *Leptospira* and provide novel insights into the pivotal steps of host adaptation in *Leptospira*. This study, which enables us to better distinguish hypervirulent *Leptospira* strains from others, may also have important public health implications such as rapid detection of P1+ isolates in clinical specimens and potential cross-protective vaccines targeting conserved antigens exposed on the surface of P1+ isolates.

Material and methods

Ethics statement

Sera were obtained from healthy donors. The blood collection was carried out in accordance with the approved French Ministry of Research and French Ethics Committee protocols by the Etablissement Français du Sang (EFS, n°18/EFS/041). Written consent was received from all participants donating blood for the study. Protocols for animal experiments conformed to the guidelines of the Animal Care and Use Committees of the Institut Pasteur (Comité d'éthique d'expérimentation animale CETEA # 220016), approved by the French Ministry of Agriculture. All animal procedures carried out in our study were performed in accordance with the

European Union legislation for the protection of animals used for scientific purposes (Directive 2010/63/EU).

Bacterial strains and culture conditions

Leptospira strains used in this study are indicated in Table 1. *Leptospira* strains were cultivated aerobically in Ellinghausen-McCullough-Johnson-Harris liquid medium (EMJH) [48,49] at 30°C with shaking at 100 rpm or onto 1% agar solid EMJH media at 30°C during one month. For all experiments, exponentially growing cultures of *Leptospira* were used.

Survival in water

Exponentially growing *Leptospira* species were centrifuged at 2,600 g for 15 min, washed three times and resuspended into filter-sterilized spring water (Volvic). All *Leptospira* species were adjusted to $5x10^8$ leptospires/ml and incubated at room temperature (RT) in the dark. At 21 days, survival was determined by enumeration of colony-forming unit (CFU) on EMJH agar plates. For staining, bacteria were fixed with 4% paraformaldehyde at RT for 15 min, stained with DAPI (1μg/ml, Thermo Fisher) for 10 min and mounted on a slide using Fluoromount mounting medium (Thermo Fisher). Quantification of DAPI staining and roundness was performed using Icy software.

In vivo animal studies

Four week-old Syrian Golden hamsters (RjHan:AURA, Janvier Labs) were infected (4 per group) by intraperitoneal injection with 10^6 or 10^8 *Leptospira* as enumerated using a Petroff-Hausser counting chamber. The animals were monitored daily and euthanized by carbon dioxide inhalation upon reaching the predefined endpoint criteria (sign of distress and morbidity). To assess leptospiral load, blood, kidney, and liver were sampled and DNA was extracted with the Tissue or Blood DNA purification kit (Maxwell, Promega). The bacterial burden and host DNA concentration were determined by qPCR with the Sso Fast EvaGreen Supermix assay (Bio-Rad) using the *flaB2* and *gapdh* genes, respectively. *Leptospira* load was expressed as genomic equivalent (GEq) per μg of host DNA.

Metabolism activity and enzymatic assays

Exponentially growing *Leptospira* ($2x10^8$) were incubated in EMJH at 30°C. Rezasurin (Alamar Blue Assay, ThermoFisher) was added, and bacteria were incubated for 8 hr. The absorbance was measured at 570 and 600 nm. Redox activity was determined based on the ability of cells to reduce rezasurin into resorufin following the manufacturer's instructions.

Exponentially growing *Leptospira* ($2x10^9$) were incubated with fluorescein diacetate (Sigma-Aldrich) at 2 mg/ml in acetone at 30°C for 10 min, then a 2:1 chloroform/methanol solution was added. After centrifugation at 5,000 g for 5 min, the aqueous phase was recovered, and the microbial activity (esterase activity) was obtained by measuring the absorbance at 490nm.

Exponentially growing *Leptospira* ($2x10^9$) were used to determine ATP concentration using the luminescent ATP detection assay (Abcam), according to the manufacturer's instructions.

Total extracts or the supernatant of *Leptospira* (5.5 μg) was used to determine the collagenase activity using the Collagenase Activity Assay Kit (Abcam), according to the manufacturer's instructions.

Comparative genomics and phylogeny

Gene acquisition and loss were assessed using an in-house developed software, MycoHIT [50]. Tblastn searches (E-value = $1e^{-10}$) were independently performed for all 69 *Leptospira* species against reference protein-coding sequences of i) *L. interrogans* str. 56601 (to evaluate gene acquisition), and ii) *L. biflexa* str. Patoc 1 (Paris) (for assessment of gene loss). Presence/absence of genes was determined by a 60% similarity threshold, observed in at least 80% of the members within the target group. The 60% threshold was set using the empirical distribution of protein similarity scores reported by the tblastn alignments against each genome relative to *L. interrogans* and calculating the mean of the 10th percentile of scores. Furthermore, to ensure specificity, presence/absence of these genes should not surpass 20% in other groups. To illustrate, for a gene to be designated as "present" in the P1+ group, it needed a similarity > 60% in a minimum of 7 out of 8 species within that group. Additionally, the presence of that gene could not exceed 20% prevalence across the species included in P1-, P2, S1, and S2 (i.e., present in < 12 species).

