Ureaplasma diversum protein interaction networks: evidence of horizontal gene transfer and evolution of reduced genomes among the Mollicutes

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Microbiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjm-2018-0688.R3</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>09-Apr-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Silva, Joana; Universidade Federal da Paraiba, Biologia Molecular Marques, Lucas; Universidade Federal da Bahia Timenetsky, Jorge; Instituto de Ciências Biomédicas, Universidade de São Paulo, Microbiologia de Farias, Savio; Universidade Federal da Paraiba, Molecular Biology</td>
</tr>
<tr>
<td>Keyword:</td>
<td>protein interaction networks, horizontal gene transfer, Ureaplasma, Mollicutes</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue?:</td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>
Ureaplasma diversum protein interaction networks: evidence of horizontal gene transfer and evolution of reduced genomes among Mollicutes

Joana Kästle Silva¹, Lucas Miranda Marques²,³, Jorge Timenetsky³, Sávio Torres de Farias¹

¹ Department of Molecular Biology, Federal University of Paraíba, João Pessoa, Brazil
² Multidisciplinary Institute of Health, Universidade Federal da Bahia, Vitória da Conquista, Brazil
³ Department of Microbiology, Institute of Biomedical Science, University of São Paulo, São Paulo, Brazil

Correspondence and reprints: Savio Torres de Farias – stfarias@yahoo.com.br

Authors correspondence

Joana Kästle – joana_kaestle@hotmail.com
Lucas Miranda Marques – lmirandamarques@gmail.com
Jorge Timenetsky – jotii@usp.br
Abstract

Ureaplasma diversum is a member of the Mollicutes class and causes urogenital tract infection in cattle and small ruminants. Studies indicate that the process of horizontal gene transfer, the exchange of genetic material among different species, has a crucial role in mollicute evolution, affecting the group’s characteristic genomic reduction process and simplification of metabolic pathways. Using bioinformatics tools and the String database of known and predicted protein interactions, we constructed the protein-protein interaction network of U. diversum and compared it with the networks of other members of the Mollicutes class. We also investigated horizontal gene transfer events in sub-networks of interest involved in purine and pyrimidine metabolism and urease function, chosen due to their intrinsic importance for host colonization and virulence. We identified horizontal gene transfer events among Mollicutes and from Ureaplasma to S. aureus and Corynebacterium, bacterial groups that colonize the urogenital niche. The overall tendency of genome reduction and simplification in the Mollicutes echoes in their protein interaction networks, which tend to be more generalistic and less “selective.” Our data suggest that the process was permitted (or enabled) by an increase in host dependence and the available gene repertoire in the urogenital tract shared via horizontal gene transfer.

Keywords: protein interaction network - horizontal gene transfer – Ureaplasma - Mollicutes.
1. Introduction

*Mollicutes* are characterized by reduced genomes and as a result, by simplified metabolic pathways (Dordet-Frisoni et al. 2014). Lacking major genes required for survival and reproduction, these cell wall-deficient bacteria rely upon the host for metabolite demand. Studies with 16S rRNA (Woese et al. 1985) determined that the *Mollicutes* class evolved from a gram-positive bacterium with a low G+C content through a genomic reduction process. The genes "lost" during evolutionary history of the *Mollicutes* are responsible for amino acid, cofactors and fatty acid biosynthesis (Razin 1985; Maniloff 1996) and can be completely absent in some species of this taxa.

The host-dependent life-style of mycoplasmas and ureaplasmas reflects the characteristic simplification of metabolic pathways that accompanies the group’s evolutionary history (Glass et al. 2017). A continuous process of genome reduction, which left the *Mollicutes* with a natural minimal set of genes, causes such simplification, showcased in incomplete or even absent metabolic pathways. The adaptation to a highly specific and stable niche of multicellular hosts permits the intake of precursors, metabolites and even finished products through the cellular membrane, making it possible for ureaplasmas and mycoplasmas to conserve energy and gene repertoire in biosynthetic pathways. The metabolic processes and gene expression patterns in mycoplasmas are curiously similar to those proposed by studies that aim at constructing artificial, minimal cells: transcription, protein translation and cellular replication through DNA and RNA precursor intake takes place, but little else (Coyle et al. 2016).

An important incomplete pathway in *Ureaplasma* and *Mycoplasma* is the de novo synthesis of purines and pyrimidines (Bizarro and Schuck 2007). This is embedded in one interesting phenomenon observed in these bacterial groups: function of metabolic pathways is observed even without the corresponding enzymes and proteins being encoded by the genome (Marques et al. 2016). In addition to this, experimental assays have shown that *Mollicutes* are capable of presenting enzyme function without the specific enzyme being present in their genomic annotation (Arraes et al. 2007). This has raised questions about alternative pathway use and metabolite intake to supply the nutritional and metabolic demand of these bacteria. How ureaplasma species use alternative pathways in their metabolisms and supply the demand for proteins and enzymes for which corresponding genes are not present in their genomes remains an open question.
As it is the case of many mycoplasmas, *Ureaplasma diversum* is a substantial pathogen of the reproductive tract in cattle, especially affecting young females (Cardoso et al. 2000). Among the symptoms caused by *U. diversum* infection are granular vulvovaginitis, salpingitis, endometritis, abortion and neonatal death, in addition to infertility and seminal vesiculitis in males (Doig et al. 1979; Sanderson et al. 2000). These characteristics make the infection caused by *U. diversum* and other mycoplasmosis an important economic factor due to its impact on livestock development and production. *U. diversum* prevalence in cattle was reported in South America (Cardoso et al. 2000; Oliveira Filho et al. 2005), Australia (Argue et al. 2013), United States (Rae et al. 1993; Sanderson et al. 2000) and Kenya (Mulira et al. 1989), reflecting damage caused by this pathogen on an international scale. One of Ureaplasma’s distinctive virulence factors is urease, an enzyme involved in urea hydrolysis that supplies the energetic demand (Ligon and Kenny 1991). Urea hydrolysis generates ammonia, which in turn is responsible for tissue damage caused in the host due to pH changes (Kokkayil and Dhawan 2015). The urease cluster in *U. diversum* and human-colonizing ureaplasmas share the same organization (Neyrolles et al. 1996) and this similarity is conserved in codifying sequence-level coding regions (Marques et al. 2016).

