G3 Genes | Genomes | Genetics

GENOME REPORT

Genome-Wide Analysis of *Mycoplasma bovirhinis* GS01 Reveals Potential Virulence Factors and Phylogenetic Relationships

Shengli Chen,[1] Huafang Hao,[1] Ping Zhao, Yongsheng Liu,[2] and Yuefeng Chu[2]
State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Xujiaping 1, Lanzhou, 730046, Gansu, People's Republic of China
ORCID ID: 0000-0001-9043-0351 (S.C.)

ABSTRACT *Mycoplasma bovirhinis* is a significant etiology in bovine pneumonia and mastitis, but our knowledge about the genetic and pathogenic mechanisms of *M. bovirhinis* is very limited. In this study, we sequenced the complete genome of *M. bovirhinis* strain GS01 isolated from the nasal swab of pneumonic calves in Gansu, China, and we found that its genome forms a 847,985 bp single circular chromosome with a GC content of 27.57% and with 707 protein-coding genes. The putative virulence determinants of *M. bovirhinis* were then analyzed. Results showed that three genomic islands and 16 putative virulence genes, including one adhesion gene enolase, seven surface lipoproteins, proteins involved in glycerol metabolism, and cation transporters, might be potential virulence factors. Glycerol and pyruvate metabolic pathways were defective. Comparative analysis revealed remarkable genome variations between GS01 and a recently reported HAZ141_2 strain, and extremely low homology with others mycoplasma species. Phylogenetic analysis demonstrated that *M. bovirhinis* was most genetically close to *M. canis*, distant from other bovine *Mycoplasma* species. Genomic dissection may provide useful information on the pathogenic mechanisms and genetics of *M. bovirhinis*.

KEYWORDS
Mycoplasma bovirhinis
genome
virulence gene
comparative analysis
phylogenetic analysis
Genome Report

Bovine respiratory disease (BRD) and cow mastitis have caused considerable economic losses worldwide, and many microbial pathogens, including viruses and bacteria, have contributed to such losses (Griffin *et al.* 2010; Mosier 2014; Fox 2012). In addition, *Mycoplasma* species play a major role in the onset of these diseases (Srikumaran *et al.* 2007). More than 12 *Mycoplasma* species, such as *M. bovis*, *M. bovigenitalium*, *M. dispar*, and *M. bovirhinis*, have been reported to be involved in BRD and mastitis (Fox 2012; Martin 1978). *M. bovirhinis* is one of the most commonly occurring species in respiratory diseases and bovine mastitis in many countries (ter Laak *et al.* 1992; Ayling *et al.* 2004). In respiratory diseases, *M. bovirhinis* is normally regarded as a secondary invader that can co-infect with other pathogens and aggravate calf pneumonia (Miles *et al.* 2004).

M. bovirhinis, belonging to *Mycoplasmataceae* under *Mollicutes*, has a relatively small genome and has no cell wall. It was first isolated by Harborne in mastitis cows in England in 1965 (Harbourne *et al.* 1965). *M. bovirhinis* concurrently infects calves with other etiologies and causes respiratory diseases (Hirose *et al.* 2003), mastitis (Higuchi *et al.* 2011), and otitis (Lamm *et al.* 2004). This pathogen is usually isolated from the nasal mucus (Hirose *et al.* 2003) and lung (Soehnlen *et al.* 2011) of pneumonic calves or from calves with mastitis (Fox 2012), otitis (Lamm *et al.* 2004), and subacute nephritis (Panangala *et al.* 1990). *M. bovirhinis* is occasionally detected in clinically healthy calves, but its detection rate in such calves is lower than that in pneumonic calves (Angen *et al.* 2009). Despite epidemiological and antimicrobial susceptibility investigations on *M. bovirhinis*, the pathogenesis of this pathogen remains unknown.

In 2016, a severely contagious respiratory disease with an incidence of more than 50% and mostly affected calves spread in a dairy farm in

Copyright © 2018 Chen *et al.*
doi: https://doi.org/10.1534/g3.118.200018
Manuscript received January 4, 2018; accepted for publication March 27, 2018; published Early Online March 30, 2018.
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.118.200018/-/DC1.
[1]These authors contributed equally to the work.
[2]Corresponding authors: State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Xujiaping 1, Lanzhou, 730046, Gansu, People's Republic of China. E-mail: chuyuefeng@caas.cn; liuyongsheng@caas.cn

Gansu Province in China. The main symptoms recorded were cough, asthma, high fever, and emaciation. Nasal swabs were sampled and examined through conventional isolation and molecular identification. One *Mycoplasma* strain named GS01 was isolated from the samples with an improved Friis medium, but other pathogenic bacteria were not detected. After cloning purification was performed thrice, 16S rRNA sequence and specific PCR analysis (Kobayashi *et al.* 1998) further confirmed that the isolate is *M. bovirhinis*.

Many *Mycoplasma* species have been identified and sequenced genomically. At present, the virulence factors and evolutionary relationships of *M. bovirhinis* are still poorly understood although the genome of a Japanese strain HAZ141_2 was recently reported (Hata *et al.* 2017). In the present study, we report the complete genomic sequence of *M. bovirhinis* strain GS01, and identify putative relevant virulence factors. Comparative and phylogenetic analyses are also conducted. The data presented in this study may improve our current understanding on the pathogenic mechanisms and genetics of *M. bovirhinis*.

MATERIALS AND METHODS

Bacterial Culture and DNA Preparation

The GS01 strain was isolated with a deep nasal swab in a modified Friis medium (Friis powder 21.4 g/L, glucose 2 g/L, 10% horse serum, 10% porcine serum, 100 mg/L ampicillin sodium, 0.01% acetic acid thallium, adding 1 M sodium hydroxide adjust to pH 7.4) at 37° for 2 days. The colony was cultured on an agar plate (adding 1.5% agar in the modified Friis medium) at 37° in a 5% CO_2 atmosphere for 5 days. The strain was purified thrice. A 500 mL mid-exponential phase culture was pelleted through centrifugation at 10,000 × g for 20 min and subjected to centrifugal washing with PBS (0.01 M, pH 7.2) twice. Total genomic DNA was extracted and treated with RNase by using a TIANamp bacterial DNA kit (Tiangen, Beijing, China) according to the manufacturer's instructions.

Library Construction and DNA Sequencing

The genomic DNA was detected through agarose gel electrophoresis and was used to prepare a 10-kb size-selected PacBio SMRTbell libraries following the manufacturer's instructions. After purification was performed, the libraries were quantified by Qubit, and the insertion size was detected. The genome of *M. bovirhinis* GS01 was sequenced using the PacBio RSII platform, resulting in a 1026-fold sequencing depth and 719× depth of coverage. Genome sequencing was conducted at the Beijing Novogene Bioinformatics Technology Co., Ltd.

