The phylogeny of Neospora caninum and Toxoplasma gondii based on ribosomal RNA sequences

O. Joakim M. Holmdahl*,a,b,c, Jens G. Mattsson b, Arvid Uggla a and Karl-Erik Johansson b

* Corresponding author. Tel.: (018) 674 000; Fax: (018) 309 162
# Part of this study was presented at the 14th International Conference of the World Association for the Advancement of Veterinary Parasitology (W.A.A.V.P.), Cambridge, UK, August 1993.

Abstract: Neospora caninum is a newly described cyst-forming coccidium which is the cause of severe neurological disease in dogs. The parasite is morphologically similar to Toxoplasma gondii, but the two species can be differentiated serologically. In order to define the phylogenetic position of N. caninum, we have determined 16S-like rRNA sequences from three members of the family of Sarcocystidae: N. caninum, T. gondii, and Sarcocystis fusiformis. The 16S-like rRNA genes from the three parasites were amplified by polymerase chain reaction and the sequences were determined by direct solid-phase sequencing. The sequences derived were computer aligned with several other 16S-like rRNA sequences from protozoan parasites to construct phylogenetic trees. The study confirmed that N. caninum should be classified as a member of the family Sarcocystidae. However, because of the close relationship to T. gondii it seems questionable that N. caninum should be placed in a new genus.

Key words: Neospora caninum; Toxoplasma gondii; 16S-like ribosomal RNA; Phylogeny; Apicomplexa; Coccidia

Introduction

In the protozoan phylum of Apicomplexa several parasites of great importance for both animals and man are found. The best-known species are the members of the genus Plasmodium, which cause the death of millions of people every year. Other species such as the coccidian parasites (enteric and cyst-forming coccidia) and members of the genera Babesia and Theileria are of immense importance in veterinary medicine. Coccidian parasites such as Toxoplasma gondii and Cryptosporidium spp. are the cause of primary infections in humans and animals but are also known as important opportunistic agents in immunocompromised individuals [1].
A decade ago, an unidentified cyst-forming coccidium was found to cause severe neurological disease in dogs [2]. The parasite showed a close morphological resemblance to *T. gondii*, but could be distinguished structurally from all known coccidia and antigenically from *T. gondii* [2,3]. The parasite was named *Neospora caninum* [3] and was placed in the phylum of Apicomplexa as a new genus in the family of Sarcocystidae, together with the genera *Toxoplasma*, *Sarcocystis*, *Hammondia*, *Besnoitia* and *Frenkelia*. The life cycle of *N. caninum* is still unknown and so far the only demonstrated route of transmission is vertical, from mother to foetus [4].

*N. caninum* has been shown to be a primary pathogen for dogs throughout the world. Neosporosis is most severe in young dogs, where it can cause paralysis and death. In cattle, *N. caninum*-like organisms have been shown to cause neonatal mortality and abortion. It has been suggested that neosporosis is one of the major causes of abortion in cattle in California. Furthermore, experimental infections with *N. caninum* in sheep have demonstrated its abortifacient effect also in this species, and *N. caninum*-like parasites have been found in aborted foetuses of goats and horses [4].

Molecular sequences are increasingly used to infer phylogenetic relationships among microorganisms. Ribosomal RNA sequences have proved particularly useful and have previously been used to study the relationships of protozoan parasites [5]. We have determined the sequences of 16S-like rRNA from *N. caninum*, *T. gondii*, and *Sarcocystis fusiformis*. These sequences were compared with 16S-like rRNA sequences from closely related coccidia within the phylum *Apicomplexa* and phylogenetic trees were constructed to determine the taxonomical status of *N. caninum*.

**Materials and Methods**

**Parasites and their purification**

*N. caninum*, NC-1 strain and *T. gondii*, RH strain were cultured on Vero cells [6]. Heavily infected cell cultures were trypsinized (0.25% (w/v) trypsin, 1:300 enzymatic activity; US Biochemical Corp., Cleveland, OH) and the organisms (tachyzoites) of the two parasite species were purified from the cell suspensions by centrifugation at 3000 × g in 30% isotonic Percoll® (Pharmacia LKB Biotechnology, Uppsala, Sweden) for 10 min. The pellets were resuspended in saline and centrifuged twice at 3000 × g for 10 min. The macroscopically visible parasitic cysts of *S. fusiformis*-containing organisms (bradyzoites) were picked out by hand from naturally infected Egyptian water buffaloes. Parasite preparations were stored as pellets at −70°C until further use.

