FULL PAPER *Public Health*

Babesia microti-Group Parasites Compared Phylogenetically by Complete Sequencing of the CCTη Gene in 36 Isolates

Rui NAKAJIMA[1)], Masayoshi TSUJI[1)]**, Kazunori ODA[1)], Aya ZAMOTO-NIIKURA[1)], Qiang WEI[1)], Takako KAWABUCHI-KURATA[1)], Atsumi NISHIDA[1)] and Chiaki ISHIHARA[1)]*

[1)]*School of Veterinary Medicine, Rakuno Gakuen University, Ebetsu 069–8501, Japan*

(Received 17 July 2008/Accepted 8 September 2008)

ABSTRACT. *Babesia microti*, the erythroparasitic cause of human babesiosis, has long been taken to be a single species because classification by parasite morphology and host spectrum blurred distinctions between the parasites. Phylogenetic analyses of the 18S ribosomal RNA gene (18S rDNA) and, more recently, the β-tubulin gene have suggested inter-group heterogeneity. Intra-group relationships, however, remain unknown. This study was conducted to clarify the intra- and inter-group phylogenetic features of the *B. microti*-group parasites with the η subunit of the chaperonin-containing *t*-complex polypeptide 1 (CCTη) gene as a candidate genetic marker for defining the *B. microti* group. We prepared complete sequences of the CCTη gene from 36 piroplasms and compared the phylogenetic trees. The *B. microti*-group parasites clustered in a monophyletic assemblage separate from the *Babesia* sensu stricto and *Theileria* genera and subdivided predominantly into 4 clades (U.S., Kobe, Hobetsu, Munich) with highly significant evolutionary distances between the clades. *B. rodhaini* branched at the base of the *B. microti*-group parasites. In addition, a unique intron presence/absence matrix not observable in 18S rDNA or β-tubulin set the *B. microti* group entirely apart from either *Babesia* sensu stricto or *Theileria*. These results have strong implications for public health, suggesting that the *B. microti*-group parasites are a full-fledged genus comprising, for now, four core species, i.e., *U.S., Kobe, Hobetsu*, and *Munich species nova*. Furthermore, the CCTη gene is an instructive and definitive genetic marker for analyzing *B. microti* and related parasites.

KEY WORDS: *Babesia microti*, *Babesia* sensu stricto, CCTη gene, *Theileria*, tiny introns.

J. Vet. Med. Sci. 71(1): 55-68, 2009

Babesia microti, an erythroparasitic protozoon found in small wild rodents, is the causative agent of human babesiosis, a potentially fatal tick-borne zoonosis. Hundreds of human cases of babesiosis due to *B. microti* have been reported in the United States, some of them fatal [17]. Relegation of parasites to the *B. microti* group, however, was often done predominantly on the basis of the parasite morphology and host spectrum. Consequently, the *B. microti* group has long been regarded as a single species (either *B. microti* sensu stricto or the U.S.-type *B. microti*), but suspicions of misnomer and dubious classification have surfaced [11, 41]. In 1999, blood transfusion-induced babesiosis was diagnosed in a resident of Kobe, Japan, and parasites (Kobe type) isolated from the patient (index case) and from rodents (*Apodemus speciosus*) captured near the residential area of a blood-transfusion donor were analyzed by 18S ribosomal RNA gene (18S rDNA) sequences [30, 39]. Phylogenetic analysis revealed substantial divergence from the U.S.-type *B. microti* [30].

Subsequently, yet a different *B. microti*-like parasite (Hobetsu type) was found in feral rodents (*A. speciosus, Clethrionomys rufocanus* and *Microtus montebelli*) and in shrews (*Sorex unguiculatus*) in Japan [38, 43]. In 2006, 18S rDNA sequence of a European rodent isolate "Munich" (GenBank accession number AB071177) disclosed marked differences from the parasites reported in either the U.S. or Japan [11] but showed high similarity to isolates from rodents (*M. arvalis* and *M. oeconomus*) and ticks (*Ixodes ricinus*) captured in Poland [26, 32]. Furthermore, novel genotypes distinct from those commonly distributed in feral rodents in the Holarctic region have been detected repeatedly in such carnivores as fox [8], skunk [8] and river otter [5] in the U.S., raccoons in Japan and the U.S. [8, 15], and dogs in Spain (*Theileria annae* or *B. annae*) [6, 42]. 18S rDNA analyses of these *B. microti*-group parasites have suggested a monophyletic nature of the *B. microti* "species" complex, leading to an increasingly large complex under the name *B. microti* [8, 41]. Intra-group relationships, however, remain unknown. The reason this mystery has not been solved is mainly that the phylogenetic signal is weak in the low-diversity sequences afforded by 18S rDNA within the group ($\leq$ 5.4% divergence) [8].

The internal topology of the *B. microti*-group "species" complex is just beginning to emerge from β-tubulin gene sequences. In a prestigious study using β-tubulin gene sequences, the *B. microti* group branched into three or more clades, including parasites common to the Holarctic rodents (clade 1, U.S. type), various carnivores (clade 2), Alaskan voles and their relatives (clade 3), and other hosts [8]. Accordingly, multiple clusters, or clades, also branched from the β-tubulin trees constructed in our laboratory, viz., the U.S., Kobe, and Hobetsu genotypes, and most of the *B. microti*-like parasites isolated from rodents in Japan, Korea, Far East Russia and northwestern China fell into one of these three genotypes [43, 44]. Cluster distinctions

* CORRESPONDENCE TO: ISHIHARA, C., School of Veterinary Medicine, Rakuno Gakuen University, 582–1 Bunkyodai Midorimachi, Ebetsu 069–8501, Japan.
e-mail: ishihara@rakuno.ac.jp

**Dr. Tsuji died December 19, 2006.

described by the two research teams, however, showed inconsistency in certain aspects, probably owing to the sampling deviations; the parasites in the respective studies were obtained mostly from North America [8] or Northeast Asia [37, 43, 44]. Today, with escalating recruitment of new genotypes from widespread geographic regions and a multiplicity of wild mammalian hosts, evidence of intra-group genetic diversity is accruing. Clear distinctions are necessary to help raise medical and environmental awareness and promote public control of the true agent of human babesiosis.

Genes other than 18S rDNA and β-tubulin, however, have not been sequenced sufficiently in the *B. microti* group. Sequencing additional protein-coding genes and aligning the resulting phylogenetic trees would presumably shed new light on relationships among the parasites. In the Munich strain, the *eta* (η) subunit of the chaperonin-containing *t*-complex polypeptide 1 (CCTη) gene has been identified as a single-copy gene [25]. The sequence of the CCTη gene from complementary DNA (cDNA) contains an open reading frame (ORF) of 1629 nucleotides (nt) (GenBank accession number AB050948) and encodes a member of the eukaryotic cytosolic chaperonin complex (CCT or TriC) comprising eight distinct subunits, i.e., $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \eta$ and θ [21, 23, 28]. The CCT subunits have been used for phylogenetic analysis in other protists [3, 4]. Given that each of the subunits is over 500 amino acids in length and is highly conserved [20], the CCTη gene would likely have properties qualifying it to serve as a phylogenetic marker. In the present study we sequenced the CCTη gene in the *B. microti*-group parasites, first, to determine whether the gene qualifies as a phylogenetic marker of *B. microti* and *B. microti*-like parasites and, second, to clarify the genetic characteristics of the parasites making up the *B. microti* "species" complex.

MATERIALS AND METHODS

Study design: To clarify the intra- and inter-group phylogenetic features of the *B. microti*-group parasites as shown by a candidate gene other than the two genes currently available for the group, we sequenced the CCTη gene using genomic DNA and/or cDNA of parasites isolated from human and other mammalian hosts in North America, Russia, and several countries in Asia and Europe. The study included 36 piroplasms, i.e., 21 from the *B. microti* group and 15 from outside the *B. microti* group (3 were obtained from the websites). The resulting sequences were aligned for comparison of intron positions and genotype branching.

Isolates from the B. microti group: The *B. microti* group comprised 9 U.S.-type parasites, 1 Kobe type, 3 Hobetsu type, and 8 others we had designated earlier on the basis of either 18S rDNA or β-tubulin gene sequences (Table 1). The U.S. type included parasites of the Gray, GI, NM69, and AK2273 strains and the Vladivostok38, Korea8, and Xinjiang1637 isolates [30, 43, 44] in addition to the HK strain [10] and an isolate obtained from *Clethrionomys rutilus* captured in Russia (Irkutsk16). The HK and Irkutsk16 isolates were provided by Jeremy Gray (University College Dublin, Ireland) and Hiroaki Kariwa (Hokkaido University, Japan), respectively, and assigned tentatively to the U.S. type on the basis of our preliminary β-tubulin gene sequences.

The Kobe-type parasite, Kobe524, was isolated from the index-case patient of babesiosis in Japan [30]. The 3 Hobetsu-type parasites included the Da116 (isolated in Daito, Shimane, Japan), Ot1 (Hirasawa, Ohtaki, Chiba, Japan) and Ho234 strains [38]. *B. rodhaini* and the Munich isolate were donated by Kenichiro Ono (The University of Tokyo, Japan). One isolate each from 2 ticks (*I. ricinus*) captured in Poland (Poland5, Poland6) and one each from 2 bank voles (*C. glareolus*) captured in the UK (UK5, UK6), donated by Marek Sawczuk and Bogumila Skotarczak (Szczecin University, Poland) and Mitsuhiko Asakawa (Rakuno Gakuen University, Japan), respectively, had a β-tubulin gene sequence identical to that of the Munich strain (GenBank accession number AB124587); hence, we included those 4 isolates in this analysis of the *B. microti* group. Other isolates were taken from raccoons and Eurasian red squirrels in Japan [15, 37].

