NOTE *Bacteriology*

The Phylogenetic Position of *Anaplasma bovis* and Inferences on the Phylogeny of the Genus *Anaplasma*

Adrian Patalinghug YBAÑEZ[1,2], Mariko SASHIKA[3] and Hisashi INOKUMA[1]

[1]*Department of Clinical Veterinary Science, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080–8555, Japan*
[2]*Visayas State University, Visca, Baybay, Leyte 6521–A, Philippines*
[3]*Graduate School of Veterinary Medicine, Hokkaido University, Hokkaido 060–0818, Japan*

(Received 15 August 2013/Accepted 18 October 2013/Published online in J-STAGE 1 November 2013)

ABSTRACT. The present study aimed to determine the complete citrate synthase (*gltA*) and heat-shock protein (*groEL*) gene sequences of *Anaplasma bovis* and to infer phylogenetic relationships within the genus *Anaplasma*. Multiple alignments from single and concatenated sequences of the 16S rRNA, *gltA* and *groEL* genes of the genus *Anaplasma* were subjected to phylogenetic analyses. Percent identities of *A. bovis* nucleotide sequences were found highest with *A. phagocytophilum* in *gltA* (65.4%) and *groEL* (79.8%). Single gene phylogenetic tree results assumed similar phylogenetic positions within the genus *Anaplasma*, except for *A. bovis*. However, consensus and concatenated sequence phylogenetic trees showed similar results, revealing 2 subgroups within the genus.

KEY WORDS: 16S rRNA gene, *Anaplasma bovis*, citrate synthase gene (*gltA*), heat-shock operon gene (*groEL*), phylogeny.

doi: 10.1292/jvms.13-0411; *J. Vet. Med. Sci.* 76(2): 307–312, 2014

The genus *Anaplasma* currently recognizes 6 distinct species: *A. phagocytophilum*, *A. platys, A. marginale, A. centrale*, *A. ovis* and *A. bovis*. The reclassification was mainly based on the phylogenetic information derived from 16S rRNA (complete representation of all member species) and heat-shock operon or *groEL* (only selected member species) gene data sequences [6]. A recent study identified a potentially novel *Anaplasma* sp. in Japan (herein provisionally referred to as *Anaplasma* sp. Japan), which revealed phylogenetic divergence in the 16S rRNA, *gltA* and *groEL* genes from any recognized *Anaplasma* spp. [21]. However, the previous studies present paucity of information on whether the use of secondary structures was employed in their phylogenetic analyses, which is known to produce better tree resolution [15]. In addition, the widely accepted 16S rRNA gene based phylogenies are sometimes inconsistent [16], which is probably due to the propensity of the 16S rRNA gene to recombination/horizontal or lateral gene transfer phenomenon [1]. Therefore, other genes like the *gltA* [12] and *groEL* [11, 13] can be alternatively used to clarify phylogenetic relationships.

A. bovis infects circulating monocytes [5]. This particular species has been mainly analyzed only using the 16S rRNA gene [6]. Previous phylogenetic analyses of the genus *Anaplasma* using the *groEL* gene did not include yet *A. bovis* [6, 21] due to the unavailability of the data sequence during the time of analyses. The present study generally aimed to molecularly characterize and analyze *A. bovis* based on *gltA* and *groEL* genes and to infer phylogenetic relationships within the genus *Anaplasma* using individual and multi-locus approach (including the 16S rRNA gene). Phylogenetic analyses were performed with or without the consideration of secondary structures, using maximum likelihood (ML) and Bayesian Inference (BI) methods.

Blood sample from a feral raccoon (*Procyon lotor*) [18] in Hokkaido, herein referred to as R499, was used. The sample was previously tested to be 16S rRNA-positive for *A. bovis* (1,387 bp; GenBank accession number GU937020) and was stored at −30°C. The DNA was extracted and stored as previously described [21].

The designing of primers, determination of the partial *gltA* and *groEL* sequences of *A. bovis* by PCR, genome walking and DNA sequencing strategies were performed as described previously [21]. Primers used in the present study are shown in Table 1. The negative control used was double distilled water. Instead of using an *A. bovis* DNA, the positive control used was *A. platys*.

