

# Analysis of the Complete *Mycoplasma hominis* LBD-4 Genome Sequence Reveals Strain-Variable Prophage Insertion and Distinctive Repeat-Containing Surface Protein Arrangements

Michael J. Calcutt, Mark F. Foecking

Department of Veterinary Pathobiology, University of Missouri, Columbia, Missouri, USA

**The complete genome sequence of *Mycoplasma hominis* LBD-4 has been determined and the gene content ascribed. The 715,165-bp chromosome contains 620 genes, including 14 carried by a strain-variable prophage genome related to *Mycoplasma fermentans* MFV-1 and *Mycoplasma arthritidis* MAV-1. Comparative analysis with the genome of *M. hominis* PG21<sup>T</sup> reveals distinctive arrangements of repeat-containing surface proteins.**

Received 26 December 2014 Accepted 15 January 2015 Published 26 February 2015

**Citation** Calcutt MJ, Foecking MF. 2015. Analysis of the complete *Mycoplasma hominis* LBD-4 genome sequence reveals strain-variable prophage insertion and distinctive repeat-containing surface protein arrangements. *Genome Announc* 3(1):e01582-14. doi:10.1128/genomeA.01582-14.

**Copyright** © 2015 Calcutt and Foecking. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Michael J. Calcutt, calcuttm@missouri.edu.

*Mycoplasma hominis* is primarily considered a urogenital tract inhabitant (1), but multiple reports of extragenital infections highlight the multifarious tropisms and clinical manifestations of this organism (2, 3). The complete genome sequence of *M. hominis* PG21<sup>T</sup> (isolated from the female lower genital tract) has been determined (4), and a draft genome sequence of 22 contigs was recently released for blood sample isolate *M. hominis* LBD-4 as part of the one thousand microbial genomes project (5). This isolate was recovered from a septicemic patient in 1965 (6). To gain further insight into overall conservation of genomic organization between *M. hominis* isolates and to determine the variations in the repeated peptide sequences that are present in multiple large *M. hominis* surface proteins, the complete genome of *M. hominis* LBD-4 was determined.

Genomic DNA was prepared from a late-logarithmic-phase culture of *M. hominis* LBD-4 (obtained from the American Type Culture Collection as ATCC 27545), further purified using a PowerClean Pro DNA cleanup kit (Mo Bio Laboratories, CA), and sequenced using Pacific Biosciences chemistry at the National Center for Genome Resources, Santa Fe, NM. A fully assembled genome of 715,165 bp, with a read depth of 616×, was obtained using HGAP version 2 (7) on the reads generated from one single-molecule real-time (SMRT) cell. The complete genome has a 26.94% G+C content and is ~31 kb larger than the previously determined Illumina data set (22 contigs), largely due to the presence of repetitive sequences that were successfully resolved using Pacific Biosciences sequencing. Open reading frames (ORFs) and RNA features were identified using the Prokka tool (8), with the subsequent manual curation of the 620 genes, including the verification of mutations and deletions associated with 16 pseudogenes. The coding sequences comprised those for 563 ORFs, 6 rRNAs (two dispersed 16S-23S rRNA operons and two unlinked 5S rRNA genes), 33 tRNAs, and the *rnpB* and *ssrA* small RNAs.

A preliminary comparative analysis with the 665,445-bp genome of *M. hominis* PG21<sup>T</sup> disclosed colinear chromosomes

punctuated by ~78 kb of variation distributed among 20 loci (>250 bp); the largest region of difference was an approximately 15.9-kb prophage genome (designated MHOV-1) related to *Mycoplasma arthritidis* MAV-1 (9) and *Mycoplasma fermentans* MFV-1 (10). In *M. hominis* LBD-4, MHOV-1 comprises 14 similarly oriented ORFs and is inserted immediately downstream of an ABC transporter gene.