The Gray, GI, NM69, AK2273, Kobe524, Da116, Ot1, Ho234, and Munich strains, in addition to *B. rodhaini*, were propagated in either hamsters or NOD/shi-*scid* (SCID) mice in our laboratory [38].

Isolates from other piroplasms: The isolates from piroplasms outside the *B. microti* group (Table 1) included *B. canis vogeli, B. bovis* Mo7, *Theileria sergenti* Chitose, and *B. ovata* Oshima (known also as *B.* sp.1) [2, 12, 33, 36]. In addition, *B. gibsoni* and *T. sergenti* Ikeda were provided by Yoshimitsu Maede and Misao Onuma (Hokkaido University, Japan), respectively. These 6 isolates were propagated in a SCID mouse model substituting the red blood cells with either bovine or canine erythrocytes as a precaution to guard against possible formation of PCR chimeras by DNA strands from the different species [2, 36].

Genomic DNA samples of *B. odocoilei* Engeling [13] and *B. divergens* Purnell [14] were donated by Patricia J. Holman (Texas A&M University, U.S.A.); and *B. caballi, B. major* and *T. equi* (*B. equi*) isolates were donated by Ikuo Igarashi (Obihiro University of Agriculture and Veterinary Medicine, Japan), Masato Ohta (National Institute of Animal Health, Japan), and Hiromi Ikadai (Kitasato University, Japan), respectively. An unclassified isolate designated the YaHam strain, was obtained from an *Apodemus speciosus* we captured in Yamanaka, Shiga Prefecture, Japan, and was propagated in hamsters. Finally, annotated genomic sequences of the CCTη gene of *T. parva* and *T. annulata* were obtained from the Institute for Genomic Research (TIGR) website (http://www.tigr.org) and the Sanger Institute website (http://www.sanger.ac.uk), respectively, and the *B. bigemina* sequence was identified by BLAST at the Sanger Institute website.

Sequencing of the CCTη gene: Genomic DNA was extracted with a whole-blood DNA extraction kit (DNA

Table 1. Parasites from which the CCTη gene was sequenced for the present analysis

Parasite (Genotype)	Place of Parasite Isolation[a)]	Host of Parasite Isolation[a)] or Laboratory Host[a)]	Molecule Type	Nucleotide Length[b)] (ORF)[c)]	CCTη Gene Accession No.
B. microti group					
GI (the U.S. type)	U.S.A.	Human	Genomic DNA	2009 (1629)	AB362581
NM69 (the U.S. type)	Hokkaido, Japan	*C. rufocanus*	Genomic DNA	2010 (1629)	AB362582
Korea8 (the U.S. type)	South Korea	*A. agrarius*	Genomic DNA	2010 (1629)	AB362583
Vladivostok38 (the U.S. type)	Vladivostok, Russia	*C. rufocanus*	Genomic DNA	2010 (1629)	AB362584
Xinjiang1637 (the U.S. type)	Xinjiang, China	*L. luteus*	Genomic DNA	2009 (1629)	AB362585
Gray (the U.S. type)	U.S.A.	Human	Genomic DNA	2009 (1629)	AB362586
AK2273 (the U.S. type)	Hokkaido, Japan	*A. speciosus*	Genomic DNA	2010 (1629)	AB362587
Munich	Uncertain	*M. musculus*	Genomic DNA	2003 (1629)	AB362588
Poland5	Poland	*I. ricinus*	Genomic DNA	2003 (1629)	–[d)]
Poland6	Poland	*I. ricinus*	Genomic DNA	2003 (1629)	–[d)]
UK5	UK	*C. glareolus*	Genomic DNA	2003 (1629)	–[d)]
UK6	UK	*C. glareolus*	Genomic DNA	2003 (1629)	–[d)]
HK (the U.S. type)	Hannover, Germany	*C. glareolus*	Genomic DNA	2008 (1629)	AB366747
Irkutsk16 (the U.S. type)	Irkutsk, Russia	*C. rutilus*	Genomic DNA	2010 (1629)	AB366748
Parasite of Squirrel	Hokkaido, Japan	Eurasian Red Squirrel	Genomic DNA	1984 (1629)	AB366749
Ho234 (the Hobetsu type)	Hobetsu, Hokkaido, Japan	*A. speciosus*	Genomic DNA	1813 (1629)	AB366750
Parasite of Raccoon	Hokkaido, Japan	Raccoon	Genomic DNA	2199 (1629)	AB366751
B. rodhaini	Uncertain	*M. musculus*	Genomic DNA	1752 (1629)	AB366752
Da116 (the Hobetsu type)	Daito, Shimane, Japan	*A. speciosus*	Genomic DNA	1813 (1629)	AB366753
Ot1 (the Hobetsu type)	Hirasawa, Ohtaki, Chiba, Japan	*A. speciosus*	Genomic DNA	1813 (1629)	AB366754
Kobe524 (the Kobe type)	Kobe, Japan	Human	Genomic DNA	1888 (1629)	AB366755
Babesia sensu stricto					
B. bovis Mo7	–	–	Genomic DNA	1605 (1605)	AB367921
B. canis vogeli	–	–	Genomic DNA	1712 (1608)	AB367922
B. gibsoni	–	–	Genomic DNA	1711 (1605)	AB367923
B. odocoilei Engeling	–	–	Genomic DNA	1705 (1608)	AB367924
B. divergens Purnell	–	–	Genomic DNA	1705 (1608)	AB367925
B. major	–	–	Genomic DNA	1742 (1608)	AB367926
B. caballi	–	–	Genomic DNA	1608 (1608)	AB367927
B. ovata Oshima	–	–	Genomic DNA	1740 (1608)	AB367928
Theileria					
T. sergenti Ikeda	–	–	Genomic DNA	2821 (1893)	AB367929
T. sergenti Chitose	–	–	Genomic DNA	2630 (1782)	AB367930
T. sergenti Ikeda	–	–	mRNA	1893 (1893)	AB367931
T. sergenti Chitose	–	–	mRNA	1782 (1782)	AB367932
T. equi /B. equi	–	–	Genomic DNA	5057 (1626)	AB367933
Unclassified Piroplasm					
YaHam	–	–	Genomic DNA	1892 (1629)	AB368373

a) Information is shown for *B. microti*-group parasites only.
b) Size (nt) of the complete sequence of the CCTη gene (from the start codon to the stop codon) in each type of molecule.
c) Size (nt) of the open reading frame (ORF) of the CCTη gene.
d) Sequence identical to that of the Munich strain, and not submitted to DNA Data Bank of Japan (DDBJ).
C. rufocanus: *Clethrionomys rufocanus*; *A. agrarius*: *Apodemus agrarius*; *L. luteus*: *Lagurus luteus*; *A. speciosus*: *Apodemus speciosus*; *M. musculus*: *Mus musculus*; *I. ricinus*: *Ixodes ricinus*; *C. glareolus*: *Clethrionomys glareolus*; *C. rutilus*: *Clethrionomys rutilus*.

Extractor WB Kit; Wako Pure Chemical Industries, Osaka, Japan) according to the instruction manual supplied by the manufacturer. Total ribonucleic acid (RNA) from *T. sergenti* Chitose and *T. sergenti* Ikeda was isolated with Trizol (Invitrogen Corp., Carlsbad, CA, U.S.A.). To synthesize the cDNA, we used ReverTra Dash (Toyobo Co., Ltd, Osaka) and a polyT primer.

For amplification of the genomic DNA by polymerase-chain reaction (PCR) assay, we designed and tested primers according to the coding sequence of the CCTη gene from cDNA of the Munich strain (GenBank accession number AB050948). Finally, for Gray, NM69, and Ho234, a combination of primers CCT-186F and CCT-1401R was used. For Kobe524 and *B. rodhaini*, no combination of the primers yielded PCR products, so we used degenerate PCR primers (CCT-390F and CCT-1327R) designed on the basis of the Munich sequence and our newly acquired sequences of Gray, NM69, and Ho234. The PCR mixtures were prepared according to the instruction manual (*Takara La Taq* polymerase, Takara Bio Inc., Otsu, Japan) and denatured for 2 min at 94°C. Thermal cycling was carried out with 30 cycles each for 30 sec at 94°C, 40 sec at 50°C, and 90 sec at

72°C in a PC320-program temperature control system (Astec Co., Ltd., Fukuoka, Japan). PCR products were purified from 1% agarose gel with MinElute Gel Extraction Kit (QIAGEN K.K., Tokyo) and cloned with the TOPO-TA cloning kit (Invitrogen).

Sequencing was performed with either the ALF DNA sequencer (GE Healthcare UK Ltd., England) or the CEQ8000 automated sequencer (Beckman Coulter, Inc., CA, U.S.A.). To determine the sequence of the 5' end and 3' end of each gene, we purified, cloned, and sequenced inverted PCR products from digested and ligated genomic DNA preparations.

For most of the other piroplasms, the CCTη gene was amplified with primer combinations of CCT-262F/TBcct-3'R1 or TBcct35F/TBcct1519R after multiple combinations were tested. Degenerate PCR products were cloned and sequenced and, on the basis of these resulting sequences, inverted PCRs were carried out; the inverted PCR products were then purified, cloned, and sequenced. PCR primers targeting the regions containing the full length of the gene were then designed according to the resulting sequences and were used to amplify the complete sequences of the gene from genomic DNA samples of all the isolates and cDNA samples of *T. sergenti* Chitose and *T. sergenti* Ikeda. Finally, complete sequences were derived by direct sequencing of the PCR fragments except those of *T. equi* (*B. equi*), for which we sequenced three independent clones. Primers and oligonucleotide sequences are shown in Table 2 (those for complete sequences of *Babesia* sensu stricto and *Theileria*, for inverted PCR, and sequencing primers are not shown).

