Ureaplasma diversum Genome Provides New Insights about the Interaction of the Surface Molecules of This Bacterium with the Host

Lucas M. Marques  
_Institute of Biomedical Science, University of São Paulo_

Izadora S. Rezende  
_Institute of Biomedical Science, University of São Paulo_

Maysa S. Barbosa  
_Institute of Biomedical Science, University of São Paulo_

Ana M. Guimarães  
_Department of Animal Health and Preventive Veterinary Medicine, College of Veterinary Medicine and Animal Science, University of São Paulo_

Hellen B. Martins  
_Multidisciplinary Institute of Health, Universidade Federal da Bahia, Vitória da Conquista, Brazil, University of Santa Cruz (UESC)_

See next page for additional authors

Follow this and additional works at: https://docs.lib.purdue.edu/cpbpubs

Part of the Medicine and Health Sciences Commons

Recommended Citation

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
RESEARCH ARTICLE

Ureaplasma diversum Genome Provides New Insights about the Interaction of the Surface Molecules of This Bacterium with the Host

Lucas M. Marques1,2,3*, Izadora S. Rezende1,2, Maysa S. Barbosa1,2, Ana M. S. Guimarães4,5, Hellen B. Martins2,3, Guilherme B. Campos1, Naíla C. do Nascimento5, Andrea P. dos Santos5, Aline T. Amorim1, Verena M. Santos1, Sávio T. Farias6, Fernanda Â. C. Barrence7, Lauro M. de Souza8, Melissa Buzinhani1, Victor E. Arana-Chavez7,9, Maria E. Zenteno10, Gustavo P. Amarante-Mendes10, Joanne B. Messick5, Jorge Timenetsky1

1 Department of Microbiology, Institute of Biomedical Science, University of São Paulo, São Paulo, Brazil, 2 Multidisciplinary Institute of Health, Universidade Federal da Bahia, Vitória da Conquista, Brazil, 3 University of Santa Cruz (UESC), Campus Soane Nazaré de Andrade, Ilhéus, Brazil, 4 Department of Animal Health and Preventive Veterinary Medicine, College of Veterinary Medicine and Animal Health and Preventive Veterinary Medicine, College of Veterinary Medicine and Animal Science, University of São Paulo, São Paulo, Brazil, 5 Department of Comparative Pathobiology, Purdue University, West Lafayette, Indiana, United States of America, 6 Evolutionary Genetics Laboratory, Department of Molecular Biology, Federal University of Paraíba, João Pessoa, Paraíba, Brazil, 7 Laboratory of Biology of Mineralized Tissues, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil, 8 Instituto de Pesquisa Pelé Pequeno Príncipe—Faculdades Pequeno Príncipe, Curitiba, Brazil, 9 Department of Dental Materials, School of Dentistry, University of São Paulo, São Paulo, Brazil, 10 Laboratory of Cellular and Molecular Biology, Institute of Biomedical Science, University of São Paulo, São Paulo, Brazil

* lucasm@ufba.br

Abstract

Whole genome sequencing and analyses of Ureaplasma diversum ATCC 49782 was undertaken as a step towards understanding U. diversum biology and pathogenicity. The complete genome showed 973,501 bp in a single circular chromosome, with 28.2% of G+C content. A total of 782 coding DNA sequences (CDSs), and 6 rRNA and 32 tRNA genes were predicted and annotated. The metabolic pathways are identical to other human ureaplasmas, including the production of ATP via hydrolysis of the urea. Genes related to pathogenicity, such as urease, phospholipase, hemolysin, and a Mycoplasma Ig binding protein (MIB)—Mycoplasma Ig protease (MIP) system were identified. More interestingly, a large number of genes (n = 40) encoding surface molecules were annotated in the genome (lipoproteins, multiple-banded antigen like protein, membrane nuclelease lipoprotein and variable surface antigens lipoprotein). In addition, a gene encoding glycosyltransferase was also found. This enzyme has been associated with the production of capsule in mycoplasmas and ureaplasma. We then sought to detect the presence of a capsule in this organism. A polysaccharide capsule from 11 to 17 nm of U. diversum was observed through electron microscopy and using specific dyes. This structure contained arabinose, xylose, mannose, galactose and glucose. In order to understand the inflammatory response against these surface molecules, we evaluated the response of murine macrophages J774 against viable and non-viable U. diversum. As with viable bacteria, non-viable bacteria were capable of promoting a significant inflammatory response by activation of Toll like receptor 2 (TLR2),...
indicating that surface molecules are important for the activation of inflammatory response. Furthermore, a cascade of genes related to the inflammasome pathway of macrophages was also up-regulated during infection with viable organisms when compared to non-infected cells. In conclusion, *U. diversum* has a typical ureaplasma genome and metabolism, and its surface molecules, including the identified capsular material, represent major components of the organism immunopathogenesis.

**Introduction**

*Ureaplasma diversum* is a bovine ureaplasma that was first isolated in 1969. Initially, it was defined as a non-pathogenic species, but recently it has been shown to cause damage to bovine tissue cells and organs [1–9]. *U. diversum* is frequently found in the genital tract of cattle and is associated with major genital disorders in these animals [5, 10, 11]. Cows infected with *U. diversum* have shown infertility, placitis, fetal alveolitis, and abortion or birth of weak calves [7, 12, 13]. In bulls, *U. diversum* may cause low sperm motility, seminal vesiculitis, and epididymitis [4, 6, 9, 13]. However, despite the description of these possible causal associations, the relationship of *U. diversum* and reproductive disorders in bovine remains controversial, mainly because high rates of positive vaginal cultures were also detected in animals with normal reproductive rates [14].

