High quality draft genome sequence of *Mycoplasma testudineum* strain BH29\(^T\), isolated from the respiratory tract of a desert tortoise

Chava L. Weitzman, Richard L. Tillett, Franziska C. Sandmeier, C. Richard Tracy and David Alvarez-Ponce*

**Abstract**

*Mycoplasma testudineum* is one of the pathogens that can cause upper respiratory tract disease in desert tortoises, *Gopherus agassizii*. We sequenced the genome of *M. testudineum* BH29\(^T\) (ATCC 700618\(^T\) = MCCM 03231\(^T\)), isolated from the upper respiratory tract of a Mojave desert tortoise with upper respiratory tract disease. The sequenced draft genome, organized in 25 scaffolds, has a length of 960,895 bp and a G + C content of 27.54%. A total of 788 protein-coding sequences, six pseudogenes and 35 RNA genes were identified. The potential presence of cytadhesin-encoding genes is investigated. This genome will enable comparative genomic studies to help understand the molecular bases of the pathogenicity of this and other *Mycoplasma* species.

**Keywords:** *Mycoplasma testudineum*, Desert tortoise, *Gopherus*, Upper respiratory tract disease, URTD

**Organism information**

**Classification and features**

*M. testudineum* infects the upper respiratory tracts of tortoises causing upper respiratory tract disease [3, 4]; however, recent investigations in wild tortoises suggest it may be present in the host without pathogenicity [11]. This microbe has been found in five tortoise species inhabiting North America—*G. agassizii*, *G. morafkai*, *G. evgoodei*, *G. berlandieri*, and *G. polyphemus* [3, 11–13]—and its presence has yet to be investigated in the sixth tortoise congener, *G. flavomarginatus* (located in north-central Mexico). From wild samples, there is some indication that *M. testudineum* may have a facilitative relationship with *M. agassizii* in tortoise hosts, but interactions with other community members are unknown [11].

*M. testudineum* is a sugar-fermenting, coccoid *Mycoplasma*, which is very similar in phenotype to the closely-related *M. agassizii* [3] (Table 1, Fig. 1).
**M. testudineum** grows in culture at 22–30°C, with an optimal growth at 30°C [3] (Table 1). These temperatures are frequently experienced in their hosts during the seasons when tortoises are found to be most active [14, 15], though tortoise body temperatures can fluctuate well above or below these temperatures within a day and over the seasons [14–16].

To determine the placement of *M. testudineum* in the mycoplasmal phylogeny, all 16S rRNA gene sequences from the type strains of *Mycoplasma* species were obtained from the SILVA database [17] and aligned using MUSCLE 3.8.31 [18], and a phylogenetic tree was constructed using the maximum likelihood method implemented in MEGA7 [19] (Fig. 2). *M. agassizii* is a sister group of *M. testudineum* in the resultant tree, and the *M. testudineum*/M. agassizii clade is a sister group of *Mycoplasma pulmonis*—the agent of murine respiratory mycoplasmosis, which also seems to be present in humans who are in contact with rodents [20]. All three species fall within the hominis group of *Mycoplasma* (see ref. [21] for group definitions). The *M. testudineum* 16S rRNA gene sequence is 93.1 and 89.2% identical to those of *M. agassizii* and *M. pulmonis*, respectively. Remarkably, these species are not closely related to *Mycoplasma testudinis*, isolated from the cloaca of a spur-thighed tortoise (*Testudo graeca*) in the UK [22], which are placed in the pneumoniae group. A previous taxonomic analysis placed *M. testudinis* within the pneumoniae group (in agreement with our results), but placed *M. testudineum* and *M. agassizii* in different hominis subgroups: the hyorhinis and the fermentans groups, respectively [23]. Our result is, however, in agreement with that by Volokhov et al. [24], which was also based on 16S rRNA data.

