Phylogenetic Relationships of *Cryptosporidium* Parasites Based on the 70-Kilodalton Heat Shock Protein (HSP70) Gene

IRSHAD M. SULAIMAN,[1] UNA M. MORGAN,[2] R. C. ANDREW THOMPSON,[2] ALTAF A. LAL,[1]
AND LIHUA XIAO[1]*

*Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention,
Public Health Services, U.S. Department of Health and Human Services, Atlanta, Georgia 30341,[1] and State
Agricultural Biotechnological Centre, Murdoch University, Murdoch, Western Australia 6150, Australia[2]*

We have characterized the nucleotide sequences of the 70-kDa heat shock protein (HSP70) genes of *Cryptosporidium baileyi*, *C. felis*, *C. meleagridis*, *C. muris*, *C. serpentis*, *C. wrairi*, and *C. parvum* from various animals. Results of the phylogenetic analysis revealed the presence of several genetically distinct species in the genus *Cryptosporidium* and eight distinct genotypes within the species *C. parvum*. Some of the latter may represent cryptic species. The phylogenetic tree constructed from these sequences is in agreement with our previous results based on the small-subunit rRNA genes of *Cryptosporidium* parasites. The *Cryptosporidium* species formed two major clades: isolates of *C. muris* and *C. serpentis* formed the first major group, while isolates of *C. felis*, *C. meleagridis*, *C. wrairi*, and eight genotypes of *C. parvum* formed the second major group. Sequence variations were also observed between *C. muris* isolates from ruminants and rodents. The HSP70 gene provides another useful locus for phylogenetic analysis of the genus *Cryptosporidium*.

Cryptosporidium is an intracellular extracytoplasmic protozoan parasite with a monoxenous life cycle, where all asexual and sexual development occurs within one host. The parasite infects the microvillus border of the gastrointestinal and respiratory epithelium of a wide range of vertebrate hosts, including humans, causing diarrheal diseases. It has been reported to cause waterborne and food-borne outbreaks worldwide (23, 32). Zoonotic infection and person-to-person transmission, however, are also known (2, 6). An understanding of its epidemiology has been hampered by poor knowledge of the species structures and public health importance of various *Cryptosporidium* species and genotypes.

Tyzzer (35, 36) was the first researcher to recognize the multispecies nature of *Cryptosporidium* parasites. He described two species in mammals, *C. muris* and *C. parvum*, based on the differences in morphology and infection sites. In 1955, a new species, *C. meleagridis*, was associated with illness and death in turkeys (31). To date, 22 species of *Cryptosporidium* have been named based on host occurrence, but only 8 are considered valid by some researchers (10). We have recently characterized the small-subunit (SSU) rRNA genes of various *Cryptosporidium* parasites for phylogenetic analysis. The results show that (i) *Cryptosporidium* parasites form a multispecies complex having at least four distinct species (*C. parvum*, *C. baileyi*, *C. muris*, and *C. serpentis*); (ii) there are two distinct genotypes of *C. muris* and various genotypes (human, bovine, dog, ferret, kangaroo, monkey, mouse, and pig) of *C. parvum* which are related to *C. felis*, *C. meleagridis*, and *C. wrairi*; and (iii) some of the *C. parvum* genotypes may be cryptic species (38, 39). These observations are in agreement with other sequence analyses of rRNA genes (24–26).

The heat shock protein (HSP) gene belongs to a multigene family that is highly conserved across the prokaryotes and eukaryotes. Under normal conditions, these proteins function as molecular chaperons for facilitating the folding of proteins in secretion and transport. Their expression, however, is upregulated under environmental stress and is involved in the protection of the cells (9, 13–15, 22). Khramtsov et al. (19) cloned and sequenced the 70-kDa HSP (HSP70) gene of an isolate of the *C. parvum* bovine genotype. Based on this sequence, several molecular diagnostic techniques have recently been designed for the detection of *Cryptosporidium* parasites in environmental samples. These techniques have been used for (i) the detection of viable *C. parvum* oocysts by reverse transcription-PCR (33), (ii) the detection of viable *C. parvum* oocysts by cell culture reverse transcription-PCR (29), and (iii) the detection of viable *C. parvum* oocysts by cell culture PCR (7). However, the polymorphic nature of the HSP70 gene sequences used as primers is not clear, which complicates the use of the assay in detecting *Cryptosporidium* in environmental and clinical samples (5). Therefore, in order to use the HSP70 gene as a diagnostic target for the analysis of clinical and environmental samples, there is a need to characterize the HSP70 genes from different species or genotypes of *Cryptosporidium*.

In this communication, we present the results of sequence characterization and phylogenetic analysis of various *Cryptosporidium* isolates from human and animal hosts at the HSP70 gene locus. Our results with the HSP70 gene confirmed our previous observations of the multispecies nature of *Cryptosporidium* parasites based on the SSU rRNA gene (38, 39). The sequence information generated from this study is also useful in the development of HSP70-based species and genotype diagnostic tools.

MATERIALS AND METHODS

Purification of oocysts and extraction of genomic DNA. Fecal samples containing oocysts of *C. baileyi* (chicken and quail), *C. felis* (cat and human), *C. meleagridis* (turkey and human), *C. muris* (cattle, camel, and mouse), *C. parvum* (human, cattle, cat, dog, ferret, monkey, mouse, kangaroo, koala, and pig), *C. serpentis* (savanah monitor and snake), *C. wrairi* (guinea pig), and an unknown *Cryptosporidium* species (desert monitor) were obtained from infected humans and animals and stored at 4°C in 2.5% potassium dichromate solution until they were used (Table 1). The oocysts were purified by the sucrose and Percoll

* Corresponding author. Mailing address: Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Building 22, Mail Stop F-12, 4770 Buford Highway, Atlanta, GA 30341-3717. Phone: (770) 488-4840. Fax: (770) 488-4454. E-mail: LAX0@CDC.GOV.