*U. diversum* is a facultative intracellular microbe, i.e. can be detected inside cells or adhered to their surfaces [15]. Recently, we have shown that the invasion of HEp-2 cells by this organism may lead to apoptosis [1], but as this phenomenon varied overtime. Thus, it is believed that *U. diversum* exerts a temporal modulation of the host programmed cell death. Invasion of bovine spermatozooids by *U. diversum* has also been linked to low sperm viability, suggesting that *U. diversum* may contribute to the death of these cells [4]. *U. diversum* also was capable of inducing significant TNF-alpha production in the uterus of experimentally infected mice [16], which indicates that the presence of this microorganism in the reproductive tract of females may significantly alter the homeostasis of the uterus microenvironment. Nevertheless, the molecular mechanisms by which this organism exerts its virulence and pathogenicity on such cells and tissues are mostly unknown [1, 15, 17, 18]. Very little genetic information of this bacterium is currently available [19, 20]. Therefore, the whole genome sequencing of *U. diversum* was undertaken as the first step towards understanding the mechanisms by which this microorganism causes disease and establishes infection, as well as to gain new insights into the biochemical pathways.

**Results and Discussion**

**General genome features**

The general genome features of *U. diversum* ATCC 49782 are summarized in Table 1 and Fig 1. The complete genome contains 973,501 bp in a single circular chromosome, with a low G+C content of 28.2%. It uses the opal stop codon (UGA) for tryptophan. A total of 782 coding DNA sequences (CDS), and 6 rRNA and 32 tRNA genes were predicted and annotated. Four hundred and seventy CDSs (60.1%) have putative functions, while 272 CDSs (35.7%) encode for hypothetical proteins. Predicted CDSs are summarized by role in Table 2.

Comparisons among *U. diversum* ATCC 49782 and the human ureaplasma serovars (10 serovars of *U. urealyticum*, and 4 serovars of *U. parvum*) [21] indicate a larger genome size in *U.
diversum (U. diversum—0.97 Mbp; U. urealyticum—0.84–0.95 Mbp; U. parvum—0.75–0.78 Mbp) and a slight higher G+C content when compared to other ureaplasmas (U. diversum—28.2%; U. urealyticum—25–27%; U. parvum—25%). U. diversum showed 782 CDSs; but U. urealyticum and U. parvum have an average of 608 CDSs. The hypothetical CDSs in U. diversum are 279, in U. urealyticum the average is 230 and in U. parvum the average is 201. These genetic differences may reflect the host specificity of U. diversum.

Gene synteny and phylogeny

In the present study, we used the gene synteny to compare the genomes of U. diversum and human ureaplasmas. CDSs that are conserved among the three species do not have the same organization, suggesting significant genomic reorganization (Fig 2A).

Moreover, phylogenetic trees based on concatenated protein sequences were constructed (Fig 2B). This protein concatenation approach has been frequently shown to increase resolution and robustness of phylogenetic analyses of mycoplasma species [22]. The phylogenetic relationship between U. diversum and human ureaplasmas remained the same compared with data from other studies with 16S rRNA and 16S-23S rRNA intergenic spacer [19, 20]. Ureaplasmas continue to branch within Mycoplasma clades [22]; the genus name Ureaplasma speaks to the fact that these organisms metabolize urea, and was coined before the era of DNA sequence based taxonomy. Interestingly, the phylogeny of Ureaplasma species corresponds to the evolutionary tree of the host animal species, suggesting coevolution between ureaplasmas and animals [20]. Although determinants for host ranges and susceptibilities to mycoplasma infections are currently unknown, Ureaplasma species have defined host ranges consisting of specific animal species [19].

Metabolic pathways

Analyzing the predicted set of U. diversum CDSs, it was observed that the metabolic pathways are identical to that of other ureaplasmas (Fig 3). The main CDSs for the urease activity was
annotated in the genome of *U. diversum* accordingly: urease (three-subunit urease + accessory proteins), an ammonia/ammonium transporter, and a F$_0$F$_1$-ATPase. Ureaplasma generates 95% of its ATP through the hydrolysis of urea by urease [23, 24]. Hydrolysis of urea generates an electrochemical gradient through accumulation of intracellular ammonia/ammonium. The gradient fosters a chemiosmotic potential that generates ATP. However, as in human ureaplasmas [25], nickel and urea transporters were not identified. Urea transporters are considered rare among bacteria and have been described in only few urease-positive organisms [26]. It is believed that urea can simply diffuse across membranes into the cytoplasm, but whether or not
this is the mechanism used by *U. diversum* to acquire this metabolite from the environment is unknown.