Sequence analysis and sampling from the databases: The coding region of the CCTη gene from cDNA of the Munich strain was obtained in its entirety from GenBank (accession number AB050948), and that of *T. sergenti* Chitose and *T. sergenti* Ikeda each was sequenced as described above. To verify the exon-intron structures, we aligned the cDNA sequences with the genomic DNA sequences. Alignment was constructed with MacVector 8.0 (MacVector, Inc., U.S.A.) and adjusted manually.

The CCTη gene-orthologous DNA sequences of other eukaryotes were sampled from the databases. Annotated genome sequences of *Toxoplasma gondii* and *Plasmodium falciparum* were obtained from Toxoplasma Gondii

Table 2. Oligonucleotide primers used for the analysis of the CCTη gene of piroplasms

Primers	Oligonucleotide Sequences (5'–3')	Use
CCT-186F	GGCCCCAGAGGGATGGACAAACTTAT	PCR[a)]
CCT-1401R	GTGGGATACATTCCAGGGCCTTAGCATA	PCR[a)]
CCT-390F	GATTTCATAATGGARGGMATGGCDCCTCAGA	PCR[a)]
CCT-1327R	TCYCTYARTATDCGTGATATTTCCATCT	PCR[a)]
CCT-262F	CARGATGAYGARGTKGGDGATGGWACBAC	PCR[a)]
TBcct-3'R1	GGCASGCKGCYTCAGTSGCTGMGTA	PCR[a)]
TBcct35F	TGAAGGARGGMACNGAYACWTCYCARGG	PCR[a)]
TBcct1519R	GTYTTYTTHACBAGGCTGGGCTCCCADATRCA	PCR[a)]
UScct-1	GGTGTGCCAGGCAAATGCCATTTTAAG	Primary PCR[b,d)]
UScct-4	GGTGTGAAAATAATAGGACTTTATGTA	Primary PCR[b)]
UScct-2	GCGATAAATCATACAGTTGTATCAC	Nested PCR[b)]
UScct-3	ATTAACAAAACATCCTGCGTTAGTCA	Nested PCR[b)]/PCR[g)]
BmCCT-5'A	AACCCAATATATCTGGCACRTA	PCR[c,g)]
HoCCT3'R1	ACCCGTGGCATTTTAGAGTTATCG	PCR[c)]
KoCCT3'R1	AATCCTGCATTTCTCATTGTTAAG	PCR[d)]
CCT-1F	GGCACGAGATATATTCGACACGTAT	Primary PCR[e)]
CCT-1735R	AGGTACAATGTAAAAAATTCAACA	Primary PCR[e)]
CCT-29F	TTGAGCGTTAGTATATAAGTTGCC	Nested PCR[e)]
CCT-1706R	CAACAAAACGTTACACGTTAGTCAC	Nested PCR[e)]
RacCCTF	GTTATAGACACAGAATTGATATGCGTAGC	Primary PCR[f)]
RacCCTR	ACTAGTCCATGCAACTAGTGGCATGATG	Primary PCR[f)]
RacCCTF2	AGACACCCATACATTATATGCCTG	Nested PCR[f)]
RacCCTR2	AAGACATGTGTAGCCAAGCACCTC	Nested PCR[f)]
rodCCT5'F	ATGGTGCCTCACAGTATCCAAC	PCR[h)]
rodCCT3'R	ATTCCCTCTTTTCTCTATTGTGGCG	PCR[h)]

a) Primers for amplification of the partial CCTη gene sequences.
b) Primers for amplification of the complete CCTη gene sequence of the U.S.-type parasites.
c) Primers for amplification of the complete CCTη gene sequence of the Hobetsu-type parasites.
d) Primers for amplification of the complete CCTη gene sequence of the Kobe-type parasites.
e) Primers for amplification of the complete CCTη gene sequence of the Munich-type parasites.
f) Primers for amplification of the complete CCTη gene sequence of the parasite from raccoon.
g) Primers for amplification of the complete CCTη gene sequence of the parasite from squirrel.
h) Primers for amplification of the complete CCTη gene sequence of *B. rodhaini*.

Genome Resource Database (ToxoDB, http://www.toxodb.org/toxo/) and the Sanger Institute website, respectively. Annotated genome sequences of *Cryptosporidium parvum* (chromosome 4, accession no. AAEE01000008), *Cryptosporidium hominis* (chromosome 4, accession no. NW_666800), *Tetrahymena pyriformis* (U46030), *Paramecium tetraurelia* (macronuclear complete genome, NC_006058), *Homo sapiens* (chromosome 2, NT_022184), *Mus musculus* (chromosome 6, NT_039353), *Arabidopsis thaliana* (chromosome 3, NC_003074), and *Oryza sativa* (chromosome 6, AP003935) were obtained from GenBank. The CCTη sequences of *Plasmodium vivax, Tetrahymena thermophila*, and *Neospora caninum* were identified with BLAST at TIGR website and the Sanger Institute website, respectively. The orthology of each of these candidate sequences was verified by multiple alignments and phylogenetic analyses.

Alignment and mapping of introns and functional domains: Amino acid sequences and ORFs were aligned with MacVector 8.0, and the intron positions at the corresponding nucleotide sequences were marked on the alignments. Only introns in exactly the same codon phase (the position of the intron within a codon) were considered to be in identical positions. For amino acid sequences, functional domains, equatorial domain (ATP-binding domain), apical domain (the region that binds substrate proteins), and intermediate domain, as predicted in mouse and other eukaryotic CCTη [3, 22], and ATP-binding motifs conserved universally in chaperonin families [16, 23] were deduced by alignment with mouse CCTη sequence (GenBank accession number P80313).

Phylogenetic analysis: DNA (ORFs) and amino acid sequences were aligned with the Clustal W Alignment program [34] in MacVector 8.0. After divergent ends of the gene or the proteins and regions of ambiguous alignment and gaps were removed from the alignments, only confidently aligned positions (nearly the full length of the ORFs or proteins from the *B. microti* group) were used for phylogenetic reconstruction. Phylogenetic trees were constructed by the neighbor-joining method [31]. Support for tree nodes was calculated with 1000 bootstrap replicates by use of the bootstrap tree algorithm.

RESULTS

CCTη gene analysis: Complete sequencing of the CCTη gene was achieved for all the isolates, and 31 of the sequences were submitted to the DNA Data Bank of Japan (DDBJ) (accession numbers shown in Table 1). In all the *B. microti*-group parasites, including *B. rodhaini*, the ORFs were the same length, 1629 nt (Table 1), encoding a peptide of 542 amino acids. In the respective putative equatorial domains, all the *B. microti*-group sequences disclosed highly conserved ATP-binding/ATP-hydrolysis motifs that are also highly conserved in universal group I and group II chaperonins (Fig. 1).

In the genus *Babesia* sensu stricto, the length of the ORFs ranged narrowly from 1605 (*B. bovis* and *B. gibsoni*) to 1608 nt (*B. canis, B. odocoilei, B. divergens, B. major, B. ovata, B. bigemina* and *B. caballi*). In *Theileria* the ORF sizes varied widely: 1626 nt for *T. equi* (*B. equi*), 1731 nt for *T. annulata*, 1740 nt for *T. parva*, 1782 nt for *T. sergenti* Chitose, and 1893 nt for *T. sergenti* Ikeda. The ORF sizes for *T. annulata* and *T. parva* were confirmed by annotated genomic sequences obtained from the websites, and those for *T. sergenti* Chitose and *T. sergenti* Ikeda were confirmed by experiments sequencing the CCTη gene from cDNAs. The ORF of YaHam was 1629 nt (Table 1).

In our sequences, substantial inter-group differences were disclosed in the *B. microti* group, with a mere 82.6% to 88.9% identity between the genotype pairs (U.S., Kobe, Hobetsu, Munich) (Table 3), as opposed to notably high intra-group similarities within each genotype itself. The U.S. genotype had 96.8% to 100% identity (Table 4). In the Hobetsu type, nucleotide sequences of the Ho234 and Da116 had 100% identity, and each of these 2 had 99.9% identity with Ot1. The amino acid sequences of the Ho234, Da116, and Ot1 isolates were all identical. Munich and its close relatives (Poland5, Poland6, UK5 and UK6) all aligned with 100% identity in both the nucleotide and amino acid sequences. *B. rodhaini* and the isolate from raccoon were distantly related to the U.S. type and to the Kobe and Hobetsu types (Table 3).

Most introns of the *B. microti* group were too short for evaluation of sequence homogeneity (Table 5), but at the 10th position both the Munich strain (73.7% identity) and the isolate from squirrel (78.2% identity) related closely to the U.S.-type Gray strain. Introns at the10th position differed widely in size among the U.S. (247 to 248 nt), Kobe (132 nt), and Hobetsu (54 nt) types (Table 5); consequently, alignment of these three types was insignificant. The isolate from raccoon, in spite of having comparable intron sizes at the 10th position, did not align with the U.S. type, thus indicating heterogeneity between the two types.

Structure of the CCTη gene: With the sole exception of *B. rodhaini*, every CCTη gene of the *B. microti*-group parasites consistently contained 6 introns (19 to 254 nt long) (Table 5, Figs. 1 and 2). All 6 introns had standard 5'-GT...AG-3' boundaries, and the coding sequences aligned readily with no size heterogeneity at the exon-intron junctions. The intron at the 2nd position and the intron at the 12th (last) position (Fig. 2) occurred after the first codon bases (phase 1), and the other 4 introns occurred between the 2 codons (phase 0) (Figs. 1 and 2). All 6 intron positions were well conserved except in *B. rodhaini*, in which the intron at the 2nd position was missing (Figs. 1 and 2).