TABLE 1. *Cryptosporidium* isolates used in this study

Isolate	Source location	Host	Species and genotype	A+T content (%)
183	Washington	Human	*C. parvum* human	58.8
497	Kenya	HIV[+] human[a]	*C. parvum* human	58.9
6	Ohio	Calf	*C. parvum* bovine	58.6
671	Australia	Calf	*C. parvum* bovine	58.5
674	Australia	Calf	*C. parvum* bovine	58.6
244	Ohio	Dog	*C. parvum* dog	48.2
715a	Maryland	Dog	*C. parvum* dog	48.1
351	Georgia	Ferret	*C. parvum* ferret	58.4
712a	Washington, D.C.	Black-footed ferret	*C. parvum* ferret	58.4
713a	Washington, D.C.	Black-footed ferret	*C. parvum* ferret	58.4
714a	Washington, D.C.	Black-footed ferret	*C. parvum* ferret	58.5
428	Australia	Red kangaroo	*C. parvum* marsupial	60.1
587	Australia	Koala	*C. parvum* marsupial	60.1
518	Georgia	Rhesus monkey	*C. parvum* monkey	58.8
359	Maryland	Mouse	*C. parvum* mouse	59.1
411	Maryland	Mouse	*C. parvum* mouse	59.2
499	Australia	Pig	*C. parvum* pig	59.8
763	Australia	Quail	*C. baileyi*	65.8
764	Maryland	Chicken	*C. baileyi*	65.8
288	Australia	Cat	*C. felis*	51.1
160	New Orleans	HIV[+] human	*C. felis*	51.0
297	New Orleans	HIV[+] human	*C. felis*	51.2
670	New Orleans	HIV[+] human	*C. felis*	51.0
295	Maryland	Turkey	*C. meleagridis*	58.1
672	Kenya	HIV[+] human	*C. meleagridis*	58.2
589	Australia	Cattle	*C. muris*	62.2
697	Czech Republic	Cattle	*C. muris*	62.2
698	Czech Republic	Cattle	*C. muris*	62.2
699	Czech Republic	Cattle	*C. muris*	62.2
700	Czech Republic	Cattle	*C. muris*	62.2
701	Czech Republic	Camel	*C. muris*	62.6
703	Czech Republic	Camel	*C. muris*	62.2
707	Czech Republic	Mouse	*C. muris*	63.5
708	Czech Republic	Mouse	*C. muris*	63.4
709	Spain	Mouse	*C. muris*	63.4
63	Maryland	Savanna monitor	*C. serpentis*	62.2
64	Maryland	Amazon tree boa	*C. serpentis*	62.2
517	Michigan	Guinea pig	*C. wrairi*	59.3
691	St. Louis	Desert monitor	*Cryptosporidium* sp.	61.0
692	St. Louis	Desert monitor	*Cryptosporidium* sp.	61.0

[a] HIV, human immunodeficiency virus.

gradient method (1). DNA was isolated from the purified oocysts as described before (34) and stored at −20°C before use. The concentration of DNA samples was measured by UV absorption at 260 nm. The identities of *Cryptosporidium* species and genotypes were established based on morphological examinations and sequence analysis of the SSU rRNA gene (38, 39).

PCR amplification. A two-step nested-PCR protocol was used to amplify the HSP70 gene fragments from various *Cryptosporidium* isolates, using primers complementary to the conserved nucleotide sequences of apicomplexan parasites downloaded from GenBank: the *C. parvum* bovine genotype (U71181), *Eimeria acervulina* (Z26134), *Plasmodium cynomolgi* (M90978), *Theileria annulata* (J04653), and *Toxoplasma gondii* (U85648). A PCR product of ~2,015 bp was amplified using forward (5′-ATG TCT GAA GGT CCA GCT ATT GGT ATT GA-3′) and reverse (5′-TTA GTC GAC CTC TTC AAC AGT TGG-3′) primers. The PCR mixture consisted of 50 ng of DNA, 200 μM (each) deoxynucleoside triphosphate, 1× PCR buffer (Perkin-Elmer, Foster City, Calif.), 3.0 mM MgCl$_2$, 5.0 U of *Taq* polymerase (GIBCO BRL, Frederick, Md.), and 200 nM (each) primer in a total volume of 100 μl. The reactions were performed for 35 cycles (each cycle was 94°C for 45 s, 55°C for 45 s, and 72°C for 60 s) in a Perkin-Elmer GeneAmp PCR 9700 thermocycler with an initial hot start (94°C for 5 min) and a final extension (72°C for 10 min). For the secondary PCR, a fragment of ~1,950 bp was amplified using 2.5 μl of primary PCR mixture and a set of nested forward (5′-TA/CT TCA TG/CT GTT GGT GTA TGG AGA AA-3′) and reverse (5′-CAA CAG TTG GAC CAT TAG ATC C-3′) primers. The conditions for the secondary PCR were identical to those for the primary PCR, except for the use of a lower annealing temperature (45°C). The PCR product was analyzed by agarose gel electrophoresis and visualized after ethidium bromide staining.

Sequencing and phylogenetic analysis. The secondary-PCR products were sequenced on an ABI 377 automated sequencer (Perkin-Elmer) using a Big Dye terminator cycle-sequencing ready-reaction kit (Perkin-Elmer). Sequence accuracy was confirmed by two-directional sequencing and by sequencing of a new PCR product if necessary. Multiple alignments of the DNA sequences were done with the Wisconsin Package version 9.0 (Genetics Computer Group, Madison, Wis.) with manual adjustment.

Two phylogenetic analyses were carried out on the aligned sequences to assess phylogenetic relationships among various species and genotypes. The first analysis was conducted to assess the evolutionary relationship between *Cryptosporidium* species and other members of the phylum Apicomplexa. In this analysis, the HSP70 gene sequences representing the *C. parvum* bovine genotype and *C. muris* were aligned with the published sequences of the *C. parvum* bovine genotype (U71181 and U69698), *E. acervulina* (Z26134), *Eimeria maxima* (Z46964), *P. cynomolgi* (M90978), *Plasmodium falciparum* (M19753), *T. annulata* (J04653), *Theileria parva* (U40190), *Theileria sergenti* (D12692), and *T. gondii* (AF045559 and U85648) obtained from GenBank. A neighbor-joining tree (30) was constructed using the program TreeconW (37), and evolutionary distances were calculated by Kimura two-parameter analysis. The sequence of *Babesia microti* (GenBank accession no. U53448) was used as an outgroup to assess the relatedness of the genus *Cryptosporidium* with other members of the phylum Apicomplexa. We chose to use *B. microti* as an outgroup because construction of unrooted trees suggested that it was the most divergent member of this group.