To have a better understanding of protein function it is important to consider that proteins rarely execute their biological role alone, being, instead, connected in functional networks. These networks are constructed from different metabolic pathways and perform their actions collectively. It has been reported that the functions fulfilled by a protein are less due to their amino acid sequence and more to the variety of interactions a specific protein carries with other proteins used by the organism, suggesting that complex cellular processes are better understood via large-scale protein-protein interaction networks (PPI) (Raman 2010). By studying these protein interaction networks, it is possible to identify key-proteins that regulate diverse cellular mechanisms and metabolic pathways. PPI networks have been used for determining unknown protein function (Deng et al. 2003) and enabling a broad vision of cellular processes when analyzed on the genomic scale. Studying PPI networks more closely, it is possible to determine the proteins occupying central positions of metabolic pathways and how their presence in the genomic repertoire controls cell function (Wuchty 2014). It is also enables the determination of functional modules in metabolic pathways (Erten et al. 2009), pinpointing essential proteins for cell survival and their integration with others, which
can then be used to indicate metabolic pathway robustness and even antibiotic treatment targets.

In microbiology, PPI networks were used in studies regarding *Bacillus licheniformis* (Han et al. 2016), *Escherichia coli* and *Saccharomyces cerevisae* (Wuchty and Uetz 2014) and the human gastric pathogen *Helicobacter pylori* (Rain et al. 2001) as target species, but to our knowledge no PPI network study was performed with ureaplasmas until this moment. With access to the genomic annotation of *Ureaplasma diversum* (Marques et al. 2016), this became possible. The comparison of PPI networks of different species on a large scale can be used to observe evolutionary processes (Erten et al. 2009), (Consortium 2011). The evolution of the *Mollicute* minimal genome has been targeted in many studies (Dordet-Frisoni et al. 2014), (Juhas et al. 2011), (Delaye and Moya 2010), but none of them considered the characteristic gene loss of this group based on the complexity of their PPI networks. We aimed to test if the protein-protein interaction networks reflect the genome reduction process in *Ureaplasma* and *Mycoplasma*, forcing them to compensate for the gene loss during their evolutionary history.

2. Materials and Methods

2.1 Construction of PPI networks in *Ureaplasma diversum*

Protein sequences obtained from the genomic annotation of *U. diversum* (Marques et al. 2016) were submitted to String (Szklarczyk et al. 2014) to construct initial PPI networks. The networks were clustered according to cellular mechanisms and metabolic pathways of their respective proteins. Still using String, we observed the general information about the nature of the predicted interactions generated by the PPI database, assembled in categories ranging from curated experimental databases to predictions based on gene neighborhood, co-occurrence and homology. The networks were visualized in Cytoscape (Shannon et al. 2003) for analysis of individual characteristics of each network and their correlated parameters. The NetworkAnalyzer plugin (Assenov et al. 2007) was used for analyzing network topology parameters. The ClusterOne plugin was used for identification and visualization of interaction clusters inside the global network and to perform sub-network arrangements.

To ensure high confidence in the predicted interactions, a 0.4 confidence interaction score cut-off was used while assembling the predicted networks in the String database.
Predicted interactions with a confidence score lower than medium were excluded from further analysis. For network robustness assessment the NetworkRandomizer plugin (Tosadori et al. 2016) was used on Cytoscape to generate randomized networks whose centrality parameters were compared to those of the predicted *U. diversum* network for network validation.

**2.2 Construction of PPI networks of species used in network complexity comparison**

*Ureaplasma urealyticum* (txid565575), *Ureaplasma parvum* (txid273119), *Mycoplasma genitalium* (txid243273), *Mycoplasma pneumoniae* (txid722438), *Mycoplasma hominis* (txid347256), *Haloplasma contractile* (txid1033810), *Mycobacterium tuberculosis* (txid419947), *Corynebacterium urealyticum* (txid504474), *Corynebacterium glucurolyticum* (txid548477) and *Staphylococcus aureus* (txid93062) were used for network comparison. These species were selected due to their phylogenetic relationship both in and outside of the *Mollicutes*. The proteomes of the species used for comparison were taken from the curated UniProt (The UniProt Consortium, 2017) database. The proteins were also visualized on the Kyoto Encyclopedia of Genes and Genomes (Kanehisa and Goto, 2000) platform and assigned to their corresponding metabolic pathways. The proteomes were submitted to String for genomic PPI network prediction and observation of general characteristics. The networks were then submitted to Cytoscape for topological and conformational analysis. The ClusterOne plugin was used for identifying interaction clusters inside the global network and for sub-network arrangement.

**2.3 Comparison of PPI networks of *U. diversum* and reference species**

The PPI networks of the reference species *U. urealyticum, U. parvum, M. genitalium, M. pneumoniae, M. hominis, H. contractile, M. tuberculosis, C. urealyticum* and *S. aureus* were loaded on Cytoscape where the topological characteristics and conformation of the PPI networks were assessed and the comparative profile of the sub-networks between these species and *U. diversum* was established.

**2.4 Construction of Mollicute phylogeny according to sub-networks complexity and interactions**

The proteins composing PPI sub-networks of *U. diversum* and reference species of interest were organized on a multi-FASTA file and submitted to Mafft (Katoh and
Standley 2013) for global alignment and submitted to ProtEST (Cuff et al. 2000) for the generation of the best-fitted evolutionary models. A maximum likelihood analysis was performed using PhyML for tree inference on the T-REX webserver (Culman et al. 2009). Still using T-REX, we assessed horizontal gene transfer rates among the studied species regarding the sub-networks of metabolic and pathogenic interest, the sub-networks of interest for *U. diversum* regard purine and pyrimidine metabolism and urease function. The phylogenetic trees obtained by these two distinct methods were then compared to a 16S rRNA based tree for validation. The 16S rRNA sequences were obtained from the genomic annotation for *U. diversum* and from the NCBI database of nucleotide sequences (Coordinators 2016) for the remaining species. The best suited evolutionary model was calculated on Mega (Tamura et al. 2007) and the parameters obtained from this analysis were inserted on T-REX (Culman et al. 2009) for the 16S rRNA tree construction and visualization.

3. Results

3.1 Characteristics and general parameters of *U. diversum* PPI network

The PPI network of *U. diversum* was constructed using a medium confidence score cut-off of 0.4. From the 10,404 predicted interactions of the global network 20.71% (2,155 interactions) obtained a very high confidence score (>0.9), 35.86% (3,732 interactions) had a high confidence score (0.6-0.899) and 43% (4,517 interactions) had a medium confidence score (0.4-0.599).