The sources of the 5% of ureaplasma ATP production not from the urea hydrolysis are most probably obtained from substrate phosphorylation [25]. CDSs coding for enzymes of the Embden-Meyerhoff-Parnas (EMP, glycolysis) pathway were present, except for the glucose-6-phosphate isomerase, which catalyzes the step leading to fructose 6-phosphate production. The absence of the enzyme was also observed in human ureaplasmas [26, 27]. Thus, it is believed that an alternative enzyme may be acting for this pathway to be functional. The pyruvate metabolism in *U. diversum* is incomplete; the *U. diversum* genome does not contain orthologs for the pyruvate dehydrogenase complex. Therefore, the production of ATP from the oxidation of pyruvate to acetate is unlikely. Other mycoplasmas and ureaplasmas also do not contain this enzyme, but this *in vitro* activity has been described [27]. Moreover, the phosphotransferase system (PTS) in *U. diversum* is incomplete. Only two putative PTS components genes were annotated (a putative phosphotransferase enzyme IIB—gudiv_313; and a phosphotransferase enzyme family - gudiv_380). The PTS is found only in bacteria which catalyzes the transport and phosphorylation of numerous monosaccharides, disaccharides, amino sugars, polyols, and other sugar derivatives [28]. Thus, it is believed that the use of glucose may have been abolished in ureaplasmas. This can be an interesting evolutionary characteristic of genomic reduction. Glycolysis is likely used only to provide substrates for the PPP pathway and for the synthesis of glycerol.

As observed in other mycoplasma species as well as with *U. parvum* [25], the pentose phosphate pathway is incomplete, with the absence of CDSs encoding glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase. However, the pentose phosphate (PP) pathway in mycoplasmas is considered to be functional [29]. Therefore, this is most likely the case for ureaplasmas as well.
The Nicotinate/Nicotinamide metabolism appears to be incomplete but functional for NADPH generation. The presence of the enzyme NAD kinase was observed in *U. diversum*. This enzyme has been related to the interconversion of NAD+ and NADP+ in nicotinate/nicotinamide metabolism, which plays a critical role in maintaining the NADH/NADPH pool balance inside the bacterial cell [30].

As expected, CDSs for the *de novo* biosynthesis of purines or pyrimidines were not identified in *U. diversum* genome. We also did not identify the enzyme ribonucleoside diphosphate reductase responsible for the conversion of ribonucleosides to deoxyribonucleosides. Like *U. urealyticum*, *U. diversum* could import all its deoxyribonucleosides and/or deoxyribonucleoside precursors, or have a different mechanism for converting ribonucleosides to deoxyribonucleosides. Moreover, the absence of the enzyme thymidine phosphorylase suggests that thymidine is the precursor imported for dTTP production [30] or another enzyme, still unknown, could supply this function.

Mycoplasmas are thought to be completely incapable of fatty acid biosynthesis from acetyl-CoA, probably due to the loss of genetic material. Phospholipids, glycolipids and sterols are the three major lipid constituents of cell membranes. This pathway has been poorly described in mycoplasmas and not all enzymes are identified [29]. Ureaplasmas have been described in the past as capable of *de novo* synthesis of saturated and unsaturated fatty acids [31]; however, our metabolic network analyses identified a very limited number of known pathways related to fatty acid biosynthesis in *U. diversum*. Among the few identified pathways, metabolic reactions from glyceraldehyde-3-phosphate to cardiolipin are mostly preserved, but one enzyme is missing, the phosphatidyl glycerophosphatase. The role of this enzyme may be replaced by the
enzyme cardiolipin synthase, which can convert cytidine 5’ diphosphate diacylglycerol directly to cardiolipin in the presence of phosphatidylglycerol. Moreover, phospholipids are of exclusive bacterial origin, therefore excluding the possibility that it is provided by the serum lipids normally present in broth media, the only possibility is that it is synthesized [32].

Transporters

Because of their genetic limitations, ureaplasmas must import more nutrients for cell growth than do most other bacteria. A total of 37 transporter-related CDSs were identified in U. diversum (Fig 3). ABC systems-related sequences represent 83.78% (31/37) of all transporter-related CDSs. Nonetheless, mycoplasmas have fewer transporters than other bacteria, and it is believed that these transporters have broader substrate specificities to compensate this limitation [32]. In U. diversum, a number of transport systems that are likely essential could not be found, such as bases/nucleotides, nickel and urea transporters. Urea transporters in bacteria are relatively rare. There are only three bacterial transporter classes, the ABC urea transporters, the Yut transporter, and the UreI transporter [33]. No homologous gene for these proteins was found in the genome of U. diversum. It is likely that CDSs of yet unknown function or other transporters are participating in the transport of such molecules.

Virulence and pathogenicity mechanisms

To date, only few features have been associated with virulence of U. diversum, such as the activity of urease and phospholipase, and its ability to invade some eukaryotic cells [8].
analyzing the genome of *U. diversum*, it was possible to find CDSs that are likely related to bacterial pathogenicity and can expand our knowledge regarding the mechanisms by which *U. diversum* causes disease. CDSs of urease (gudiv_255, gudiv_254, gudiv_253, gudiv_252, gudiv_251, gudiv_250, and gudiv_249), phospholipase (gudiv_472), hemolysin (gudiv_91), and a Mycoplasma Ig binding protein (MIB) — Mycoplasma Ig protease (MIP) system (gudiv_161/gudiv_162 and gudiv_612/gudiv_613) were found and are described below in more detail. In addition, CDSs that are likely related to the production of a capsular structure (gudiv_216) and surface molecules were observed. A scheme of these virulence factors is found in Fig 4A.