In the second analysis, a neighbor-joining tree was constructed for all the isolates of *Cryptosporidium* to assess the relationship among various *Cryptosporidium* species and within *C. parvum* genotypes. The tree was rooted using *P. cynomolgi* (GenBank accession no. M90978) and *P. falciparum* (GenBank acces-

TABLE 2. Evolutionary genetic distances among *Cryptosporidium* species and eight genotypes of *C. parvum*

Species and genotype	No. of nucleotide differences per 100 bases[a]																	
	P. falciparum	*P. cynomolgi*	*C. muris* bovine	*C. muris* murine	*C. serpentis*	*C. baileyi*	Desert monitor	*C. felis*	*C. Tparvum* dog	*C. parvum* pig	*C. parvum* marsupial	*C. meleagridis*	*C. parvum* ferret	*C. wrairi*	*C. parvum* human	*C. parvum* monkey	*C. parvum* mouse	*C. parvum* bovine
P. falciparum	0.00																	
P. cynomolgi	21.03	0.00																
C. muris bovine	34.84	40.51	0.00															
C. muris murine	32.73	38.79	1.76	0.00														
C. serpentis	34.55	40.43	5.16	4.34	0.00													
C. baileyi	29.27	36.74	15.50	15.60	15.74	0.00												
Cryptosporidium sp.	32.04	37.25	20.44	20.40	21.11	14.80	0.00											
C. felis	42.57	38.79	27.74	27.84	27.97	26.59	19.42	0.00										
C. parvum dog	43.51	38.19	29.77	30.02	29.64	28.37	20.86	15.66	0.00									
C. parvum pig	33.87	37.28	20.47	20.07	20.25	16.17	11.53	18.29	18.51	0.00								
C. parvum marsupial	32.28	37.22	20.92	21.15	20.61	15.97	12.15	18.36	19.23	8.00	0.00							
C. meleagridis	33.08	36.97	21.23	21.17	21.22	17.35	13.67	18.58	16.54	9.61	5.79	0.00						
C. parvum ferret	34.38	37.82	22.08	21.77	21.90	16.73	13.75	18.65	17.17	8.86	5.31	4.02	0.00					
C. wrairi	32.60	36.80	21.21	20.85	20.62	16.21	12.46	18.24	17.87	8.12	4.03	4.04	2.57	0.00				
C. parvum human	33.77	37.07	21.71	21.31	21.72	16.37	13.55	18.45	17.49	8.55	4.90	4.47	3.68	2.29	0.00			
C. parvum monkey	34.20	37.10	21.78	21.38	21.59	16.51	13.44	18.41	17.51	8.73	4.96	4.58	3.73	2.35	0.25	0.00		
C. parvum mouse	33.19	36.44	21.79	21.45	21.47	16.38	13.22	18.68	17.42	8.44	4.96	4.25	2.93	2.18	2.05	2.22	0.00	
C. parvum bovine	34.19	37.69	21.78	21.38	21.74	16.79	14.22	18.31	16.78	8.62	4.97	4.48	3.14	1.92	2.88	2.47	1.51	0.00

[a] The values are nucleotide changes per 100 bp calculated by the Kimura two-parameter method using the Wisconsin Package, version 9.0 (Genetics Computer Group).

sion no. M19753). The second phylogenetic analysis also included the construction of a neighbor-joining tree based on the deduced amino acid sequences. The reliabilities of these trees were assessed by the bootstrap method (11) with 1,000 pseudoreplicates. We used 95% as the statistically significant value (8); however, values greater than 70% are reported, since the bootstrap method (11) may be a conservative estimate for the reliability of a clade (16).

Nucleotide sequence accession numbers. The nucleotide sequences of the HSP70 genes of *C. baileyi*, *C. felis*, *C. meleagridis*, *C. muris*, *C. serpentis*, *C. wrairi*, the unknown *Cryptosporidium* sp., and eight genotypes of *C. parvum* (human, bovine, dog, ferret, marsupial, monkey, mouse, and pig) have been deposited in the GenBank database under accession no. AF221528 to AF221543.

RESULTS

We sequenced ~1,950 bp of the HSP70 genes from 17 *C. parvum* isolates, 2 *C. baileyi* isolates, 4 *C. felis* isolates, 2 *C. meleagridis* isolates, 10 *C. muris* isolates, 2 *C. serpentis* isolates, 1 *C. wrairi* isolate, and 2 isolates from an unknown *Cryptosporidium* species from a desert monitor. The *C. parvum* isolates, represented the following genotypes of *C. parvum*: two isolates of the *C. parvum* human genotype, three isolates of the *C. parvum* bovine genotype, two isolates of the *C. parvum* dog genotype, four isolates of the *C. parvum* ferret genotype, two isolates of the *C. parvum* marsupial genotype, one isolate of the *C. parvum* monkey genotype, two isolates of the *C. parvum* mouse genotype, and one isolate of the *C. parvum* pig genotype. HSP70 gene sequences of an unknown *Cryptosporidium* species from two desert monitors were also obtained. The HSP70 gene of *Cryptosporidium* parasites was AT rich (58.4 to 65.8%), except for the isolates of the *C. parvum* dog genotype and *C. felis* (48.1 to 51.2%). However, within each *Cryptosporidium* species and *C. parvum* genotype the A+T contents of different isolates were quite consistent (Table 1).