Within the order Piroplasmida, *Theileria* parasites carried 7 spliceosomal introns, as verified by the extensive cDNA analyses (*T. sergenti* Chitose and *T. sergenti* Ikeda) and annotated genome sequences (*T. annulata* and *T. parva*). With the sole exception of the 2nd-position intron, which is universal to the order Piroplasmida, however, none of the introns of the genus *Theileria* occurred in the same position as those found in the *B. microti*-group parasites (Fig. 2).

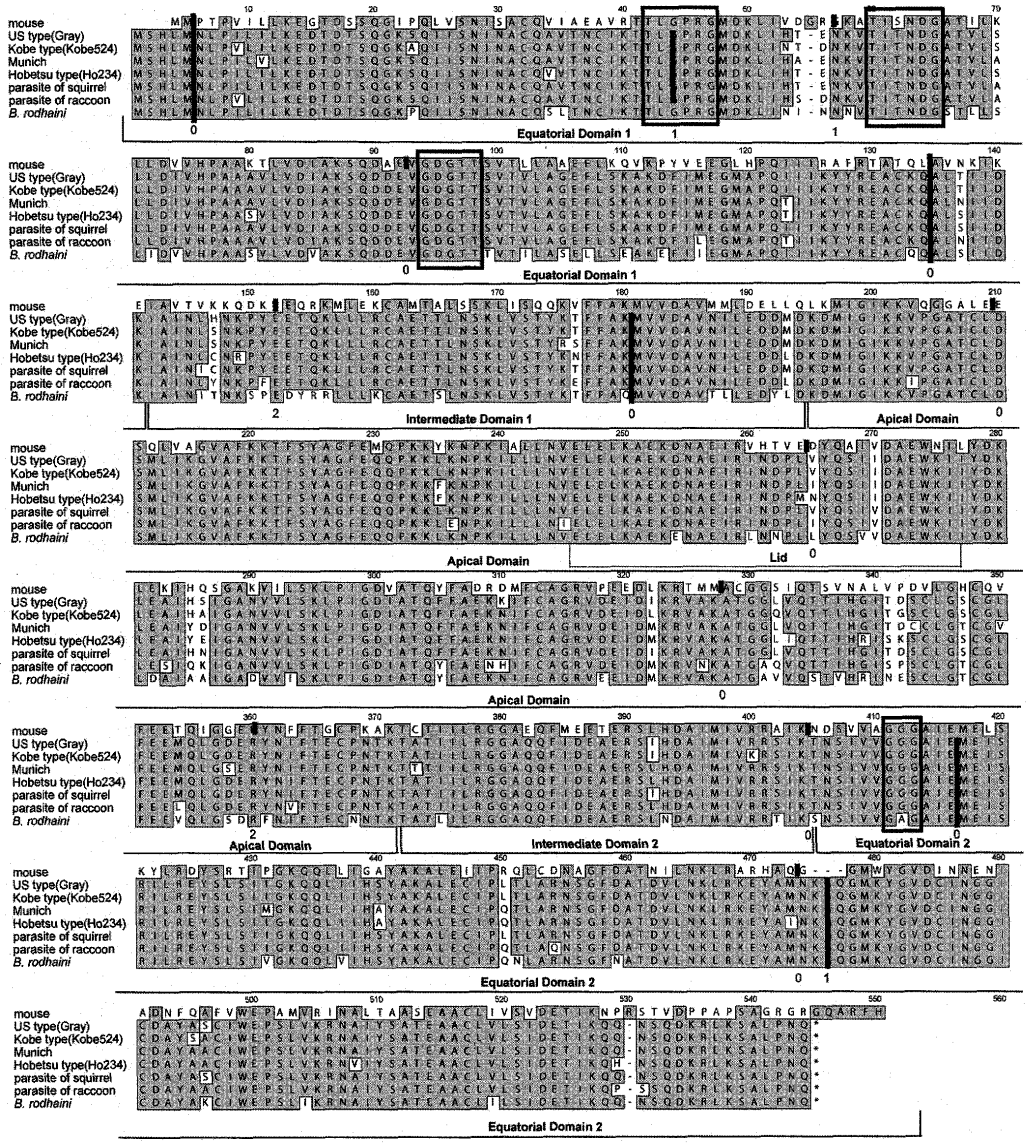

Fig. 1. Intron positions and alignment of the predicted amino acid sequences of CCTη from the *B. microti*-group parasites. Of the 21 sequences, 7 representative sequences are shown (Gray, Kobe524, Munich, Ho234, parasite from squirrel, parasite from raccoon, and *B. rodhaini*). An amino acid sequence of mouse (*Mus musculus*) CCTη (accession number P80313) is also shown. Functional domains predicted in mouse and other eukaryotic CCTη (equatorial, intermediate, apical and lid) are noted. Gray shading: conserved amino acids; Vertical red lines: intron positions; Boxes: ATP-binding motifs conserved universally in chaperonin families; Numerals below the red lines: intron phases (the position of the intron within a codon).

The 7 intron positions found in the *Theileria* parasites were well conserved except the one at the 11th position (Fig. 2), where intra-group diversity was evident in the presence/absence matrix.

In stark contrast to the parasites of the intron-rich *B. microti* group and to those of *Theileria*, the parasites in *Babesia* sensu stricto carried either no universal 2nd-position intron (*B. bovis* and *B. caballi*) or only 1 in that position (7 of the 9 species tested) (Fig. 2). Of the 12 introns in the order Piroplasmida (Fig. 2), 3 were exclusive to *Theileria* (3rd, 8th and 11th positions), and the other 9 occurred outside the piroplasmids, in the coccidian parasites *Toxoplasma gondii* and *Neospora caninum* (Fig. 3). As many as 4/6 of the intron positions in the *B. microti* group and 2/7 in *Theileria* were found to be conserved in mammals (*Homo sapiens* and *Mus musculus*) and/or flowering plants (*Arabidopsis thaliana* and *Oryza sativa*) (Fig. 3).

Tiny introns, the smallest known (~19 to 23 nt in length), were observed frequently in the *B. microti*-group parasites (Table 5), particularly at the 1st, 4th and 12th positions (100%, 21/21) and at the 2nd (95%, 20/21), and 1 at the 6th (5%) but none at the 10th position. In *Babesia* sensu stricto and *Theileria*, however, the introns were all of typical length even at the 2nd position, where a universal but tiny intron

Table 3. Pairwise comparison of CCT η genes (ORFs) and amino acids

Parasite (Strain)	Percentage of amino acid sequence identity								
	B. microti group							*Babesia* sensu stricto	*Theileria*
	U.S.[a] (Gray)	Munich[a] (Munich)	Kobe[a] (Kobe524)	Hobetsu[a] (Ho234)	Parasite of Squirrel	Parasite of Raccoon	*B. rodhaini*	*B. ovata* (Oshima)	*T. sergenti* (Chitose)
U.S.[a] (Gray)	–	95.6	97.6	95.2	98.7	93.2	85.4	65.1	64.2
Munich[a] (Munich)	**88.9**	–	94.6	95.0	95.6	93.2	84.7	65.7	64.4
Kobe[a] (Kobe524)	**86.4**	**83.2**	–	94.3	97.0	93.2	85.4	65.5	64.4
Hobetsu[a] (Ho234)	**85.0**	**85.1**	**82.6**	–	95.2	92.6	85.1	65.7	64.8
Parasite of Squirrel	**93.4**	**87.8**	**85.9**	**85.3**	–	93.4	86.0	65.5	63.8
Parasite of Raccoon	**76.8**	**76.1**	**78.1**	**78.8**	**76.6**	–	84.3	65.3	64.4
B. rodhaini	**75.0**	**73.7**	**73.9**	**73.4**	**74.7**	**73.9**	–	64.0	63.3
B. ovata (Oshima)	**61.3**	**62.0**	**61.0**	**60.4**	**62.0**	**59.8**	**59.4**	–	67.5
T. sergenti (Chitose)	**61.3**	**61.3**	**61.7**	**61.5**	**61.8**	**61.8**	**60.1**	**65.1**	–
	Percentage of nucleotide sequence identity								

ORFs: open reading frames.
Percent shown in regular typefont indicate identity of amino acid sequences.
Percent shown in boldface indicate identity of nucleotide sequences (ORFs).
a) A genotype in the *B. microti* group.

Table 4. Comparison of CCT η genes (ORFs) and amino acids within the U.S.-type parasites

Parasite	Percentage of amino acid sequence identity								
	NM69	AK2273	Korea8	Vladivostok38	Irkutsk16	Gray	GI	HK	Xinjiang1637
NM69	–	100.0	99.8	99.6	99.6	99.6	99.6	99.4	99.6
AK2273	**100.0**	–	99.8	99.6	99.6	99.6	99.6	99.4	99.6
Korea8	**99.6**	**99.6**	–	99.8	99.8	99.8	99.8	99.6	99.8
Vladivostok38	**99.5**	**99.5**	**99.8**	–	100.0	99.8	99.8	99.4	99.6
Irkutsk16	**99.4**	**99.4**	**99.7**	**99.8**	–	99.8	99.8	99.4	99.6
Gray	**98.7**	**98.7**	**98.8**	**98.7**	**98.6**	–	100.0	99.4	99.6
GI	**98.7**	**98.7**	**98.8**	**98.7**	**98.6**	**100.0**	–	99.4	99.6
HK	**97.2**	**97.2**	**97.3**	**97.3**	**97.1**	**97.4**	**97.4**	–	99.8
Xinjiang1637	**96.9**	**96.9**	**97.0**	**97.0**	**96.8**	**97.1**	**97.1**	**99.3**	–
	Percentage of nucleotide sequence identity								

ORFs: open reading frames.
Percent shown in regular typefont indicate identity of amino acid sequences.
Percent shown in boldface indicate identity of nucleotide sequences (ORFs).

occupied the *B. microti* group. Notably, *T. equi* (*B. equi*) carried 2 extraordinarily large (>1000 nt) introns not reported in the other piroplasms, in addition to 4 relatively short (34 to 65 nt) introns (Table 5). The YaHam strain was also characterized by relatively short (37 to 64 nt) introns (Table 5).