The 301 nodes of the PPI network of *U. diversum* were categorized and colored according to protein function: proteins involved in gene expression are shown in blue; proteins involved inter-cellular processes, protein transport, protein fate and cellular envelope construction and maintenance shown in purple; proteins with regulatory function, proteins involved in the metabolism of cofactors and carriers, fatty acid and energy metabolism related functions are depicted in orange; the proteins and enzymes shown in yellow are related to purine and pyrimidine metabolism and the proteins involved with the intermediary metabolism and central metabolism of nitrogen (including the urease cluster) are illustrated in pink (Figure 1).
Connectivity distribution, also known as node degree ($k$), refers to the number of nodes to which each node in a determined network is connected. It can also be seen as a representation of the interconnections inside the network. The node degree of the PPI network of *U. diversum* follows a power law distribution, a characteristic behavior of biological networks (Figure 2A). When distribution of nodes inside a network follows a power law, the majority of the nodes interact with few proteins and a low number of central nodes (the so-called hub proteins) concentrate the bulk of the network’s connections (Raman 2010). In a biological context, this means that few proteins in *U. diversum* interact with many nodes (proteins), providing the foundation of cellular processes, while the majority of the proteins agglomerate in smaller groups compromising specific interactions. This behavior can also be observed in the remaining topological parameters. Closeness centrality increases according to the number of neighbors to which a node is connected (Figure 2B). Proteins with a high connection to their proteomic neighborhood will automatically have a higher closeness centrality than proteins with few interactions. Similarly, the average clustering coefficient also follows the number of neighboring nodes connected to a given node (Figure 2C). The shortest path length refers to the velocity in which the information flows inside a PPI network, representing the shortest path between two nodes. In the *U. diversum* PPI network, the most frequent shortest path length is two nodes, showing that the information flux is very quick inside the PPI network (Figure 2D). In a metabolic context, this means that the PPI network of *U. diversum* shows a tendency to a high connection among the integrating proteins, suggesting a rapid message flux in the metabolism (Raman 2010).

Network robustness can be assessed by comparing the conformation and centrality parameters of a PPI network to networks randomly generated by its data (Tosadori et al. 2016). This can also be used to validate biological networks as a representation of organic processes. The Cytoscape NetworkRandomizer plugin generates random networks using a simple shuffle algorithm in the studied network (Tosadori et al. 2016). To validate the PPI network of *U. diversum*, 10 random networks were generated using their interactions and nodes as base and comparison of the subsequent centrality measures were performed. Only the clustering coefficient and network centrality scores were used due to the rest of the parameters remaining the same since the random networks were constructed based on the predicted network data (Supplementary information – Table 1). The clustering coefficient was considerably lower in the random networks compared to the predicted PPI network of *U. diversum* (0.613 in *U. diversum* and an average of 0.230 in the random
networks). This translates into a lower tendency among the nodes of the random networks to form clusters. The network centrality scores of the random networks are also smaller, showing a decrease from 0.359 in *U. diversum* to an average of 0.074 in the random networks. These scores show that the random networks are poorly connected, which reinforces their classification as non-representative of a biological system and scores the PPI network of *U. diversum* as a robust organic network.

### 3.2 Central Proteins of the *U. diversum* PPI network

The two most used central parameters to determine the relevance of a node to its surrounding PPI network are 1) the node degree (*k*) which calculates the number of connections a given node has to other nodes in the network, being therefore a local centrality measurement, and 2) the node intersection, a parameter that represents the number of shortest paths of the network crossing it (Azevedo, Moreira-Filho, 2015), depicting the information flow inside the network. Organic networks such as PPI networks respond poorly to the removal of nodes with the high centrality measurements, often referred to as hub proteins, since much of the global network’s foundation relies on them and their role in “stitching” the neighborhood proteins together. Proteins and enzymes with critical function in specific metabolic pathways and cellular processes, such as urease and its role in supplying the energetic demand of ureaplasmas, are also prone to cause network collapse and cell death upon removal from the PPI network.

To identify protein hubs in the *U. diversum* PPI network, node degree (*k*) and node intersection of all nodes composing the network were accessed. To pinpoint the cellular processes in which they are involved we searched for experimental studies in the literature and in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa, GOTO, 2000) (Table 1).

### 3.3 Comparison of the global PPI network of *U. diversum* and correlated species

We aimed to determine whether the genomic reduction, simplification and their related generalization of the metabolic pathways in *Ureaplasma* (and *Mycoplasma*) are reflected in the conformation and general parameters of their PPI networks. To do so, we compared the PPI networks of these *Mollicutes* to the networks of distantly related species with a higher metabolism complexity, such as *S. aureus* and *Corynebacterium* species (Table 2).
The scores analyzed in this study show a tendency among the *Mollicutes* to interact more in clusters than phylogenetically distant species, which is shown by their higher clustering coefficient rates. The average clustering coefficient among the *Mollicutes* is 0.550, while *Corynebacterium* species show an average of 0.357 and the analyzed *S. aureus* populations, averaged 0.361. This can be due to the generalization of function in the remaining proteins left after the genomic reduction process in the *Mollicutes*, forcing them to interact at higher frequency rates. The studied *Mycoplasma* and *Ureaplasma* species also show a higher centrality index in their PPI networks and increased connectivity scores (depicted by their higher average number of neighbor scores). The genomic reduction and simplification of the metabolic networks in the *Mollicutes* class appears to be followed by a generalization process in protein function and interaction.

### 3.4 Horizontal Gene Transfer in Ureaplasma

Our horizontal gene transfer results based on the PPI network of *U. diversum* shows gene transfer events in the sub-networks related to urease function and purine and pyrimidine metabolism (Figure 3). The bioinformatic analysis indicates two events of horizontal gene transfer in the urease metabolic pathway (Figure 3A), depicting a flow of the urease operon between *Ureaplasma* and *Corynebacterium* and *Ureaplasma* and *S. aureus*. In the purine subnetwork, analysis showed the horizontal transfer between *Mycoplasma* and *Ureaplasma* (Figure 3B), two *Mollicutes* genera sharing the same ecological niche in their multicellular hosts. In the pyrimidine metabolism subnetworks, two events of horizontal gene transfer were indicated by our analysis (Figure 3C).