### Genome sequencing information

#### Genome project history

The type strain of *M. testudineum*, strain BH29<sup>T</sup>, was selected for sequencing. This strain was isolated from a nasal flush of the choana of a Mojave desert tortoise, which was filtered through a 0.45 μm filter and then grown in SP4 broth [2, 3]. Sequencing was conducted in October 2016. The Whole Genome Shotgun project was deposited at DDBJ/ENA/GenBank under the accession number NNCE00000000. The version described in this paper is the first version, NNCE01000000. A summary of the project information in compliance with MIGS version 2.0 [25] is shown in Table 2.

#### Growth conditions and genomic DNA preparation

Freeze-dried *M. testudineum*, strain BH29<sup>T</sup>, was obtained from the ATCC in November 2014 (ATCC 700618<sup>T</sup>) and had been cultured by the ATCC on Spiroplasma SP4 medium at 30°C in aerobic conditions. Genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit protocol for Gram-negative bacteria and eluted with ultra-pure water. Extracted DNA was quantified on a Qiagen QIAxpert system and with Picogreen analysis.

#### Genome sequencing and assembly

Genome sequencing was conducted using the Illumina Nextera XT DNA Library Preparation Kit (Illumina, Inc., San Diego, USA) with the Illumina NextSeq500 platform (150 bp, paired-end) and 2 ng of starting genomic DNA at the Nevada Genomics Center (University of Nevada, Reno). Sequencing was performed in multiplex with multiple samples, using dual index sequences from the Illumina Nextera XT Index Kit, v2 (index 1, N701; index 2, S502). A total of 455,422 read pairs were obtained. Using Trimmomatic,
version 0.36 [26], reads were trimmed to remove Nextera adapter sequences and low quality nucleotides from either end (average Phred score Q ≤ 5, four bp sliding window), and sequences trimmed to < 35 bp were removed. After trimming, 412,763 read pairs and 36,907 single-reads (the pairs of which were removed) remained. De novo genome assembly was performed using SPAdes 3.10.1 [27], using as inputs the trimmed paired reads, and the trimmed single reads (assembly k-mer sizes 21, 33, 55, and 77; with read error-correction enabled and ‘-careful’ mode mismatch correction). After removing scaffolds of less than 500 bp, the final assembly consisted of 25 scaffolds, with a total length of 960,895 bp, an average length of 38,435 bp, and an N50 of 130,815 bp. The coverage was 64×.

**Genome annotation**

Gene prediction was carried out using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) 4.2 [28]. For each predicted protein, (i) families were identified using the Pfam 31.0 [29] batch search tool (“gathering threshold” option), (ii) COG categories were assigned using eggNOG-mapper [30] based on eggNOG 4.5.1 data [31], (iii) signal peptides were identified using the SignalP server 4.1 [32], and (iv) transmembrane helices were inferred using the TMHMM server v. 2.0 [33]. CRISPR repeats were identified using PGAP and CRISPRFinder [34].

**Genome properties**

The properties of the draft genome are summarized in Table 3. The final assembly consisted of 25 scaffolds, with a total length of 960,895 bp and a G + C content of 27.54%. The small genome size and low G + C content is consistent with those of other *Mycoplasma* genomes sequenced [35, 36]. PGAP [28] identified a total of 788 protein-coding genes, 6

| MIGS ID | Property | Term | Evidence code<sup>a</sup> |
|---------|----------|------|--------------------------|
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |
|         |          |      |                          |

<sup>a</sup>Evidence codes - IDA Inferred from Direct Assay, TAS Traceable Author Statement (i.e., a direct report exists in the literature), NAS Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [55]
Fig. 2 (See legend on next page.)
pseudogenes, and 35 RNA genes. The identified RNA genes include 3 rRNAs (one 5S, one 16S and one 23S), 3 ncRNAs and 29 tRNAs. PGAP identified 4 CRISPR repeats, and CRISPRFinder [34] identified 4 “confirmed” repeats, and another 3 that were flagged as “questionable” by the server. The numbers of protein-coding genes in each COG category [37] are summarized in Table 4.