Phylogenetic analysis: In both the nucleotide-based tree and the amino acid-based tree, all the *B. microti*-group parasites clustered consistently in a monophyletic assemblage with a highly significant bootstrap value (1000), as did also *Babesia* sensu stricto and *Theileria* in their respective groups (Fig. 4). This inter-group diversity appears to rank *B. microti* as a genus along with the other two genera. On the intra-group level, the *B. microti* group branched mostly into four clades, i.e., U.S., Kobe, Hobetsu, and Munich. Munich along with Poland5, Poland6, UK5 and UK6 formed a common clade in both the nucleotide- and amino acid-based trees (Fig. 4). The isolate from Eurasian red squirrel in Japan fell close to but distinctly separate from the U.S. clade, and the isolate from raccoon in Japan branched outside all of these four or five divisions, further bearing out intra-group heterogeneity of the *B. microti*-group parasites (Fig. 4). *B. rodhaini* branched at the base of all the *B. microti*-group parasites (Fig. 4).

The nucleotide-based tree exhibited high resolution extending to the terminal branches. The U.S.-type parasites clustered according to geographic origin, e.g., Japan (NM69 and AK2273, bootstrap value 996), Far East Russia and East Asia including Japan (NM69, AK2273, Korea8 and Vladivostok38, bootstrap value 998), United States (Gray and GI, bootstrap value 1000), Europe and northwestern China (HK and Xinjiang1637, bootstrap value 1000) (Fig. 4A). Such geographic clustering was not evident in the amino acid-based tree, as indicated by the substitutions,

Table 5. Nucleotide length of introns in the CCTη gene of the parasites in the *B. microti* group and other piroplasms

Parasite (Genotype)	Intron Positions											
	1	2	3	4	5	6	7	8	9	10	11	12
T. annulata	–	123	88	–	129	–	121	98	137	–	196	–
T. parva	–	172	106	–	79	–	93	95	175	–	135	–
T. sergenti Chitose	–	258	121	–	100	–	86	187	96	–	–	–
T. sergenti Ikeda	–	254	121	–	128	–	91	237	97	–	–	–
T. equi/B. equi	–	1349	36	–	1912	–	35	34	65	–	–	–
YaHam	–	43	37	–	40	–	38	41	64	–	–	–
Gray (U.S.)	21	22	–	20	–	49	–	–	–	247	–	21
GI (U.S.)	21	22	–	20	–	49	–	–	–	247	–	21
NM69 (U.S.)	21	22	–	20	–	49	–	–	–	248	–	21
AK2273 (U.S.)	21	22	–	20	–	49	–	–	–	248	–	21
Vladivostok38 (U.S.)	21	22	–	20	–	49	–	–	–	248	–	21
Irkutsk16 (U.S.)	21	22	–	20	–	49	–	–	–	248	–	21
Korea8 (U.S.)	21	22	–	20	–	49	–	–	–	248	–	21
Xinjiang1637 (U.S.)	21	22	–	20	–	49	–	–	–	247	–	21
HK (U.S.)	21	22	–	20	–	48	–	–	–	247	–	21
Munich (Munich)	21	22	–	20	–	49	–	–	–	241	–	21
Poland5 (Munich)	21	22	–	20	–	49	–	–	–	241	–	21
Poland6 (Munich)	21	22	–	20	–	49	–	–	–	241	–	21
UK5 (Munich)	21	22	–	20	–	49	–	–	–	241	–	21
UK6 (Munich)	21	22	–	20	–	49	–	–	–	241	–	21
Kobe524 (Kobe)	20	22	–	20	–	45	–	–	–	132	–	20
Ho234 (Hobetsu)	21	22	–	20	–	46	–	–	–	54	–	21
Da116 (Hobetsu)	21	22	–	20	–	46	–	–	–	54	–	21
Ot1 (Hobetsu)	21	22	–	20	–	46	–	–	–	54	–	21
Parasite of Squirrel	21	22	–	20	–	47	–	–	–	224	–	21
Parasite of Raccoon	20	22	–	20	–	254	–	–	–	231	–	23
B. rodhaini	20	–	–	21	–	19	–	–	–	43	–	20
B. canis	–	104	–	–	–	–	–	–	–	–	–	–
B. gibsoni	–	106	–	–	–	–	–	–	–	–	–	–
B. odocoilei	–	97	–	–	–	–	–	–	–	–	–	–
B. divergens	–	97	–	–	–	–	–	–	–	–	–	–
B. major	–	134	–	–	–	–	–	–	–	–	–	–
B. bigemina	–	134	–	–	–	–	–	–	–	–	–	–
B. ovata	–	132	–	–	–	–	–	–	–	–	–	–
B. caballi	–	–	–	–	–	–	–	–	–	–	–	–
B. bovis	–	–	–	–	–	–	–	–	–	–	–	–

most of which occurred at the third nucleotide of the triplet genetic codon (Fig. 4B).

In the nucleotide-based tree, the stratification showed the U.S. type at the uppermost branches, then Munich (bootstrap value 956), followed by Kobe (bootstrap value 821) and, most distantly, Hobetsu at the lower branches (bootstrap value 1000) (Fig. 4A). This branching order was compatible with the intron sequences at the 10th position (described above). In the amino acid-based tree, on the other hand, the branching of Kobe and Munich was reversed, the Kobe type branching next to the U.S. type, followed by Munich (Fig. 4B). Hobetsu branched the most distant from the U.S. type in both trees (Fig. 4).

DISCUSSION

This study had three major findings. First, on the inter-group level, the parasites clustered into three overarching categories of what appear to be equal rank, i.e., *B. microti*, *Babesia* sensu stricto, and *Theileria*, as shown by the intron positions and phylogenetic trees. This suggests a *B. microti genus nova* ranking apart from but equal to the other 2 recognized genera. Second, on the intra-group level, a well-demarcated four-way division was seen repeatedly in the *B. microti* group, as demonstrated by the unique intron sizes and by both the nucleotide and amino acid phylogenetic trees. These unequivocal divisions seem to point to the formation of the *U.S., Kobe, Hobetsu*, and *Munich species nova* within the possible *B. microti* genus. Third, the study bears out that the CCTη gene is an instructive and valuable genetic marker for differentiating the *B. microti* and related parasites on the basis of inter-group and intra-group genetic attributes. By reexamining the outdated practice of classifying the parasites by their morphology, the more robust genome-based approach used in this study has strong implications for public health with special reference to human babesiosis. Reclassification of the *B. microti*-group parasites according to their genetic attributes could help raise

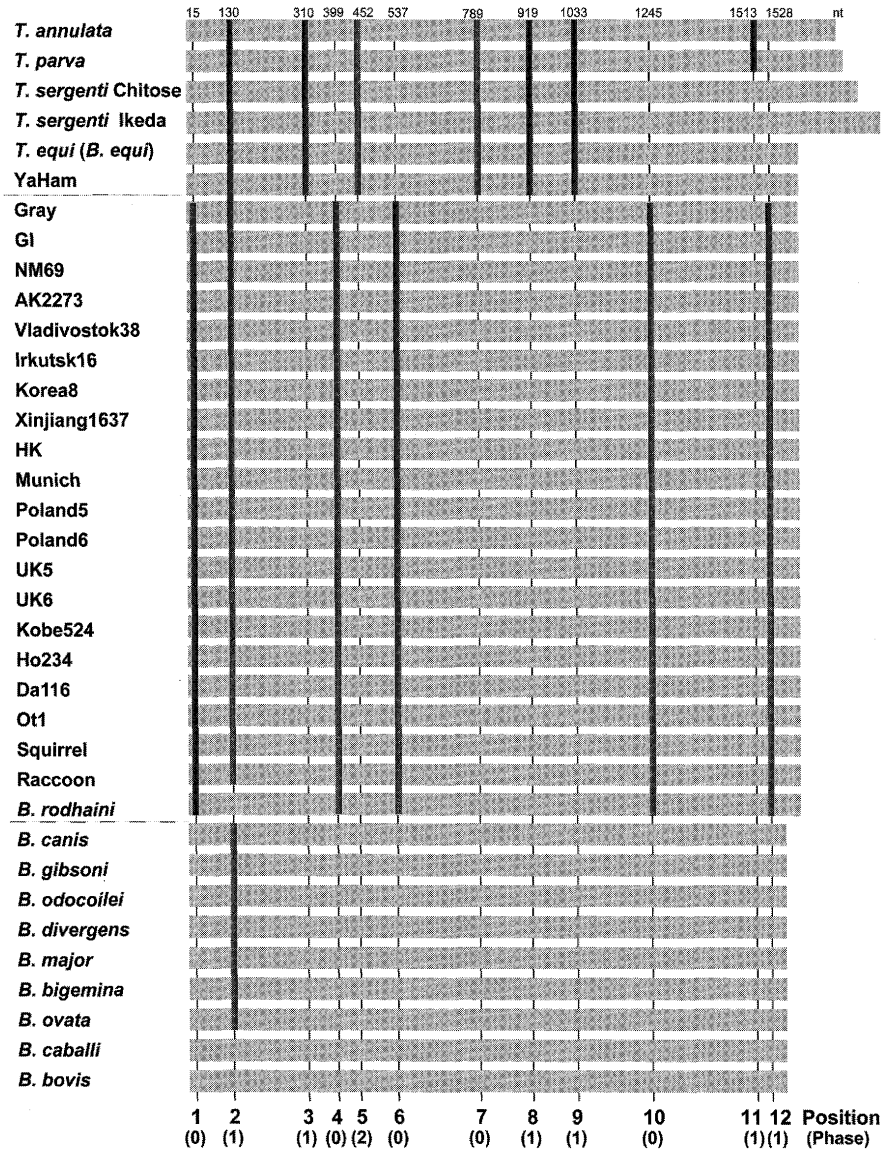

Fig. 2. Intron positions occupied in the CCTη gene from the 36 parasites (21 from the *B. microti* group and 15 other piroplasms) in the order Piroplasmida. Numerals across the top of the schema indicate the number of nucleotides in the CCTη gene open reading frame (ORF) of *T. annulata* at the respective intron positions. Numerals across the bottom indicate the intron positions, and the intron phases (the position of the intron within a codon) appear in parentheses. nt: nucleotides; horizontal gray shading: ORF of the CCTη gene; thick vertical line: position occupied by an intron; squirrel: parasite from squirrel; raccoon: parasite from raccoon.