Genome Assembly and Annotation

Low-quality raw data were filtered by SMRT v2.3.0 analysis software suite to obtain clean data. A total of 83,447 reads totaling 1,023,037,568 bases (mean read length: 12,259 bp) was obtained. The N50 read length was 15,812 bp and mean read score was 0.82. Read length distribution for the sequenced genome is shown in Figure S1 in File S1. Genome assembly was performed using the SMRT portal, included in SMRT v2.3.0 and the *de novo* assembly was conducted following the hierarchical genome-assembly process (HGAP) assembly protocol with Quiver polishing (Chin *et al.* 2013). Finally, one polished contig without gap was generated. The depth of coverage distribution map for genome assembly is shown in Figure S2 in File S1.

The genome component prediction was conducted as follows: gene prediction was conducted using GeneMarkS v4.17 (Besemer *et al.* 2001) with default parameters. Transfer RNAs (tRNAs) were identified using tRNAscan-SE v1.3.1 (Lowe and Eddy 1997) with default parameters. Ribosomal RNAs (rRNAs) were detected using RNAmmer v1.2 (Lagesen *et al.* 2007) with default parameters. Small nuclear RNAs (snRNAs) were identified using BLAST against the Rfam v12.1 (Gardner *et al.* 2009). Pseudogenes were predicted by the NCBI Prokaryotic Genome Annotation Pipeline on the GenBank database. Genomic islands, insertion sequences, interspersed repetitive sequences, and tandem repeats were predicted with IslandPath-DIOMB program (Hsiao *et al.* 2003), ISfinder (https://www-is.biotoul.fr/), RepeatMasker v4.0.5 (Saha *et al.* 2008), and TRF v4.0.7b (Benson 1999) with default parameters, respectively.

Functional annotations were conducted by BLASTP algorithm similarity search against the non-redundant (NR) protein database (release 2016-04), Swiss-Prot (release 2016-04), Clusters of Orthologous Groups (COG) (release 2015-12-14), Gene Ontology (GO) (release 2014-10-19), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (release 2016-04). BLASTP algorithm were set as E-value less than 1e-5, and the homology identity and minimal alignment length percentage were larger than 40% for the above functional annotation tools. The results were subsequently filtered by selecting the highest score of the alignment. The potential virulence genes were predicted through gene annotation and reference studies (Bürki *et al.* 2015; Chen *et al.* 2011; Hames *et al.* 2009; Groisman *et al.* 2013; Zheng *et al.* 2015). Secretory proteins were predicted with the SignalP v4.1 database (Petersen *et al.* 2011) with default parameters, and type III effector proteins secreted by *M. bovirhinis* GS01 were predicted with EffectiveT3 (Eichinger *et al.* 2016) with default parameters.

Comparative and Phylogenetic Analysis

The genome comparisons between *M. bovirhinis* GS01 and HAZ141_2 were conducted using the above genome annotation methods. Genomic synteny was analyzed on the basis of the results of the alignment, which was conducted using MUMmer v3.23 (Delcher *et al.* 2003) and LASTZ v1.03.54 (Chiaromonte *et al.* 2002) tools between GS01 and referenced HAZ141_2 genome under default parameters. Briefly, the large scale co-linear relationship of the target GS01 genome and the referenced HAZ141_2 genome was determined using mummer program under MUMmer v3.23 software package with default parameters. The output of results were then processed regionally by LASTZ v1.03.54 with best-chain alignments following the manufacturer's instructions, and the local arrangement of relationships (collinear, translocation, inversion and translocation + inversion) were determined. The presented results are derived from a combination of MUMmer and LASTZ analyses. The overall sequence similarity of the two genomes was calculated using BLASTN v2.2.26 alignment method. Core genes were identified by CD-HIT software v4.6 (Li and Godzik 2006), with the following parameters: 50% pairwise identity threshold and 0.7 length difference cutoff in amino acid. Multiple sequences of single-copy core genes among 19 *Mycoplasma* strains were aligned using MUSCLE v3.8.31 (Edgar 2004). Phylogenetic trees based on the single-copy core genes were constructed by TreeBeST v1.9.2 (Nandi *et al.* 2010) using the maximum likelihood model with 1000 bootstrap replicates under default parameters and by MrBayes v.3.2.6 (Ronquist *et al.* 2012) using Bayesian inference method. Bayesian inference phylogenyetic construction was conducted using mixed model under the following set conditions: mcmc ngen = 2×10^4, samplefreq = 100, printfreq = 1000, diagnfreq = 1000, and to obtain standard deviation of split frequencies below 0.01. After discarding the burn-in samples, a Bayesian phylogenyetic tree was generated based on the remaining data and was shown by using iTOL tool

(http://itol.embl.de/). The genome sequence of other *Mycoplasma* species were obtained from the NCBI database.

Data availability

Strains are available upon request. The genome sequence data were deposited in GenBank with the accession number CP024049. Supplemental Material, Figure S1 in File S1 shows PacBio read length distribution for the sequenced *M. bovirhinis* GS01 genome. Figure S2 in File S1 shows the depth of coverage distribution map for *M. bovirhinis* GS01 genome assembly. Table S1 in File S1 shows prediction of the genome component of *M. bovirhinis* GS01. Table S2 in File S1 shows the statistical results of repetitive sequences of *M. bovirhinis* GS01. Table S3 in File S1 shows functional category in COG of *M. bovirhinis*. Table S4 in File S1 shows genes predicted involving in transporter system of *M. bovirhinis* GS01. Table S5 in File S1 shows the predicted genes involving in metabolism in the GS01 genome. Table S6 in File S1 shows proteins involved in secretion system of *M. bovirhinis* GS01. Table S7 in File S1 shows the effective proteins of T3SS in *M. bovirhinis* GS01 genome. Table S8 in File S1 shows genes in the genomic islands of *M. bovirhinis* GS01. Table S9 in File S1 shows the genome comparisons between *M. bovirhinis* GS01 and HAZ141_2 strains. Table S10 in File S1 shows the genes of the 53.4-kb deletion in GS01 relative to HAZ141_2. Table S11 in File S1 shows list of 14 single-copy core genes of 19 selected *Mycoplasma* strains.

RESULTS

General Genome Features

The complete genome of *M. bovirhinis* GS01 is composed of a 847,985 bp single circular chromosome with a 27.57% GC content (Figure 1). A total of 83,447 reads with an average length of 12,259 bp were produced. Then, 707 protein-encoding genes with an average length 1,072 bp were identified in the genome, and the coding percentage of the genome was 89.34%. The non-coding RNA of this organism consists of 31 tRNAs and 8 rRNAs (2 5S rRNA, 3 16S rRNA, and 3 23S rRNA) (Table S1 in File S1). Moreover, 67 tandem repeats and 68 interspersed nuclear elements were identified. Table S2 in File S1 presents the statistical results of the repetitive sequences of *M. bovirhinis* GS01.