Protein-coding genes within each species were classified based on the Clusters of Orthologous Genes (COG) database using eggNOG mapper (options:—evalue 0.001—score 60— pident 40—query_cover 20—subject_cover 20—target_orthologs all) [51]. The representativeness of COG categories per species was calculated as the ratio of protein-coding genes within each specific category, normalized by the total number of protein-coding genes in the respective species.

Pseudogene prediction was performed using Pseudofinder version 1.1.0 (option—annotate) [52]. The search procedure involved sequence alignment against the UniProt/TrEMBL protein database through DIAMOND (option—diamond within the Pseudofinder command line) [53]. Functional characterization of identified pseudogenes was performed based on COG classification, as previously described.

Representative genomes of *Leptospira* species are indicated in Table 1. The gene orthology was determined using GET_HOMOLOGUES version 2021 using parameters of minimum protein coverage of $\geq$ 70%, E-value = 0.01.

Protein motifs were detected using InterProscan v5.30_69.0 (https://doi.org/10.1093/bioinformatics/btu031) for each genomic reference of the 69 *Leptospira* species.

Ancestral reconstruction of the phenotypes of the *Leptospira* genus was performed using PastML [26]. First, a core genome alignment of 69 species of the *Leptospira* genus, encompassing members from P1+, P1-, P2, S1 and S2, was performed using MAFFT through Roary v3.11.2. [54]. The alignment parameters included a 60% identity cut-off and required gene prevalence in at least 95% of the analyzed genomes (Roary's options: -e–mafft -I 60 -cd 95). The resulting alignment comprised a total of 624 soft-core genes, which was used for subsequent phylogenetic analysis. The best evolutionary model was determined using ModelFinder within IQ-TREE version 1.6.11 [55]. The best-fit model was identified as GTR + F + R7 by Bayesian Information Criterion (BIC). Maximum-likelihood phylogenetic analysis was performed with IQ-TREE, incorporating 10,000 ultrafast bootstrap replicates [56]. The final phylogenetic tree was visualized with FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and rooted with the saprophytic group (S, comprising S1 and S2 as shown in Vincent et al. [9]). This rooted phylogenetic tree, along with a csv file associating species with taxonomic groups (P1+, P1-, P2, S1 and S2), served as inputs for PastML. Ancestry prediction was performed by using a combination of all maximum-likelihood methods (maximum a posteriori, MAP; joint; marginal posterior probabilities approximation, MPPA), and the F81 model.

Positive selection of P1+ specific genes

The dN/dS ratios were calculated using codeML though the PoSeiDON pipeline (10.1093/bio-informatics/btaa695). This pipeline performed in-frame alignment of each protein-coding sequence, phylogenetic reconstructions and detection of positively selected sites in the full alignment. Maximum-likelihood tests to detect positive selection under varying site models are performed using M7 versus M8 by codeML with three independent codon models F1X4, F3X4 and F6. Then, we used an empirical Bayes approach to calculate posterior probabilities that a codon coming from a site class with dN/dS>1. Genes were considered to be positively selected when the p-value is <0.05 and the dN/dS ratio exceeds one.

Complement system activity

Leptospira strains were incubated with 20% human serum from healthy individual donors or 20% heat-inactivated human serum diluted in PBS. At the indicated times, bacteria were harvested by centrifugation at 2,600 g for 15 min and fixed with 4% paraformaldehyde for 15 min at RT. Bacteria were incubated with anti-C5b9 (Thermo Fisher) or with anti-C3b (Thermo Fisher) for 2 hr at RT. Bacteria were washed three times with PBS and incubated with Alexa Fluor 555 secondary antibody (Thermo Fisher) for 2 hr at RT. Fluorescence was analyzed using a CytoFLEX Flow Cytometer (Beckman Coulter). The analysis was performed using the FlowJo software.

Macrophages infection model

THP-1 (ATCC TIB-202) cells were cultured in RPMI 1640 medium (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (Sigma) and 2 mM L glutamine (Gibco). For all experiments, THP-1 cells were differentiated/activated into macrophages by a treatment with 50 nM PMA for 2 days following by a 24 hr incubation without PMA. Activated macrophages were inoculated with *Leptospira* at a multiplicity of infection (MOI) of 100 bacteria-per-cell during 2 hr following by 1 hr of gentamicin treatment (Sigma) at 100 μg/ml. Before and after gentamicin treatment, cells were washed three times with PBS.

For Carboxyfluorescein succinimidyl ester (CFSE) labelling, leptospires were resuspended in PBS with 5 μM CFSE (Sigma-Aldrich) for 30 min at RT and then washed three times in PBS. After infection, cells were incubated with 100 nM LysoTracker DND-99 (Thermo Fisher) for 1 hr at 37°C. THP-1 cells were washed three times with PBS and then fixed with 4% para-formaldehyde at room temperature for 15 min. Nuclei were stained with DAPI (1 μg/ml, Thermo Fisher) during 10 min and mounted on a side using Fluoromount mounting medium (Thermo Fisher). Fluorescence was analyzed using a Leica TCS SP8 Confocal System. Quantification of CFSE-positive cell was performed using Icy software.