medical and environmental awareness of potentially serious infection and, at the same time, stabilize the taxonomy.

Our main point of departure from other studies was the exploration of a gene other than the two genes commonly used so far in defining the large *B. microti* group of parasites. To date, eukaryotic multicopy genes such as rDNA (e.g., 18S rDNA) have been the most widely used genetic markers for investigating ecology and evolution, but because of intra-genomic variations in multicopy genes the rDNA can potentially confound evaluation [35]. Conversely, single-copy genes do not confound the phylogenetic reconstruction. Apicomplexan CCTη sequences offer a single-copy orthologous gene, as established by Nishisaka *et al.* [25] and recent genome projects. Moreover, whereas a paucity of material is available from protein-coding genes in the *B. microti*-group parasites, the CCTη gene in the present study afforded construction of the amino acid-based tree as well as the nucleotide-based tree (Fig. 4). This capability makes the gene especially reliable and useful for phylogenetic analysis. While informed by valuable studies relying on 18S rDNA and the β-tubulin gene [8, 9, 26, 30, 32, 37, 38, 41, 43, 44], if we are to identify the true character

	Intron Positions 1	2	3	4	5	6	7	8	9	10	11	12		
Theileria		+	+		+		+	+	+		(+)		7	PI
***B. microti* group**	+	(+)		+		+				+		+	6	
***Babesia* sensu stricto**		(+)											1	
(a) *Toxoplasma* and *Neospora*	+	+		+	+	+	+		+	+		+	9/10	AP
(b) *Cryptosporidium*													0/0	
(c) *Plasmodium*	+												1/1	
(d) *Tetrahymena*	+	+			+								3/5	AL
(e) *Paramecium*													0/4	
(f) Mammals	+			+	+		+						4/11	PR
(g) Flowering Plants	+			+						+		+	4/13	AN / PL

Fig. 3. Distribution of the 12 introns shown in the CCTη gene of the order Piroplasmida (boldface) that are shared by the CCTη gene in other eukaryotes (regular typeface). Numerals across the top indicate the 12 intron positions disclosed in the piroplasmids in this study. +: the presence of an intron; (+): the presence of an intron that exhibits intra-group presence/absence diversity. Numerals shown vertically on the right of the 3 piroplasmid groups indicate the number of introns found in the respective parasites in the group (in total, all piroplasms listed in Fig. 2 are represented). Numerals listed vertically on the right of the other eukaryotes indicate the number of introns shared with the piroplasmids per total number of introns in the respective eukaryotes represented. The eukaryotic species represented are (a) *T. gondii* and *N. caninum*; (b) *C. parvum* and *C. hominis*; (c) *P. falciparum* and *P. vivax*; (d) *T. pyriformis* and *T. thermophila*; (e) *P. tetraurelia*; (f) *Homo sapiens* and *Mus musculus*; (g) *Arabidopsis thaliana* and *Oryza sativa*. Only introns in exactly the same codon phase were considered to be in the same position. PI: piroplasms; AP: apicomplexans; AL: alveolates; PR: protists; AN: animals; PL: plants.

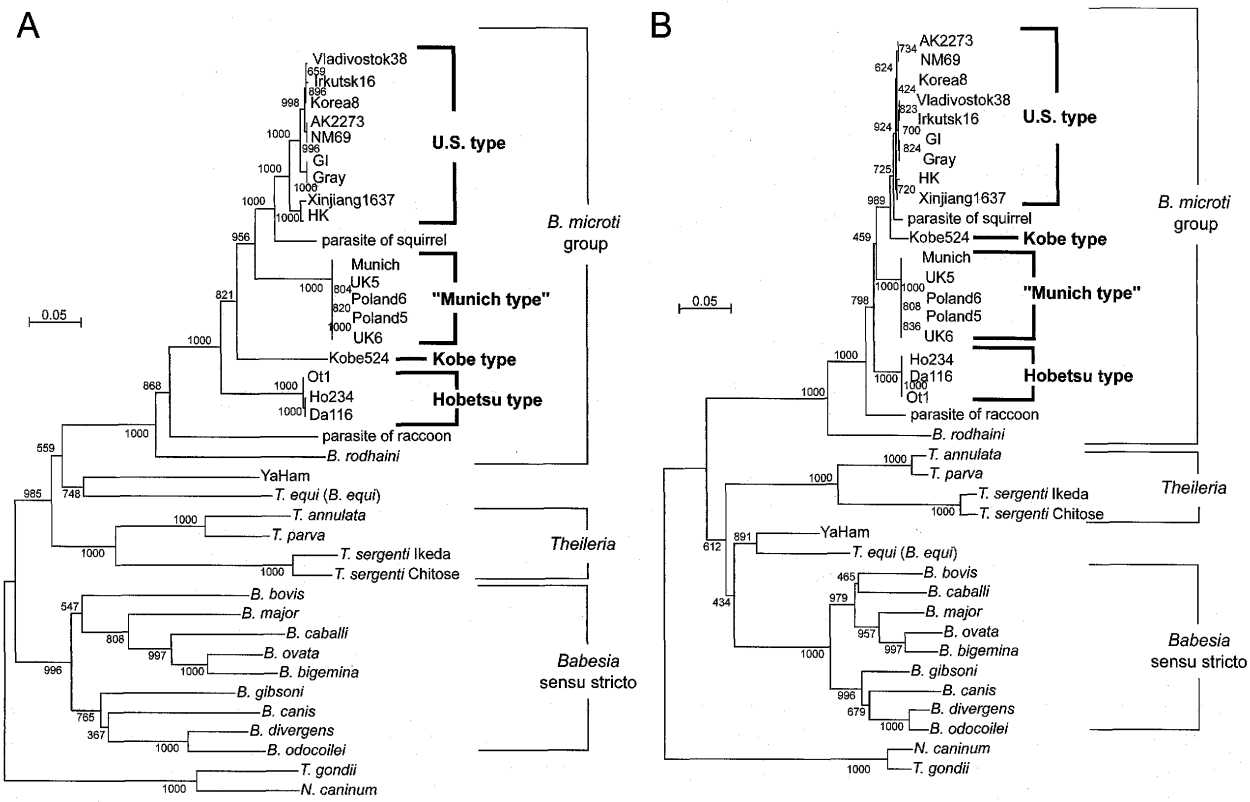

Fig. 4. Phylogenetic trees derived from CCTη. (A) nucleotide-based tree and (B) amino acid-based tree. The trees were constructed by the neighbor-joining method. Numbers at the nodes indicate bootstrap support from 1000 replications. *T. gondii*: *Toxoplasma gondii* (used as outgroup); *N. caninum*: *Neospora caninum* (used as outgroup). Scale bar: the inferred number of substitutions per site.

of the infectious agent *B. microti* and separate the "related" but noninfectious culprits, then the research arsenal must make use of additional genes such as CCTη in the target parasites.

Our second point of departure from some of the studies to date was the complete sequencing of the gene rather than a selected segment assumed to be representative of the whole. Obtaining a full view of the intron/exon structure is important for eliminating false assumptions derived from a narrow portion that fails to provide a view of the 1st and last intron positions.

The CCTη gene unveiled patterns not described in the literature before and dramatically expanded the level of scrutiny enabled by either the 18S rDNA or the β-tubulin gene, thus affording a far more detailed look into the genetic makeup of the parasites. The rich abundance and instructive array of tiny introns (~19 to 23 nt) occurred exclusively in the *B. microti* group, whereas the larger, more typical introns characterized *Babesia* sensu stricto and *Theileria* (Table 5).