The *gltA* and *groEL* sequences were translated into deduced amino acids (dAA) and were manually trimmed to include only the sequence of interest (generally from the start to stop codon). Percent identities were computed as previously described [21]. Multiple sequence alignments (MSA) were performed as suggested by Hall [7] or by using PROMALS3D [15], which considers secondary structures for protein coding genes. Subsequent analyses with and without using the secondary structure information were performed using raxmlGUI [19] by general time reversible (GTR) model. Analyses by ML with prior best model testing using MEGA 5 [20] and by BI using MrBayes 3.2 [17] were also employed. For the protein coding genes, ML analyses were performed using MEGA 5 with prior best model test-

*CORRESPONDENCE TO: INOKUMA, H., Department of Clinical Veterinary Science, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080–8555, Japan.
e-mail: inokuma@obihiro.ac.jp

Table 1. Oligonucleotide sequences of primers used in this study

Primer name	Oligonucleotide (5'→3')	Reference
PCR amplification for a partial *gltA* sequence		
CS7F2	ATG**R***TAGAAAA**W***GCTGTTTT	[21]
HG1085R	ACTATACC**K***GAGTAAAAGTC	[10]
F1b	GATCATGA**R***CA**R***AATGCTTC	[9]
AnaCS1085R1	ACTATACC**K***GAGTAAAAR*TC	This study
PCR amplification for a partial *groEL*sequence		
EEgro1F	GAGTTCGACGGTAAGAAGTTCA	[3]
Anagro712R	CCGCGATCAAACTGCATACC	[21]
Anagro122F	AAATACGGT**W***GTCACGGG	This study
Anagro649R	CTTTCTTC**R***ACAGTTATAAG	This study
Genome walker gene-specific primers (*gltA*)		
ABgl-46R1	AATGCAGCTGCTCCCGCACTTAAGCAAGTA	This study
ABgl-1R2	GTTGAGCCCACCATTCTTACTGTAGATGTA	This study
ABgl-307F1	CAATATAGCGATCGCTATGGAAGAAATAGC	This study
ABgl-338 F2	TTGCAGGATGAATACTTCATAGAGAGGAAG	This study
Genome walker gene-specific primers (*groEL*)		
ABgr-361F1	TAAAGGCAAAAGAGGCTGTTCTTACGG	This study
ABgr-385F2	CGGC TCTT ATGT CCAT GAGA CGTG AA	This study
ABgr-1010GRF3	ACAGTGCATCTTCCAGCATAGAAAGTAG	This study
ABgr-1117F4	AGCTTTCAGGTGGTGTGGCTGTGCTGAAAG	This study
ABgr-1132F5	TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA	This study
ABgr-96R1	TTAGGTCCAGCAGTACATCCAACAGCA	This study
ABgr-30R2	GGATGTGTACTATCTCCCTAATGGCTTA	This study
ABgr-25R3	TCTCCCTAATGGCTTTATCCAACTTCTC	This study

*Degenerate primers: R=A or G, W=A or T, K=G or T

ing, while BI was performed using MrBayes 3.2 guided by the prior best model test results from MEGA 5. All analyses utilizing MEGA 5 were estimated using 100 bootstrap replications. Selected representative sequences from species which had available information on the 3 different genes were concatenated (herein referred to as "supermatrix"). Supermatrices from sequences were analyzed using MrBayes 3.2. Tree results from MrBayes and raxmlGUI were viewed using the FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). Phylogenetic tree inputs were generated from each MSA (with secondary structure guidance) that were concatenated for the supermatrix by BI and were used to construct of a majority-rule consensus tree using the Dendrosope [9]. Using the SplitsTree4 [8], the supermatrix was subjected to phylogenetic network analysis (by NeighborNet method) and was tested for the presence of recombination phenomenon using the PHI test.

Initially, partial *gltA* (382 bp) and *groEL* (385 bp) sequences of *A. bovis* from R499 were obtained. After the genome walking procedure and sequencing, the complete *glt*A (1,239 bp; JN588561) and *groE*L (1,644 bp; JN588562) sequences were determined. In comparison with other *Anaplasma* species, the *gltA* and *groEL* sequences of *A. bovis* were found closest to *A. phagocytophilum* (Table 2). Within the genus *Anaplasma*, species shared at least 60.8 and 60.4% identities in the nucleotide and dAA sequences of the *gltA* gene, respectively and 77.2 and 89.4% identities in the nucleotide and dAA of the *groEL* gene, respectively (data not shown).