### 3.5 Horizontal Gene Transfer and network structure

The urease PPI network conformation is highly similar in the analyzed species, showing slight differences in the proteins that interact with the urease structural subunits and accessory proteins. All the investigated species present the structural urease subunits (*ureA, ureB, ureC*) as well as the accessory proteins (*ureE, ureF, ureG*) in their PPI network. This can be observed in the urease intersection network (Supplementary information – Figure 1). The confidence index of the interactions among the urease proteins was predicted as very high (>0.9) in the studied species, depicted by their colored edges. The comparison of the urease subnetwork in *U. diversum, U. urealyticum* and *U. parvum* (Supplementary information - Figure 2) illustrates a high similarity among the interacting partners, edges and positions inside the PPI network. The network
conformation is also similar, reinforced by proximal interaction parameters among the
ureaplasma species that differ in the proteins comprising the individual networks. Figure
4 shows the urease subnetwork of the species targeted in Ureaplasma horizontal gene
transfer events: U. diversum (A), C. glucurolyticum (B), C. urealyticum (C) and S. aureus
(D). The interactions of both networks are similar, differentiating in the proteins that
interact with the urease structural and accessory proteins. The comparison of these
networks and their associated topological parameters (Supplementary information - Table
2) suggest that the urease subnetwork is conserved structurally in these species.

Our analysis also showed a horizontal gene transfer event during the evolution of
Mycoplasma and Ureaplasma in proteins involved in purine metabolism, driving us to
focus our studies on their related subnetworks and corresponding network parameters. In
figure 5, the purine subnetworks of the species with suggested gene flow by lateral gene
transfer are shown, U. diversum (A), M. genitalium (B), M. pneumoniae (C) and M.
hominis (D). One of the differences in the PPI network composition in the studied
Mycoplasma species is the absence of the deoD gene in the subnetwork of M. hominis
(the only representative of the hominis group in this study), which occurs in the
subnetworks of M. genitalium and M. pneumoniae. The studied Ureaplasma species (U.
diversum, U. urealyticum, U. parvum) present similar purine metabolism PPI networks
in structure and composition (Supplementary information – Figure 3). Mycoplasma and
Ureaplasma share some of the proteins in their purine metabolism PPI network. As we
did in the analysis of the urease PPI networks, we also calculated the topological
parameters in the constructed purine metabolism subnetworks of the species involved in
the horizontal gene transfer events (Supplementary information - Table 3). Again, the
correlated network parameters remained very similar between the analyzed species.

The pyrimidine subnetwork analysis suggested two horizontal gene transfer events:
the first from Mycoplasma to Ureaplasma and the second from Mycoplasma to
Corynebacterium. In Figure 6, the pyrimidine subnetworks of the species involved in
horizontal gene transfer events are depicted.

The PPI networks are similar between these groups, which is reflected in their
corresponding network parameters (Supplementary information - Table 4). The centrality
measurements of the pyrimidine metabolism PPI networks in Mycoplasma, Ureaplasma
and Corynebacterium share similarities regarding clustering coefficient, network
centrality and characteristic path length parameters, which are almost identical in the
analyzed species (Supplementary information - Table 4). Similar to the purine
metabolism networks, the pyrimidine PPI networks of the studied species share protein composition and topological parameters.

The horizontal gene transfer events in the urease, purine and pyrimidine metabolism subnetworks appointed by our bioinformatic analysis is reflected in a highly similar PPI network composition and organization between the studied groups. The centrality parameters suggest a conservation in network topology taking place during Ureaplasma, Mycoplasma, Corynebacterium and S. aureus evolution. This can be due to the similar lifestyles shared by these urogenital tract colonizers.

4. Discussion

4.1 U. diversum reduced metabolism

Mollicutes are known for their reduced genomes and simplified metabolic pathways resulting from gene loss in the course of their evolutionary history. Gene loss is often regarded as a deleterious or detrimental process to the organism undergoing it or a neutral factor resulting simply from genetic drift, but this is not always the case. In fact, the loss of specific gene groups can prove to be a major evolutionary mechanism, driving species to colonize new environments, saving energy by ‘throwing out’ dispensable genes and adjusting the metabolism to accommodate niche adaptation. Spontaneous loss of genes was observed to increase bacterial fitness in different growth conditions in previous studies (Koskiniemi et al. 2012). Thus, gene loss might play a major role in the transition to new lifestyles such as host colonization and pathogenicity development in a similar way to the acquisition of specific genes through processes such as, for instance, horizontal gene transfer events (Albalat and Cañestro 2016). Both gene loss and acquisition act as valuable sources of genetic diversity, which in turn can contribute to adaptive phenotypic variety.

Regarding the gene losses of the mycoplasma group, some species lost almost all genes involved in specific biosynthetic pathways. This distinctive characteristic also applies to M. pneumoniae and M. genitalium. Both lost all genes of the amino acid production pathway, relying upon the host or environment for meeting their amino acid needs (Razin 2006). The group also lost the majority of genes compromised with fatty acid synthesis and cofactor production (Weisburg et al. 1989). Comparatively, U.
*U. diversum* also has an incomplete fatty acid metabolism and the majority of *U. diversum* annotated proteins are involved in gene expression and regulation.

The genomic reduction in *U. diversum* can be easily visualized in the number of nodes observed in the global PPI network carrying all of its known proteins. The process of genomic reduction in *U. diversum* was accompanied by a rearrangement in its protein interaction networks. Our data suggests an increase in connectivity between the proteins left behind after the progressive gene loss process during the evolutionary history of this bacterium. The indicated connectivity increase can be related to the maintenance of essential metabolic pathways where remaining proteins had to partially or completely assume the function executed by lost proteins during the genomic reduction process. Our data indicates a mechanism of adaptation to genomic reduction by rearrangement of protein-protein interaction networks.

### 4.2 Central Proteins of the *U. diversum* PPI network: signs of horizontal gene transfer importance for minimal genomes

It is interesting to note that all proteins with the highest connections in the global network of *U. diversum* perform functions in DNA replication and transcription, contributing to the flow of genetic information inside and outside of the bacterial cell. Among these, RecA (recombinase A), dnaK and a DNA helicase (UUR10_0577 also known as PcrA) are also involved in stress response pathways, activation of bacterial S.O.S response, homologous recombination and horizontal gene transfer.