The tiny introns occupied not only the 2nd position but 5 of the total 6 positions (1st, 2nd, 4th, 6th, 12th) in *B. microti* (Table 5). Curiously, although the 2nd position is universal for the order Piroplasmida, our 2nd-position introns in *Babesia* and *Theileria* were 97 to 1349 nt long except for the relatively short 43-nt intron of YaHam in that position (Table 5). The reason the tiniest introns occurred preferentially in the *B. microti*-group parasites is not clear. Other than in piroplasmids, such tiny introns have not been demonstrated in the CCTη gene from alveolates such as the apicomplexans *Toxoplasma, Neospora*, and *Plasmodium* and the ciliate *Tetrahymena*. Comparatively small (25 to 29 nt) introns reside in the CCTη gene of ciliate *Paramecium* (GenBank accession number NC_006058), and the tiny introns (~19 to 23 nt) have been reported in macronuclear genome of *Paramecium* [29] and nucleomorph genome of chlorarachniophytes [7]. Gilson *et al.* surmised that the tiny introns in chlorarachniophytes may originate as miniaturizations of the larger introns present in the endosymbiont at the time of capture [7]. They based their reasoning on the finding that 77% of the 137 introns they analyzed in 44 highly conserved nucleomorph genes occupied the same position as did the large introns in homologs from either *Chlamydomonas* or *Arabidopsis* [7]. Curiously, in the present study the 6 intron positions found in *B. microti* are also well conserved in the CCTη gene from apicomplexans other than piroplasmids, i.e., *Toxoplasma* and *Neospora* (Fig. 3), whose homologous introns are 218 to 978 nt long (ToxoDB and the Sanger Institute website). Both by size and position, the tiny introns in our study lead us to conjecture that the tiny introns may originate from longer spliceosomal introns present in the common ancestor of all piroplasmids. Further study is warranted.

Within the *B. microti* group, the CCTη gene provided sharp precision of details not only in the introns (Table 5) but also in the phylogenetic topologies (Fig. 4). In the phylogenetic trees derived from the CCTη gene sequences and protein sequences, the evolutionary distances among the U.S., Kobe, Hobetsu, and Munich groups were highly significant and were comparable to distances well established between species in the order Piroplasmida, such as between *T. parva* and *T. annulata* and between *B. odocoilei* and *B. divergens* (Fig. 4). These phenomena give rise to the strong possibility that U.S., Kobe, Hobetsu, and Munich are species themselves. Taken together, the introns and phylogenetic trees in the present study not only reinforce the intra-group tri-classification proposed previously (U.S., Kobe, Hobetsu) [37, 38, 43, 44] but, more important, identify Munich as a fourth distinctive group making up the *B. microti* complex. This finding is in keeping with suspicions raised by 18S rDNA [11, 26, 32] and β-tubulin [37] gene sequencing analyses, but the CCTη gene adds a new dimension dramatically expanding and enriching the work achieved so far. Surprisingly, in the CCTη gene sequences the 5 Munich isolates had 100% identity among themselves (intra-Munich) in spite of having geographically diverse origins (UK, Poland and Germany), but they had negligible identity with the U.S. (88.9%), Kobe (83.2%) and Hobetsu (85.1%) parasites (Table 3). These findings seem to dictate division of the large *B. microti* complex into four different species. The size variations of the 10th-position introns found in these four types also provide a crucial line of evidence in support of the genotyping demonstrated by the CCTη trees (Table 5).

Particularly striking is the finding that the intron presence/absence matrix for the *B. microti* group differs vastly from that of either *Babesia* sensu stricto or *Theileria*. This demarcation attests to the lack of commonality among these three overarching groups (or genera) and undermines the long-held convention of equating the *B. microti* group with *Babesia* sensu stricto, a persistent but ungrounded complication associated with the simplistic classification of the parasites by morphology. The paramount advantage that the CCTη gene demonstrates over the β-tubulin gene is the notably clear array of introns in the CCTη analyses as compared to the lack of introns in *B. microti* β-tubulin. Distinct patterns in the intron positions of the three overarching groups were disclosed here for the first time. Only 2 intron positions are in the β-tubulin gene in the order Piroplasmida [44]; but in the CCTη gene, among the 12 intron positions in our 36 isolates of the order Piroplasmida, 5 positions (1st, 4th, 6th, 10th and 12th) were conserved exclusively in the *B. microti* group and 6 other positions (3rd, 5th, 7–9th and 11th) were specific for *Theileria*. The remaining position (2nd), the only intron found in *Babesia* sensu stricto, is distributed universally across all three genera (Fig. 2). This intron presence/absence matrix is entirely conserved except at the 2 positions that exhibit intra-group presence/absence diversity (2nd position for both *B. microti* and *Babesia* sensu stricto, and the 11th position for *Theileria*) (Fig. 2). In addition to delineating the genetic features of the parasites, the intron orientations disclosed in the *B. microti* group should prove phylogenetically meaningful in describing rare genomic changes such as indels [18, 19, 24], for temporal

landmarks of evolution.

Furthermore, *T. equi* was found to be a *de facto* member of the genus *Theileria*, as borne out by the intron presence/absence matrix (Fig. 2). For years, ongoing debate has argued whether *T. equi* (*B. equi*) ought to be assigned to *Babesia* sensu stricto or assigned to *Theileria*, but *T. equi* has eluded placement because the introns were not available in the genes analyzed so far and because phylogenetic topologies for this parasite have been inconsistent, depending on the gene [1, 27, 40] and type of sequence (nucleotide or amino acid) [40]. In spite of branching in slightly different locations in our CCTη tree topologies (Fig. 4), *T. equi* exhibited a distinctive intron orientation identifying *T. equi* with *Theileria*. This finding reinforces the demarcation shown by the distances at the genus level.

Moreover, the intron presence/absence pattern of the heretofore unclassified isolate YaHam was also identical to that of the *Theileria* genus at all intron positions except the 11th (Fig. 2). YaHam is most closely related to the unclassified Piroplasmida spp. (accession numbers AB188086 and AB242140) and, next, to *Cytauxzoon felis* (accession number AY679105), according to 18S rDNA sequences in the current GenBank database (the 18S rDNA sequence for YaHam is unpublished). Further study is warranted to determine whether the CCTη intron presence/absence matrix of YaHam concurs with that of *C. felis*.

Taken together, the present results bring the study of *B. microti*-related parasites to the brink of an imminent paradigm shift and facilitate a new takeoff point from which the large conglomerate of parasites can be clarified by way of complete sequences of more than one gene. The CCTη gene adds a dramatic intron presence/absence dimension not observed in the parasites before. From discovery of the protozoon in wild rodents 4 decades ago on Nantucket Island, off the coast of Massachusetts, to the current multiplicity of wild mammalian hosts around the world, the *B. microti* group of parasites has evaded official classification. Today the availability of genome sequence databases and sophisticated laboratory technology may help bring about consistency in the nomenclature and replace vernacular names that overlap and confuse. Accruing evidence strongly suggests that reclassification of the so-called "*B. microti* group" would contribute favorably toward raising medical and environmental awareness of the true parasites constituting what may be the *B. microti* genus and could conceivably lead to improved management and control of human babesiosis. This vantage point is of crucial concern today when babesiosis infection could be life-threatening to more and more persons whose health is compromised by immunosuppression such as that associated with organ transplant or HIV. As a step toward reaching taxonomic consensus, multicenter studies are warranted for validating and further delineating the *B. microti* genus and related parasites with complete sequencing of the CCTη gene from parasites once dubbed members of the large *B. microti* "species" complex at a time when classification was fraught with guesswork and uncertainties before the more robust genome-based analyses were possible.

In conclusion, this study of 36 piroplasms has strong implications for public health, documenting, first, that the *B. microti* "species" complex is not a single species as has been inadvertently purported for four decades but that this complex is a heterogeneous conglomerate made up of at least four species at this time. Second, the study establishes that the CCTη gene is an instructive and definitive genetic marker for identifying intra-group and inter-group characteristics of the *B. microti* and related parasites. Facilitating both the retrieval of new information and the validation (or refutation) of what has been reported before, this gene shows promise as the current gene of choice for bringing to light the growing gap between *B. microti* parasites and "related" but not *de facto B. microti* parasites. In the absence of official nomenclature, the literature is necessarily replete with, and confounded by, vernacular names (e.g., type, group, strain, "species" complex, *B. microti*-like, *B. microti*-related, assemblage, congregate, category, class). Today any study that addresses the genetic heterogeneity of the parasites, therefore, must inescapably address the nomenclature and taxonomy as well. To stabilize the taxonomy and promote the protection of public health from human babesiosis, reclassification of the "*B. microti* group" is an expedient and necessary step. In light of the present results, our proposal is that the parasite group be named the *B. microti* genus and that at least the *U.S., Kobe, Hobetsu,* and *Munich spp.* be recognized as four species known to constitute the genus at this time.

ACKNOWLEDGMENTS. This work was supported in part by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology, Japan; by Health Science Grants for Research on Emerging and Re-emerging Infectious Diseases from the Ministry of Health, Labor and Welfare of Japan; and by Gakujutsu Frontier Cooperative Research and High Technological Research Centers at Rakuno Gakuen University. We express our debt of gratitude to all who donated parasites used in the study and to all who painstakingly helped capture the host animals and isolate the parasites. With heartfelt appreciation, we also thank Dr. Nell L. Kennedy, Rakuno Gakuen University, for critical reading and help with the manuscript. In addition, we wish to thank Dr. Danny H.-Kauffmann Jokl, Department of Ophthalmology, Columbia University, New York, New York, and New York Medical College, Valhala, New York, U.S.A. for valuable discussions.