**DNA extraction and amplification of the 16S-like rRNA gene**

Genomic DNA from the three different parasites was isolated according to standard procedures [7] and analysed by agarose gel electrophoresis. The 16S-like rRNA genes were amplified by the polymerase chain reaction (PCR) with the forward primer EUPB-21, 5’ AAC CTG GTT GAT CCT GCC AGT 3’, and the reverse primer EUPB-24, 5’ TGA TCC TTC TGC AGG TTC ACC TAC 3’, complementary to universally conserved regions [7]. The forward primer was synthesized with a biotin amidite at its 5’-end. PCR mixtures (100 μl each) contained 10 mM Tris·HCl, pH 8.3, 50 mM KCl, 2 mM MgCl2, approximately 0.5 μg DNA, 1 μM of each primer and 200 μM of each deoxynucleotide. The mixture was heated to 94°C for 5 min and thereafter 0.5 U of Taq DNA polymerase (Promega, Madison, WI) was added and the samples were amplified for 30 cycles in a model 480 thermocycler (Perkin-Elmer Cetus, Norwalk, CT). Each cycle consisted of a 2-min denaturation at 94°C, 2 min at the annealing temperature of 37°C and a 6-min extension at 72°C. The final extension step was continued for an additional 9 min at 72°C. The PCR products were analysed by agarose gel electrophoresis and excess amounts of biotinylated primer were removed by purification with Magic PCR Preps™ (Promega).

**Immobilization and sequencing reactions**

The biotinylated amplicons were immobilized on streptavidin-coated para-magnetic beads (Dynabeads M280-Streptavidin; Dynal AS, Oslo,
Norway) and the DNA strands were separated as described previously [8]. The nucleotide sequences were determined by the dideoxy chain termination procedure with Sequenase (US Biochemical Corp.) and [$^{35}$S]dATP. Both strands were sequenced with primers complementary to universal regions of 16S-like rRNA

Sequence analysis

The sequences from 16S-like rRNA of *N. caninum*, *T. gondii* and *S. fusiformis* were aligned together with other 16S-like rRNA sequences previously deposited in GenBank. These sequences were from the closely related parasites *Babesia bovis* (accession number L19077), *Theileria parva* (accession number L02366), *Sarcocystis muris*, (accession number M34846), *Cryptosporidium muris* (accession number L19069), and *Cryptosporidium parvum* (accession number L16997). The alignment was performed with the aid of the

Sequence Analysis Software Package from the Genetics Computer Group at the University of Wisconsin [9]. The phylogenetic relationship between the aligned sequences was analysed with the phylogenetic inference package, PHYLIP [10]. The sequence data were subjected to parsimony, distance matrix and bootstrap analysis with the dinoflagellate *Dictyostelium discoideum* (accession number X00134) as an outgroup. Bootstrap analysis [10] was performed in order to check the rigorousness of the tree.

Results

Sequences and sequence alignments

The complete 16S-like rRNA sequences of *N. caninum*, *T. gondii* and *S. fusiformis* were determined with very few ambiguities. The sequences have been deposited in GenBank under the accession numbers U03069, U03070 and U03071, respectively. Alignment between the 16S-like rRNA sequences from *S. fusiformis* and *S. muris*
revealed a large number of sequence differences, which were particularly pronounced in the V4 region \[11,12\]. The alignment of the sequences of the two species in the genus *Sarcocystis* exhibited substantially more sequence differences than to the alignment between the sequences from *T. gondii* and *N. caninum*. This was also true when comparing the number of differences between the latter genera and the alignment with the rRNA sequences from *C. muris* and *C. parvum*.

**Phylogenetic analysis**

The sequence alignment between the nine species of parasites studied here was adjusted in order to subject the data to phylogenetic analysis. The major parts of the highly variable regions V4 and V8 were excluded from the phylogenetic analysis. The homology in these regions is too low among the most distant species included in this study to allow accurate alignment. Generally the most informative sites are located in semivariable and less variable regions. Corrections made a total of 1629 sites liable for further analysis. The entire dataset was subjected to parsimonial analysis by using DNAPARS from PHYLIP \[10\] and all the possible trees for the nine taxa were evaluated. The data revealed a single most parsimonious tree with *D. discoideum* used as an outgroup (see Fig. 1). The tree required 905 nucleotide substitutions. Monophyletic groups were formed within the genera of *Cryptosporidium* and *Sarcocystis* and this was also true for *N. caninum* and *T. gondii* as well as for the species *B. bovis* and *Th. parva*. All the parasites in the family of Sarcocystidae formed a monophyletic clade.

Bootstrapping the data 100 times gave the results shown in Fig. 1. The analysis supported the monophilies among the parasites as described above. The only exception was the monophyly between *B. bovis* and *Th. parva*, but in that case monophyly was still true for a majority of the bootstraps (62%).

The sequence data were also used to calculate a distance matrix (Table 1) by using DNADIST \[10\] to construct a distance matrix tree as shown in Fig. 2. The results from the calculation confirmed the monophyly between the species as described above. Furthermore, in the tree inferred with the distance matrix analyses, the close relationships between the two groups of parasites from the family of Sarcocystidae were supported. The two pairs seem to be more closely related than the species in the genus of *Cryptosporidium*.