The role of RecA, a DNA-dependent ATPase and recombinase, in homologous recombination has positioned it among the main factors involved in the capture of exogenous genetic material, especially in the processes of conjugation and transformation, known mechanisms of bacterial horizontal gene transfer (Cox, 1991). The process of DNA transfer mediated by RecA requires the participation of several proteins, about 40 in *Bacillus subtilis* (Saito; Akamatsu, 2006), which may explain the high connectivity of this node in the PPI network of *U. diversum*. In *B. subtilis* and *E. coli*, RecA interacts with exogenous single-stranded DNA (ssDNA), forming nucleofilaments capable of recognizing recombination sites in the bacterial chromosome, where it promotes the exchange of genetic information (Prentiss et al., 2015). This process has been observed in a variety of bacterial species, including *Neisseria gonorrhoeae*, *Hemophillus influenzae*, *Streptococcus pneumoniae*, *Streptococcus mutans* and...
Helicobacter pylori, suggesting that the mechanism for incorporation of exogenous DNA acquired through RecA-mediated horizontal gene transfer is of great importance for introducing genetic diversity in these bacteria (Michod et al., 2008). Ishag et al. (2017) experimentally determined that RecA activity increases homologous recombination rates in Mycoplasma hyorhinis and considering the high connectivity of RecA in the PPI network of U. diversum, which may also be the case in ureaplasmas and other mycoplasmas.

RecA is a highly conserved protein in biological systems with widespread functional homologs among eubacteria, archaebacteria and eukaryotes (Brendel et al. 1997), illustrating the crucial role it plays in cell-life maintenance. In eubacteria, invariant RecA segments compromise ATP, DNA and LexA-binding sites, where a high amino acid residue conservation assures proper RecA function and regulation (Karlin and Brocchieri 1996). Recent studies have expanded known RecA functions in an array of cellular processes, pointing out its role in signaling pathways leading to programmed cell death (PCE) in bacteria (Peeters and de Jonge 2017) and the underlying molecular mechanisms involved in its contribution to genomic integrity (Bell and Kowalczykowski 2016).

PcrA is a helicase responsible for relieving tension in the DNA strands generated by replication and transcription processes. Studies indicate that helicases, including PcrA, regulate the rate of genetic recombination in B. subtilis (Ehrlich; Petit, 2002) promoted by proteins such as RecA. A similar regulatory system may take place in U. diversum, with the RecA-mediated homologous recombination activity being under the control of PcrA.

4.3 Horizontal gene transfer in Ureaplasma

Horizontal gene transfer has been discussed as a possible explanation for the distinct metabolic and phylogenetic characteristics among the Mollicutes. Jain et al. (1999) reported that the abundant frequency of horizontal transfer events taking place between prokaryotes occur at higher rates among operational (housekeeping) genes than genes involved in genome expression. The degenerative evolution mechanism of the Mollicutes does not contribute to addition of new genes (Cordova et al. 2016), suggesting that horizontal gene transfer may play an important part in gene acquisition and genomic diversity in this group. Xiao et al. (2011) observed horizontal gene transfer among
different serotypes of human ureaplasmas. García-Castillo et al. (2008) showcased the horizontal acquisition of biofilm formation genes by ureaplasma strains from other bacterial groups and previous studies reported horizontal gene transfer occurring in species occupying the same ecological niche (Sirand-Pugnet et al. 2007). We also know that M. pulmonis strains share genes horizontally (Teachman et al. 2002). Horizontal gene transfer also seems to be more prevalent in some Mollicutes than other bacteria taxa. As an example, 18% of the M. agalactiae genome was acquired by this mechanism from the Mycoides group (Dordet-Frisoni et al. 2014). Even though the exact mechanisms by which gene transfer among Mollicutes occur are still unclear, this process appears to not only guarantee gene acquisition and exchange among Mollicutes (Cordova et al. 2016) but also upgrades the metabolic machinery of Mollicute species for an increased survivability against host defense mechanisms and could also enable the colonization of new hosts (Sirand-Pugnet et al. 2007). During the analysis and interpretation of horizontal gene transfer events, the observation of the niches occupied by the species taking part in it is an important tool revealing the ecological contextualization of the obtained data. Species that occupy the same niche, such as an abiotic environment or a multicellular host, are in direct contact and therefore contribute to the local gene flow and genetic repertoire.

4.4 Urease cluster horizontal gene transfer events

The urease cluster in the PPI network of U. diversum is among the sub-networks with the highest confidence scores detected in this study. Analyses have shown that urea hydrolysis is responsible for generating the majority (95%) of ATP demand in Ureaplasmas (Smith et al. 1993; Neyrolles et al. 1996). No urea and nickel transporters were found in U. diversum genome annotation (Marques et al. 2016) and the same occurs in other Ureaplasmas (Pollack 2001). Urea could be acquired directly from the environment by diffusion across the cell membrane, but it is still not clear if this process is responsible for urea intake in U. diversum.

The frequency and impact of horizontal gene transfer are high among Corynebacterium, especially regarding pathogenicity and virulence factors (Oliveira et al. 2017). Urease can supply the energetic demand of an organism through urea hydrolysis, but it is also an important virulence factor of pathogenic bacteria. Horizontal gene transfer has already been detected in Corynebacterium species occurring among classic pathogenicity islands, such as adhesins, secreted toxins and proteins involved in
iron uptake from the host (Ruiz et al. 2011). The Corynebacterium species analyzed in this study were C. glucuronolyticum and C. urealyticum. The 16S rRNA trees generated as a control can be visualized in Figure 5 of the Supplementary information.

*Corynebacterium glucuronolyticum* is an opportunistic pathogen known for causing urogenital infections in humans under certain clinical conditions (Gherardi et al. 2015) and it is also the causative agent of urologic infections and reproductive disorders among swine (Devriese et al. 2000) and ovine (Takahashi et al. 1997). *Corynebacterium urealyticum* is a common isolated species from urological infection samples (Salem et al. 2015), and studies suggest that co-infection by *Ureaplasma urealyticum* can occur in these cases (Trinchieri 2014), aggravating the infection and promoting kidney stone formation.