Multiple alignment of the HSP70 gene sequences revealed distinct sequences for the eight species of *Cryptosporidium* (*C. baileyi*, *C. parvum*, *C. meleagridis*, *C. muris*, *C. serpentis*, *C. felis*, *C. wrairi*, and the unknown *Cryptosporidium* sp.) analyzed in

the study. Distinct interspecies variations were also noticed throughout the entire HSP70 gene, including in the regions of PCR primers utilized by Stinear et al. (33), Rochelle et al. (29), and Di Giovanni et al. (7) (Table 2 and Fig. 1). Eight genotypes of *C. parvum* and two genotypes of *C. muris* were found by HSP70 gene analyses, in concordance with the results of analysis at the SSU rRNA gene locus (38, 39). The extent of genetic variation in the genus *Cryptosporidium* was also assessed by comparing the *C. parvum* bovine genotype nucleotide sequence with other HSP70 gene sequences of different *Cryptosporidium* species and *C. parvum* genotypes. The variation between the *C. parvum* bovine genotype and other *C. parvum* genotype isolates was low (1.4 to 7.4%), except for the *C. parvum* dog genotype isolates (12.5%). A significant difference was observed among non-*parvum* species, such as *C. baileyi*, *C. muris*, *C. serpentis*, *C. felis*, and the unknown *Cryptosporidium* sp. (14.7 to 18.7%). The genetic differences between the *C. parvum* bovine genotype and *C. meleagridis* and *C. wrairi* isolates, however, was substantially lower (1.9 to 4.0%) (Table 3). Compared to the *C. parvum* bovine genotype, the majority of mutations in the HSP70 genes of other *Cryptosporidium* parasites were synonymous. However, the percentages of nonsilent mutations were higher at the interspecies level and lower at the intergenotype level, except for isolates of *C. parvum* from humans and a monkey and *C. baileyi* (Table 3).

Phylogenetic analysis of *C. parvum* and *C. muris* together with published HSP70 sequences of various members of the phylum Apicomplexa, including *C. parvum* (U71181 and U69698), *E. acervulina* (Z26134), *E. maxima* (Z46964), *P. cynomolgi* (M90978), *P. falciparum* (M19753), *T. annulata* (J04653), *T. parva* (U40190), *T. sergenti* (D12692), *T. gondii* (AF045559 and U85648), and *B. microti* (U534448), revealed a close relationship between the genera *Cryptosporidium* and *Plasmodium* (Fig. 2). A neighbor-joining tree showed that the *Cryptosporidium* clade and the *Plasmodium* clade clustered to-

A

Species or Genotype	cphsp2386F (5'→3')	cphsp2672R (5'→3')
Rochelle et al. (1997)	**CTGTTGCTTATGGTGCTGCTG**	**CCTCTTGGTGCTGGTGGAATA**
C. serpentis	...................A.	..A.G...A...........
C. muris bovine	...................A.	..A.G...A...........
C. muris rodent	...................A.	..A.G...A...........
C. felis	.C..G..A........C....	.....G..A..C..G..G..T
C. baileyi	.......A............	.....A..............
Cryptosporidium sp.	...................A.	...........A........
C. meleagridis	.C.....C............	....
C. wrairi	....	....
C. parvum dog genotype	.......C..........C.	............C.......
C. parvum pig genotype	....	....
C. parvum marsupial genotype	....	..C.................
C. parvum ferret genotype	....	....
C. parvum human genotype	....	....
C. parvum monkey genotype	....	....
C. parvum mouse genotype	....	....
C. parvum bovine genotype	....	....

B

Species or Genotype	cphsp2423F (5'→3')	cphsp2764R (5'→3')
Rochelle et al. (1997)	**AAATGGTGAGCAATCCTCTG**	**CTTGCTGCTCTTACCAGTAC**
C. serpentis	.........A.....A..A.	.........T.....G....
C. muris bovine	.........A..G..A..A.	...............G....
C. muris rodent	.........A..G..A..A.	...............G....
C. felis	G..C.......G..T....	......A.....C.....G.
C. baileyi	......A..A.....A..A.	T..A................
Cryptosporidium sp.	G........A.....T....	T.....A.............
C. meleagridis	....	.................G.
C. wrairi	....	....
C. parvum dog genotype	G..C...............C.	...............G..G.
C. parvum pig genotype	...C.....A..G..T....	......A.............
C. parvum marsupial genotype	..............T....	....
C. parvum ferret genotype	....	....
C. parvum human genotype	....	.................G.
C. parvum monkey genotype	....	.................G.
C. parvum mouse genotype	....	....
C. parvum bovine genotype	....	....

C

Species or Genotype	Chsp1 (5'→3')	Chsp4 (5'→3')
Stinear et al. (1996)	**AGCAATCCTCTGCCGTACAGG**	**AAGAGCATCCTTGATCTTCT**
C. serpentis	.A.....A..A.TA.....A.	...T....T...A..T..AG
C. muris bovine	.A..G..A..A.TA..G..A.	...TA.C.....A..T..AG
C. muris rodent	.A..G..A..A.TA.....A.	...TA.C.....A..T..AG
C. felis	....G..T.....G..C..A.	G..G..T.............
C. baileyi	.A.....A..A..T.......	...T.....T..A.......
Cryptosporidium sp.	.A.....T.....T..T..A.	....
C. meleagridis	...............T....	.............A......
C. wrairi	...............T....	....
C. parvum dog genotype	..........C.....C....	...G..G.............
C. parvum pig genotype	.A..G..T.....T..T..A.	....
C. parvum marsupial genotype	.......T.....T..T..A.	....
C. parvum ferret genotype	................T....	....
C. parvum human genotype	....	....
C. parvum monkey genotype	....	....
C. parvum mouse genotype	....	....
C. parvum bovine genotype	....	....

FIG. 1. Variation in the HSP70 gene nucleotide sequences in the primer regions of diagnostic tools by Rochelle et al. (29) (A and B), Di Giovanni et al. (7) (A and B), and Stinear et al. (33) (C).