Insights from the genome sequence

Brown et al. [3] sequenced most of the 16S rRNA gene of *M. testudineum* strain BH297 (GenBank ID: AY366210). They had previously sequenced the homologous region for *M. testudineum* strain H3110, which differed only in one nucleotide position (GenBank ID: U19768, ref. [23]). Comparison of their BH297 sequence and that obtained by us revealed 5 point differences and an indel of 14 nucleotides (present in Brown et al.’s sequence but not in ours) (Fig. 3). Remarkably, 4 of the 5 point differences were located toward the ends of Brown et al.’s sequence, and thus may represent sequencing errors. The other differences probably represent mutations accumulated since the isolation of the strain in 1995. Our 16S rRNA gene sequence is identical to that generated by Volokhov et al. [24], with the exception of the first nucleotide of Volokhov et al.’s sequence. Nevertheless, the placement of *M. testudineum* in the tree (Fig. 2) is not affected by the particular sequence used.

In general, *Mycoplasma* cells need to adhere to mucosal epithelial cells of the hosts as a pre-requisite for pathogenesis. The mechanisms of adhesion are relatively well understood in *Mycoplasma pneumoniae* and its close relatives, but much less so in other *Mycoplasma* groups [38]. We used BLASTP and TBLASTN (E < 10\(^{-5}\); low-complexity regions filtered out) to search for homologs of *M. pneumoniae* cytadhesins P1, P30, P65, P40 and P90 —proteins involved in adhesion—and cytadhesin accessory proteins Hmw1, Hmw2 and Hmw3 in all available *Mycoplasma* genomic data (nr database). We only found homologs in species closely related to *M. pneumoniae* (*Mycoplasma genitalium*, *Mycoplasma gallisepticum*, *Mycoplasma pirum*, *Mycoplasma alvi*, *Mycoplasma imitans*, and *M. testudinis*), as previously noted [38, 39].

---

### Table 2

| Attribute                  | Value     | % of Total |
|----------------------------|-----------|------------|
| Genome size (bp)           | 960,895   | 100.00     |
| DNA coding (bp)            | 865,251\(^a\) | 90.05\(^c\) |
| DNA G+C (bp)               | 264,678   | 27.54\(^c\) |
| DNA scaffolds              | 25        | 100.00     |
| Total genes                | 829       | 100.00     |
| Protein coding genes       | 788       | 95.05\(^d\) |
| RNA genes                  | 35        | 4.22\(^d\) |
| Pseudo genes               | 6         | 0.72\(^d\) |
| Genes in internal clusters | –         | –          |
| Genes with function prediction | 370\(^b\) | 46.95\(^e\) |
| Genes assigned to COGs     | 539       | 68.40\(^e\) |
| Genes with Pfam domains    | 558       | 70.81\(^e\) |
| Genes with signal peptides | 78        | 9.90\(^e\) |
| Genes with transmembrane helices | 217 | 27.54\(^e\) |
| CRISPR repeats             | 4         | –          |