*S. aureus* is one of the most studied human pathogens and is the causative agent of many infections of clinical importance (Tong et al. 2015). *S. aureus*, as well as *Corynebacterium* and *Ureaplasma*, colonizes the urogenital tract niche. Differing from *Ureaplasma* and *Corynebacterium* infections, the urogenital tract diseases caused by *S. aureus* are recurrent in immunosuppressed hospital patients with urinary catheters, which facilitate bacterial colonization and dissemination. Classic studies show that *Staphylococcus* urease has similar biochemical properties to those present in other bacterial groups and there is topological conservation in the urease operon among urease producing bacteria (Gatermann and Marre 1989).

*S. aureus* infections in cattle are of economic relevance since they affect product quality (Guinane et al. 2011). The fact that specific *S. aureus* clones are capable of infecting animal hosts such as poultry and cattle, but are not isolated from human samples, suggests a necessary adaptation to new niches enabled by the acquisition of crucial genes in these bacterial populations (Guinane et al. 2011). Specific genes have been discovered as being transferred horizontally from human-colonizing *S. aureus* populations and regarded as facilitators in the colonization of poultry (Lowder et al. 2009) and cattle hosts (Herron-Olson et al. 2007). Guinane et al. (2011) using population genetics, comparative genomics and a functional analytical approach estimated that the horizontal gene transfer event (or events) enabling *S. aureus* populations to colonize domestic animals such as poultry and small and big ruminants occurred 11 thousand years ago, when domestication of these animals was in its early stages (Zeder 2008). The colonizing “jump” made by specific *S. aureus* populations (later referred to as serovars) to ruminants and poultry continued in the adaptation process enabled by allelic diversification, genetic loss and
mobile genetic elements acquisition (Guinane et al. 2011). The latter can compromise 15
to 20% of the S. aureus genome (Lindsay 2014). Compared to human-colonizing S.
aureus, the serovars infecting poultry and ruminants present an increase and diversification in virulence factors, adhesins and metabolic machinery that ensure survivability in these new hosts. These data suggest that during evolution, S. aureus had to adapt to new hosts, which is also true for Ureaplasma serovars that exclusively colonize domestic animals.

In U. diversum, poorly characterized hypothetical proteins occupy the urease network (UUR10_0212, UUR10_0168 and purines UUR10_0577) (Figure 4), while in U. urealyticum and U. parvum the same spots are occupied (including the interactions with urease accessory proteins) by DNA directed RNA polymerase (rpoB), fructose biphosphate aldolase (fba), acetate kinase (ackA) (Supplementary information – Figure 2), suggesting that the functions of the hypothetical proteins in U. diversum may be similar to those performed by these proteins. When placed in the same network, they continue to interact with the remaining ureaplasmal protein repertoire, indicating that the biological function configured by the interactions between these proteins is similar. C. urealyticum have DNA polymerase in their network (cu1143), a gmp synthase (guaA) and a carboxylate dehydrogenase (cu1505), while these locations in C. glucurolyticum are occupied by hypothetical proteins and gatA, a transcriptional factor. As in U. diversum, it may be possible that the hypothetical proteins in C. glucurolyticum have functions similar to the characterized proteins in the C. urealyticum PPI network.

Mycoplasmas and ureaplasmas are known for low genomic GC content compared to more distantly related bacterial groups. Relying on the host for intake of metabolites and precursors, these bacteria were able to eliminate genes related to biosynthetic pathways mainly through genetic drift, a process facilitated by restriction to specific hosts and consequent lack of recombination between different strains (McCutcheon and Moran 2012). The genomic GC content of different bacterial species has been linked to lineage-specific mutational patterns and to selection of diverse genome-wide properties (Sueoka 1962; Muto and Osawa 1987; Rocha and Feil 2010). The high AT bias found in the Mollicutes also characterizes other minimal and reduced genomes. As discussed by McCutcheon and Moran (2012), the AT bias in small genomes can be explained by the characteristic loss of DNA repair genes in these groups, leading to an accumulation of AT mutations. This is the case, for example, of cytosine deamination and guanine oxidation, leading to (G or C)→(A or T) changes in the genome, especially in symbiotic
bacteria (Lind and Andersson 2008). The accumulation of deleterious substitutions can increase the rate of genomic evolution at nucleotide and amino acid levels (McCutcheon and Moran 2012).

During evolution, changes in the so-called “universal genetic code” are more common than initially assumed, leading to coding divergence among organisms (Knight et al. 2001). Amid these changes, the reassignment of UGA from a stop codon to a tryptophan codon in bacterial lineages such as the *Mollicutes* is one of the most common (Maniloff 2002). This coding property could be a significant obstacle for horizontal gene transfer events in which mollicute species act as donors, since a transferred gene could lead to a premature translation stop in organisms with universal codon usage. On the other hand, transfer of mollicute genes to more distantly related species with higher GC content is expected to leave a low GC “fingerprint” behind. In our study, this was observed in the urease cluster, determined by our analysis as being transferred from *Ureaplasma* species to *S. aureus*, *C. urealyticum* and *C. glucuronolyticum*. In table 3, the average GC content of the transferred genes among the bacterial species and the overall mean genomic GC percentage are listed. As can be seen, while *S. aureus* GC content does not change significantly between the urease cluster and the whole genome, this is not the case for the *Corynebacterium* species analyzed in our study (Riegel et al. 1995). Both show a decrease in GC content in the urease cluster related to the genome wide percentage, reinforcing the genetic transfer taking place between these groups.

4.5 Purine and Pyrimidine metabolism horizontal gene transfer events

*Ureaplasma*, as well as the majority of mycoplasmas, is unable to realize *de novo* biosynthesis of purines and pyrimidines. To compensate for the lack of metabolic genes, *Ureaplasma* relies on the host to import a high number of precursors and even finished metabolites. In the PPI network of *U. diversum*, 79 proteins are involved in transport and binding processes and the purine and pyrimidine pathway is incomplete (Marques et al. 2016). Studies have shown that incorporation of nucleotides, nucleosides and free bases can occur in many prokaryotic organisms (Pandey 1984; Fraser et al. 1995). The demand of purines and pyrimidines can be supplied by intracellular nucleotides in the host cells, and Razin and Avigan (1963) observed that mollicute strains use the host tissues and medium as nutritional source of nucleotide precursors (Razin 1985). This is most probably also the case in *U. diversum*. 

https://mc06.manuscriptcentral.com/cjm-pubs
In species such as the mammalian intestinal tract parasite Cryptosporidium parvum, alternative enzymes acquired through horizontal gene transfer events assist in purine and pyrimidine synthesis using precursors imported from the host (Striepen et al. 2004). U. diversum has three of these alternative enzymes that assist in nucleotide salvage: uridine kinase, uracil phosphoribosyltransferase and thymidine kinase. In C. parvum thymidine kinase was acquired through horizontal gene transfer from a proteobacteria (Striepen et al. 2004).