Of the 707 coding genes of *M. bovirhinis* GS01, 259 were assigned to 20 functional categories in COG database (Table S3 in File S1) and 81 (31.27%) were involved in transport and metabolism. Biological functions were defined for 462 (65.34%) genes, and 127 genes encoded hypothetical proteins with unknown functions. Genomic sequence data were submitted to GenBank with the accession number CP024049.

Transporter and Metabolism

A total of 56 *M. bovirhinis* genes were annotated as being related to transporter systems (Table S4 in File S1). These mainly belonged to the ATP-binding cassette (ABC) transporter system and the phosphotransferase system (PTS). Furthermore, 42 genes were encoded for the ABC transporter system, including 17 ATP-binding proteins, 16 permease

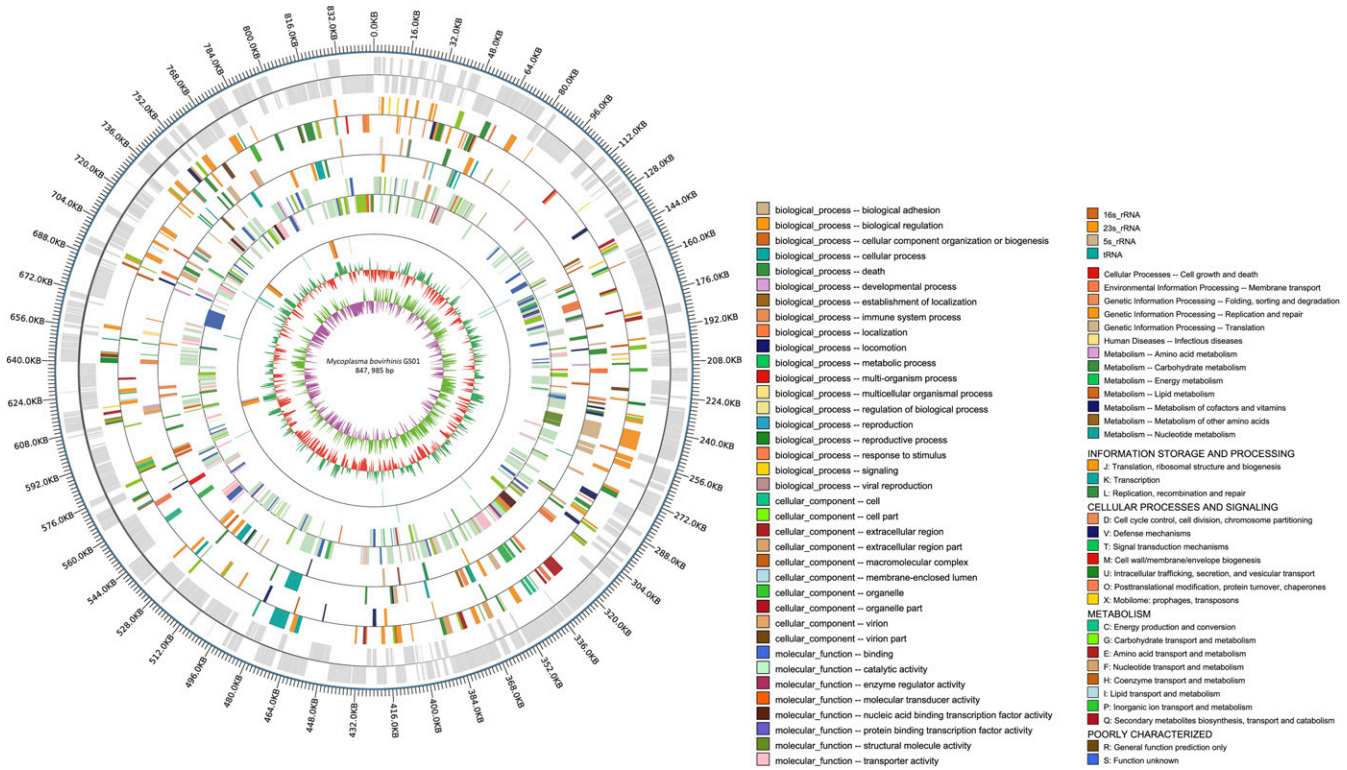


Figure 1 Chromosome atlas of *Mycoplasma bovirhinis* strain GS01. *dnaA* at position 1. From outer to inner circles, the first circle represents the genome position coordinates; the second circle shows the locations of predicted coding genes on plus and minus strands; the third, fourth, and fifth circles show the results of color-coded by COG categories, KEGG, and GO annotation (see the description at the bottom-right corner), respectively; the sixth circle represents the ncRNA in GS01 genome; the seventh circle represents the mean centered G+C content whose baseline is average GC, and outward and inward projections express higher and lower values than the average, respectively; and the eighth circle illustrates the GC (G+C) skew plot: above and below zero are highlighted in green and purple, respectively.

proteins, two substrate-binding proteins, and seven other proteins. These transporters are mainly involved in maltose/maltodextrin, cobalt/nickel, spermidine/putrescine, oligopeptide, sugar, phosphate/phosphonate, and cation transport systems. The PTS system contained six genes, including *ptsI* and *ptsH*, and four genes encoded carbohydrate-specific EII complexes, which catalyze concomitant carbohydrate translocation and phosphorylation. The usually three-component EII complex was composed of hydrophilic IIA and IIB and carbohydrate-selective transmembrane IIC domains. The GS01 genome has one fructose-specific, one lichenan-specific, and two glucose-specific EIIABC components, which indicated that the PTS system is involved in the saccharide absorbance of *M. bovirhinis*.

Given their small genomes, *Mycoplasma* have lost most biosynthetic activities and lack many metabolic pathways, forcing them to obtain nourishment from their host. Although 91 genes were predicted to participate in the metabolic system of *M. bovirhinis* (Table S5 in File S1), the metabolic ability was defective. The glycerol glycerophosphocholine importer GlpU, oxidase GlpO, and glycerol import system GtsABC were not found in the genome. Glycerol uptake may be facilitated by the glycerol uptake facilitator protein GlpF (Mbr-GS01GM000491) and phosphorylated by glycerol kinase GlpK (Mbr-GS01GM000492) to become glycerol-3-phosphate (G3P), which is transformed into glyceraldehyde-3-phosphate by glycerol-3-phosphate dehydrogenase GlpD (Mbr-GS01GM000493) and triosephosphate isomerase (Mbr-GS01GM000512). Glyceraldehyde-3-phosphate is an important product of the glycolysis pathway and is metabolized into pyruvate. The genes required to transform glucose into pyruvate and pyruvate into lactate were identified in *M. bovirhinis*. However, the genes involved in tricarboxylic acid cycle were missing, such as pyruvate dehydrogenase (PDH) enzyme complex genes and phosphate acetyltransferase. Besides, two genes (Mbr-GS01GM000577 and Mbr-GS01GM000638) in the pentose phosphate pathway were also annotated in the genome.