\(^a\)Protein-coding sequences, not including stop codons  
\(^b\)Proteins not annotated as "hypothetical protein" by PGAP  
\(^c\)Relative to genome size  
\(^d\)Relative to total number of genes  
\(^e\)Relative to protein-coding genes
against the *M. testudineum* BH29<sup>T</sup> proteome detected no hits, and none of the 788 predicted *M. testudineum* proteins contained any of the Pfam domains present in the *M. pneumoniae* cytadhesins and accessory proteins (domains “CytadhesinP1”, “Adhesin_P1”, “Cytadhesin_P30”, “MgpC” and “EAGR_box”). These observations may have at least three alternative explanations: (i) the adhesion proteins used by *M. pneumoniae* may be specific to its group, (ii) adhesion proteins evolve very fast, perhaps due to co-evolutionary races, thus hindering the detection of distant homologs, or (iii) *M. testudineum* may exhibit limited adhesion capabilities. In support of the first possibility, *M. pulmonis*, the most closely related species to the *M. testudineum/M. agassizii* clade (Fig. 2), is known to have adhesion mechanisms different from *M. pneumoniae*: *M. pneumoniae* exhibits a specialized attachment organelle, whereas *M. pulmonis* adhesion takes place by generalized interaction of the pathogen and the host cell membranes [40]. The adhesins of *M. pulmonis* are unknown. In support of the second scenario, putative cytadhesins identified in *M. pirum* and *M. gallisepticum* are only 26–29% identical at the amino acid level to those of *M. pneumoniae* [41, 42]. To extend our search, we obtained a list of known *Mycoplasma* adhesins from the UniProt database [43] (search: “Mycoplasma adhesin”). Again, BLASTP and TBLASTN searches (E < 10<sup>−5</sup>; low-complexity regions filtered out) against the *M. testudineum* BH29<sup>T</sup> proteome/genome did not identify any significant hits. *M. pneumoniae* proteins GAPDH and EF-Tu and *M. hominis* protein OppA have been reported to be adhesins in addition to their traditional functions [44–46]. We found homologs of all three proteins in *M. testudineum*. It should be noted, however, that this does not guarantee that these proteins act as adhesins in *M. testudineum*. For instance, whereas *M. pneumoniae* EF-Tu binds fibronectin [45], *M. genitalium* EF-Tu, which is 96% identical, does not [47]. The *M. testudineum* protein is only 70% identical to that of *M. pneumoniae*, and serine 343, proline 345, and threonine 357 (replacement of which significantly reduces the fibronectin binding of EF-Tu in *M. pneumoniae*;

---

**Table 4 Number of genes associated with general COG functional categories**

| Code | Value<sup>b</sup> | %age  | Description                                     |
|------|------------------|-------|------------------------------------------------|
| J    | 102              | 12.94 | Translation, ribosomal structure and biogenesis |
| A    | 0                | 0.00  | RNA processing and modification                |
| K    | 19               | 2.41  | Transcription                                  |
| L    | 52               | 6.60  | Replication, recombination and repair           |
| B    | 0                | 0.00  | Chromatin structure and dynamics                |
| D    | 4                | 0.51  | Cell cycle control, Cell division, chromosome partitioning |
| V    | 18               | 2.28  | Defense mechanisms                              |
| T    | 4                | 0.51  | Signal transduction mechanisms                  |
| M    | 8                | 1.02  | Cell wall/membrane biogenesis                   |
| N    | 0                | 0.00  | Cell motility                                  |
| U    | 10               | 1.27  | Intracellular trafficking and secretion         |
| O    | 24               | 3.05  | Posttranslational modification, protein turnover, chaperones |
| C    | 35               | 4.44  | Energy production and conversion                |
| G    | 66               | 8.38  | Carbohydrate transport and metabolism           |
| E    | 29               | 3.68  | Amino acid transport and metabolism             |
| F    | 25               | 3.17  | Nucleotide transport and metabolism             |
| H    | 14               | 1.78  | Coenzyme transport and metabolism               |
| I    | 8                | 1.02  | Lipid transport and metabolism                  |
| P    | 37               | 4.70  | Inorganic ion transport and metabolism          |
| Q    | 1                | 0.13  | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 0                | 0.00  | General function prediction only                |
| S    | 89               | 11.29 | Function unknown                                |
| –    | 249              | 31.60 | Not in COGs                                    |

Percentages are based on the total number of protein coding genes in the genome

*a* COG category code  

*b* Number of genes in the category
ref. [47]) are not conserved in *M. testudineum*. Additional work will be required to understand the mechanisms of adhesion in *M. testudineum* and its close relatives.