Mollicutes rely upon intake of pre-existing nucleosides and bases from the host and environment, using either nucleobases or nucleosides to synthesize required nucleoside triphosphates and deoxynucleoside triphosphates (Arraes et al. 2007). Purine nucleoside phosphorylase is an enzyme important to both purine and pyrimidine metabolism and has a nucleoside ribosyltransferase function (Mombach et al. 2006). M. genitalium and M. pneumoniae do not have nucleobase or nucleoside transporters (Paulsen et al. 2000) in their genomes. Other transporters with a broad variety of substrates identified in both species could be active in the transport of nucleotide precursor metabolites, such as the family of ATP-binding cassette (ABC) proteins and the Major Facilitator Superfamily (MSF) primary active transporter. Minion et al. (1993) tested 20 mycoplasma species for external membrane-associated nuclease activity, which could be critical for nucleotide precursor transport across bacterial membranes and found these reactions to take place. A few years later, mnuA, a membrane nuclease gene, was first cloned and isolated in M. pulmonis (Jarvill-Taylor et al. 1999). Membrane nucleases can be involved in virulence, but also contribute to the degradation of host RNA and DNA, which could also be a source of free bases and nucleotide precursors for Mycoplasma species (Razin et al. 1998).

In the pyrimidine pathway, uracil phosphoribosyltransferase is responsible for the phosphorylation of uracil or cytidine (Tu and Turnbough 1997; Lundegaard and Jensen 1999) and is present in the genome annotation of U. diversum. Uridylate kinase catalyzes the phosphorylation of UMP (Schultz et al. 1997) and, since cytidylate kinase is absent in the genome annotation of U. diversum, it should be the protein responsible for UMP phosphorylation (Bucurenci et al. 1996). Ribonucleotide reductase is a ubiquitous enzyme responsible for deoxyribonucleotide synthesis and as expected, is also present in U. diversum. Reichard (2002) showed that this single protein is capable of reducing all four common ribonucleotides. Thymidylate kinase is part of the nucleoside monophosphate kinase super family and is responsible for the reaction of thymidine monophosphate phosphorylation, using ATP as its preferred phosphoribosyl donor. As in other organisms,
the regulation of thymidylate kinase in *U. diversum* is essential due to its participation in the overall control of DNA synthesis adjustments in the dTTP pool.

In the purine metabolic enzymes, *U. diversum* relies on guanylate kinase in its metabolic machinery. Guanylate kinase is responsible for the reversible phosphoryl transfer form ATP to GMP (Li et al. 1996). The genomic annotation of *U. diversum* showed the absence of thymidine phosphorylase, which points to thymidine as the precursor imported for dTTP production or alternatively an unknown enzyme could be responsible for this function (Santos et al. 2011).

Horizontal gene transfer has an important role in the evolution of the *Mollicutes*, especially when contextualized with the genomic reduction process that characterizes this group. A high flow of genetic information is expected, as well as a horizontal gene transfer rate not limited to more distantly related species. A continuous gene flow between *Mycoplasma, Ureaplasma* and *Corynebacterium* could have been facilitated by the direct interaction among these groups in the urogenital tract of their hosts, since these bacteria are prevalent in urogenital tract infections (Russo et al. 1981).

5. Conclusion

*Ureaplasma* and *Mycoplasma* are among the microorganisms with the smallest known genomes resulting from a genome reduction process that led these groups throughout their evolution to present natural “almost minimal” gene repertoires. This reflects in their protein-protein interaction networks. While phylogenetically distant bacteria containing elaborate metabolic pathways such as *S. aureus* and *Corynebacterium urealyticum* have global PPI networks with higher compartmentation compromised with complete metabolic pathways, our data suggests that *U. diversum*, other ureaplasmas and mycoplasmases go in the opposite direction. Even though they present protein clusters that interact at a higher frequency, these groups have proteins that connect with each other in a more simple and general way, performing the role occupied by many proteins in more distantly related species. Urease is a niche gene crucial for energy production during urogenital tract colonization by *Ureaplasma, Corynebacterium* and *S. aureus*, and this enzyme is shared between these groups via horizontal gene transfer. The process of genome reduction among the *Mollicutes* therefore seems to have favored more loosely connected PPI networks, in which higher interaction among the remaining proteins compensates for the gene loss suffered by these groups throughout their evolutionary history.
Conflict of interest

- The authors declare that there is no conflict of interest.

Acknowledgements

Joana Kästle Silva received a Masters degree fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil. Jim Hesson edited the manuscript (https://www.academicenglishsolutions.com).

6. References


Escherichia coli is structurally related to other nucleoside monophosphate kinases.

J. Biol. Chem. 271(5): 2856–2862. ASBMB.


Culman, S.W., Bukowski, R., Gauch, H.G., Cadillo-Quiroz, H., and Buckley, D.H.


Jain, R., Rivera, M.C., and Lake, J.A. 1999. Horizontal gene transfer among genomes:


Acad Sciences.


ACS Publications.


Tu, A.H., and Turnbough, C.L. 1997. Regulation of upp expression in Escherichia coli by UTP-sensitive selection of transcriptional start sites coupled with UTP-


---

**Figure Caption**

**Figure 1.** General view of the PPI network in *U. diversum*. The colored modules correspond to the following metabolic pathways and cellular processes: Blue – Gene expression, Purple – Cellular processes, Transport proteins and Protein fate, Cellular envelope; Orange – Regulatory functions, fatty acid and energy metabolism, cofactors and carriers; Yellow – Purine and Pyrimidine metabolism; Pink – Intermediary metabolism and central Nitrogen metabolism (urease cluster).
Figure 2. Node degree distribution (A) Closeness centrality (B); Average clustering coefficient (C) and Shortest path length (D) distribution in the *U. diversum* PPI network.

Figure 3. Horizontal gene transfer events. In (A) urease function subnetwork, (B) purine metabolism subnetwork and (C) pyrimidine metabolism subnetwork. The red arrow indicates the direction of the horizontal gene transfer.

