JB Journals.ASM.org

GENOME ANNOUNCEMENT

Genome Sequence of "*Candidatus* Mycoplasma haemolamae" Strain Purdue, a Red Blood Cell Pathogen of Alpacas (*Vicugna pacos*) and Llamas (*Lama glama*)

Ana M. S. Guimaraes,[a,b] Balazs Toth,[c] Andrea P. Santos,[a] Naíla C. do Nascimento,[a] Janice E. Kritchevsky,[c] and Joanne B. Messick[a]

Department of Comparative Pathobiology, Purdue University, West Lafayette, Indiana, USA[a]; CAPES-Fulbright Program, Ministério da Educação, Brasília, Brazil[b]; and Department of Veterinary Clinical Sciences, Purdue University, West Lafayette, Indiana, USA[c]

We report the complete genome sequence of "*Candidatus* Mycoplasma haemolamae," an endemic red-cell pathogen of camelids. The single, circular chromosome has 756,845 bp, a 39.3% G+C content, and 925 coding sequences (CDSs). A great proportion (49.1%) of these CDSs are organized into paralogous gene families, which can now be further explored with regard to antigenic variation.

The mollicute "*Candidatus* Mycoplasma haemolamae" (8) is a highly endemic red blood cell pathogen of llamas (*Lama glama*) and alpacas (*Vicugna pacos*). "*Candidatus* Mycoplasma haemolamae" has been found in camelids from Europe and North and South America, with a herd prevalence that can be as high as 29% (1, 3, 6, 11). Both chronic and acute infections have been observed in the field. Although the majority of infected camelids have the asymptomatic chronic form of the infection (3, 6), crias and animals experiencing stress, such as parturition and gestation, may develop severe, life-threatening anemia (acute disease) (4–6). Moreover, "*Candidatus* Mycoplasma haemolamae" is known to persist in the host despite antibiotic treatment and/or an immune response (10). Mechanisms by which "*Candidatus* Mycoplasma haemolamae" evades the immune system and causes chronic disease, as well as those that trigger development of acute disease, are poorly understood. Therefore, the complete sequencing and assembly of the "*Candidatus* Mycoplasma haemolamae" genome was undertaken to gain new insights into the metabolism and pathogenesis of this bacterium.

"*Candidatus* Mycoplasma haemolamae" organisms were isolated from the blood of a naturally infected alpaca at peak of bacteremia following a splenectomy as previously described (7). "*Candidatus* Mycoplasma haemolamae" genomic DNA was extracted using a Quick-gDNA midiprep kit (Zymo Research Corp., Irvine, CA), and whole-genome sequencing was performed at the University of Notre Dame's Genomics Core Facility using Roche 454 GS-FLX and Titanium chemistry to sequence a genomic library according to rapid library prep protocol. Reads were assembled using gs-Assembler (v.2.3) (454 Life Sciences, Roche Applied Science), resulting in 7 scaffolds with GC content and/or homology similar to that of *Mycoplasma* spp. (from 1,118 to 731,011 bp). To facilitate gap closure, we used the same sample to obtain Illumina reads from a paired-end library (TruSeq DNA sample preparation kit; Illumina, San Diego, CA) using 20% of an Illumina v3 chemistry lane (HiScanSQ) at Purdue University's Genomics Core Facility. Illumina reads were subsequently assembled using ABySS-PE v1.2.7 utilizing 20% of the reads, with "kmer" set to 80 bases. Predicted scaffolds were used to manually close the gaps of the "*Candidatus* Mycoplasma haemolamae" genome. First-pass annotation was obtained using the NCBI pipeline.

The final "*Candidatus* Mycoplasma haemolamae" genome is typical of a mycoplasma and is composed of a single, circular chromosome of 756,845 bp with a GC content of 39.3%. A total of 925 protein coding sequences (CDSs) (961 genes) were identified, and the rRNA genes were found in single copies and unlinked; the 23S and the 5S are organized as an operon, and the 16S rRNA is located 131 kb downstream of these genes. Only 280 (30.3%) of the CDSs have known functions; 645 (69.7%) were classified as hypothetical proteins. Also, as observed in other hemoplasmas, a great proportion of these CDSs are organized in paralogous gene families (454/925; 49.1%). It is proposed that some of these families may be related to antigenic variation and bacterial persistence within the host (2, 9). Further detailed analyses of this genome will provide information about the "*Candidatus* Mycoplasma haemolamae" biology and pathogenicity mechanisms.