**Urease.** In addition to the hydrolysis of urea, the enzyme urease can also be a virulence factor of ureaplasmas. The toxicity of urease is mediated by the amount of ammonium ions or free ammonia generated from urea [34]. The urease activity of human ureaplasmas produces ammonia, which can damage host tissues due to changes in the pH [35]. To date, there is no data in the literature on the pathogenic effects of the expression of the urease gene by *U. diversum*. Nevertheless, the analysis of the organization of the *U. diversum* urease gene cluster
revealed the same organization of the human ureaplasma urease gene clusters [36]. There are three CDSs encoding structural subunits, ureA (gudiv_255), ureB (gudiv_254), and ureC (gudiv_253). Furthermore, downstream of ureC, four CDSs coding for accessory proteins (ureE—gudiv_252, ureF—gudiv_251, ureG—gudiv_250, and ureD—gudiv_249) were found (Fig 4B). A high identity of these CDSs when compared to CDSs of human ureaplasmas was also observed (above 90%).

**Phospholipase.** An array of potent hydrolytic enzymes has been identified in mycoplasmas, including phospholipases, proteases and nucleases [37]. Phospholipase C, A1, and A2 (PLC, PLA1, PLA2) activities have been reported in human ureaplasmas, however, no genes showed significant similarity to known sequences of PLC, PLA1, or PLA2 [38]. Orthologs to these CDSs, however, were not found in the *U. diversum* genome. This is somewhat surprising, considering a recent report showing that *U. diversum* reference strains, including the ATCC 49782 sequenced herein, and clinical isolates have high phospholipase activity [8]. It has been speculated that this phospholipase activity could have helped the *in vitro* invasion of Hep-2 cells by *U. diversum* [8]. Furthermore, it has been shown that a similar phospholipase activity of *U. diversum* are able to interfere with prostaglandin E2 and prostaglandin F2a production by bovine endometrial cells and induce premature labor [18]. In *U. diversum*, genes encoding phospholipase D family protein (PLD) (gudiv_472) and triacylglycerol lipase (Lipase family, such as phospholipases) (gudiv_748) were identified. Therefore, in contrast to human ureaplasmas, the PLD is likely the true virulence factor of *U. diversum* linked to cell invasion and premature labor in intrauterine infection, suggesting the phospholipase activity detected in our previous study [8] was PLD rather than phospholipase C activity. Experimental studies should be conducted with PLD to confirm this hypothesis.

**Hemolysin.** Hemolysins cause lysis of erythrocytes by forming pores of varying diameters in the host cell membrane [39]. There is one CDS (gudiv_91) coding a putative hemolysin protein in the genome of *U. diversum*. The predicted protein sequence has 63.1% identity with the hemolysin found in human ureaplasmas. In *U. parvum* serovar 3, this hemolysin has been characterized as an α-hemolysin (hlyA) with both hemolysin and cytotoxic activities [25]. Other mycoplasmas also have hemolysins [40, 41], but no orthologs to these proteins were found in the *U. diversum* genome. It seems that the mycoplasma or ureaplasma hemolysins belong to a unique class of bacterial toxins; more studies are needed to better understand the role of this enzyme in the pathogenesis of bovine ureaplasmas.

**MIB-MIP system.** Recently, a new virulence factor has been described in *Mollicutes*: the MIB-MIP system [42]. MIB acts as a high-affinity IgG binding protein, whereas MIP specifically cleaves the IgG heavy chain at an unconventional site, located between the VH and CH3 domains, and inactivate this molecule. This MIB–MIP system is encoded by a pair of genes often found in several copies in a wide range of mycoplasmas that infect various hosts [42]. Accordingly, two copies of the MIB-MIP system (gudiv_161/gudiv_162 and gudiv_612/gudiv_613) were found in the *U. diversum* genome. However, experimental studies should be conducted to better understand the role of this enzyme in the pathogenesis of bovine ureaplasmas.

**Surface molecules.** Although surface molecules of mycoplasmas are thought to play a crucial role in interactions with their hosts, very few had their biochemical function defined [43]. CDSs linked to lipoproteins (n = 37), multiple banded antigen like protein (MBA-like) (n = 1), membrane nuclease lipoprotein (n = 1) and variable surface antigen lipoproteins (possible VsA) (n = 1) were found in the genome of *U. diversum* (Fig 4A). Possibly due to the absence of a cell wall, mycoplasmas posses a larger number of membrane bound lipoproteins when compared to other eubacteria [44]. These molecules can act as virulence factors and/or be targets of humoral immunity [45]. They serve as potent cytokine inducers for monocytes/macrophages
and have cytolytic activity [46]. This immunogenicity of lipoproteins is probably caused by their surface exposure and the presence of the amino-terminal lipoylated structure [47].