Figure 4. Urease subnetworks in (A) *U. diversum*, (B) *C. glucurolyticum*, (C) *C. urealyticum* and (D) *S. aureus*

Figure 5. Purine subnetworks in (A) *U. diversum*, (B) *M. genitalium*, (C) *M. pneumoniae* and (D) *M. hominis*

Figure 6. PPI networks compromised with pyrimidine metabolism in (A) *U. diversum*, (B) *M. hominis*, (C) *M. genitalium*, (D) *M. pneumoniae* and (E) *C. urealyticum*
General view of the PPI network in U. diversum. The colored modules correspond to the following metabolic pathways and cellular processes: Blue – Gene expression, Purple – Cellular processes, Transport proteins and Protein fate, Cellular envelope; Orange – Regulatory functions, fatty acid and energy metabolism, cofactors and carriers; Yellow – Purine and Pyrimidine metabolism; Pink – Intermediary metabolism and central Nitrogen metabolism (urease cluster).
Node degree distribution (A) Closeness centrality (B); Average clustering coefficient (C) and Shortest path length (D) distribution in the U. diversum PPI network.
Horizontal gene transfer events. In (A) urease function subnetwork, (B) purine metabolism subnetwork and (C) pyrimidine metabolism subnetwork. The red arrow indicates the direction of the horizontal gene transfer.
Urease subnetworks in (A) U. diversum, (B) C. glucurolyticum, (C) C. urealyticum and (D) S. aureus
Purine subnetworks in (A) U. diversum, (B) M. genitalium, (C) M. pneumoniae and (D) M. hominis
PPI networks compromised with pyrimidine metabolism in (A) U. diversum, (B) M. hominis, (C) M. genitalium, (D) M. pneumoniae, and (E) C. urealyticum.
<table>
<thead>
<tr>
<th>Proteins</th>
<th>Node degree ($k$)</th>
<th>Node Intersection</th>
<th>Cellular process</th>
</tr>
</thead>
<tbody>
<tr>
<td>recA</td>
<td>176</td>
<td>0.021</td>
<td>Genome integrity maintenance</td>
</tr>
<tr>
<td>dnaK</td>
<td>172</td>
<td>0.019</td>
<td>Transcription</td>
</tr>
<tr>
<td>topA</td>
<td>164</td>
<td>0.040</td>
<td>Replication/Transcription</td>
</tr>
<tr>
<td>ligA</td>
<td>154</td>
<td>0.022</td>
<td>Replication/DNA repair</td>
</tr>
<tr>
<td>UUR10_0577</td>
<td>69</td>
<td>0.023</td>
<td>Replication/DNA repair</td>
</tr>
</tbody>
</table>
Table 2. Topological parameters of the PPI networks in the analyzed species.

<table>
<thead>
<tr>
<th></th>
<th>N. of nodes</th>
<th>N. of edges</th>
<th>Clust. Coeff.</th>
<th>Centrality</th>
<th>Avg. n. of neighbors</th>
<th>Path length</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>U. diversum</em></td>
<td>301</td>
<td>10.404</td>
<td>0.613</td>
<td>0.359</td>
<td>69.130</td>
<td>1.970</td>
</tr>
<tr>
<td><em>M. pneumoniae FH</em></td>
<td>629</td>
<td>14.418</td>
<td>0.460</td>
<td>0.253</td>
<td>45.844</td>
<td>2.406</td>
</tr>
<tr>
<td><em>M. hominis ATCC 23114</em></td>
<td>523</td>
<td>12.706</td>
<td>0.558</td>
<td>0.276</td>
<td>48.589</td>
<td>2.512</td>
</tr>
<tr>
<td><em>M. genitalium G37</em></td>
<td>461</td>
<td>14.197</td>
<td>0.533</td>
<td>0.307</td>
<td>61.592</td>
<td>2.179</td>
</tr>
<tr>
<td><em>U. urealyticum ATCC 33699</em></td>
<td>646</td>
<td>16.964</td>
<td>0.576</td>
<td>0.259</td>
<td>52.520</td>
<td>2.616</td>
</tr>
<tr>
<td><em>U. parvum ATCC 700970</em></td>
<td>614</td>
<td>17.062</td>
<td>0.564</td>
<td>0.253</td>
<td>55.577</td>
<td>2.484</td>
</tr>
<tr>
<td><em>C. urealyticum DSM 7109</em></td>
<td>2.022</td>
<td>30.153</td>
<td>0.364</td>
<td>0.224</td>
<td>29.825</td>
<td>3.091</td>
</tr>
<tr>
<td><em>C. glucurolyticum ATCC 51867</em></td>
<td>2.645</td>
<td>53.828</td>
<td>0.351</td>
<td>0.181</td>
<td>40.702</td>
<td>2.981</td>
</tr>
<tr>
<td><em>S. aureus MSHR1132</em></td>
<td>2.510</td>
<td>41.222</td>
<td>0.367</td>
<td>0.190</td>
<td>32.846</td>
<td>3.105</td>
</tr>
<tr>
<td><em>S. aureus COL</em></td>
<td>2.615</td>
<td>52.435</td>
<td>0.377</td>
<td>0.185</td>
<td>40.103</td>
<td>2.995</td>
</tr>
<tr>
<td><em>S. aureus Mu50</em></td>
<td>2.730</td>
<td>57.290</td>
<td>0.340</td>
<td>0.204</td>
<td>41.971</td>
<td>2.918</td>
</tr>
</tbody>
</table>
Table 3. the organisms used in the HGT analysis and their respective genome CG content.

<table>
<thead>
<tr>
<th>Species</th>
<th>GC%</th>
<th>DNA length</th>
<th>Mean GC content %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ureaplasma urealyticum</td>
<td>30.489362</td>
<td>5.700</td>
<td>25-27</td>
</tr>
<tr>
<td>Ureaplasma parvum</td>
<td>30.279305</td>
<td>5.687</td>
<td>25</td>
</tr>
<tr>
<td>Ureaplasma diversum</td>
<td>35.878301</td>
<td>5.226</td>
<td>28.2</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>35.308254</td>
<td>5.025</td>
<td>32.7</td>
</tr>
<tr>
<td>Corynebacterium urealyticum</td>
<td>56.420846</td>
<td>5.085</td>
<td>64.2</td>
</tr>
<tr>
<td>Corynebacterium glucuronolyticum</td>
<td>52.500913</td>
<td>5.478</td>
<td>59</td>
</tr>
</tbody>
</table>