**Discussion**

Molecular sequences have been used since the mid-1960s as documents with which to trace evolutionary relationships. A decade later, rRNA was recognized as being the most suitable molecule for this purpose. The use of rRNA sequences, especially 16S-like rRNA \[13\], offers several advantages. The molecule is present in all

<table>
<thead>
<tr>
<th></th>
<th>Cmu</th>
<th>Cpa</th>
<th>Tgo</th>
<th>Nca</th>
<th>Smu</th>
<th>Sfu</th>
<th>Thp</th>
<th>Bbo</th>
<th>Ddi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cmu</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cpa</td>
<td>0.0154</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tgo</td>
<td>0.1062</td>
<td>0.1184</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nca</td>
<td>0.1062</td>
<td>0.1185</td>
<td>0.0006</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smu</td>
<td>0.1069</td>
<td>0.1219</td>
<td>0.0259</td>
<td>0.0272</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sfu</td>
<td>0.1148</td>
<td>0.1198</td>
<td>0.0330</td>
<td>0.0337</td>
<td>0.0233</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thp</td>
<td>0.1141</td>
<td>0.1277</td>
<td>0.1000</td>
<td>0.1007</td>
<td>0.0996</td>
<td>0.1062</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bbo</td>
<td>0.1842</td>
<td>0.1942</td>
<td>0.1873</td>
<td>0.1873</td>
<td>0.1871</td>
<td>0.1911</td>
<td>0.1467</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Ddi</td>
<td>0.2645</td>
<td>0.2659</td>
<td>0.2870</td>
<td>0.2872</td>
<td>0.2992</td>
<td>0.2985</td>
<td>0.3057</td>
<td>0.3463</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

*Cmu*, *C. muris*; *Cpa*, *C. parvum*; *Tgo*, *T. gondii*; *Nca*, *N. caninum*; *Smu*, *S. muris*; *Sfu*, *S. fusiformis*; *Thp*, *Th. parva*; *Bbo*, *B. bovis* and *Ddi*, *D. discoideum.*
organisms, and there is a constraint of functional constancy which ensures a good clocklike behaviour. Furthermore, 16S-like rRNA has a mosaic organization with an evolutionary conserved core interspersed with regions of moderate and high sequence variability, which allows for a spectrum of phylogenetic distances to be compared [13].

In this study we have determined the 16S-like rRNA nucleotide sequences from three coccidian parasites. The sequences were determined by solid-phase sequencing of in vitro amplified DNA. As compared with direct rRNA sequencing, this method produces high quality data with few sequencing ambiguities. In approaches where the amplified gene is cloned before sequencing, there is a risk of PCR-introduced errors and usually several clones have to be sequenced in order to verify the data. This risk is minimized by direct sequencing of the PCR product [14].

The tree inferred from the sequence data presented here, obtained by parsimony analysis, is in general agreement with previously published trees, including Cryptosporidium, Sarcocystis and Toxoplasma [7,15–18]. The data from the bootstrap strongly supported the monophyly observed for the different groups. This latter method randomly samples subsets of data and calculates whether nodes on the tree supported by the entire dataset are also supported by subsets of the data. The optimal tree in a parsimony analysis is the one that requires the fewest nucleotide substitutions and normally this method is quite justified as long as the evolutionary changes between the species are the same. However, in a distance matrix analysis, rapidly evolving sequences will be recognized, since they appear as very long branches. The Fitch-Margoliash tree shared the same topology as the parsimony tree and the only longer branch was that for the outgroup, D. discoideum. All the inter-node distances were short and the distance between T. gondii and N. caninum was minimal. It is also worth noting that the sequence differences between the two studied members of the genus Cryptosporidium were notably higher than for T. gondii and N. caninum. The classification of N. caninum as a member of family Sarcocystidae is thus well supported in this study, but not the assignment as being a member of a new genus.

The taxonomic affiliation of N. caninum was previously based mainly on two criteria as compared with T. gondii. Firstly, there are antigenic differences as visualized by immunohistochemistry and serology, and secondly, N. caninum differs morphologically as regards cyst wall morphology, number of rhoptries of tachyzoites and the absence of a parasitophorous vacuole in the host cell [3]. Due to the close phylogenetic relationship between T. gondii and N. caninum, as revealed in this study, it seems questionable to place these species in separate genera. This conclusion has been drawn by two independently working groups using similar strategies [19,20].

Acknowledgements

We are grateful to Dr. J.P. Dubey for valuable discussions on cyst-forming coccidia and for supplying the N. caninum NC1 strain. We are also indebted to Dr. A.J. Trees for constructive support. This study was supported by the Ivar and Elsa Sandberg Foundation and by the Swedish Council for Forestry and Agricultural Research.

References

7 Gajadhar, A.A., Marquardt, W.C., Hall, R., Gunderson, J., Ariztia-Camona, E.V. and Sogin, M.L. (1991) Ribosomal RNA sequences of Sarcocystis muris, Theileria annulata and Cryptocodinium cohnii reveal evolutionary rela-