Host immune response

Quantification of cytokines present in cell culture supernatants was performed by ELISA using the DuoSet ELISA kit (R&D Systems) for TNF-α, IL-1β and IL-6, according to the manufacturer's instructions.

Analysis of NF-KB activation was performed at 6 hr post-infection. Cells were fixed with 4% paraformaldehyde at RT for 15 min, incubated for 5 min with 0.5% saponin (Sigma-Aldrich) in PBS and then incubated for 30 min in 1% BSA (Sigma-Aldrich) and 0.1% saponin (Sigma-Aldrich) in PBS, to permeabilize the cells and to block nonspecific binding. Cells were incubated with anti-NF-kB antibody (Thermo Fisher) overnight at 4°C. Cells were washed and incubated with Alexa Fluor 555 secondary antibody (Thermo Fisher) for 2 hr. Nuclei were

stained with DAPI (1 µg/ml) for 10 min. After labeling, coverslips were set in Fluoromount G medium (Thermo Fisher). Fluorescence was analyzed using a Leica TCS SP8 Confocal System and the NF-kB translocation analysis was performed using Icy software.

Measure of gene expression

Total RNAs from *Leptospira* species or macrophages were extracted using QIAzol lysis reagent (Qiagen) and purified with RNeasy columns (Qiagen). Reverse transcription of mRNA to cDNA was carried out using the iScript cDNA Synthesis kit (Bio-Rad), followed by cDNA amplification using the SsoFast EvaGreen Supermix (Bio-Rad). All primers used in this study are listed in S1 Table. Reactions were performed using the CFX96 real-time PCR detection system (Bio-Rad). The relative gene expression was assessed according to the $2^{-\Delta Ct}$ method using *flaB2* (*LIMLP_09410*) or 16S RNA (*LIMLP_04870*) as reference genes for *Leptospira* or *gadph* as reference gene for human macrophages.

Genetic manipulations of *Leptospira*

Heterologous expression of *LIMLP_09380*, *LIMLP_03665* and *LIMLP_11655 L. interrogans* genes in *L. adleri* and *L. yasudae* was performed by cloning the genes in pMaGRO [57] and introducing the conjugative plasmids in *Leptospira* as previously described [39]. *Leptospira* conjugants were selected on EMJH agar plates containing 50 µg/ml spectinomycin.

Gene silencing in *L. interrogans* was performed using the leptospiral replicative vector pMaOri.dCas9 [58]. The sgRNA cassettes targeting *LIMLP_09380* (`5'TGTGGTACTATG AGATATAC3'`), *LIMLP_03665* (`5'TCGGTTATGACTTTCGGACC3'`) and *LIMLP_11655* (`5'TTTGGACTGCTGACGGAAAG3'`) were cloned into the replicative plasmid pMaori.dcas9 and plasmids pMaori.dcas9_sgRNA-*LIMLP_09380*, pMaori.dcas9_sgRNA-*LIMLP_03665* and pMaori.dCas9_sgRNA-*LIMLP_11655* were introduced into *L. interrogans* by conjugation. *Leptospira* conjugants were selected on EMJH agar plates containing 50 µg/ml spectinomycin.

Quantification and statistical analysis

Data are expressed as means ± standard deviations (SD). Statistical analyses were performed with Prism software (GraphPad Software Inc.), using the t test and one-way analysis of variance (ANOVA) as indicated in the figure legends.

Supporting information

S1 Table. List of primers used in this study.
(XLSX)

S2 Table. List of virulence-associated genes in Leptospira.
(DOC)

S3 Table. Gene gain and loss in P (sheets 1 and 2, respectively), P1- and (sheet s3 and 4, respectively) P1+ groups (sheets 5 and 6, respectively). Presence/absence of genes was determined by a 60% similarity threshold, observed in at least 80% of the members within the target group. Furthermore, to ensure specificity, presence/absence of these genes should not surpass 20% in other groups. The "P" corresponds to P2, P1- and P1+ species.
(XLSX)

S4 Table. Functional KEGG annotation of pseudogenes found in P1+ species.
(XLSX)

S5 Table. List of specific and lost genes in P1+ group.
(XLSX)

S6 Table. Positive selection genes found in the P1+-specific set of genes.
(XLSX)

S7 Table. Expression of P1+-specific genes under positive selection that are deregulated under *in vivo-like* conditions in previous studies [30–32]. Data shown in the column "serum" correspond to deregulated genes of *L. interrogans* serovar Copenhageni incubated with 50% of normal guinea pig serum at 37˚C for to 2 hr; heat-inactivated serum was used as a control. Data shown in the column "DMC" correspond to *L. interrogans* serovar Copenhageni grown in dialysis membrane chambers (DMC) into the peritoneal cavities of rats for 10 days; *in vitro*-cultivated leptospires at 30˚C in EMJH supplemented with 1% rabbit serum was used as a control. Data shown in the column "physiological osmolarity" correspond to *L. interrogans* serovar Copenhageni cultivated in EMJH complemented with 120 nM NaCl for 20 hr at 30˚C; *L. interrogans* cultivated in EMJH without 120 nM NaCl was used as control. The absence of significant deregulation of gene expression was annotated by "-".
(PDF)