In the 16S rRNA phylogenetic analyses, 2 subclades were seen: (1) a subclade containing *A. marginale, A. centrale* and *A. ovis* and (2) a subclade containing *A. phagocytophilum, A. platys, A. bovis* and *Anaplasma* sp. Japan (Fig. 1). *A. bovis* also frequently formed a cluster with *Anaplasma* sp. Japan. In the *gltA* phylogenetic analyses (Fig. 2), topologies revealed the 2 subclades observed in the 16S rRNA trees. In the *groEL* phylogenetic analyses, some positions within the genus *Anaplasma* changed depending on whether nucleotide or dAA sequences were used, but the 2 subclades were still frequently observed (Fig. 3). On the other hand, trees generated from the supermatrix also revealed the 2 subclades (Fig. 4).

A. bovis consistently formed a cluster with *A. phagocytophilum, A. platys* and *Anaplasma* sp. Japan in the 16S rRNA phylogenetic trees. This finding varied from the tree results of Dumler *et al.* [6], in which their 16S rRNA phylogenetic analysis placed *A. bovis* closer to *A. centrale* and *A. ovis*, but was similar to that of Ooshiro *et al.* [14] and Doan *et al.* [4], in which *A. bovis* formed a cluster with *A. phagocytophilum* and *A. platys*.

For the *gltA* and *groE*L gene phylogenetic analyses, the subclade groupings of the different taxa appear to be consistent when protein secondary structures were considered in the MSA construction, than when nucleotide sequences were used. The *groEL*-based trees generated in the present study also varied from those of Dumler *et al.* [6] as sequences of *A. bovis, A. ovis, A. centrale* and *A. platys* were not yet

Table 2. Percent identities of *Anaplasma bovis* from R499 with other *Anaplasma* spp. based on 16S rRNA, *gltA* and *groEL* gene sequences

Organism	% Identity*				
	16S rRNA	*gltA*		*groEL*	
	Nucleotide	*Nucleotide*	*Amino Acids*	*Nucleotide*	*Amino Acids*
Anaplasma phagocytophilum	96.7	65.4	62.4	79.8	91.4
	NC007797				
Anaplasma platys	97	60.8	61.1	79.6	92.9
	AY077619	AY077620		AF478129	
Anaplasma ovis	95.5	----		77.4	90.9
	AF309865			AF441131	
Anaplasma marginale	95.4	74.6	62.4	77.5	90.9
	NC012026				
Anaplasma centrale	95.5	73.6	60.4	77.2	90.9
	NC013532				
Anaplasma sp. Japan	96.5	63.7	61.6	79	89.4
	JN055357	JN55361		JN55359	

*Gaps were not considered in the computation. Note: the accession numbers of the sequences compared are indicated below the cells.