Among the detected lipoproteins, one CDS was identified as a membrane nuclease lipoprotein (gudiv_93). Membrane nuclease activity in mycoplasmas was first reported in _M. pulmonis_ [48]. Nuclease activity in _Mollicutes_ has been proposed as the mechanism by which these organisms acquire the precursors for nucleic acid production [49], but it can also function as a virulence factor. The characterization of nucleases in _Mollicutes_ reveals that these enzymes differ from each other [50], and have been implicated in host cytotoxicity. The membrane nuclease MGA_0676, for example, is a remarkable pathogenic compound for the colonization and infection persistence of _M. gallisepticum_ [51]. Additionally, _M. pneumoniae_ nuclease (Mpn133) acts as a virulence determinant by binding to and internalizing within human airway cells [50].

Another important lipoprotein that has been identified is a putative variable surface antigen lipoprotein (gudiv_179) with 36% identity to the VsaA protein of _M. pulmonis_. Antigenic variation of surface proteins is thought to be a survival strategy for many mycoplasmas [52]. _M. bovis_ possess a large family of variable surface lipoproteins designated as Vsps, which are encoded by 13 different genes. The Vsps undergo phase variation in the respiratory tract of infected calves, which limits the microorganism elimination [53]. The experiments with variable surface proteins support the hypothesis that the generation of antigenic variation in mycoplasmas is critical for the bacterial survival for extended periods within the host, but do not establish a link with disease severity [45].

Although with borderline similarity (33%) a multiple banded antigen like protein (MBA-like) was annotated in the genome of _U. diversum_ ATCC 49782 (gudiv_653) (Fig 4C). This protein showed the conserved domain of the Multiple Banded Antigen-MBA of human ureaplasmas. The MBA is a surface exposed lipoprotein that can undergo size and phase variation _in vitro_ and _in vivo_ [54, 55]. The MBA is predicted to be a major human ureaplasma virulence factor and is the predominant antigen recognized by sera during infections in humans [56]. Analyzing the MBA locus composition among different human ureaplasmas [21], we can verify that the MBA-like locus in _U. diversum_ features a similar organization with the MBA locus of _U. parvum_ serovar 14 (UPA14) (Fig 4C). In particular, this area has two inverted hypothetical proteins and a transposase recognition site. Furthermore, unlike other ureaplasmas, in UPA14 and _U. diversum_ loci, the DNA pol III alpha subunit is not located adjacent to the locus. More studies are needed to understand the biology and role of MBA-like in the pathogenicity of _U. diversum_. Interestingly, the MBA-like (gudiv_653) and the variable surface antigen lipoproteins (gudiv_179) are in the same paralogous gene family in _U. diversum_. These two proteins, added to fourteen other paralogous proteins (annotated as hypothetical proteins), may play a critical role in modulating the immune response against _U. diversum_ (S1 Table).

**Capsule.** In the analysis of the genome of _U. diversum_, a glycosyltransferase enzyme (gudiv_216) was found [57]. Glycosyltransferase has also been described in the genome of human ureaplasmas. This enzyme is known to connect lipids and sugars that constitute the capsule of certain mycoplasmas [58]. In _M. mycoides_ subsp. _mycoides_, the presence of glycosyltransferase is believed to be related to the attachment of galactan to the capsule [57]. Thus, we decided to evaluate if a capsular structure is also present in _U. diversum_. By using electron microscopy and the red ruthenium dye, a polysaccharide capsule extending from 11 to 17 nm outside of _U. diversum_ cells was observed for the first time (Fig 5A). In a previous study with _U. urealyticum_, using the same dye utilized herein, a capsular structure was also observed, indicating that this may be a common feature of _Ureaplasma_ species [13]. Increased virulence and adhesion to host cells are associated with the presence of a capsule in certain bacterial species [59]. For instance, the _M. gallisepticum_ and _M. hyopneumoniae_ capsules have been suggested...
as additional structures that assist the cytoadherence to host cells [60] in addition to the “bleb” (a mycoplasma structure containing adhesins and accessory proteins) [61]. In *M. mycoides* subsp. *mycoides* strains, the presence of a polysaccharide capsule causes host injury [62]. Other studies have correlated the presence of a capsule and the resistance of mycoplasma to phagocytic cells. In *M. dispar*, the capsule has been considered the most important structure to interfere the phagocytosis [63]. Therefore, the detection of a capsular structure in *U. diversum* opens venues for new studies related to the virulence and pathogenicity of this bacterium.

The chemical composition of the capsular material of *U. diversum* ATCC 49782 was subsequently evaluated. Arabinose, xylose, mannose, galactose and glucose were the main components detected in the capsular material of this bacterium (Fig 5B). The biochemical composition of capsular polymers of *Mollicutes* has been little studied. The capsules of *M.*
mycoides subsp. mycoides SC, M. mycoides subsp. capri and M. dispar are the only ones described in the literature and are mainly composed of 1,6-galactose, polyglucan, and galacturonic acid polymers, respectively [64]. Probably, the capsular biosynthesis of M. mycoides subsp. mycoides is related to a glucose-dependent synthetic pathway due to the presence of a UDP-glucose 4-epimerase (GalE) and a glucose-1-phosphate uridylyltransferase (GalU) that uses glucose-1-phosphate to generate UDP-glucose [57]. In our in silico analysis, we did not observe genes related to the biosynthesis of polysaccharides. However, the presence of two genes of a sugar transporter was observed (Ribose/galactose ABC transporter — gudiv_307 and gudiv_308). We believe it is likely that these sugars are captured by U. diversum and are incorporated using glycosyltransferase to form the capsule.