S8 Table. Gene expression of LIMLP_09380, LIMLP_11655 and LIMLP_03665 in *L. interrogans*, *L. adleri* and *L. yasudae*.
(XLSX)

S1 Fig. Metabolic evolution of the genus *Leptospira*. (a) Generation time of S, P2, P1- and P1+ species was calculated from growth curves in EMJH medium at 30˚C. **(b)** Evaluation of the redox activity of *Leptospira* using the Alamar Blue assay. **(c)** Microbial activity of *Leptospira* was determined using the fluorescein diacetate assay. Bacteria were incubated with fluorescein diacetate during 10 min and then microbial activity was assessed through the ability of cells to hydrolyze fluorescein diacetate into fluorescein. **(d)** Determination of the ATP concentration of *Leptospira* spp. S: *L. biflexa*; P2: *L. licerasiae*, *L. fluminis*; P1- group: *L. adleri*, *L. gomenensis*, *L. tipperyarensis*, *L. yasudae*; P1+ group: *L. interrogans*, *L. noguchii*, *L. weilii*, *L. santarosai*, *L. mayottensis*. **(e)** Distribution of mobile elements in the genus *Leptospira*. **(f)** Distribution of pseudogenes in the genus *Leptospira* (left panel). Pseudogenes are distinguished using the COG classification (right panel). **(g)** Protein abundance distribution for the category Metabolism in the genus *Leptospira*. Each circle represents one *Leptospira* species. The abundance of proteins associated with the Metabolism category was calculated by the number of metabolic proteins divided by the total number of proteins for each *Leptospira* species. **(h)** Pan-genome distribution of metabolism-associated genes for groups S, P2, P1- and P1+ into four categories (core, softcore, shellcore and cloud). Unpaired two-tailed Student's t test was used. *p< 0.01, **p<0.001, ***p<0.0001, ns: not significant.
(TIF)

S2 Fig. Distribution of putative virulence factors among P1+ and P1- isolates. The presence/absence CDSs of all *Leptospira* species with percentage of similarity ⩾60% according to the protein sequence in *L. interrogans*. Grey and white squares denote the presence and absence of genes, respectively.
(TIF)

S3 Fig. C3b protein deposition and opsonization on macrophages. (a) Detection of C3b protein deposition in *Leptospira* by indirect immunofluorescence. *Leptospira* were stained with CFSE and then incubated with human serum for 30 min. Fixed cells were incubated with an anti-C3b antibody. Indirect immunofluorescence was quantified by flow cytometry. **(b)**

Effect of human serum opsonization on macrophage uptake of *Leptospira* (MOI 1:100). CFSE-labelled *Leptospira* were incubated with human serum during 30 min. Bacteria were then added to macrophages (THP-1) during 2 hr following with 1 hr of gentamicin (100 μg/ml) treatment. Macrophages were lysed and intracellular bacteria were fixed. Number of CFSE bacteria was determined by flow cytometry. Opsonization was assessed by the ratio of human serum to inactivated-human serum. S: *L. biflexa*; P2: *L. licerasiae*, *L. fluminis*; P1- group: *L. adleri*, *L. gomenensis*, *L. tipperyarensis*, *L. yasudae*; P1+ group: *L. interrogans*, *L. noguchii*, *L. weilii*, *L. santarosai*, *L. mayottensis*. Unpaired two-tailed Student's t test was used. **$p < 0.001$, ***$p < 0.0001$.
(TIF)

S4 Fig. Genes lost in P1+ group. COG classification of genes lost in P1+ species.
(TIF)

S5 Fig. Expression of *hyp/LIMLP_09380*, *colA/LIMLP_03665* and *VM/LIMLP_11655* genes in host-like conditions. Relative expression of the genes in *L. interrogans* after overnight incubation in EMJH or MEME medium at 30°C or 37°C. Relative expression levels were normalized to the 16S RNA gene.
(TIF)

S6 Fig. Collagenase activity of *L. adleri* and *L. yasudae* isolates. Measurement of collagenase activity in the supernatant or in total extracts of *Leptospira*.
(TIF)

S7 Fig. Silencing of *hyp/LIMLP_09380*, *colA/LIMLP_03665* and *VM/LIMLP_11655* genes expression using the CRISPR-dcas9 knockdown in *L. interrogans*. Total RNA was extracted from exponentially growing cultures of *L. interrogans*. Relative expression of genes was measured by RT-qPCR. Log$_2$FC levels were normalized to the *flaB2* gene and compared to *L. interrogans* WT.
(TIF)

