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

Adrian Patalinghug YBAÑEZ1,2, Mariko SASHIKA3 and Hisashi INOKUMA1

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.


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-
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 genus Anaplasma gene-specific primers (gltA).

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Oligonucleotide (5'→3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABgl-361F1</td>
<td>TAAAGGCCAAGAGAGGCTGTTCTTACGG</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-385F2</td>
<td>CGGC TCTT ATGT CCAT GAGA CGTG AA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-433R1</td>
<td>AATGCAGCTGCTCCCGCACTTAAGAAGTA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-1117F4</td>
<td>ACAGTAAAAGTC [10]</td>
<td></td>
</tr>
<tr>
<td>ABgl-96R1</td>
<td>GTCACGGG [21]</td>
<td></td>
</tr>
<tr>
<td>ABgl-338F2</td>
<td>TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-1132F5</td>
<td>TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-338F2</td>
<td>TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-361F1</td>
<td>TAAAGGCCAAGAGAGGCTGTTCTTACGG</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-433R1</td>
<td>AATGCAGCTGCTCCCGCACTTAAGAAGTA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-1117F4</td>
<td>ACAGTAAAAGTC [10]</td>
<td></td>
</tr>
<tr>
<td>ABgl-96R1</td>
<td>GTCACGGG [21]</td>
<td></td>
</tr>
<tr>
<td>ABgl-338F2</td>
<td>TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-1132F5</td>
<td>TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA</td>
<td>This study</td>
</tr>
<tr>
<td>ABgl-96R1</td>
<td>GTCACGGG [21]</td>
<td></td>
</tr>
<tr>
<td>ABgl-338F2</td>
<td>TGGCTGTGCTGAAAGTTGGTGGATCAAGTGA</td>
<td>This study</td>
</tr>
</tbody>
</table>

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

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 groEL 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
included in their analyses. The groEL sequences of *A. centerale*, *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 \( \text{gltA} \) 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. \textit{Rickettsia prowazekii} was set as the outgroup.

Fig. 3. Phylogenetic trees based on \( \text{groEL} \) 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. \textit{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