Secretion Systems

A total of 47 secreted proteins with N-terminal signal peptides were predicted in the GS01 genome, and the peptides were 19–31 amino acids long. Our results revealed the components of the secretion machinery, which included the signal recognition particle receptor FtsY and subunit Ffh; components SecA, SecE, SecG, SecY, and SecDF in the major translocation pathway and the general chaperone trigger factor DnaK, LepA; and the competence protein ComEA (Table S6 in File S1). There was one putative inner membrane protein translocase component YidC that is involved in the insertion of hydrophobic sequences into the lipid bilayer either independently or via the SecYEG translocase complex (Scotti *et al.* 2000). Moreover, one signal peptidase I gene and one signal peptidase II gene, which respectively encode the enzymes for the cleavage of the common protein and lipoprotein signal peptides, were found in the GS01 genome.

Furthermore, the effector proteins of type III secretion/translocation systems were predicted and 16 proteins (Table S7 in File S1), including ATP synthase, ribosomal protein, transporter protein, and a conversed hypothetical protein, were found. These proteins were secreted into the extracellular environment or host, which may be related to the survival and pathogenicity of bacteria.

Virulence Factors

Three genomic islands were found in the GS01 genome (from 463,870 to 481,877, from 561,137 to 565,821, and from 597,644 to 607,572) with a total length of 32,622 bp, containing a total of 23 genes (Table S8 in File S1). No complete insertion sequence element was found in the genome, and seven transposes were found to locate outside the genomic islands. Although the particular function of proteins cited in the genomic islands were unclear, pathogenicity islands were often considered essential for bacteria virulence (Schmidt and Hensel 2004).

Adherence to host cells is a key step in *Mycoplasma* colonization and infection, and adherence proteins are regarded as virulence factors. Enolase is considered an adherence factor that contributes to adherence by binding a chicken plasminogen in *M. gallinaceum* (Chen *et al.* 2011). An enolase gene (Mbr-GS01GM000688) was identified in the GS01 genome (Table 1), which showed 87% amino acid identity with the α-enolase gene of *M. gallinaceum*, and might be associated with *M. bovirhinis* virulence. The capsule is also considered an important virulence factor in microorganisms, such as bacteria (Boyce and Adler 2000). Only one capsule synthesis-related gene (Mbr-GS01GM000446) was annotated in the genome and might be involved in the virulence of *M. bovirhinis*.

Lipoproteins on the *Mycoplasma* surface play a crucial role in interactions between pathogen and eukaryotic cells, antigenic variation, and immunity evasion. Thus, they are responsible for *Mycoplasma* virulence (Bürki *et al.* 2015). Seven surface or membrane-associated lipoproteins (Table 1), including P60-like proteins and seven other lipoproteins, were found. P60 is regarded as a virulence factor of *M. hyopneumoniae* (Seymour *et al.* 2012). These lipoproteins may be virulence factors of *M. bovirhinis*.

Glycerol metabolism and its metabolic product H_2O_2 contributes to the virulence of *Mycoplasma* (Blötz and Stülke 2017; Hames *et al.* 2009). The *glpF-glpK-glpD* gene cluster was found in the *M. bovirhinis* genome (Table 1). The gene cluster *gtsABC* is an efficient active glycerol import system and is found in many *Mycoplasma* genomes, such as those in *M. mycoides* subsp. *mycoides* SC (Blötz and Stülke 2017), *M. pneumoniae* (Hames *et al.* 2009), and *M. capricolum* subsp. *capripneumoniae* (Mccp) (Chen *et al.* 2017), but no gene cluster *gtsABC* was identified in the GS01 genome.

Magnesium transporters MgtA and MgtE are considered virulence factors in some bacteria (Groisman *et al.* 2013). The potassium transporter TrkA is related to the virulence of *Salmonella* (Su *et al.* 2009). In the GS01 genome, three genes, namely, *mgtA* (Mbr-GS01GM000528), *mgtE* (Mbr-GS01GM000661), and *trkA* (Mbr-GS01GM000385), were predicted, and their encoding proteins may be related to the virulence of *M. bovirhinis*.

A 5′-nucleotidase that utilizes host nucleotides and can enhance macrophage death is considered a virulence factor in *Streptococcus pyogenes* (Zheng *et al.* 2015). A 5′-nucleotidase gene (Mbr-GS01GM000652) was identified in the genome and may be associated with the virulence of *M. bovirhinis*.

Comparative and Evolutionary Analysis

The genome comparisons between *M. bovirhinis* GS01 and HAZ141_2 strains were conducted and are shown in Table S9 in File S1. The HAZ141_2 genome size is 948,039 bp, 100,054 bp longer than the GS01 genome. The overall sequence similarity of GS01 and HAZ141_2 was calculated to be 97.56%. GS01 has 707 protein-encoding genes, and HAZ141_2 was annotated to have 821 protein-encoding genes using the same genome annotation method as GS01. 16 and 38 pseudogenes were separately identified in GS01 and HAZ141_2 according to the annotation by the NCBI Prokaryotic Genome Annotation Pipeline on the GenBank database. The functional category in COG of *M. bovirhinis* GS01 and HAZ141_2 were compared and are shown in Table S3 in File S1. The collinearity analysis between GS01 and HAZ141_2 was also conducted. The origin of HAZ141_2 genome

Table 1 Predicted virulence genes in the GS01 genome

Locus	Product	Gene	Protein length (aa)	Position
Mbr-GS01GM000045	membrane-associated lipoprotein	—	834	38324...40828
Mbr-GS01GM000101	predicted lipoprotein	—	345	113136...114173
Mbr-GS01GM000385	potassium transporter TrkA	*trkA*	223	431400...432071
Mbr-GS01GM000400	membrane-associated lipoprotein	—	833	456200...458701
Mbr-GS01GM000446	glycosyl transferase	—	340	522199...523221
Mbr-GS01GM000464	surface protein	—	713	545544...547685
Mbr-GS01GM000465	membrane-associated lipoprotein	—	851	547762...550317
Mbr-GS01GM000491	glycerol uptake facilitator protein	*glpF*	248	581842...582588
Mbr-GS01GM000492	glycerol kinase	*glpK*	505	582597...584114
Mbr-GS01GM000493	glycerol-3-phosphate dehydrogenase	—	384	584124...585278
Mbr-GS01GM000528	magnesium transporter	*mgtE*	524	621734...623308
Mbr-GS01GM000652	5′-nucleotidase	—	707	765862...767985
Mbr-GS01GM000661	magnesium-transporting ATPase (P-type)	*mgtA*	904	779505...782219
Mbr-GS01GM000688	enolase	*eno*	453	816615...817976
Mbr-GS01GM000690	P60-like lipoprotein	*p60*	413	818435...819676
Mbr-GS01GM000691	putative membrane protein P80	*p80*	722	819676...821844

was 752 bp upstream of the RNase J family beta-CASP ribonuclease gene, while the *dnaA* gene was at position 1 for GS01 as majority *Mycoplasma*. Their genome structures did not have very high synteny (Figure 2), with 130 blocks in the comparison. Approximately 12.3-kb translocation + inversion was found at the beginning of genome GS01 and 60-kb inversion at the end and start. The genome size of GS01 was 100-kb smaller than that of HAZ141_2. Several insertions and deletions were observed, and the largest deletion was 53,407 bp long. The 53.4-kb deletion in GS01 relative to HAZ141_2 was sited from 871,457 to 924,505, correspond to the 53.5-kb insertion identified in HAZ141_2 by Hata *et al.* (2017). The corresponding segment encoded 47 proteins, mainly including phage-related proteins, hypothetical proteins, and others (Table S10 in File S1).