TABLE 3. Divergence of *Cryptosporidium* spp. and other *C. parvum* genotypes from the *C. parvum* bovine genotype

Isolate	Species and genotype	Variation at nucleotide level (%)	No. of amino acids changed	Nonsilent mutations/total mutations (%)
589	*C. muris*	18.7	41	11.5
707	*C. muris*	18.4	38	11.5
64	*C. serpentis*	17.6	43	12.1
764	*C. baileyi*	14.7	12	4.3
288	*C. felis*	15.3	13	4.3
295	*C. meleagridis*	4.0	6	7.7
517	*C. wrairi*	1.9	1	2.7
692	*Cryptosporidium* sp.	12.5	15	6.0
183	*C. parvum* human	1.5	5	16.6
244	*C. parvum* dog	14.7	7	2.5
428	*C. parvum* marsupial	4.7	3	3.2
351	*C. parvum* ferret	3.0	2	3.3
518	*C. parvum* monkey	1.7	6	18.1
359	*C. parvum* mouse	1.4	0	0.0
499	*C. parvum* pig	7.4	7	4.8

gether with full statistical reliability. Other intestinal coccidian parasites traditionally associated with *Cryptosporidium* parasites were placed in a different cluster in this analysis (Fig. 2).

In a second phylogenetic analysis, a neighbor-joining tree was constructed from aligned HSP70 gene sequences of various *Cryptosporidium* isolates, using nucleotide sequences of *P. cynomolgi* (M90978) and *P. falciparum* (M19753) as an outgroup (Fig. 3A). The genus *Cryptosporidium* formed two distinct clusters in this phylogenetic analysis: the first consisted of two genotypes of *C. muris* and *C. serpentis* and *C. baileyi* isolates, and the second cluster contained isolates of *C. felis*, *C. meleagridis*, and *C. wrairi*, eight different genotypes of *C. parvum*, and two isolates of the unknown *Cryptosporidium* species. Within the first cluster, the *C. muris* bovine type and the *C. muris* rodent murine isolates formed distinct clades. In the second major cluster, the unknown *Cryptosporidium* species, *C. felis*, and the *C. parvum* dog genotype were separated from the remaining member of the cluster (seven *C. parvum* genotypes, *C. meleagridis*, and *C. wrairi*). Significant intraspecies diversity was seen in *C. parvum*, as reflected by the presence of eight genotypes. Similar phylogenetic structure was also observed with analysis of deduced amino acid sequences. In the latter, however, *C. baileyi* clustered together with the broad *C. par-*
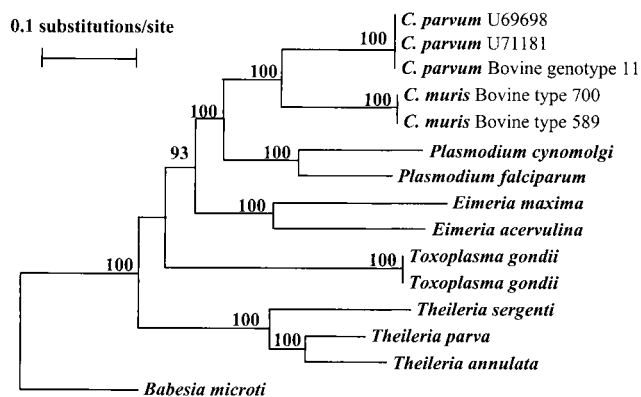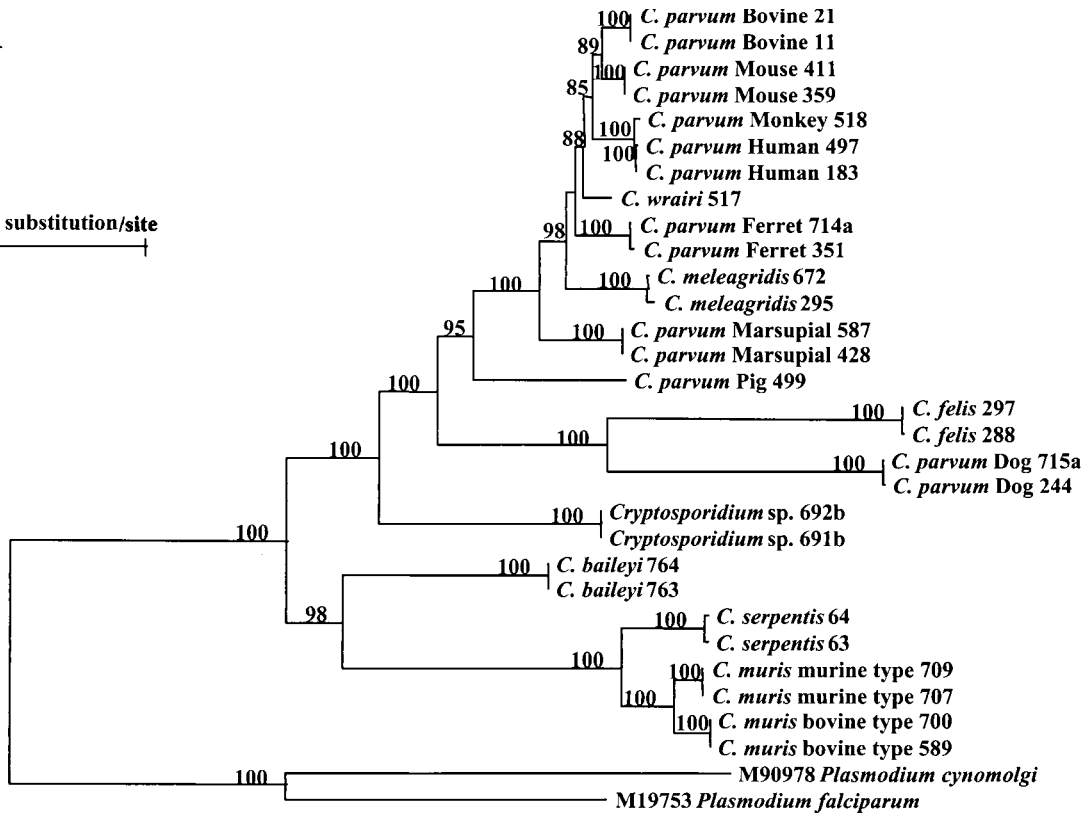

FIG. 2. Phylogenetic relationship of *Cryptosporidium* parasites to other apicomplexan parasites inferred from neighbor-joining analysis of HSP70 gene nucleotide sequences.

vum group, and the *C. parvum* dog genotype and *C. felis* did not form a clade (Fig. 3B).