**Host immune response against surface molecules.** In order to better understand the immune response against U. diversum surface molecules, we infect murine macrophages J774 with viable and nonviable microorganisms. In the analysis of cytokine production, there was an increased production of pro-inflammatory cytokines (IL-1β, IL-6, TNF-α) in both conditions compared to the negative control (p < 0.05), but there was no statistical difference between the profile triggered by the viable and nonviable ureaplasm (Fig 6A). Therefore, it is suggested that surface molecules are the main determinants of immune system activation in U. diversum. Surface molecules are composed of bacterial pathogen-associated molecular patterns (PAMPs) that, when recognized by the innate immune system, trigger the production of pro-inflammatory cytokines from manifold cells, causing inflammation [65–67]. The ability of surface molecules of Mollicutes to induce the secretion of inflammatory cytokines such as IL-1β, TNF-α and IL-6 have been reported in studies with different species, e.g. U. diversum [5], U. urealyticum [68, 69], and M. fermentans [65].

The expression of Toll-like receptors (TLRs) and inflammasome pathway genes were also evaluated using murine macrophages J774 infected with viable U. diversum cells. It can be observed that U. diversum induced a higher gene expression of TLR2 when compared to the uninfected cells (Fig 6B) (p < 0.05). The gene expression of other TLRs are presented in S1 Fig. The downregulation of TLR5, TLR8, TLR9 and TLR11 was also observed. These findings are supported by the literature, which indicates that TLR2 is the main receptor for the immune response of Mollicutes [45]. However, a study of U. urealyticum infecting macrophages demonstrated the activation of TLR2 and TLR4 [70]. Moreover, the absence of TLR 2 was related to the not recognition of bacterial lipoproteins of M. fermentans [71, 72].

The results of the inflammasome expression assay demonstrate the up-regulation of genes related to chemokine binding domains, interleukin-6, inhibitors of kinases and B cells, NF-kB, and interferon regulatory prostaglandin endoperoxide (Fig 6C). In contrast, caspase 12, CD40 ligand, interferon gamma, interleukin 12A, and 33 families of ligand to NLR and tumor necrosis factor genes were downregulated. Several studies have demonstrated the importance of the inflammasomes in response to mycoplasma infection [73]; however, the literature does not define mechanisms involved in the inflammasome response against U. diversum. A study of the development of gastric tumors using monocytes exposed to M. hyorhinis confirmed a high production of IL-1β and IL-18-induced NLRP3-dependent mechanisms [74]. It has been shown that monocytes and macrophages infected with live or inactivated by heat Acholeplasma laidlawii have the capacity to promote the release of IL-1β by the presence of cytosolic ASC (apoptosis-associated speck-like protein containing a carboxy-terminal CARD), promoting inflammasome activation [75]. In the present study, the up-regulation of NLRPs was not observed. Therefore, more studies are necessary to elucidate the involvement of inflammasome activation in the immune response induced by U. diversum. This study is the first report of the activation mechanisms involved in the immune response against Ureaplasma diversum. These findings can contribute to the studies on macrophage activation for the regulated secretion of
Fig 6. Host immune response against surface molecules. (A) IL-1β, IL-6, TNF-α and IL-10 levels in supernatant of macrophage J774 culture after infection with viable and nonviable *U. diversum* ATCC 49782 strain compared to uninfected cells. Statistical significance (p<0.05) is represented by the asterisk (*) (non-parametric Mann-Whitney analysis—One-tailed test, GraphPad Prism® version 6.01). (B) Gene expression of Toll-like receptors 2 in macrophage J774 culture after infection with *U. diversum* ATCC 49782 strain compared to the uninfected cells. Statistical significance (p<0.05) is represented by the
cytokines to prevent tissue damage caused by an exacerbated immune response. The results described herein will also support further research to understand the possible effector mechanisms that lead to effective protection against *U. diversum* infection.

**Conclusion**

The whole genome sequence described herein represents a valuable new resource for the study of *U. diversum* on a genetic basis. Furthermore, extensive analyses of these data helped us to detect important biological features of this mollicute and its mechanisms of pathogenicity. The results of this study will help to better understand this microbe, and can direct future research on the pathogenicity of this bacterium in cattle. Moreover, it will also aid the development of new strategies for treatment, prevention and control of ureaplasma infections.

**Material and Methods**

**Bacterial strain culture conditions and DNA extraction**

*Ureaplasma diversum* ATCC 49782 was originally isolated from a cow with clinically acute granular vulvitis in 1978 [10]. The virulence of this strain has been confirmed by many studies [1–5, 7, 8, 76]. *U. diversum* ATCC 49782 was first cultured in 2 mL of ureaplasma medium (UB) at 37°C, followed by propagation in 3,000 mL of the same broth. At the logarithmic growth phase (based on colorimetric changes), the culture was centrifuged at 20,600 x g for 30 minutes at 25°C. The DNA was extracted using a PureLink™ Genomic DNA Mini Kit (Life Technologies, Brazil) following the manufacturer's instructions.