The use of long-read sequencing enabled the assembly of regions encoding large, peptide-repeat-harboring surface proteins. Size and sequence variations were evident for the multiple LMP repeat-containing surface proteins (4, 11) present in each strain. Furthermore, an expanded array of four genes encoding repeat-containing lipoproteins is present at a locus in *M. hominis* LBD-4; the corresponding site in strain PG21<sup>T</sup> encodes a single, albeit larger, repeat-containing lipoprotein gene (MHO\_2260).

The availability of a second complete *M. hominis* genome further informs our understanding of strain divergence in this taxon, with variable mobile genetic element repertoires and genes that are predicted to confer distinctive surface peptide displays encoded within a largely conserved genomic framework.

**Nucleotide sequence accession number.** This complete genome sequence has been deposited at DDBJ/EMBL/GenBank under the accession no. [CP009652](https://ncbi.nlm.nih.gov/nucl/CP009652).

## ACKNOWLEDGMENTS

This project was supported by funding from the University of Missouri.

We acknowledge the excellent data set provided by Nicholas Devitt and colleagues at the NCGR and thank Kim Wise and Qijing Zhang for fruitful discussions during this project. We thank Nikos Kyrpides and the Joint Genome Institute for their release policy for genomic data sets from the KMG project.

## REFERENCES

1. Taylor-Robinson D, McCormack WM. 1980. The genital mycoplasmas. *N Engl J Med* 302:1003–1010. <http://dx.doi.org/10.1056/NEJM198005013021805>.

2. Luttrell LM, Kanj SS, Corey GR, Lins RE, Spinner RJ, Mallon WJ, Sexton DJ. 1994. *Mycoplasma hominis* septic arthritis: two case reports and review. *Clin Infect Dis* 19:1067–1070. <http://dx.doi.org/10.1093/clinids/19.6.1067>.
3. Mitsani D, Nguyen MH, Silveira FP, Bermudez C, Toyoda Y, Pasculle AW, Clancy CJ. 2010. *Mycoplasma hominis* pericarditis in a lung transplant recipient: review of the literature about an uncommon but important cardiothoracic pathogen. *Transpl Infect Dis* 12:146–150. <http://dx.doi.org/10.1111/j.1399-3062.2009.00457.x>.
4. Pereyre S, Sirand-Pugnet P, Beven L, Charron A, Renaudin H, Barré A, Avenaoud P, Jacob D, Couloux A, Barbe V, de Daruvar A, Blanchard A, Bébéar C. 2009. Life on arginine for *Mycoplasma hominis*: clues from its minimal genome and comparison with other human urogenital mycoplasmas. *PLoS Genet* 5:e1000677. <http://dx.doi.org/10.1371/journal.pgen.1000677>.
5. Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, Whitman WB, Hugenholtz P, Klenk HP. 2014. Genomic encyclopedia of type strains, phase I: the one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci* 9:1278–1284. <http://dx.doi.org/10.4056/sigs.5068949>.
6. Tully JG, Brown MS, Sheagren JN, Young VM, Wolff SM. 1965. Septicemia due to *Mycoplasma hominis* type 1. *N Engl J Med* 273:648–650. <http://dx.doi.org/10.1056/NEJM196509162731207>.
7. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <http://dx.doi.org/10.1038/nmeth.2474>.
8. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.
9. Voelker LL, Dybvig K. 1999. Sequence analysis of the *Mycoplasma arthritidis* bacteriophage MAV1 genome identifies the putative virulence factor. *Gene* 233:101–107. [http://dx.doi.org/10.1016/S0378-1119\(99\)00150-X](http://dx.doi.org/10.1016/S0378-1119(99)00150-X).
10. Röske K, Calcutt MJ, Wise KS. 2004. The *Mycoplasma fermentans* prophage phiMFV1: genome organization, mobility and variable expression of an encoded surface protein. *Mol Microbiol* 52:1703–1720. <http://dx.doi.org/10.1111/j.1365-2958.2004.04087.x>.
11. Ladefoged SA, Jensen LT, Brock B, Birkelund S, Christiansen G. 1996. Analysis of 0.5-kilobase-pair repeats in the *Mycoplasma hominis* *lmp* gene system and identification of gene products. *J Bacteriol* 178:2775–2784.