S8 Fig. Effect of heterologous expression of hyp/LIMLP_09380 on MAC deposition. MAC deposition in P1- species (*L. adleri* and *L. yasudae*) expressing or not *LIMLP_09380* was detected by indirect immunofluorescence. *Leptospira* were stained with CFSE and then incubated with human serum for 30 min. Fixed cells were incubated with an anti-MAC antibody (C5b9). Indirect immunofluorescence of MAC was quantified by flow cytometry. Unpaired two-tailed Student's t test was used. *$p < 0.01$, **$p < 0.001$, ***$p < 0.0001$. Error bars represent the mean ± SD.
(TIF)

S9 Fig. Effect of silencing of *L. interrogans* *hyp/LIMLP_09380* and *VM/LIMLP_11655* in host responses. (a) The CRISPR-*dcas9* knockdown of *hyp/LIMLP_09380* in *L. interrogans* was incubated in 20% of normal or inactivated human serum for 2 hr. After incubation, the bacteria were enumerated by CFU and the percentage of surviving bacteria was normalized by *Leptospira* exposed to inactivated-human serum. Unpaired two-tailed Student's t test was used. ***$p < 0.0001$. Error bars represent the mean ± SD. **(b)** Detection by indirect immunofluorescence of MAC deposition in the CRISPR-dcas9 knockdown of *hyp/LIMLP_09380* in *L. interrogans*. CFSE-labelled *Leptospira* were incubated with human serum for 30 min. Fixed cells were incubated with an anti-MAC antibody (C5b9). Indirect immunofluorescence of MAC was quantified by flow cytometry. Unpaired two-tailed Student's t test was used. **$p < 0.001$. Error bars represent the mean ± SD. **(c)** Ratio between nuclear and cytosolic NF-κB fluorescence intensity (n > 100 cells per condition, two-way

ANOVA test; ****p<0,001; ns: not significant) in the CRISPR-dcas9 knockdown of *VM/LIMLP_11655* in *L. interrogans*.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Alexandre Giraud-Gatineau, Frédéric J. Veyrier, Mathieu Picardeau.

Formal analysis: Alexandre Giraud-Gatineau, Cecilia Nieves, Luke B. Harrison, Frédéric J. Veyrier, Mathieu Picardeau.

Funding acquisition: Frédéric J. Veyrier, Mathieu Picardeau.

Methodology: Alexandre Giraud-Gatineau, Cecilia Nieves, Luke B. Harrison.

Supervision: Nadia Benaroudj, Frédéric J. Veyrier, Mathieu Picardeau.

Writing – original draft: Alexandre Giraud-Gatineau, Frédéric J. Veyrier, Mathieu Picardeau.

Writing – review & editing: Alexandre Giraud-Gatineau, Cecilia Nieves, Luke B. Harrison, Nadia Benaroudj, Frédéric J. Veyrier, Mathieu Picardeau.

References

1. Costa F, Hagan JE, Calcagno J, Kane M, Torgerson P, Martinez-Silveira MS, et al. Global Morbidity and Mortality of Leptospirosis: A Systematic Review. PLoS Negl Trop Dis 2015; 9:e0003898. https://doi.org/10.1371/journal.pntd.0003898 PMID: 26379143

2. Goarant C, Dellagi K, Picardeau M. Ending the Neglect of Treatable Bacterial Zoonoses Responsible for Non-Malaria Fevers. Yale J Biol Med. 2021; 94:351–60. https://doi.org/10.1038/nature12060 PMID: 34211354

3. Davignon G, Cagliero J, Guentas L, Bierque E, Genthon P, Gunckel-Grillon P, et al. Leptospirosis: Towards a Better Understanding of the Environmental Lifestyle of Leptospira. Frontiers in Water. 2023; 5:1195094. https://doi.org/10.3389/frwa.2023.1195094.

4. Picardeau M. Virulence of the zoonotic agent of leptospirosis: still terra incognita? Nat Rev Microbiol 2017; 15:297–307. https://doi.org/10.1038/nrmicro.2017.5 PMID: 28260786

5. Huete SG, Benaroudj N. The Arsenal of Leptospira Species against Oxidants. Antioxidants (Basel). 2023; 12:1273. https://doi.org/10.3390/antiox12061273 PMID: 37372003

6. San Martin F, Fule L, Iraola G, Buschiazzo A, Picardeau M. Diving into the complexity of the spirochetal endoflagellum. Trends Microbiol 2023; 31:294–307. https://doi.org/10.1016/j.tim.2022.09.010 PMID: 36244923

7. Fouts DE, Matthias MA, Adhikarla H, Adler B, Berg DE, Bulach D, et al. What Makes a Bacterial Species Pathogenic?: Comparative Genomic Analysis of the Genus Leptospira PLoS Negl Trop Dis. 2016; 10: e0004403.