**DNA sequencing, sequence assembly, gap closure and validation**

The whole genome was sequenced from a paired-end library using Illumina HiSeq 2000 (Illumina, Inc., San Diego, CA) at the Purdue University Genomics Core Facility. Average reads of about 100 bases were assembled using ABySS 1.2.7. After assembly resulting from a 4,552X genome coverage, eight remaining gaps were closed using conventional PCR, followed by sequencing using Sanger method in both directions. The genome sequence of *U. diversum* strain ATCC 49782 has been deposited in the GenBank database under the accession no. CP009770.

**Identification of CDS and annotation**

First-pass annotation was achieved using the NCBI (National Center for Biotechnology Information) prokaryotic genome annotation pipeline (PGAP). An initial set of CDSs was identified using PRODIGAL [77] and analyzed using the annotation engine MANATEE from the Institute for Genomic Sciences (University of Maryland, School of Medicine, Baltimore, MD) [78]. Individual CDS were manually curated; evidence generated by the pipeline (including BER, HMMs, PROSITE matches, TMHMM, and SignalP) was used to infer annotations. CDSs with an HMM score below the trusted value, less than 40% identity, or only local similarities to known protein sequences were called hypothetical proteins. tRNAs were located using tRNA-scan-SE [79]. Lipoproteins and signal peptides were identified using LipoP and SignalP algorithms, respectively [80, 81], and paralogous gene families were recognized using...
BLASTCLUST analysis (National Center for Biotechnology Information, NCBI, Bethesda, MD) using 30% sequence identity and 70% covered length thresholds. Comparative analyses with other bacterial genomes were performed based on genome annotations deposited in the NCBI Genome database. Considering that the genome annotation of human ureaplasmas deposited in GenBank, in particular of *U. parvum* serovars 3, may be outdated (*U. parvum* serovars 3 genome was one of the first bacterial genomes to be sequenced and annotated), we used tblastn to search for possible misannotated sequences when an ortholog was not found. A diagram of the overall structure of genome was generated using a DNAPlotter version 1.4 from Artemis 12.0, Sanger Institute [82].

**Comparative and phylogenetic analyses**

Whole genome synteny comparisons among *U. diversum*, *U. urealyticum* and *U. parvum* were performed using Sybil [83]. In Sybil, orthologs (homologous protein sequences from different bacterial species) were used to define syntenic relationships between species/strains.

Phylogenetic analyses of the 41 mollicutes were also performed using a multiple sequence alignment of 32 concatenated protein sequences from each organism (S2 Table). These proteins were chosen based on previous reports of phylogenomic analyses in prokaryotes [84, 85] according to their presence in all selected species, absence of additional fused domains, no subjection to HGT, and completeness [88]. Following protein concatenation using the UNION tool from EMBOSS [86], the sequences were aligned using MAFFT version 7 [87]. The resulting alignment was employed to build a phylogenetic tree using the neighbor-joining method [88] and maximum likelihood, with 1,000 bootstrap replicates from MEGA 6 [89].

**Capsule detection**

The ureaplasma was cultured in 50 mL of UB media. Cells in the logarithmic growth phase were collected by centrifugation at 20,600 x g for 30 minutes, and the pellets were suspended in 3% of glutaraldehyde and 0.2% of ruthenium red in 0.1M-cacodylate buffer (pH 7.4). Ruthenium red is a polycationic dye that binds polysaccharides and is frequently used to detect bacterial capsules [59]. After five washes with 0.05 M cacodylate buffer, the ureaplasmas were post-fixed in 1% (w/v) osmium tetroxide, then dehydrated in a graded series of ethanol and embedded in Spurr resin (Electron Microscopy Sciences, Fort Washington, PA, USA). Ultrathin sections collected on 200-mesh copper grids were stained with uranyl acetate and lead citrate before examination in a JEOL 1010 Transmission Electron Microscope.

To extract and purify the capsular component, the bacterial strain was cultured in 2 liters of UB media and then centrifuged (20,600 xg for 50 minutes) to retain the pellet. The pellet was washed twice in PBS (137 mM NaCl; 2.7 mM KCl; 4.3 mM Na2HPO4; 1.47 mM KH2PO4) (to avoid contamination of the polysaccharides derived from the culture medium), and resuspended in 10 mL of the same buffer. The extraction of capsular polysaccharide continued by the phenol-chloroform method [64]. To assess the monosaccharide composition, each polysaccharide was hydrolyzed with 300 μL of 2 mol L⁻¹ of trifluoroacetic acid, at 100°C for 8 h. The samples were dried under nitrogen stream and then dissolved in water 300 μL and the pH was elevated to 8 by addition of NH4OH. The samples were reduced with NaBH₄ at room temperature for 4 hours, then the cationic resin was added to remove salts, the samples were dried under nitrogen stream and followed by dissolution/dryness process in methanol (500 μL, x 3) to remove the residual borate. The alditol products were acetylated by addition of 200 μL of acetic anhydride-pyridine (1:1, v/v), held overnight at room temperature. Methanol (200 μL) was added to stop reaction and the samples were dried under nitrogen stream. The alditol acetates were analyzed by gas chromatography coupled to mass spectrometry (GC-MS—Varian, model Saturn 2000) with Ion
Trap analyzer. The chromatography was developed in a DB-225-MS fused silica capillary column (30 m, 0.25 mm, i.d. and 0.25 m of film thickness). The temperatures were: injector 250°C and the oven programmed from 50°C to 220°C at 40°C min⁻¹. The monosaccharides (as alditol acetates) were identified on the basis of their mass spectra at electron ionization (EI, 70eV) and authentic standards (Sigma-Aldrich).