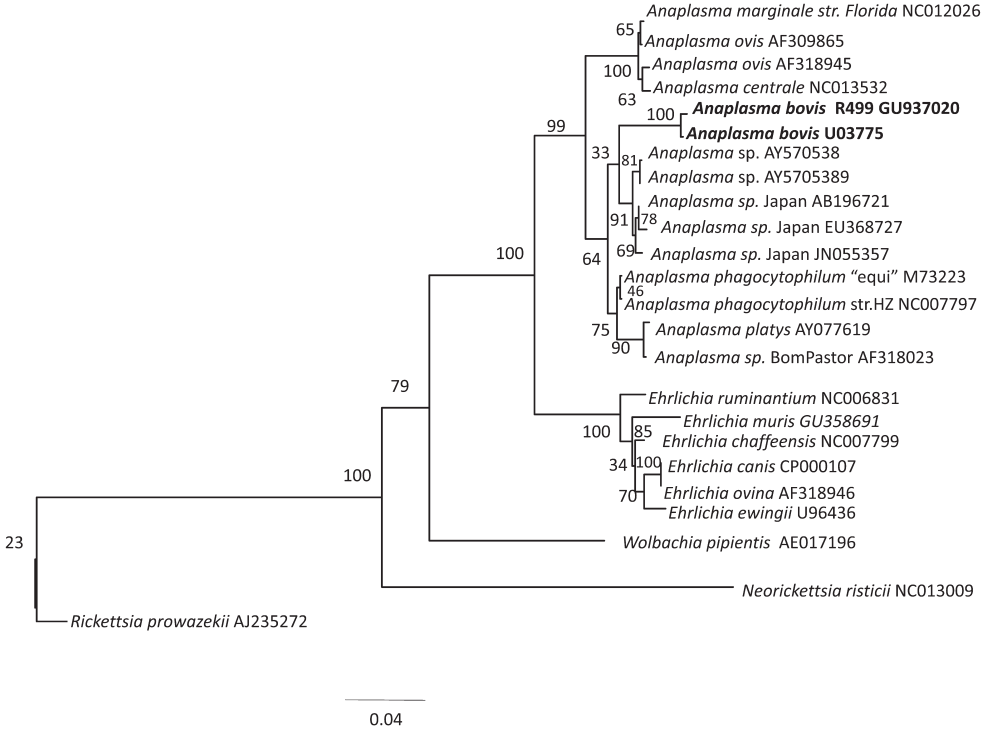

Fig. 1. Phylogenetic trees based on the 16S rRNA gene. Analyses were performed by maximum likelihood (general time reversible model) using 100 bootstrap replications employed in raxmlGUI [18]. *Rickettsia prowazekii* was set as the outgroup.

included in their analyses. The *groEL* sequences of *A. centrale, A. ovis* [13] and *A. platys* [12] were only determined at a later time. Dumler *et al.* [6] pointed out the ambiguities among Anaplasma spp. and the arbitrary position of *A. bovis* within the *Anaplasma* species clade in the various phylogenetic analyses they performed.

Comparing the single gene or the multi-loci phylogenetic trees, the consistently observed result was the formation of the 2 subclades when secondary structures were considered. Moreover, the resulting topologies corroborated with our

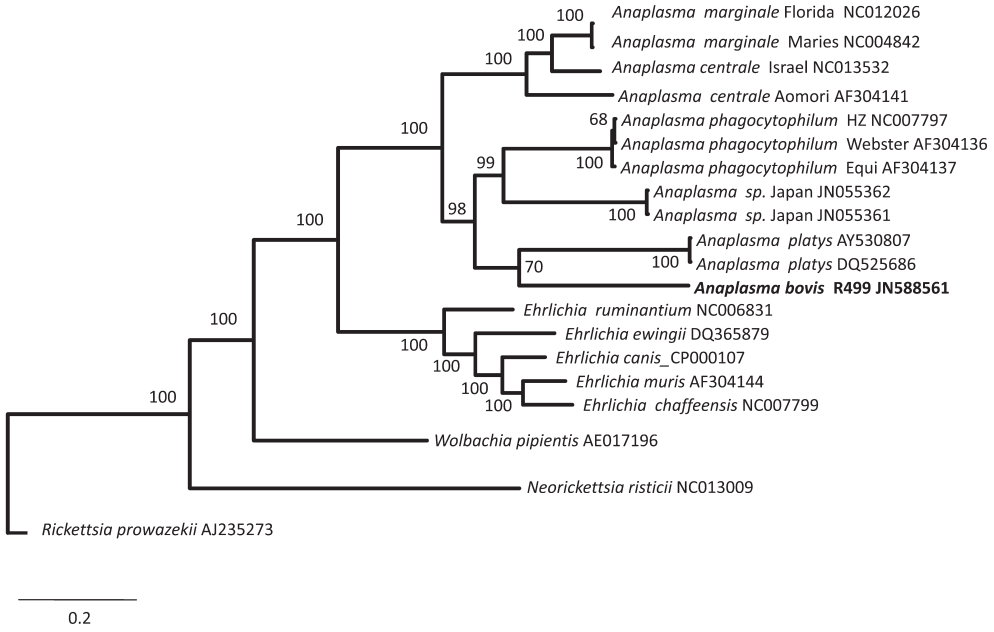

Fig. 2. Phylogenetic trees based on *gltA* with consideration of the protein secondary structures. Analyses were performed by the Bayesian method (Jones-Taylor-Thornton model) employed in MrBayes 3.2 [16]. Values in the nodes represent posterior probability values expressed in percent. *Rickettsia prowazekii* was set as the outgroup.