8. Xu Y, Zhu Y, Wang Y, Chang YF, Zhang Y, Jiang X, et al. Whole genome sequencing revealed host adaptation-focused genomic plasticity of pathogenic Leptospira. Sci Rep. 2016; 6:20020. https://doi.org/10.1038/srep20020 PMID: 26833181

9. Vincent AT, Schiettekatte S, Goarant C, Neela VK, Bernet E, Thibeaux R, et al. Revisiting the taxonomy and evolution of pathogenicity of the genus Leptospira through the prism of genomics. PLoS Negl Trop Dis. 2019; 13:e0007270. https://doi.org/10.1371/journal.pntd.0007270 PMID: 31120895

10. Guglielmini J, Bourhy P, Schiettekatte O, Zinini F, Brisse S, Picardeau M. Genus-wide Leptospira core genome multilocus sequence typing for strain taxonomy and global surveillance. PLoS Negl Trop Dis. 2019; 13:e0007374. https://doi.org/10.1371/journal.pntd.0007374 PMID: 31026256

11. Thibeaux R, Iraola G, Ferrés I, Bierque E, Girault D, Soupé-Gilbert ME, et al. Deciphering the unexplored Leptospira diversity from soils uncovers genomic evolution to virulence. Microb Genom 2018; 4 (e000144). https://doi.org/10.1099/mgen.0.000144 PMID: 29310748

12. Fernandes LGV, Stone NE, Roe CC, Goris MGA, vanderLinden H, Sahl JW, et al. Leptospira sanjuanensis sp. nov., a pathogenic species of the genus Leptospira isolated from soil in Puerto Rico. Int J Syst Evol Microbiol. 2022; 72:005560.

13. Balamurugan V, Gangadhar NL, Mohandoss N, Thirumalesh SR, Dhar M, Shome R, et al. Characterization of leptospira isolates from animals and humans: phylogenetic analysis identifies the prevalence of intermediate species in India. Springerplus. 2013; 2:362. https://doi.org/10.1186/2193-1801-2-362 PMID: 23961424

14. Chiriboga J, Barragan VA, Arroyo G, Sosa A, Birdsell DN, España K, et al. High Prevalence of Intermediate Leptospira spp. DNA in Febrile Humans from Urban and Rural Ecuador. Emerg Infect Dis. 2015; 21:2141–7. https://doi.org/10.3201/eid2112.140659 PMID: 26583534

15. Puche R, Ferrés I, Caraballo L, Rangel Y, Picardeau M, Takiff H, et al. Leptospira venezuelensis sp. nov., a new member of the intermediate group isolated from rodents, cattle and humans. Int J Syst Evol Microbiol. 2018; 68:513–7. https://doi.org/10.1099/ijsem.0.002528 PMID: 29239713

16. Schiff SJ, Kiwanuka J, Riggio G, Nguyen L, Mu K, Sproul E, et al. Separating putative pathogens from background contamination with principal orthogonal decomposition: evidence for Leptospira in the Ugandan neonatal septisome. Front Med. 2016; 3:22. https://doi.org/10.3389/fmed.2016.00022 PMID: 27379237

17. Zakeri S, Khorami N, Ganji ZF, Sepahian N, Malmasi AA, Gouya MM, et al. Leptospira wolffii, a potential new pathogenic Leptospira species detected in human, sheep and dog. Infect Genet Evol. 2010; 10:273–7. https://doi.org/10.1016/j.meegid.2010.01.001 PMID: 20074666

18. Bulach DM, Zuerner RL, Wilson P, Seemann T, McGrath A, Cullen PA, et al. Genome reduction in Leptospira borgpetersenii reflects limited transmission potential. Proc Natl Acad Sci USA. 2006; 103:14560–5.

19. Smythe L, Adler B, Hartskeerl RA, Galloway RL, Turenne CY, Levett PN. Classification of Leptospira genomospecies 1, genomospecies 3, genomospecies 4 and genomospecies 5 as Leptospira alstonii sp. nov., Leptospira vanthielii sp. nov., Leptospira terpstrae sp. nov., Leptospira yanagawae sp. nov., respectively. Int J Syst Evol Microbiol. 2012; 63:1859–62.

20. Nally JE, Arent Z, Bayles DO, Hornsby RL, Gilmore C, Regan S, et al. Emerging Infectious Disease Implications of Invasive Mammalian Species: The Greater White-Toothed Shrew (Crocidura russula) Is Associated With a Novel Serovar of Pathogenic Leptospira in Ireland. PLoS Negl Trop Dis. 2016; 10: e0005174. https://doi.org/10.1371/journal.pntd.0005174 PMID: 27935961

21. Fernandez C, Lubar AA, Vinetz JM, Matthias MA. Experimental Infection of Rattus norvegicus by the Group II Intermediate Pathogen, Leptospira licerasiae. Am J Trop Med Hyg. 2018; 99(2): 275–280. https://doi.org/10.4269/ajtmh.17-0844 PMID: 29943708.