Nucleotide sequence accession number. The "*Candidatus* Mycoplasma haemolamae" strain Purdue sequence was deposited in GenBank under the accession number CP003731.1.

ACKNOWLEDGMENTS

This work was supported by an Indiana Clinical Translational Science Institute (Indiana CTSI) Core Facilities Pilot grant and the Morris Animal Foundation (grant D10FE-004).

We are grateful to the Genomic Core Facilities at the University of Notre Dame and Purdue University.

REFERENCES

1. **Crosse P, et al.** 2012. First detection of '*Candidatus* Mycoplasma haemolamae' infection in alpacas in England. Vet. Rec. **171**:71–74.
2. **Guimaraes AMS, et al.** 2011. Complete genome sequence of *Mycoplasma suis* and insights into its biology and adaption to an erythrocyte niche. PLoS One. **6**:e19574. doi:10.1371/journal.pone.0019574.
3. **Kaufmann C, Meli ML, Hofmann-Lehmann R, Riond B, Zanolari P.** 2011. Epidemiology of '*Candidatus* Mycoplasma haemolamae' infection in South American camelids in Central Europe. J. Camel. Sci. **4**:23–29.

Received 23 August 2012 Accepted 29 August 2012

Address correspondence to Ana M. S. Guimaraes, anamsg@purdue.edu, or Joanne B. Messick, jmessic@purdue.edu.

doi:10.1128/JB.01557-12

4. **Lascola K, Vandis M, Bain P, Bedenice D.** 2009. Concurrent infection with *Anaplasma phagocytophilum* and *Mycoplasma haemolamae* in a young alpaca. J. Vet. Intern. Med. **23**:379–382.
5. **McLaughlin BG, et al.** 1990. An Eperythrozoon-like parasite in llamas. J. Am. Vet. Med. Assoc. **197**:1170–1175.
6. **Meli ML, et al.** 2010. Development and application of a real-time TaqMan qPCR assay for the detection and quantification of '*Candidatus* Mycoplasma haemolamae' in South American camelids. Vet. Microbiol. **146**:290–294.
7. **Messick JB, Smith G, Berent L, Cooper S.** 2000. Genome size of *Eperythrozoon suis* and hybridization with 16S rRNA gene. Can. J. Microbiol. **46**:1082–1086.
8. **Messick JB, Walker PG, Raphael W, Berent L, Shi X.** 2002. '*Candidatus* Mycoplasma haemodidelphidis' sp. nov., '*Candidatus* Mycoplasma haemolamae' sp. nov. and *Mycoplasma haemocanis* comb. nov., haemotrophic parasites from a naturally infected opossum (*Didelphis virginiana*), alpaca (*Lama pacos*) and dog (*Canis familiaris*): phylogenetic and secondary structural relatedness of their 16S rRNA genes to other mycoplasmas. Int. J. Syst. Evol. Microbiol. **52**:693–698.
9. **Santos AP, et al.** 2011. Genome of *Mycoplasma haemofelis*, unraveling its strategies for survival and persistence. Vet. Res. **42**:102–118.
10. **Tornquist SJ, Boeder LJ, Cebra CK, Messick JB.** 2009. Use of a polymerase chain reaction assay to study response to oxytetracycline treatment in experimental '*Candidatus* Mycoplasma haemolamae' infection in alpacas. Am. J. Vet. Res. **70**:1102–1107.
11. **Tornquist SJ, Boeder L, Rios-Phillips C, Alarcon V.** 2010. Prevalence of *Mycoplasma haemolamae* infection in Peruvian and Chilean llamas and alpacas. J. Vet. Diagn. Invest. **22**:766–769.