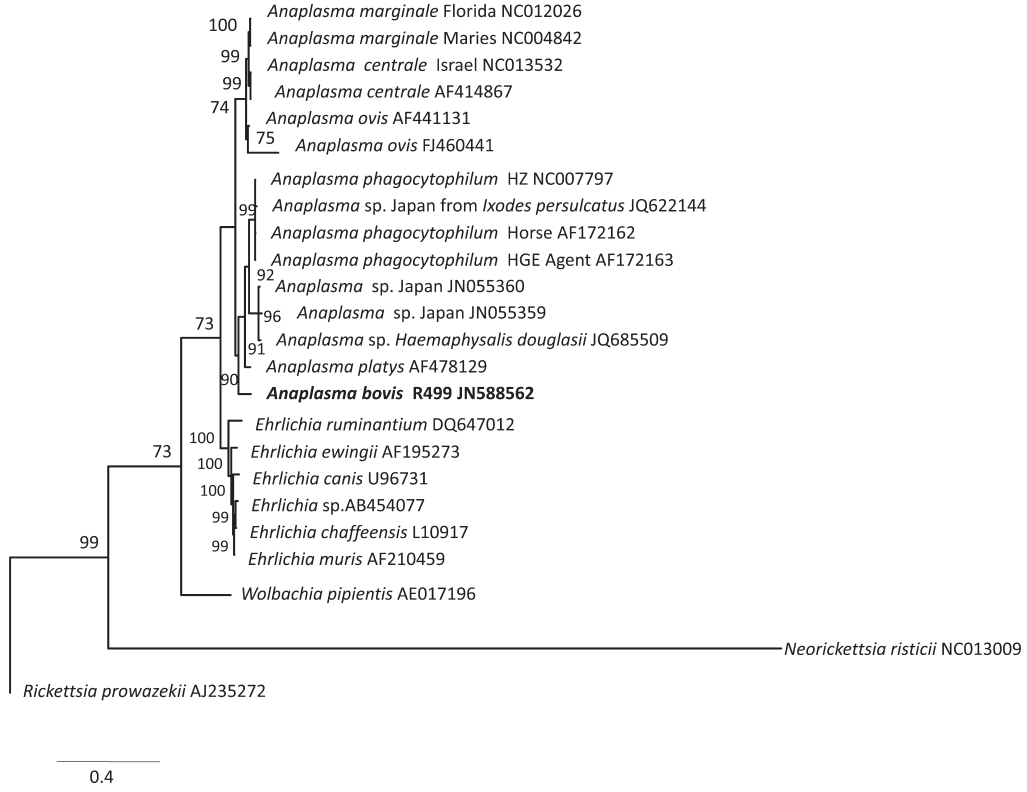

Fig. 3. Phylogenetic trees based on *groEL* genes with consideration of the protein secondary structures. Analyses were performed by the Bayesian method (Jones-Taylor-Thornton model) employed in MrBayes 3.2 [16]. Values in the nodes represent posterior probability values expressed in percent. *Rickettsia prowazekii* was set as the outgroup.