DISCUSSION

In the present study, nucleotide sequences of the HSP70 gene were obtained from eight *Cryptosporidium* species and eight different *C. parvum* genotypes to reassess the species structure of this genus as well as the evolutionary relationship of *Cryptosporidium* parasites to other members of the phylum Apicomplexa. Results of sequence and phylogenetic analyses are in agreement with our previous observations, based on the SSU rRNA gene, of the presence of multiple species within the genus *Cryptosporidium*, two genotypes of *C. muris*, and multiple genotypes of *C. parvum* (38, 39). The various *Cryptosporidium* species are placed in different clades and, with the exception of *C. wrairi* and *C. meleagridis*, showed interspecies genetic distances comparable to those between different apicomplexan parasites. A close relatedness of *Cryptosporidium* parasites to *Plasmodium* parasites was also revealed by phylogenetic analysis of HSP70 sequences, an observation previously suggested by analysis of the structural organization of the rRNA gene (21).

Analysis of the HSP70 gene sequence further suggests that the unknown *Cryptosporidium* parasite from desert monitors, *C. felis*, and the *C. parvum* dog genotype may be valid species. The *Cryptosporidium* sp. from desert monitors was consistently placed at the bottom of the broad *C. parvum* cluster. This is also reflected in its genetic distance from other *Cryptosporidium* parasites, which was >11.53 nucleotide changes per 100 bp. This genetic distance was greater than the distance between *C. muris* and *C. serpentis* (4.34 to 5.16) or the *C. parvum* bovine genotype and *C. meleagridis* (4.48) and was comparable to the distance between *C. baileyi* and other *Cryptosporidium* parasites (14.80 to 17.35). It remains unclear whether this unknown *Cryptosporidium* sp. is *C. saurophilum*, a new species identified in desert monitors and other lizards (20). Although, *C. felis* and the *C. parvum* dog genotype clustered together with the broad *C. parvum* clade, their genetic distances from *C. muris*, *C. serpentis*, and *C. baileyi* (26.59 to 30.02 nucleotide changes/100 bp) were far greater than those of other *Cryptosporidium* spp. (<22 nucleotide changes/100 bp). These isolates also diverged significantly from the rest of the *C. parvum* genotypes, *C. meleagridis*, and *C. wrairi*, with genetic distances (16.54 to 19.23) comparable to those between *C. baileyi* and other *Cryptosporidium* species (14.80 to 17.35). This is also reflected in the low G+C contents of HSP70 nucleotide sequences in these two parasites. Previous analysis at the SSU rRNA gene locus also suggested that *C. felis*, the *C. parvum* dog genotype, and the unknown *Cryptosporidium* parasite from desert monitors might be cryptic species (39).

The results of this study suggest that the HSP70 gene offers several advantages over the SSU rRNA gene for phylogenetic studies of *Cryptosporidium* parasites. Although this gene is under selection pressure (as reflected in the presence of a high percentage of synonymous mutations), the HSP70 gene is apparently more permissive of nucleotide changes. As a result, higher heterogeneity was seen in the HSP70 gene nucleotide sequences than in SSU rRNA gene sequences, which makes it a better target for genotyping. This lower selection pressure in the HSP70 gene is also reflected in the location of nucleotide mutations. Unlike those in the SSU rRNA gene, which restricts nucleotide changes to a certain region of the gene, mutations in the HSP70 gene are spread over the entire sequence. Because deletions and insertions are limited in the HSP70 gene, the alignment of sequences from very different organisms is much easier. Therefore, phylogenetic analysis of *Cryptosporidium*
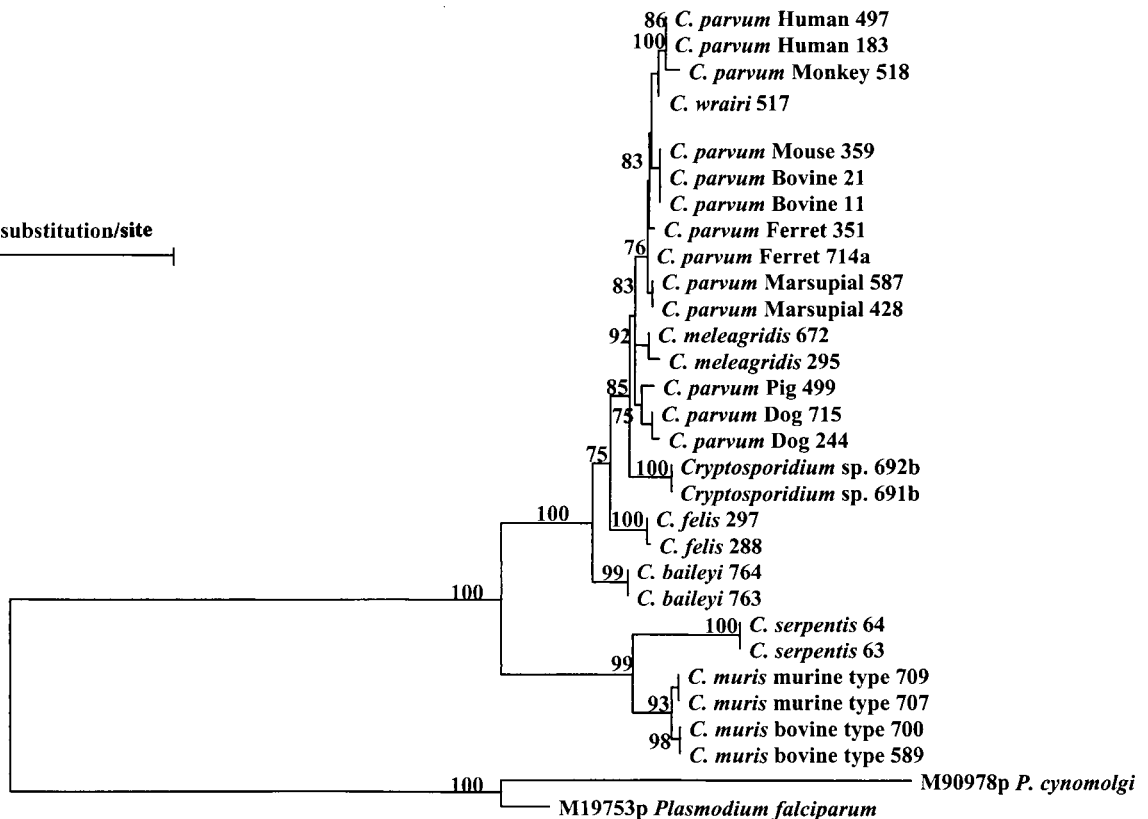
FIG. 3. Phylogenetic relationships among *Cryptosporidium* parasites inferred from neighbor-joining analysis of nucleotide sequences (A) and deduced amino acid sequences (B).

