JOURNAL OF BACTERIOLOGY, Jan. 2010, p. 379–380 Vol. 192, No. 1
0021-9193/10/$12.00 doi:10.1128/JB.01330-09

GENOME ANNOUNCEMENT

Complete Genome Sequence of *Anaplasma marginale* subsp. *centrale*▽

David R. Herndon,[1] Guy H. Palmer,[2] Varda Shkap,[3] Donald P. Knowles, Jr.,[1,2] and Kelly A. Brayton[2]*

Animal Diseases Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Pullman, Washington 99164[1]; Programs in Vector-Borne Diseases and Genomics, Department of Veterinary Microbiology and Pathology and School for Global Animal Health, Washington State University, Pullman, Washington 99164[2]; and Division of Parasitology, Kimron Veterinary Institute, Bet Dagan 50250, Israel[3]

Received 8 October 2009/Accepted 16 October 2009

***Anaplasma marginale* subsp. *centrale* is a naturally attenuated subtype that has been used as a vaccine for a century. We sequenced the genome of this organism and compared it to those of virulent *senso stricto A. marginale* strains. The comparison markedly narrows the number of outer membrane protein candidates for development of a safer inactivated vaccine and provides insight into the diversity among strains of *senso lato A. marginale*.**

Sir Arnold Theiler described *Anaplasma marginale* as the "cause of a specific tick-borne disease of cattle" in 1908 (14), providing the first identification of a rickettsial pathogen. Two years later, Theiler isolated a less virulent organism, which he designated *A. marginale* subtype *centrale* (15). This naturally attenuated strain has been used as a live vaccine to prevent severe disease due to *A. marginale senso stricto* strains for 100 years. Understanding the genetic similarities and differences between the vaccine strain and wild-type *A. marginale* strains will provide clues as to how the vaccine provides protection. To that end, we have sequenced the *A. marginale* subsp. *centrale* vaccine strain using a whole-genome shotgun sequencing strategy.

Genomic DNA, obtained from Kimron Veterinary institute, was fragmented by hydroshearing and ligated into pSmartLCKan (Lucigen). A total of 10,752 paired-end sequence reads (~6.5× coverage) were generated. Assembly with Phrap (www.phrap.org) resulted in 148 contigs. Closure was achieved by applying the genome walking method across gap-spanning subclones and genomic DNA amplicons. For polymorphic loci, the most frequently observed subclone sequence was selected.

Coding sequences (CDSs) in the single, circular, 1,206,806-bp chromosome were predicted using Glimmer2 and Glimmer3 (4, 5, 12). Annotation was as described previously for *A. marginale senso stricto* genomes (2, 3). There are 925 predicted CDSs, 19 pseudogenes, 37 tRNA genes, and a single set of rRNA genes in the genome. *A. marginale* subsp. *centrale* contains 10 putative genes not found in the closed-core genomes of *senso stricto* strains (3). Similarly, 18 genes found in *senso stricto* strains are absent from *A. marginale* subsp. *centrale*. This divergence is consistent with the subspecies nomenclature (15), but the findings do not resolve whether these genetic differences warrant classification of the vaccine strain as a distinct species within the genus *Anaplasma* (6).

The ability of live *A. marginale* subsp. *centrale* to protect against a diversity of *A. marginale* strains indicates that epitopes critical for protective immunity are broadly conserved (11). As immunity against *A. marginale* can be induced by immunization with purified outer membrane protein (OMP) complexes (8–10, 13), identification of OMPs conserved between *A. marginale* subsp. *centrale* and *senso stricto A. marginale* may narrow the vaccine candidate list. *A. marginale* OMPs cluster predominately into two protein superfamilies, major surface protein 1 (Msp1) and Pfam01617/Msp2 (2). Members of the Msp1 superfamily from *senso stricto* strains (1, 2) are not well conserved (e.g., Msp1a, Msp1b-1, Msp1b-2, and Mlp2 to Mlp4; 13 to 48% amino acid identity) or are nonexistent (e.g., the products of Msp1b partial genes 1 to 3) in *A. marginale* subsp. *centrale*, suggesting that immunity induced by the live vaccine strain is unlikely to be associated with the Msp1 superfamily.