The core genes between *M. bovirhinis* GS01 and 18 other *Mycoplasma* genomes were identified using CD-HIT, and 14 single-copy core genes were found (Table S11 in File S1). Phylogenetic trees were performed based on the single-copy core genes of 19 *Mycoplasma* strains using maximum likelihood and bayesian inference methods, and the topology of the phylogeny were consistent (Figure 3). According to the phylogenetic trees, the *M. bovirhinis* GS01 and HAZ141_2 strains have the nearest relationship as expected. *M. bovirhinis* was closely related with *M. canis*, followed by *M. cynos*, but was distant from *M. bovis*, which is also an important bovine respiratory and mastitis pathogen.

DISCUSSION

Enolase, which catalyzes the conversion of 2-phosphoglycerate (2-PGA) to phosphoenolpyruvate (PEP) during glycolysis and a reverse reaction during glycogen synthesis (Spring and Wold 1971), is a plasminogen-binding protein and is closely related to the adherence to the host cell in many mycoplasma species, such as *M. gallisepticum* (Chen *et al.* 2011), *M. bovis* (Song *et al.* 2012) and *M. synoviae* (Bao *et al.* 2014). An enolase was identified in the GS01 genome, also in HAZ141_2, and may be considered an adherence and virulence factor in *M. bovirhinis*, but its role in this *Mycoplasma* species needs further investigations.

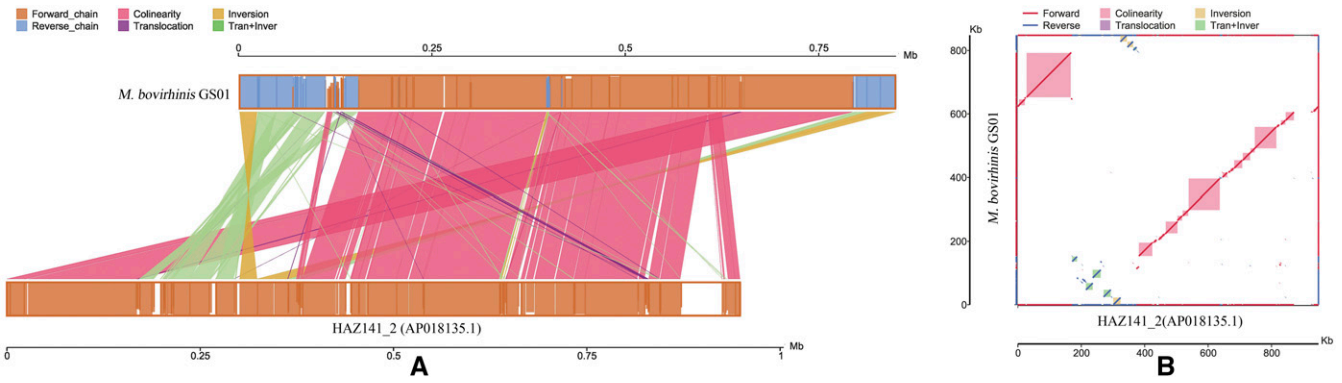


Figure 2 Collinearity analysis of the complete genome between *M. bovirhinis* strain GS01 and HAZ141_2. (a): Two genomes on parallel collinearity. The upper and lower shafts represent the sequenced and reference genomes, respectively. The orange box in the axes indicates the genomic forward strand, and the blue one corresponds to the reverse chain. The filled color in the box denotes alignment similarity, and the complete filling suggests 100% similarity. The color of the link graph between the upper and lower axes represents alignment type: pink, purple, yellow, and green denote collinear, translocation, inversion, and translocation+inversion, respectively. (b): Two-dimensional comparison between two genomes. The vertical and horizontal axes represent GS01 genome and reference HAZ141_2 genome, respectively. The blue line shows the reverse chain, and the red line corresponds to forward alignment. The pink and green modules indicate collinear and translocation+inversion between them, respectively.

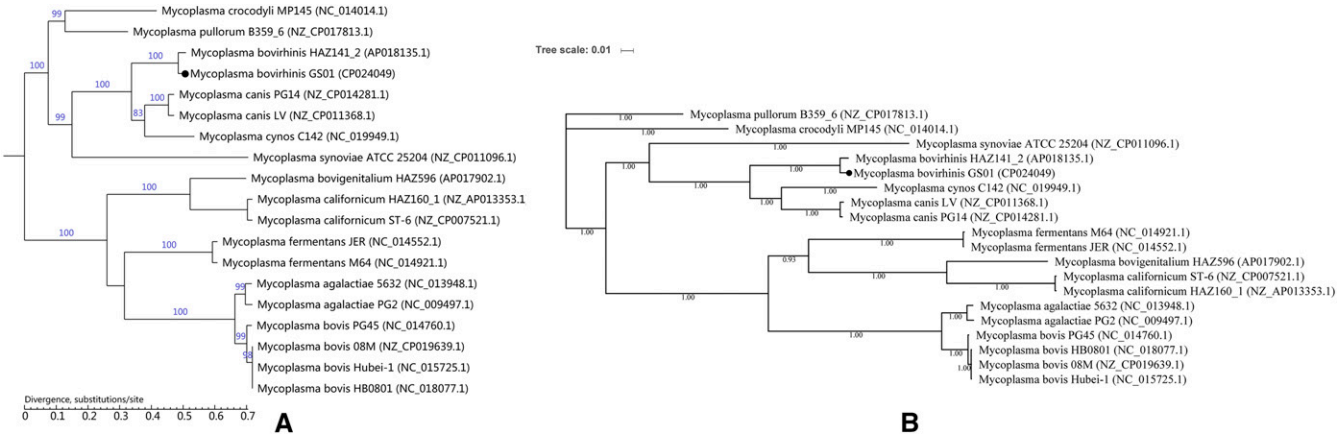


Figure 3 Phylogenetic trees of 14 single-copy core genes of 19 selected *Mycoplasma* strains. (a): Phylogenetic tree was constructed using TreeBeST with the maximum likelihood model. Bootstrap replicates were 1000, and the number of bootstraps for each node is shown. The tree is displayed to scale, with branch lengths measured in the number of substitutions per site. *M. bovirhinis* strain GS01 is indicated by a black dot. (b): Phylogenetic tree was constructed using MrBayes with bayesian inference method. The numbers upon each node indicate bayesian posterior probabilities. The tree is displayed to scale, with branch lengths measured in the number of substitutions per site. *M. bovirhinis* strain GS01 is indicated by a black dot.