parasites based on HSP70 gene sequences is much more robust than those based on the SSU rRNA gene, with significantly higher bootstrap values. The minor dissimilarity between the nucleotide and amino acid trees may be explained by the lesser heterogeneity in amino acid sequences due to the predominance of synonymous mutations in the nucleotide sequences. In other eukaryotic systems, the HSP70 gene has also become a useful alternative in the study of molecular evolution (3, 4, 12, 15, 17, 18).

The HSP70 gene sequence generated in this study reveals problems in the current HSP70-gene-based PCR diagnosis tools designed for the use of environmental samples (7, 29, 33). As shown in Fig. 1, the primers used in these protocols matched only sequences from the *C. parvum* bovine, human, and mouse genotypes. We have found dissimilarities with the *C. parvum* dog, pig, and marsupial genotypes, *C. meleagridis*, and *C. felis*. It is likely that the efficiencies of these primers in amplifying DNA from these organisms may be compromised due to the heterogeneity in the primer regions, especially in the case of environmental samples, which usually have small numbers of organisms. This is a matter of concern, because many of the *Cryptosporidium* parasites, such as the *C. parvum* dog genotype, *C. felis*, and *C. meleagridis*, have been found in patients with AIDS (27, 28) and in children (L. Xiao, C. Bern, and A. A. Lal, unpublished observation). Indeed, all six *C. parvum* HSP70 genotypes recently described in water samples (7) belong to the bovine, human, and mouse genotypes in the present study. The minor differences among some of the water genotypes are much smaller than those among other genotypes; thus, they may represent intragenotype diversity or artifacts. The nucleotide sequences of the HSP70 gene generated in this study will be useful in the improvement of these diagnostic tools and in the development of new molecular tools for *Cryptosporidium* species and genotype differentiation.