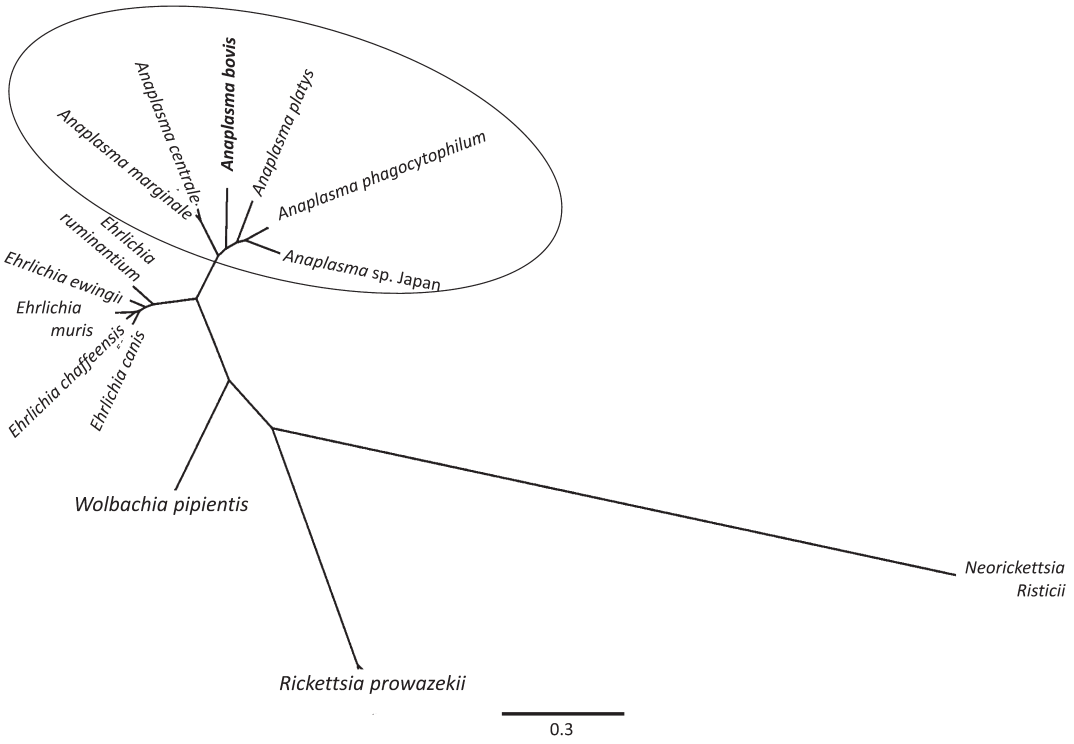

Fig. 4. Results of the concatenation approach by the Bayesian method using MrBayes 3.2 [16]. All nodes in the figure revealed posterior probability values of 1 or 100%. *Rickettsia prowazekii* was set as the outgroup.

previous findings [20], in which the *Anaplasma* sp. Japan was found to be a potentially novel species. The absence of statistical evidence of recombination (using PHI test) and the subsequent result of the phylogenetic network analysis (by NeighborNet method) on the concatenated alignment also supported the reliability of the tree results. PHI tests are used to test MSAs for the presence of recombination, which can obscure the results of phylogenetic analyses [2].

The present study documented the first molecular analyses of *A. bovis* based on complete *groEL* and *gltA* gene sequences and inferred phylogenetic relationships within the genus *Anaplasma* with the inclusion of new sequence data. Results clarified the phylogenetic position of *A. bovis* and established the existence of 2 subclades within the genus *Anaplasma*. This information can serve as a guide to future phylogenetic studies using the same genus.

ACKNOWLEDGMENTS. This study was supported by a grant for Research on Emerging and Re-emerging Infectious Diseases (H21-Shinkou-Ippan-06) from the Ministry of Health, Labor and Welfare, Japan. The authors also thank Mr. Fujisawa, T., Mr. Angeles, J. Ma. M., Dr. Hakimi, H. and Ms. Ybañez, R. H. D. for their technical assistance.

REFERENCES

1. Brayton, K. A., Palmer, G. H., Lundgren, A., Yi, J. and Barbet, A. F. 2002. Antigenic variation of *Anaplasma marginale* msp2 occurs by combinatorial gene conversion. *Mol. Microbiol.* **43**: 1151–1159. [Medline] [CrossRef]
2. Bruen, T. C., Philippe, H. and Bryant, D. 2006. A Simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**: 2665–2681. [Medline] [CrossRef]
3. Chae, J. S., Foley, J. E., Dumler, J. S. and Madigan, J. E. 2000. Comparison of the nucleotide sequences of 16S rRNA, 444 Ep-ank, and groESL heat shock operon genes in naturally occurring *Ehrlichia equi* and human granulocytic ehrlichiosis agent isolates from Northern California. *J. Clin. Microbiol.* **38**: 1364–1369. [Medline]
4. Doan, H. T., Noh, J. H., Choe, S. E., Yoo, M. S., Kim, Y. H., Reddy, K. E., Van Quyen, D., Nguyen, L. T., Nguyen, T. T., Kweon, C. H., Jung, S. C., Chang, K. Y. and Kang, S. W. 2013. Molecular detection and phylogenetic analysis of *Anaplasma bovis* from *Haemaphysalis longicornis* feeding on grazing cattle in Korea. *Vet. Parasitol.* **196**: 478–481. [Medline] [CrossRef]
5. Donatien, A. and Lestoquard, F. 1936. *Rickettsia bovis*, novelle espece pathogene pour le boeuf. *Bull. Soc. Pathol. Exot.* **29**: 1057–1061.
6. Dumler, J. S., Barbet, A. F., Bekker, C. P., Dasch, G. A., Palmer, G. H., Ray, S. C., Rikihisa, Y. and Rurangirwa, F. R. 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. [Medline] [CrossRef]
7. Hall, B. G. 2011. Phylogenetic Trees Made Easy: A How-To Manual, 4th ed., Sinauer Associates, Sunderland.