22. Matthias MA, Ricaldi JN, Cespedes M, Diaz MM, Galloway RL, Saito M, et al. Human leptospirosis caused by a new, antigenically unique leptospira associated with a rattus species reservoir in the peruvian Amazon. PLoS Negl Trop Dis. 2008; 2:e213. https://doi.org/10.1371/journal.pntd.0000213 PMID: 18382606

23. Ganoza CA, Matthias MA, Saito M, Cespedes M, Gotuzzo E, Vinetz JM. Asymptomatic renal colonization of humans in the peruvian Amazon by Leptospira. PLoS Negl Trop Dis 2010; 4:e612. https://doi.org/10.1371/journal.pntd.0000612 PMID: 20186328

24. Perolat P, Chappel RJ, Adler B, Baranton G, Bulach DM, Billinghurst ML, et al. Leptospira fainei sp. nov., isolated from pigs in Australia. Int J Syst Bacteriol. 1998; 48:851–8.

25. Cilia G, Bertelloni F, Cerri D, Fratini F. Leptospira fainei Detected in Testicles and Epididymis of Wild Boar (Sus scrofa). Biology (Basel). 2021; 10:193. https://doi.org/10.3390/biology10030193 PMID: 33806519

26. Ishikawa SA, Zhukova A, Iwasaki W, Gascuel O. A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios. Mol Biol Evol. 2019; 36(9):2069–2085. doi: 10.1093/molbev/msz131 PMID: 31127303

27. Bonhomme D, Werts C. Host and Species-Specificities of Pattern Recognition Receptors Upon Infection With Leptospira interrogans. Front Cell Infect Microbiol 2022; 12:932137. https://doi.org/10.3389/fcimb.2022.932137 PMID: 35937697

28. Kassegne K, Hu W, Ojcius DM, Sun D, Ge Y, Zhao J, et al. Identification of collagenase as a critical virulence factor for invasiveness and transmission of pathogenic leptospira species. J Infect Dis. 2014; 209 (7):1105–15. https://doi.org/10.1093/infdis/jit659 PMID: 24277745

29. Narayanavari SA, Sritharan M, Haake DA, Matsunaga J. Multiple leptospiral sphingomyelinases (or are there?). Microbiology. 2012; 158:1137–46. https://doi.org/10.1099/mic.0.057737-0 PMID: 22422753

30. Chaurasia R, Marroquin AS, Vinetz JM, Matthias MA. Pathogenic Leptospira Evolved a Unique Gene Family Comprised of Ricin B-Like Lectin Domain-Containing Cytotoxins. Front Microbiol 2022; 13:859680. https://doi.org/10.3389/fmicb.2022.859680 PMID: 35422779

31. Caimano MJ, Sivasankaran SK, Allard A, Hurley D, Hokamp K, Grassmann AA, et al. A model system for studying the transcriptomic and physiological changes associated with mammalian host-adaptation by Leptospira interrogans serovar Copenhageni. PLoS Pathog. 2014; 10:e1004004. https://doi.org/10.1371/journal.ppat.1004004 PMID: 24626166

32. Matsunaga J, Lo M, Bulach DM, Zuerner RL, Adler B, Haake DA. Response of Leptospira interrogans to physiologic osmolarity: relevance in signaling the environment-to-host transition. Infect Immun. 2007; 75:2864–74. https://doi.org/10.1128/IAI.01619-06 PMID: 17371863

33. Patarakul K, Lo M, Adler B. Global transcriptomic response of Leptospira interrogans serovar Copenhageni upon exposure to serum. BMC Microbiol 2010; 10:31. https://doi.org/10.1186/1471-2180-10-31 PMID: 20113507

34. Pětrošová H, Mikhael A, Culos S, Giraud-Gatineau A, Gomez AM, Sherman ME, et al. Lipid A structural diversity among members of the genus Leptospira. Front Microbiol. 2023; 14:1181034. https://doi.org/10.3389/fmicb.2023.1181034 PMID: 37303810

35. Smith DFQ. Pathogen Evolution: Exploring accidental virulence. eLife. 2024; 13:e94556.

36. Boritsch EC, Brosch R. Evolution of Mycobacterium tuberculosis: New Insights into Pathogenicity and Drug Resistance. Microbiol Spectr. 2016; 4. https://doi.org/10.1128/microbiolspec.TBTB2-0020-2016 PMID: 27787194

37. Reuter S, Connor TR, Barquist L, Walker D, Feltwell T, Harris SR, et al. Parallel independent evolution of pathogenicity within the genus Yersinia Proc Natl Acad Sci U S A. 2006; 111:6768–73.

38. Yu Y, Kim HS, Chua HH, Lin CH, Sim SH, Lin D, et al. Genomic patterns of pathogen evolution revealed by comparison of Burkholderia pseudomallei, the causative agent of melioidosis, to avirulent Burkholderia thailanensis. BMC Microbiology. 2006; 6:46.

39. Picardeau M. Conjugative transfer between Escherichia coli and Leptospira spp. as a new genetic tool. Appl Environ Microbiol 2008; 74:319–22. https://doi.org/10.1128/AEM.02172-07 PMID: 17993560

40. Nieves C, Vincent AT, Zarantonelli L, Picardeau M, Veyrier FJ, Buschiazzo A. Horizontal transfer of therfbcluster inLeptospirais a genetic determinant of serovar identity. Life Science Alliance. 2023; 6(2). https://doi.org/10.26508/lsa.202201480 PMID: 36622346

41. Huang L, Zhu W, He P, Zhang Y, Zhuang X, Zhao G, et al. Re-characterization of an extrachromosomal circular plasmid in the pathogenic Leptospira interrogans serovar Lai strain 56601. Acta Biochim Biophys Sin (Shanghai). 2014; 46(7):605–11. Epub 20140529. https://doi.org/10.1093/abbs/gmu033 PMID: 24874103.