Glycerol metabolism and H_2O_2 production influence *Mycoplasma* virulence (Hames *et al.* 2009; Vilei and Frey 2001). Glycerol uptake usually occurs through an efficient active glycerol import system GtsABC, and possibly a bypass pathway via the glycerol facilitator factor GlpF. In the GS01 genome, the *gtsABC* gene cluster were absent. Glycerol metabolism may be supported by the *glpF-glpK-glpD* gene cluster, but need verification. In HAZ141_2, *gtsABC* and *glpF-glpK-glpD* gene clusters existed, indicating that glycerol metabolism may vary among *M. bovirhinis* strains.

In addition to 16 putative virulence genes predicted in this study, ClpB (Chastanet *et al.* 2004), ClpC (Nair *et al.* 2000), hemolysins (Goebel *et al.* 1988), PDH enzyme complex (Gates *et al.* 2008; Gründel *et al.* 2015), and lipoate-protein ligase LplA (O'Riordan *et al.* 2003) are regarded as relevant virulence factors in bacteria. These genes are not found in *M. bovirhinis* GS01 and HAZ141_2 genomes, but were present in genomes of *M. bovis* (Li *et al.* 2011) and Mccp (Chen *et al.* 2017), important contagious pleuropneumonia pathogens for cattle and goats, respectively. The absence of these crucial virulence factors may be associated with the relatively lower pathogenicity of *M. bovirhinis* than that of *M. bovis* as expected.

Signal peptidase I (SPase I) is for general protein secretion, whereas signal peptidase II, also called lipoprotein signal peptidase (LspA), functions by releasing signal peptides from bacterial prolipoproteins. Both signal peptidase I and II genes are found in the *M. bovirhinis* GS01 and HAZ141_2 genomes, *M. synoviae* (Vasconcelos *et al.* 2005), *M. conjunctivae* (Calderon-Copete *et al.* 2009), and *M. hyopneumoniae* (Moitinho-Silva *et al.* 2012). For the genomes of other *Mycoplasma* species, such as *M. bovis* and Mccp, only the signal peptidase II gene has been identified (Li *et al.* 2011; Chen *et al.* 2017). This result indicates that the mechanism of protein secretion in *M. bovirhinis* may be the same as that in *M. synoviae*, *M. conjunctivae*, and *M. hyopneumoniae*, but different from that in Mccp and *M. bovis*.

Two genomic sequences of *M. bovirhinis* have now been determined. GS01 had the nearest relationship with HAZ141_2 based on the phylogenetic tree and comparative analysis. Moderate genomic synteny and a large-scale deletion in GS01 relative to HAZ141_2 indicated remarkable genome variations of *M. bovirhinis*. *M. bovirhinis* was most genetically close to *M. canis*, which is usually considered a commensal or opportunistic cofactor in respiratory or urogenital tract diseases of dogs (Ulgen *et al.* 2006) and can be recovered from pneumonic calves (ter Laak *et al.* 1993). The genetics and phylogenetic of *M. bovirhinis* will be revealed more clearly with the aid of more strains sequenced in the future.

ACKNOWLEDGMENTS

This work was partially funded by the National Natural Science Foundation of China (Grant No. 31602088 and 31402223), the National Key Research and Development Plan (Grant No. 2017YFD0500905 and 2016YFD0500907), the National Key Technologies R&D Program (Grant No. 2015BAD12B02), and the Fundamental Research Funds for CAAS (Grant No. 1610312016027).

Author Contributions: Conceived and designed the experiments: Shengli Chen, Huafang Hao. Performed the experiments: Huafang Hao. Analyzed the data: Shengli Chen, Huafang Hao. Contributed reagents/materials/analysis tools: Yuefeng Chu, Yongsheng Liu, Ping Zhao. Wrote the paper: Shengli Chen, Huafang Hao. Critically revised the manuscript: Huafang Hao, Shengli Chen. Supervised all work: Yuefeng Chu, Yongsheng Liu. All authors read and approved this version of manuscript. The authors declare no competing interests regarding the publication of this article. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

LITERATURE CITED

Angen, Ø., J. Thomsen, L. E. Larsen, J. Larsen, B. Kokotovic *et al.*, 2009 Respiratory disease in calves: Microbiological investigations on trans-tracheally aspirated bronchoalveolar fluid and acute phase protein response. Vet. Microbiol. 137(1-2): 165–171. https://doi.org/10.1016/j.vetmic.2008.12.024

Ayling, R. D., S. E. Bashiruddin, and R. A. Nicholas, 2004 Mycoplasma species and related organisms isolated from ruminants in Britain between 1990 and 2000. Vet. Rec. 155(14): 413–416. https://doi.org/10.1136/vr.155.14.413

Bao, S., X. Guo, S. Yu, J. Ding, L. Tan *et al.*, 2014 Mycoplasma synoviae enolase is a plasminogen/fibronectin binding protein. BMC Vet. Res. 10(1): 223. https://doi.org/10.1186/s12917-014-0223-6

Benson, G., 1999 Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27(2): 573–580. https://doi.org/10.1093/nar/27.2.573

Besemer, J., A. Lomsadze, and M. Borodovsky, 2001 GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 29(12): 2607–2618. https://doi.org/10.1093/nar/29.12.2607

Blötz, C., and J. Stülke, 2017 Glycerol metabolism and its implication in virulence in Mycoplasma. FEMS Microbiol. Rev. 41(5): 640–652. https://doi.org/10.1093/femsre/fux033

Boyce, J. D., and B. Adler, 2000 The capsule is a virulence determinant in the pathogenesis of Pasteurella multocida M1404 (B:2). Infect. Immun. 68(6): 3463–3468. https://doi.org/10.1128/IAI.68.6.3463-3468.2000

Bürki, S., J. Frey, and P. Pilo, 2015 Virulence, persistence and dissemination of Mycoplasma bovis. Vet. Microbiol. 179(1-2): 15–22. https://doi.org/10.1016/j.vetmic.2015.02.024

Calderon-Copete, S. P., G. Wigger, C. Wunderlin, T. Schmidheini, J. Frey *et al.*, 2009 The Mycoplasma conjunctivae genome sequencing, annotation and analysis. BMC Bioinformatics 10(Suppl 6): S7. https://doi.org/10.1186/1471-2105-10-S6-S7

Chastanet, A., I. Derre, S. Nair, and T. Msadek, 2004 clpB, a novel member of the Listeria monocytogenes CtsR regulon, is involved in virulence but not in general stress tolerance. J. Bacteriol. 186(4): 1165–1174. https://doi.org/10.1128/JB.186.4.1165-1174.2004

Chen, H., S. Yu, X. Shen, D. Chen, X. Qiu *et al.*, 2011 The Mycoplasma gallisepticum α-enolase is cell surface-exposed and mediates adherence by binding to chicken plasminogen. Microb. Pathog. 51(4): 285–290. https://doi.org/10.1016/j.micpath.2011.03.012

Chen, S., H. Hao, P. Zhao, F. Thiaucourt, Y. He *et al.*, 2017 Genome-Wide Analysis of the First Sequenced Mycoplasma capricolum subsp. capripneumoniae Strain M1601. G3 (Bethesda) 7(9): 2899–2906. https://doi.org/10.1534/g3.117.300085

Chiaromonte, F., V. B. Yap, and W. Miller, 2002 Scoring pairwise genomic sequence alignments. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing. 7: 115.