Comparative analysis of the Pfam01617/Msp2 superfamily (2, 8) reveals both conservation and diversity. OpAG1 to OpAG3 and Msp4 are generally well conserved, while the family comprising Omp1 to Omp15 found in *senso stricto* strains (2, 3, 8) is reduced in *A. marginale* subsp. *centrale*: genes for the closely related proteins Omp7 to Omp9 are collapsed into a single CDS, and genes for homologs of Omp2, Omp3, Omp6, and Omp15 are missing. The OMP complex capable of inducing protective immunity contains 11 proteins (7, 8). By excluding those without homologs in the vaccine strain and the highly variable Msp2 and Msp3, the number of candidates is narrowed to six: four Msp2 superfamily members (Msp4, Omp1, Omp7, and OpAG2) and two non-superfamily members (AM779/ACIS557 and AM854/ACIS486). The degree of identity among these candidates from the vaccine strain and *senso stricto A. marginale* strains ranges from 63% (for OpAG2 proteins) to 88% (for Msp4 homologs). While the next steps in

* Corresponding author. Mailing address: Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA 99164-7040. Phone: (509) 335-6340. Fax: (509) 335-8529. E-mail: kbrayton@vetmed.wsu.edu.

▽ Published ahead of print on 23 October 2009.

vaccine development will require strain analysis for epitope conservation in these candidates and immunization trials to test in vivo efficacy, progress will be accelerated using the minimal candidate list defined by the comparative genomics approach.

Nucleotide sequence accession number. The genome sequence and annotation were deposited in GenBank with accession number CP001759.

This research was supported by BARD grant no. US-3315-02C, NIH grant no. R01 AI44005, USDA ARS-CRIS grant no. 5348-32000-027-03S, and the Wellcome Trust grant no. GR075800M.

REFERENCES

1. **Barbet, A. F., G. H. Palmer, P. J. Myler, and T. C. McGuire.** 1987. Characterization of an immunoprotective protein complex of *Anaplasma marginale* by cloning and expression of the gene coding for polypeptide Am105L. Infect. Immun. **55:**2428–2435.
2. **Brayton, K. A., L. S. Kappmeyer, D. R. Herndon, M. J. Dark, D. L. Tibbals, G. H. Palmer, T. C. McGuire, and D. P. Knowles, Jr.** 2005. Complete genome sequencing of *Anaplasma marginale* reveals that the surface is skewed to two superfamilies of outer membrane proteins. Proc. Natl. Acad. Sci. U. S. A. **102:**844–849.
3. **Dark, M. J., D. R. Herndon, L. S. Kappmeyer, M. P. Gonzales, E. Nordeen, G. H. Palmer, D. P. Knowles, Jr., and K. A. Brayton.** 2009. Conservation in the face of diversity: multistrain analysis of an intracellular bacterium. BMC Genomics **10:**16.
4. **Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23:**673–679.
5. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
6. **Dumler, J. S., A. F. Barbet, C. P. Bekker, G. A. Dasch, G. H. Palmer, S. C. Ray, Y. Rikihisa, and F. R. Rurangirwa.** 2001. Reorganization of genera in the families *Rickettsiaceae* and *Anaplasmataceae* in the order *Rickettsiales*: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int. J. Syst. Evol. Microbiol. **51:**2145–2165.
7. **Noh, S. M., K. A. Brayton, W. C. Brown, J. Norimine, G. R. Munske, C. M. Davitt, and G. H. Palmer.** 2008. Composition of the surface proteome of *Anaplasma marginale* and its role in protective immunity induced by outer membrane immunization. Infect. Immun. **76:**2219–2226.
8. **Noh, S. M., K. A. Brayton, D. P. Knowles, J. T. Agnes, M. J. Dark, W. C. Brown, T. V. Baszler, and G. H. Palmer.** 2006. Differential expression and sequence conservation of the *Anaplasma marginale msp2* gene superfamily outer membrane proteins. Infect. Immun. **74:**3471–3479.
9. **Palmer, G. H., A. F. Barbet, W. C. Davis, and T. C. McGuire.** 1986. Immunization with an isolate-common surface protein protects cattle against anaplasmosis. Science **231:**1299–1302.
10. **Palmer, G. H., F. R. Rurangirwa, K. M. Kocan, and W. C. Brown.** 1999. Molecular basis for vaccine development against the ehrlichial pathogen *Anaplasma marginale*. Parasitol. Today **15:**281–286.
11. **Pipano, E.** 1995. Live vaccines against hemoparasitic diseases in livestock. Vet. Parasitol. **57:**213–231.
12. **Salzberg, S. L., A. L. Delcher, S. Kasif, and O. White.** 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Res. **26:**544–548.
13. **Tebele, N., T. C. McGuire, and G. H. Palmer.** 1991. Induction of protective immunity by using *Anaplasma marginale* initial body membranes. Infect. Immun. **59:**3199–3204.
14. **Theiler, A.** 1910. Report of the government veterinary bacteriologist, 1908–1909, p. 7–64. Department of Agriculture, Union of South Africa, Johannesburg, South Africa.
15. **Theiler, A.** 1911. First report of the director of veterinary research, Union of South Africa, Johannesburg, South Africa, p. 7–46. Department of Agriculture, Union of South Africa, Johannesburg, South Africa.