Macrophage interaction

Inoculations in murine macrophages J774 (ATCC TIB-67™) were performed with viable bacterial strain and bacterial strain inactivated by heat (100°C for 30 minutes) for 24 hours (MOI 1:100). The inactivation was confirmed by absence of a positive culture on ureaplasma medium (UB). Negative controls without bacteria inoculation were also used. After 24 hours, the supernatant of the bottles was collected and frozen at -80°C to perform the ELISA. The cells were washed with trypsin and the cell suspension was placed in microtubes with RNA later for RNA extraction procedures.

The dosage of the cytokines TNF-α, IL-1β, IL-6, IL-10 was set using Ready-SET-GO enzyme-linked immunosorbent assay kit (eBioscience, San Diego, USA). The mRNA from cells was extracted using TRIzol Plus RNA purification kit (Invitrogen, USA), following the protocol supplied by the manufacturer. The cDNA was obtained using a retro-transcription (RT) from the mRNA using the SuperScript III Reverse Transcriptase kit. cDNA was used in a qPCR reaction to determine gene expression of TLR's 1–9 and 11 according to a previously described protocol [90]. The RT-qPCR was run three independent times with at least five samples per group. Analysis of relative gene expression data was performed using the 2(-ΔΔC(T)) Method [91]. GAPDH served as housekeeping gene to assess the overall cDNA content.

Gene expression of the inflammasome pathway was verified by quantitative PCR (qPCR) methodology. The cDNA obtained was subjected to analysis using Mouse Inflammasomes RT² Profiler™ qPCR Array kit (Qiagen-SABioscience, Brazil) for the expression of 84 key genes involved in the function of inflammasomes, protein complexes involved in innate immunity, as well as general NOD-like receptor (NLR) signaling (S3 Table). This array includes genes encoding inflammasome components as well as genes involved in downstream signaling and inhibition of the inflammasome function. In addition, this array includes other NLR family members, which may potentially form additional inflammasomes and their downstream signaling genes. All procedures were performed according to the manufacturer’s instructions.

Supporting Information

S1 Fig. Gene expression of Toll-like receptors in macrophage J774 culture after infection with U. diversum ATCC 49782 strain compared to uninfected cells. Statistical significance (p<0.05) is represented by the asterisk (*) (non-parametric Mann-Whitney analysis—One-tailed test, GraphPad Prism® version 6.01).

S1 Table. Analyses of paralogous gene families in Ureaplasma diversum ATCC49782.

S2 Table. Cluster of orthologous groups (COG)/proteins used for phylogenetic analysis.

S3 Table. Genes evaluated using Mouse Inflammasomes RT² Profiler™ qPCR Array kit.

S1 Table. Analyses of paralogous gene families in Ureaplasma diversum ATCC49782.

S2 Table. Cluster of orthologous groups (COG)/proteins used for phylogenetic analysis.

S3 Table. Genes evaluated using Mouse Inflammasomes RT² Profiler™ qPCR Array kit.
Acknowledgments
We are extremely grateful to the Genomics Core Facility at Purdue University for constructing the library, running the HiSeq 2000 platform, and performing the ABySS assembly. We thank Aricelma P. França for invaluable technical assistance and AcademicEnglishSolutions.com for revising the English.

Author Contributions
Conceptualization: LMM AMSG JBM JT.
Data curation: LMM AMSG NCN APS STF VEAC GPAM JBM JT.
Formal analysis: LMM AMSG NCN APS STF VEAC GPAM JBM JT.
Funding acquisition: LMM JT.
Investigation: LMM AMSG ISR MSB HBM GBC APS NCN ATA VMS STF FACB LMS MB MEZ.
Methodology: LMM AMSG ISR MSB HBM GBC APS NCN ATA VMS STF FACB LMS MB MEZ.
Project administration: LMM JBM JT.
Resources: STF VEAC GPAM JBM JT.
Software: STF.
Supervision: JBM JT.
Validation: LMM AMSG NCN APS STF VEAC GPAM.
Visualization: LMM AMSG NCN APS STF VEAC GPAM JBM JT.
Writing – original draft: LMM AMSG NCN APS STF VEAC GPAM JBM JT.
Writing – review & editing: LMM AMSG JBM JT.

References


Almeida RA, Rosenbusch RF. Capsulelike surface material of Mycoplasma dispar induced by in vitro growth in culture with bovine cells is antigenically related to similar structures expressed in vivo. Infection and immunity. 1991; 59(9):3119–25. PMID: 1715319; PubMed Central PMCID:PMC258142.

Tajima M, Yagihashi T, Miki Y. Capsular material of Mycoplasma gallisepticum and its possible relevance to the pathogenic process. Infection and immunity. 1982; 36(2):830–3. PMID: 6177640; PubMed Central PMCID:PMC351303.


Garcia J, Lemercier B, Roman-Roman S, Rawadi G. A Mycoplasma fermentans-derived synthetic lipopeptide induces AP-1 and NF-kappaB activity and cytokine secretion in macrophages via the activation...
Ureaplasma diversum Genome Provides New Insights about the Interaction with the Host