Chin, C. S., D. H. Alexander, P. Marks, A. A. Klammer, J. Drake *et al.*, 2013 Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods 10(6): 563–569. https://doi.org/10.1038/nmeth.2474

Delcher, A. L., S. L. Salzberg, and A. M. Phillippy, 2003 Using MUMmer to identify similar regions in large sequence sets. Curr. Protoc. Bioinformatics Chapter 10: Unit 10.3. https://doi.org/10.1002/0471250953.bi1003s00

Edgar, R. C., 2004 MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5(1): 113. https://doi.org/10.1186/1471-2105-5-113

Eichinger, V., T. Nussbaumer, A. Platzer, M. A. Jehl, R. Arnold *et al.*, 2016 EffectiveDB–updates and novel features for a better annotation of bacterial secreted proteins and Type III, IV, VI secretion systems. Nucleic Acids Res. 44(D1): D669–D674. https://doi.org/10.1093/nar/gkv1269

Fox, L. K., 2012 Mycoplasma mastitis: causes, transmission, and control. Vet. Clin. North Am. Food Anim. Pract. 28(2): 225–237. https://doi.org/10.1016/j.cvfa.2012.03.007

Gardner, P. P., J. Daub, J. G. Tate, E. P. Nawrocki, D. L. Kolbe *et al.*, 2009 Rfam: updates to the RNA families database. Nucleic Acids Res. 37(Database): D136–D140. https://doi.org/10.1093/nar/gkn766

Gates, A., S. Frasca, A. T. Nyaoke, L. Silbart, and S. Geary, 2008 Comparative assessment of a metabolically attenuated Mycoplasma gallisepticum mutant as a live vaccine for the prevention of avian respiratory mycoplasmosis. Vaccine 26(16): 2010–2019. https://doi.org/10.1016/j.vaccine.2008.02.010

Goebel, W., T. Chakraborty, and J. Kreft, 1988 Bacterial hemolysins as virulence factors. Antonie van Leeuwenhoek 54(5): 453–463. https://doi.org/10.1007/BF00461864

Griffin, D., M. M. Chengappa, J. Kuszak, and D. S. McVey, 2010 Bacterial pathogens of the bovine respiratory disease complex. Vet. Clin. North Am. Food Anim. Pract. 26(2): 381–394. https://doi.org/10.1016/j.cvfa.2010.04.004

Groisman, E. A., K. Hollands, M. A. Kriner, E. J. Lee, S. Y. Park *et al.*, 2013 Bacterial Mg2+ homeostasis, transport, and virulence. Annu. Rev. Genet. 47(1): 625–646. https://doi.org/10.1146/annurev-genet-051313-051025

Gründel, A., K. Friedrich, M. Pfeiffer, E. Jacobs, and R. Dumke, 2015 Subunits of the Pyruvate Dehydrogenase Cluster of Mycoplasma pneumoniae Are Surface-Displayed Proteins that Bind and Activate Human Plasminogen. PLoS One 10(5): e0126600. https://doi.org/10.1371/journal.pone.0126600

Hames, C., S. Halbedel, M. Hoppert, J. Frey, and J. Stulke, 2009 Glycerol metabolism is important for cytotoxicity of Mycoplasma pneumoniae. J. Bacteriol. 191(3): 747–753. https://doi.org/10.1128/JB.01103-08

Harbourne, J. F., D. Hunter, and R. H. Leach, 1965 The isolation of mycoplasma from bovine lungs and nasal swabs. Res. Vet. Sci. 6: 178–188.

Hata, E., K. Nagai, and K. Murakami, 2017 Complete Genome Sequence of Mycoplasma bovirhinis Strain HAZ141_2 from Bovine Nasal Discharge in Japan. Genome Announc. 5: e01000–e01017. https://doi.org/10.1128/genomeA.01000-17

Higuchi, H., H. Iwano, K. Kawai, T. Ohta, T. Obayashi *et al.*, 2011 A simplified PCR assay for fast and easy mycoplasma mastitis screening in dairy cattle. J. Vet. Sci. 12(2): 191–193. https://doi.org/10.4142/jvs.2011.12.2.191

Hirose, K., H. Kobayashi, N. Ito, Y. Kawasaki, M. Zako *et al.*, 2003 Isolation of Mycoplasmas from nasal swabs of calves affected with respiratory diseases and antimicrobial susceptibility of their isolates. J. Vet. Med. B Infect. Dis. Vet. Public Health 50(7): 347–351. https://doi.org/10.1046/j.1439-0450.2003.00681.x

Hsiao, W., I. Wan, S. J. Jones, and F. S. Brinkman, 2003 IslandPath: aiding detection of genomic islands in prokaryotes. Bioinformatics 19(3): 418–420. https://doi.org/10.1093/bioinformatics/btg004

Kobayashi, H., K. Hirose, A. Worarach, P. Paugtes, N. Ito *et al.*, 1998 In vitro amplification of the 16S rRNA genes from Mycoplasma bovirhinis, Mycoplasma alkalescens and Mycoplasma bovigenitalium by PCR. J. Vet. Med. Sci. 60(12): 1299–1303. https://doi.org/10.1292/jvms.60.1299

Lagesen, K., P. Hallin, E. A. Rodland, H. H. Staerfeldt, T. Rognes *et al.*, 2007 RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35(9): 3100–3108. https://doi.org/10.1093/nar/gkm160

Lamm, C. G., L. Munson, M. C. Thurmond, B. C. Barr, and L. W. George, 2004 Mycoplasma otitis in California calves. J. Vet. Diagn. Invest. 16(5): 397–402. https://doi.org/10.1177/104063870401600505

Li, W., and A. Godzik, 2006 Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658–1659.

Li, Y., H. Zheng, Y. Liu, Y. Jiang, J. Xin *et al.*, 2011 The complete genome sequence of Mycoplasma bovis strain Hubei-1. PLoS One 6(6): e20999. https://doi.org/10.1371/journal.pone.0020999

Lowe, T. M., and S. R. Eddy, 1997 tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25(5): 955–964. https://doi.org/10.1093/nar/25.5.0955

Martin, W. B., 1978 Respiratory Diseases in Cattle, pp. 279–283 in Mycoplasmas Involved in Bovine Pneumonia, edited by Ernø, H., Springer, Dordrecht, Netherlands.