**Conclusions**

We have obtained a draft genome sequence of *M. testudineum* BH297 isolated from the upper respiratory tract of a desert tortoise with URTD in the Mojave Desert. Our analysis revealed some features typical of *Mycoplasma* genomes: a very small size and low G+C content. The new genome will enable comparative genomic studies to help understand the molecular bases of the pathogenicity of this and other *Mycoplasma* species.

**Abbreviations**

ATCC: American Type Culture Collection; BLAST: Basic local alignment search tool; COG: Clusters of Orthologous Groups; EF-Tu: Elongation factor Tu; GAPDH: Glyceraldehyde-3-phosphate dehydrogenase; MIGS: Minimum information on the genome sequence; NCBI: National Center for Biotechnology Information; OppA: Substrate-binding domain of the oligopeptide permease

**Acknowledgements**

The authors are very grateful to Kris Kruse from the Nevada Genomics Center for technical assistance, and to Marco Fondi for helpful discussions. They are also grateful to the Nevada Genomics Center for providing sequencing services for free.

**Funding**

This work was made possible by a grant from the National Institute of General Medical Sciences (P20GM103440) from the National Institutes of Health. The funder did not play any role in the study.
Authors’ contributions
CLW, FCS, CRT and DAP conceived the work. CLW conducted laboratory work. RLT and DAP conducted bioinformatic analyses. CLW and DAP drafted the manuscript. All authors contributed to interpreting data and improving the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Biology Department, University of Nevada, Reno, USA. 2Nevada Center for Bioinformatics, University of Nevada, Reno, USA. 3Biology Department, Colorado State University, Pueblo, USA.

Received: 20 October 2017 Accepted: 21 March 2018
Published online: 11 April 2018