42. Bourhy P, Salaun L, Lajus A, Medigue C, Boursaux-Eude C, Picardeau M. A genomic island of the pathogen Leptospira interrogans serovar Lai can excise from its chromosome. Infect Immun. 2007; 75(2):677–83. Epub 20061121. https://doi.org/10.1128/IAI.01067-06 PMID: 17118975; PubMed Central PMCID: PMC1828511.

43. Schiettekatte O, Vincent AT, Malosse C, Lechat P, Chamot-Rooke J, Veyrier FJ, et al. Characterization of LE3 and LE4, the only lytic phages known to infect the spirochete Leptospira. Sci Rep. 2018; 8(1):11781. Epub 20180806. https://doi.org/10.1038/s41598-018-29983-6 PMID: 30082683; PubMed Central PMCID: PMC6078989.

44. Eshghi A, Gaultney RA, England P, Brûlé S, Miras I, Sato H, et al. An extracellular Leptospira interrogans leucine-rich repeat protein binds human E- and VE-cadherins. Cell Microbiol. 2019; 21(e12949). https://doi.org/10.1111/cmi.12949 PMID: 30171791

45. Daroz BB, Fernandes LGV, Cavenague MF, Kochi LT, Passalia FJ, Takahashi MB, et al. A Review on Host-Leptospira Interactions: What We Know and Future Expectations. Front Cell Infect Microbiol. 2021; 11:777709. https://doi.org/10.3389/fcimb.2021.777709 PMID: 34900757

46. Santecchia I, Bonhomme D, Papadopoulos S, Escoll P, Giraud-Gatineau A, Moya-Nilges M, et al. Alive Pathogenic and Saprophytic Leptospires Enter and Exit Human and Mouse Macrophages With No Intracellular Replication. Front Cell Infect Microbiol. 2022; 2:936931. https://doi.org/10.3389/fcimb.2022.936931 PMID: 35899053

47. Giraud-Gatineau A, Ayachit G, Nieves C, Dagbo KC, Bourhy K, Pulido F, et al. Inter-species transcriptomic analysis reveals a constitutive adaptation against oxidative stress for the highly virulent Leptospira species. Mol Biol Evol. 2024; 41(4):msae066. https://doi.org/10.1093/molbev/msae066

48. Ellinghausen HC, McCullough WG. Nutrition of Leptospira pomona and growth of 13 other serotypes: fractionation of oleic albumin complex and a medium of bovine albumin and polysorbate 80. Am J Vet Res. 1965; 26:45–51.

49. Johnson RC, Harris VG. Differentiation of pathogenic and saprophytic leptospires. J Bacteriol. 1967; 94:27–31.

50. Veyrier F, Pletzer D, Turenne C, Behr MA. Phylogenetic detection of horizontal gene transfer during the step-wise genesis of Mycobacterium tuberculosis. BMC Evol Biol. 2009; 9:196. https://doi.org/10.1186/1471-2148-9-196 PMID: 19664275

51. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, et al. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Molecular Biology and Evolution. 2017; 34:2115–22. https://doi.org/10.1093/molbev/msx148 PMID: 28460117

52. Syberg-Olsen MJ, Garber AI, Keeling PJ, McCutcheon JP, Husnik F. Pseudofinder: Detection of Pseudogenes in Prokaryotic Genomes. Molecular Biology and Evolution. 2022; 39. https://doi.org/10.1093/molbev/msac153 PMID: 35801562

53. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nature Methods. 2015; 12:59–60. https://doi.org/10.1038/nmeth.3176 PMID: 25402007

54. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis Bioinformatics. 2015; 31:3691–83.

55. Nguyen LT, Schmidt HA, vonHaeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015; 32:268–74. https://doi.org/10.1093/molbev/msu300 PMID: 25371430

56. Hoang DT, Chernomor O, vonHaeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol. 2018; 35:518–22. https://doi.org/10.1093/molbev/msx281 PMID: 29077904

57. Gaultney RA, Vincent AT, Lorioux C, Coppée JY, Sismeiro O, Varet V, et al. 4-methylcytosine DNA modification is critical for global epigenetic regulation and virulence in the human pathogen Leptospira interrogans. Nucleic Acids Res. 2020; 48(21):12102–12115. https://doi.org/10.1093/nar/gkaa966 PMID: 33301041

58. Fernandes LGV, Guaman LP, Vasconcellos S, Heinemann MB, Picardeau M, Nascimento ALTO. Gene silencing based on RNA-guided catalytically inactive Cas9 (dCas9): a new tool for genetic engineering in Leptospira. Sci Rep. 2019; 9:1839.