Miles, K., L. McAuliffe, R. D. Ayling, and R. A. Nicholas, 2004 Rapid detection of Mycoplasma dispar and M. bovirhinis using allele specific polymerase chain reaction protocols. FEMS Microbiol. Lett. 241(1): 103–107. https://doi.org/10.1016/j.femsle.2004.10.010

Moitinho-Silva, L., B. L. Heineck, L. A. Reolon, J. A. Paes, C. S. Klein *et al.*, 2012 Mycoplasma hyopneumoniae type I signal peptidase: expression and evaluation of its diagnostic potential. Vet. Microbiol. 154(3-4): 282–291. https://doi.org/10.1016/j.vetmic.2011.07.009

Mosier, D., 2014 Review of BRD pathogenesis: the old and the new. Anim. Health Res. Rev. 15(02): 166–168. https://doi.org/10.1017/S1466252314000176

Nair, S., E. Milohanic, and P. Berche, 2000 ClpC ATPase Is Required for Cell Adhesion and Invasion of Listeria monocytogenes. Infect. Immun. 68(12): 7061–7068. https://doi.org/10.1128/IAI.68.12.7061-7068.2000

Nandi, T., C. Ong, A. P. Singh, J. Boddey, T. Atkins *et al.*, 2010 A genomic survey of positive selection in Burkholderia pseudomallei provides insights into the evolution of accidental virulence. PLoS Pathog. 6(4): e1000845. https://doi.org/10.1371/journal.ppat.1000845

O'Riordan, M., M. A. Moors, and D. A. Portnoy, 2003 Listeria intracellular growth and virulence require host-derived lipoic acid. Science 302(5644): 462–464. https://doi.org/10.1126/science.1088170

Panangala, V. S., J. S. Stringfellow, J. L. Oliver, and D. F. Wolfe, 1990 Isolation of Mycoplasma bovirhinis from the kidneys of a bull with urinary obstruction and subacute nephritis. J. Am. Vet. Med. Assoc. 197: 381–382.

Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen, 2011 SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8(10): 785–786. https://doi.org/10.1038/nmeth.1701

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling *et al.*, 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61(3): 539–542. https://doi.org/10.1093/sysbio/sys029

Saha, S., S. Bridges, Z. V. Magbanua, and D. G. Peterson, 2008 Empirical comparison of ab initio repeat finding programs. Nucleic Acids Res. 36(7): 2284–2294. https://doi.org/10.1093/nar/gkn064

Schmidt, H., and M. Hensel, 2004 Pathogenicity islands in bacterial pathogenesis. Clin. Microbiol. Rev. 17(1): 14–56. https://doi.org/10.1128/CMR.17.1.14-56.2004

Scotti, P. A., M. L. Urbanus, J. Brunner, J. W. de Gier, H. G. Von *et al.*, 2000 YidC, the Escherichia coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase. EMBO J. 19(4): 542–549. https://doi.org/10.1093/emboj/19.4.542

Seymour, L. M., C. Jenkins, A. T. Deutscher, B. B. Raymond, M. P. Padula *et al.*, 2012 Mhp182 (P102) binds fibronectin and contributes to the recruitment of plasmin(ogen) to the Mycoplasma hyopneumoniae cell surface. Cell. Microbiol. 14(1): 81–94. https://doi.org/10.1111/j.1462-5822.2011.01702.x

Soehnlen, M. K., S. Kariyawasam, J. A. Lumadue, T. A. Pierre, D. R. Wolfgang *et al.*, 2011 Molecular epidemiological analysis of Mycoplasma bovis isolates from the Pennsylvania Animal Diagnostic Laboratory showing genetic diversity. J. Dairy Sci. 94(4): 1893–1899. https://doi.org/10.3168/jds.2010-3967

Song, Z., Y. Li, Y. Liu, J. Xin, X. Zou *et al.*, 2012 alpha-Enolase, an adhesion-related factor of Mycoplasma bovis. PLoS One 7(6): e38836. https://doi.org/10.1371/journal.pone.0038836

Spring, T. G., and F. Wold, 1971 The purification and characterization of Escherichia coli enolase. J. Biol. Chem. 246: 6797–6802.

Srikumaran, S., C. L. Kelling, and A. Ambagala, 2007 Immune evasion by pathogens of bovine respiratory disease complex. Anim. Health Res. Rev. 8(02): 215–229. https://doi.org/10.1017/S1466252307001326

Su, J., H. Gong, J. Lai, A. Main, and S. Lu, 2009 The potassium transporter Trk and external potassium modulate Salmonella enterica protein secretion and virulence. Infect. Immun. 77(2): 667–675. https://doi.org/10.1128/IAI.01027-08

ter Laak, E. A., J. G. Tully, H. H. Noordergraaf, D. L. Rose, P. Carle *et al.*, 1993 Recognition of Mycoplasma canis as part of the mycoplasmal flora of the bovine respiratory tract. Vet. Microbiol. 34(2): 175–189. https://doi.org/10.1016/0378-1135(93)90171-3

ter Laak, E. A., J. H. Noordergraaf, and R. P. Dieltjes, 1992 Prevalence of mycoplasmas in the respiratory tracts of pneumonic calves. Zentralbl. Veterinarmed. B. 39: 553–562.

Ulgen, M., C. Cetin, S. Sentürk, A. E. Ozel, and U. Ozdemir, 2006 Urinary tract infections due to Mycoplasma canis in dogs. J. Vet. Med. A Physiol. Pathol. Clin. Med. 53: 379–382. https://doi.org/10.1111/j.1439-0442.2006.00858.x

Vasconcelos, A. T., H. B. Ferreira, C. V. Bizarro, S. L. Bonatto, M. O. Carvalho *et al.*, 2005 Swine and poultry pathogens: the complete genome sequences of two strains of Mycoplasma hyopneumoniae and a strain of Mycoplasma synoviae. J. Bacteriol. 187(16): 5568–5577. https://doi.org/10.1128/JB.187.16.5568-5577.2005

Vilei, E. M., and J. Frey, 2001 Genetic and biochemical characterization of glycerol uptake in mycoplasma mycoides subsp. mycoides SC: its impact on H(2)O(2) production and virulence. Clin. Diagn. Lab. Immunol. 8: 85–92. https://doi.org/10.1128/CDLI.8.1.85-92.2001

Zheng, L., A. Khemlani, N. Lorenz, J. M. Loh, R. J. Langley *et al.*, 2015 Streptococcal 5′-Nucleotidase A (S5nA), a Novel Streptococcus pyogenes Virulence Factor That Facilitates Immune Evasion. J. Biol. Chem. 290(52): 31126–31137. https://doi.org/10.1074/jbc.M115.677443

Communicating editor: D. J. de Koning