References
1. Brown DR. Mycoplasmosis and immunity of fish and reptiles. Front Biosci. 2002;7:1338–46.
2. Brown MB, Schumacher IM, Klein PA, Harris K, Correll T, Jacobson ER. Mycoplasma agassizii causes upper respiratory tract disease in the desert tortoise. Infect Immun. 1994;62(10):4580–6.
3. Brown D, Merritt J, Jacobson E, Klein P, Tully J, Brown M. Mycoplasma testudinum sp. nov., from a desert tortoise (Gopherus agassizii) with upper respiratory tract disease. Int J Evid Microbiol. 2004;54(5):1527–9.
4. Jacobson ER, Berry KH. Mycoplasma testudinum in free-ranging desert tortoises, Gopherus agassizii. J Wildl Dis. 2012;48(4):1063–8.
5. Brown MB, McLaughlin GS, Klein PA, Crenshaw BC, Schumacher IM, Brown DR, Jacobson ER. Upper respiratory tract disease in the gopher tortoise is caused by Mycoplasma agassizii. J Clin Microbiol. 1999;37:2262–9.
6. Desert Tortoise Recovery Team. Desert tortoise (Mojave population): recovery plan. Portland: US Fish and Wildlife Service; 1994.
7. Seigel RA, Smith RB, Seigel NA. Swine flu or 1918 pandemic? Upper respiratory tract disease and the sudden mortality of gopher tortoises (Gopherus polyphemus) on a protected habitat in Florida. J Herpetol. 2003;37(1):137–44.
8. Enge KM, Berish JE, Bolt R, Dzieworski A, Mushinsky HR. Biological status report: gopher tortoise. Tallahassee: Florida Fish and Wildlife Conservation Commission; 2006.
9. Sandmeier FC, Tracy CR, Hunter K. Upper respiratory tract disease (URTD) as a threat to desert tortoise populations: a reevaluation. Biol Conserv. 2009;142(7):1255–64.
10. Diemer Berish JE, Wendland LD, Kitte RA, Garrison EP, Gates CA. Effects of mycoplasmal upper respiratory tract disease on morbidity and mortality of gopher tortoises in northern and Central Florida. J Wildl Dis. 2010;46(3):695–705.
11. Weltzmann CL, Gov R, Sandmeier FC, Snyder SJ, Tracy CR. Co-infection does not predict disease in Gopherus tortoises. Royal Soc Open Sci. 2017;4:201710303.
12. Berry KH, Brown MB, Vaughn M, Gowan TA, Hasskamp MA, Torres MCM. Mycoplasma agassizii in Morafka's desert tortoise (Gopherus agassizii) in Florida. J Herpetol. 2003;37(1):137–44.
13. McGinnis SM, Voigt WG. Thermoregulation in the desert tortoise, Gopherus agassizii. Comp Biochem Physiol A Physiol. 1971;40(1):119–26.
14. Anderson NJ. The thermal biology of the gopher tortoise (Gopherus polyphemus) and the importance of microhabitat selection. MS dissertation, Hammond: Southeastern Louisiana University; 2001.
15. McGinnis SM, Voigt WG. Thermoregulation in the desert tortoise, Gopherus agassizii. Comp Biochem Physiol A Physiol. 1971;40(1):119–26.
16. Snyder SJ. Effects of fire on desert tortoise (Gopherus agassizii) thermal ecology. PhD dissertation. Reno: University of Nevada; 2014.
17. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41(Database issue):D590–6.
42. Keeler C, Hnatow LL, Whetzel PL, Dohms JE. Cloning and characterization of a putative cytadhesin gene (mgc1) from Mycoplasma gallisepticum. Infect Immun. 1996;64(5):1541–7.
43. Uniprot Consortium. UniProt: a hub for protein information. Nucleic Acids Res. 2015;43(Database issue):D204–12.
44. Dumke R, Hausner M, Jacobs E. Role of Mycoplasma pneumoniae glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in mediating interactions with the human extracellular matrix. Microbiology. 2011;157(Pt 8):2328–38.
45. D allo SF, Kannan TR, Blaylock MW, Baseman JB. Elongation factor Tu and E1 beta subunit of pyruvate dehydrogenase complex act as fibronectin binding proteins in Mycoplasma pneumoniae. Mol Microbiol. 2002;46:1041–51.
46. Henrich B, Hopfe M, Kitzerow A, Hadding U. The adherence-associated lipoprotein P100, encoded by an Opp operon structure, functions as the oligopeptide-binding domain OppA of a putative oligopeptide transport system in Mycoplasma hominis. J Bacteriol. 1999;181:4873–8.
47. Balasubramanian S, Kannan TR, Hart PJ, Baseman JB. Amino acid changes in elongation factor Tu of Mycoplasma pneumoniae and Mycoplasma genitalium influence fibronectin binding. Infect Immun. 2000;67(9):3533–41.
48. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87(12):4576–9.
49. Murray RGE. The higher taxa, or, a place for everything...? In: Krieg NR, Holt JG, editors. Bergey’s manual of systematic bacteriology, vol. 1. Baltimore: Williams & Wilkins; 1984. p. 31–4.
50. Edward DGF, Freundt E. Proposal for the class Mollicutes as name of the order established for the order Mycoplasmatales. Int J Syst Evol Microbiol. 1967;17(3):267–8.
51. Edward DGF, Freundt E. Type strains of species of the order Mycoplasmatales, including designation of new types for Mycoplasma mycoides subsp. mycoides, Mycoplasma agalactiae subsp. agalactiae, and Mycoplasma arthritidis. Int J Syst Evol Microbiol. 1973;23(1):55–61.
52. Freundt E. The classification of the pleuropneumonia group of organisms (Borellomyctes). Int J Syst Evol Microbiol. 1955;5(2):67–78.
53. Nowak J. Morphologie, nature et cycle évolutive du microbe de la pérpneumonie des bovidés. Ann Inst Pasteur (Paris). 1929;43(1330–52.
54. Freundt EA. The mycoplasmas. In: Buchanan RE, Gibbons NE, editors. Bergey’s manual of determinative bacteriology. 8th ed. Baltimore: Williams and Wilkins; 1974. p. 929–54.
55. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000;25(1):25–9.