8. Huson, D. H. and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**: 254–267. [Medline] [CrossRef]
9. Huson, D. H., Richter, D. C., Rausch, C., Dezulian, T., Franz, M. and Rupp, R. 2007. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8**: 460. [Medline]
10. Inokuma, H., Brouqui, P., Drancourt, M. and Raoult, D. 2001. Citrate synthase gene sequence: a new tool for phylogenetic analysis and identification of *Ehrlichia*. *J. Clin. Microbiol.* **39**: 3031–3039. [Medline] [CrossRef]
11. Inokuma, H., Terada, Y., Kamio, T., Raoult, D. and Brouqui, P. 2001. Analysis of the 16S rRNA gene sequences of *Anaplasma centrale* and its phylogenetic relatedness to other *Ehrlichiae*. *Clin. Diag. Lab. Immunol.* **8**: 241–244.
12. Inokuma, H., Fujii, K., Okuda, M., Onishi, T., Beaufils, J. P., Raoult, D. and Brouqui, P. 2002. Determination of the nucleotide sequences of heat shock operon (*groESL*) and the citrate synthase gene (*gltA*) of *Anaplasma* (*Ehrlichia*) *platys* for phylogenetic and diagnostic studies. *Clin. Diagn. Lab. Immunol.* **9**: 1132–1136. [Medline]
13. Lew, A. E., Gale, K. R., Minchin, C. M., Shkap, V. and de Waal, D. T. 2003. Phylogenetic analysis of the erythrocytic *Anaplasma* species based on 16S rDNA and GroEL (HSP60) sequences of *A. marginale, A. centrale,* and *A. ovis* and the specific detection of *A. centrale* vaccine strain. *Vet. Microbiol.* **92**: 145–160. [Medline] [CrossRef]
14. Ooshiro, M., Zakimi, S., Matsukawa, Y., Katagiri, Y. and Inokuma, H. 2008. Detection of *Anaplasma bovis* and *Anaplasma phagocytophilum* from cattle on Yonaguni Island, Okinawa, Japan. *Vet. Parasitol.* **154**: 360–364. [Medline] [CrossRef]
15. Pei, J., Kim, B. H. and Grishin, N. V. 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**: 2295–2300. [Medline] [CrossRef]
16. Reischl, U., Feldmann, K., Naumann, L., Gaugler, B. J. M., Ninet, B., Hirschel, B. and Emler, S. 1998. 16S rRNA sequence diversity in *Mycobacterium celatum* strains caused by presence of two different copies of 16S rRNA gene. *J. Clin. Microbiol.* **36**: 1761–1764. [Medline]
17. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A. and Huelsenbeck, J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**: 539–542. [Medline] [CrossRef]
18. Sashika, M., Abe, G., Matsumoto, K. and Inokuma, H. 2011. Molecular survey of *Anaplasma* and *Ehrlichia* infections of feral raccoons (*Procyon lotor*) in Hokkaido, Japan. *Vector Borne Zoonotic Dis.* **11**: 349–354. [Medline] [CrossRef]
19. Silvestro, D. and Michalak, I. 2012. raxmlGUI: a graphical front-end for RAxML. *Org. Divers. Evol.* **12**: 335–337.
20. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**: 2731–2739. [Medline] [CrossRef]
21. Ybañez, A. P., Matsumoto, K., Kishimoto, T. and Inokuma, H. 2012. Molecular analyses of a potentially novel *Anaplasma* species closely related to *Anaplasma phagocytophilum* detected in sika deer (*Cervus nippon yesoensis*) in Japan. *Vet. Microbiol.* **157**: 232–236. [Medline] [CrossRef]