REFERENCES

1. **Arrowood, M. J., and C. R. Sterling.** 1987. Isolation of *Cryptosporidium* oocysts and sporozoites using discontinuous sucrose and isopycnic Percoll gradients. J. Parasitol. **73:**314–319.
2. **Awad-El-Kariem, F. A., D. C. Warhurst, and V. McDonald.** 1994. Detection of *Cryptosporidium* oocysts using a system based on PCR and endonuclease restriction. Parasitology **109:**19–22.
3. **Budin, K., and H. Philippe.** 1998. New insights into the phylogeny of eukaryotes based on ciliate Hsp70 sequences. Mol. Biol. Evol. **15:**943–956.
4. **Bui, E. T., P. J. Bradley, and P. J. Johnson.** 1996. A common evolutionary origin of the mitochondria and hydrogenosomes. Proc. Natl. Acad. Sci. USA **93:**9651–9656.
5. **Champliaud, D., P. Gobet, M. Naciri, O. Vagner, J. Lopez, J. C. Buisson, I. Varga, G. Harly, R. Mancassola, and A. Bonnin.** 1998. Failure to differentiate *Cryptosporidium parvum* from *C. meleagridis* based on PCR amplification of eight DNA sequences. Appl. Environ. Microbiol. **64:**1454–1458.
6. **Cordell, R. L., and D. G. Addiss.** 1994. Cryptosporidiosis in child care settings: a review of the literature and recommendations for prevention and control. Pediatr. Infect. Dis. J. **13:**311–317.
7. **Di Giovanni, G. D., F. H. Hashemi, N. J. Shaw, F. A. Abrams, M. W. LeChevallier, and M. Abbaszadegan.** 1999. Detection of infectious *Cryptosporidium parvum* oocysts in surface and filter backwash water samples by immunomagnetic separation and integrated cell culture-PCR. Appl. Environ. Microbiol. **65:**3427–3432.
8. **Efron, B., E. Halloran, and S. Holmes.** 1996. Bootstrap confidence levels for phylogenetic trees. Proc. Natl. Acad. Sci. USA **93:**13429–13434.
9. **Ellis, R. J., and S. M. van der Vies.** 1991. Molecular chaperones. Annu. Rev. Biochem. **60:**321–347.
10. **Fayer, R., C. A. Spear, and J. P. Dubey.** 1997. The general biology of *Cryptosporidium*, p. 1–41, *In* R. Fayer (ed.), *Cryptosporidium* and cryptosporidiosis. CRC Press, Boca Raton, Fla.
11. **Felsenstein, J.** 1985. Confidence limits on the phylogenies: an approach using bootstrap. Evolution **39:**783–791.
12. **Germot, A., and H. Philippe.** 1999. Critical analysis of eukaryotic phylogeny: a case study based on the HSP70 family. J. Eukaryot. Microbiol. **46:**116–124.
13. **Gething, M. J., and J. Sambrook.** 1992. Protein folding in the cell. Nature **355:**33–45.
14. **Gupta, R. S., M. Bustard, M. Falah, and D. Singh.** 1997. Sequencing of heat shock protein 70 (DnaK) homologs from *Deinococcus proteolyticus* and Archaebacteria, Eubacteria, and *Thermomicrobium roseum* and their integration in a protein-based phylogeny of prokaryotes. J. Bacteriol. **179:**345–357.
15. **Gupta, R. S., and G. B. Golding.** 1993. Evolution of the Hsp70 gene and its implications regarding relationships between Archaebacteria, Eubacteria, and Eukaryotes. J. Mol. Evol. **37:**573–582.
16. **Hills, D. M., and J. J. Bull.** 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. **8:**189–191.
17. **Karasev, A. V., A. S. Kashina, V. I. Gelfand, and V. V. Dolja.** 1992. Hsp70-related 65 kDa protein of beet yellows closterovirus is a microtubule-binding protein. FEBS Lett. **304:**12–14.
18. **Karlin, S., and L. Brocchieri.** 1998. Heat shock protein 70 family: multiple sequence comparisons, function, and evolution. J. Mol. Evol. **47:**565–577.
19. **Khramtsov, N. V., M. Tilley, D. S. Blunt, B. A. Montelone, and S. J. Upton.** 1995. Cloning and analysis of a *Cryptosporidium parvum* gene encoding a protein with homology to cytoplasmic form HSP70. J. Eukaryot. Microbiol. **42:**416–422.
20. **Koudela, B., and D. Modry.** 1998. New species of *Cryptosporidium* (Apicomplexa, Cryptosporidiidae) from lizards. Folia Parasitol. **45:**93–100.
21. **Le Blancq, S. V., N. V. Khramtsov, F. Zamani, S. J. Upton, and T. W. Wu.** 1997. Ribosomal RNA gene organization in *Cryptosporidium parvum*. Mol. Biochem. Parasitol. **90:**463–478.
22. **Lindquist, S., and E. A. Craig.** 1988. The heat shock proteins. Annu. Rev. Genet. **22:**631–677.
23. **Millard, P. S., K. F. Gensheimer, D. G. Addis, D. M. Sosin, G. A. Beckett, A. Houck-Jankoski, and A. Hudson.** 1994. An outbreak of cryptosporidiosis from fresh-pressed apple cider. JAMA **272:**1592–1596.
24. **Morgan, U. M., A. P. Sturdee, G. Singleton, M. S. Gomez, M. Gracenea, J. Torres, S. G. Hamilton, D. P. Woodside, and R. C. A. Thompson.** 1999. The *Cryptosporidium* "Mouse" genotype is conserved across geographic areas. J. Clin. Microbiol. **37:**1302–1305.
25. **Morgan, U. M., J. R. Buddle, A. Armson, A. Elliot, and R. C. A. Thompson.** 1999. Molecular and biological characterization of *Cryptosporidium* in pigs. Aust. Vet. **77:**44–47.
26. **Morgan, U. M., P. Monis, R. Fayer, P. Deplazes, and R. C. A. Thompson.** 1999. Phylogenetic relationships among isolates of *Cryptosporidium*: evidence for several new species. J. Parasitol. **85:**1126–1133.
27. **Morgan, U. M., L. Xiao, I. Sulaiman, R. C. A. Thompson, A. Lal, and P. Deplazes.** 1999. Which genotypes/species of *Cryptosporidium* are humans susceptible to? J. Eukaryot. Microbiol. **46:**42S–43S.
28. **Pieniazek, N. J., F. J. Bornay-Llinares, S. B. Slemenda, A. J. da Silva. I. N. Moura, M. J. Arrowood, O. Ditrich, and D. G. Addiss.** 1999. New cryptosporidium genotypes in HIV-infected persons. Emerg. Infect. Dis. **5:**444–449.
29. **Rochelle, P. A., D. M. Ferguson, T. R. Handojo, R. D. Leon, M. H. Stewart, and R. L. Wolfe.** 1997. An assay combining cell culture with reverse transcriptase PCR to detect and determine the infectivity of waterborne *Cryptosporidium parvum*. Appl. Environ. Microbiol. **63:**2029–2037.
30. **Saitou, N., and M. Nei.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
31. **Slavin, D.** 1955. *Cryptosporidium meleagridis* (sp. nov.). J. Comp. Pathol. **65:**262–265.
32. **Smith, H. V., and J. B. Rose.** 1998. Water borne cryptosporidiosis: current status. Parasitol. Today **14:**14–22.
33. **Stinear, T., A. Matusan, K. Hines, and M. Sandery.** 1996. Detection of a single viable *Cryptosporidium parvum* oocyst in environmental water concentrates by reverse transcription-PCR. Appl. Environ. Microbiol. **62:**3385–3390.
34. **Sulaiman, I. M., L. Xiao, C. Yang, A. Moore, C. B. Beard, M. J. Arrowood, and A. A. Lal.** 1998. Differentiating human from animal isolates of *Cryptosporidium parvum*. Emerg. Infect. Dis. **4:**681–685.
35. **Tyzzer, E. E.** 1907. A sporozoon found in the peptic glands of the common mouse. Proc. Soc. Exp. Biol. Med. **5:**12–13.
36. **Tyzzer, E. E.** 1910. An extracellular coccidium, *Cryptosporidium muris* (gen. & sp. nov.) of the gastric glands of the common mouse. J. Med. Res. **1:**487–590.
37. **Van de Peer, Y., and R. De Wachter.** 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput. Appl. Biosci. **10:**569–570.
38. **Xiao, L., L. E. Escalante, C. Yang, I. M. Sulaiman, A. A. Escalante, R. Montali, R. Fayer, and A. A. Lal.** 1999. Phylogenetic analysis of *Cryptosporidium* parasites on the small-subunit ribosomal RNA gene locus. Appl. Environ. Microbiol. **65:**1578–1583.
39. **Xiao, L., U. Morgan, J. Limor, A. Escalante, M. Arrowood, W. Shulaw, R. C. A. Thompson, R. Fayer, and A. A. Lal.** 1999. Genetic diversity within *Cryptosporidium parvum* and related species of *Cryptosporidium*. Appl. Environ. Microbiol. **65:**3386–3391.