REFERENCES

1. Allsopp, M. T. and Allsopp, B. A. 2006. Molecular sequence evidence for the reclassification of some *Babesia* species. *Ann. New York Acad. Sci.* **1081**: 509–517.
2. Arai, S., Tsuji, M., Kim, S. J., Nakade, T., Kanno, Y. and Ishihara, C. 1998. *Babesia canis* infection in canine-red blood cell-substituted SCID mice. *Int. J. Parasitol.* **28**: 1429–1435.
3. Archibald, J. M., Logsdon, J. M., Jr. and Doolittle, W. F. 2000. Origin and evolution of eukaryotic chaperonins: phylogenetic

evidence for ancient duplications in CCT genes. *Mol. Biol. Evol.* **17**: 1456–1466.
4. Archibald, J. M., O'Kelly, C. J. and Doolittle, W. F. 2002. The chaperonin genes of jakobid and jakobid-like flagellates: implications for eukaryotic evolution. *Mol. Biol. Evol.* **19**: 422–431.
5. Birkenheuer, A. J., Harms, C. A., Neel, J., Marr, H. S., Tucker, M. D., Acton, A. E., Tuttle, A. D. and Stoskopf, M. K. 2007. The identification of a genetically unique piroplasma in North American river otters (*Lontra canadensis*). *Parasitology* **134**: 631–635.
6. Camacho, A. T., Pallas, E., Gestal, J. J., Guitian, F. J., Olmeda, A. S., Goethert, H. K. and Telford, S. R. 2001. Infection of dogs in north-west Spain with a *Babesia microti*-like agent. *Vet. Rec.* **149**: 552–555.
7. Gilson, P. R., Su, V., Slamovits, C. H., Reith, M. E., Keeling, P. J. and McFadden, G. I. 2006. Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 9566–9571.
8. Goethert, H. K. and Telford, S. R. III. 2003. What is *Babesia microti*? *Parasitology* **127**: 301–309.
9. Goethert, H. K., Cook, J. A., Lance, E. W. and Telford, S. R. III. 2006. Fay and Rausch 1969 revisited: *Babesia microti* in Alaskan small mammals. *J. Parasitol.* **92**: 826–831.
10. Gray, J., von Stedingk, L. V., Gurtelschmid, M. and Granstrom, M. 2002. Transmission studies of *Babesia microti* in *Ixodes ricinus* ticks and gerbils. *J. Clin. Microbiol.* **40**: 1259–1263.
11. Gray, J. S. 2006. Identity of the causal agents of human babesiosis in Europe. *Int. J. Med. Microbiol.* **296**: 131–136.
12. Hines, S. A., McElwain, T. F., Buening, G. M. and Palmer, G. H. 1989. Molecular characterization of *Babesia bovis* merozoite surface proteins bearing epitopes immunodominant in protected cattle. *Mol. Biochem. Parasitol.* **37**: 1–9.
13. Holman, P. J., Waldrup, K. A. and Wagner, G. G. 1988. *In vitro* cultivation of a Babesia isolated from a white-tailed deer (*Odocoileus virginianus*). *J. Parasitol.* **74**: 111–115.
14. Holman, P. J., Madeley, J., Craig, T. M., Allsopp, B. A., Allsopp, M. T., Petrini, K. R., Waghela, S. D. and Wagner, G. G. 2000. Antigenic, phenotypic and molecular characterization confirms *Babesia odocoilei* isolated from three cervids. *J. Wildl. Dis.* **36**: 518–530.
15. Kawabuchi, T., Tsuji, M., Sado, A., Matoba, Y., Asakawa, M. and Ishihara, C. 2005. *Babesia microti*-like parasites detected in feral raccoons (*Procyon lotor*) captured in Hokkaido, Japan. *J. Vet. Med. Sci.* **67**: 825–827.
16. Kim, S., Willison, K. R. and Horwich, A. L. 1994. Cytosolic chaperonin subunits have a conserved ATPase domain but diverged polypeptide-binding domains. *Trends Biochem. Sci.* **19**: 543–548.
17. Kjemtrup, A. M. and Conrad, P. A. 2000. Human babesiosis: an emerging tick-borne disease. *Int. J. Parasitol.* **30**: 1323–1337.
18. Krauss, V., Pecyna, M., Kurz, K. and Sass, H. 2005. Phylogenetic mapping of intron positions: a case study of translation initiation factor eIF2γ. *Mol. Biol. Evol.* **22**: 74–84.
19. Kriegs, J. O., Churakov, G., Kiefmann, M., Jordan, U., Brosius, J. and Schmitz, J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. *PLoS Biol.* **4**: e91.
20. Kubota, H. 2002. Function and regulation of cytosolic molecular chaperone CCT. *Vitam. Horm.* **65**: 313–331.
21. Kubota, H., Hynes, G. and Willison, K. 1995. The eighth *Cct* gene, *Cctq*, encoding the theta subunit of the cytosolic chaperonin containing TCP-1. *Gene* **154**: 231–236.
22. Kubota, H., Yokota, S., Yanagi, H. and Yura, T. 1999. Structures and co-regulated expression of the genes encoding mouse cytosolic chaperonin CCT subunits. *Eur. J. Biochem.* **262**: 492–500.
23. Kubota, H., Hynes, G., Carne, A., Ashworth, A. and Willison, K. 1994. Identification of six *Tcp-1*-related genes encoding divergent subunits of the TCP-1-containing chaperonin. *Curr. Biol.* **4**: 89–99.
24. Nikaido, M., Matsuno, F., Hamilton, H., Brownell, R. L., Jr., Cao, Y., Ding, W., Zuoyan, Z., Shedlock, A. M., Fordyce, R. E., Hasegawa, M. and Okada, N. 2001. Retroposon analysis of major cetacean lineages: the monophyly of toothed whales and the paraphyly of river dolphins. *Proc. Natl. Acad. Sci. U.S.A.* **98**: 7384–7389.
25. Nishisaka, M., Yokoyama, N., Xuan, X., Inoue, N., Nagasawa, H., Fujisaki, K., Mikami, T. and Igarashi, I. 2001. Characterisation of the gene encoding a protective antigen from *Babesia microti* identified it as η subunit of chaperonin containing T-complex protein 1. *Int. J. Parasitol.* **31**: 1673–1679.
26. Pieniazek, N., Sawczuk, M. and Skotarczak, B. 2006. Molecular identification of *Babesia* parasites isolated from *Ixodes ricinus* ticks collected in northwestern Poland. *J. Parasitol.* **92**: 32–35.
27. Reichard, M. V., Van Den Bussche, R. A., Meinkoth, J. H., Hoover, J. P. and Kocan, A. A. 2005. A new species of *Cytauxzoon* from Pallas' cats caught in Mongolia and comments on the systematics and taxonomy of piroplasmids. *J. Parasitol.* **91**: 420–426.
28. Rommelaere, H., Van Troys, M., Gao, Y., Melki, R., Cowan, N. J., Vandekerckhove, J. and Ampe, C. 1993. Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven related subunits. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 11975–11979.
29. Russell, C. B., Fraga, D. and Hinrichsen, R. D. 1994. Extremely short 20–33 nucleotide introns are the standard length in *Paramecium tetraurelia*. *Nucleic Acids Res.* **22**: 1221–1225.
30. Saito-Ito, A., Tsuji, M., Wei, Q., He, S., Matsui, T., Kohsaki, M., Arai, S., Kamiyama, T., Hioki, K. and Ishihara, C. 2000. Transfusion-acquired, autochthonous human babesiosis in Japan: isolation of *Babesia microti*-like parasites with hu-RBC-SCID mice. *J. Clin. Microbiol.* **38**: 4511–4516.
31. Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
32. Sinski, E., Bajer, A., Welc, R., Pawelczyk, A., Ogrzewalska, M. and Behnke, J. M. 2006. *Babesia microti*: prevalence in wild rodents and *Ixodes ricinus* ticks from the Mazury Lakes District of North-Eastern Poland. *Int. J. Med. Microbiol.* **296**: 137–143.
33. Takahashi, K. 1976. Studies on the Infection and Immunity of *Theileria sergenti*. *J. Coll. Dairying* **6**: 179–248 (in Japanese).
34. Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
35. Thornhill, D. J., Lajeunesse, T. C. and Santos, S. R. 2007. Measuring rDNA diversity in eukaryotic microbial systems: how intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates. *Mol. Ecol.* **16**: 5326–5340.
36. Tsuji, M., Terada, Y., Arai, S., Okada, H. and Ishihara, C.

1995. Use of the Bo-RBC-SCID mouse model for isolation of a *Babesia* parasite from grazing calves in Japan. *Exp. Parasitol.* **81**: 512–518.
37. Tsuji, M., Zamoto, A., Kawabuchi, T., Kataoka, T., Nakajima, R., Asakawa, M. and Ishihara, C. 2006. *Babesia microti*-like parasites detected in Eurasian red squirrels (*Sciurus vulgaris orientis*) in Hokkaido, Japan. *J. Vet. Med. Sci.* **68**: 643–646.
38. Tsuji, M., Wei, Q., Zamoto, A., Morita, C., Arai, S., Shiota, T., Fujimagari, M., Itagaki, A., Fujita, H. and Ishihara, C. 2001. Human babesiosis in Japan: epizootiologic survey of rodent reservoir and isolation of new type of *Babesia microti*-like parasite. *J. Clin. Microbiol.* **39**: 4316–4322.
39. Wei, Q., Tsuji, M., Zamoto, A., Kohsaki, M., Matsui, T., Shiota, T., Telford, S. R. III and Ishihara, C. 2001. Human babesiosis in Japan: isolation of *Babesia microti*-like parasites from an asymptomatic transfusion donor and from a rodent from an area where babesiosis is endemic. *J. Clin. Microbiol.* **39**: 2178–2183.
40. Yamasaki, M., Inokuma, H., Sugimoto, C., Shaw, S. E., Aktas, M., Yabsley, M. J., Yamato, O. and Maede, Y. 2007. Comparison and phylogenetic analysis of the heat shock protein 70 gene of *Babesia* parasites from dogs. *Vet. Parasitol.* **145**: 217–227.
41. Zahler, M., Rinder, H. and Gothe, R. 2000. Genotypic status of *Babesia microti* within the piroplasms. *Parasitol. Res.* **86**: 642–646.
42. Zahler, M., Rinder, H., Schein, E. and Gothe, R. 2000. Detection of a new pathogenic *Babesia microti*-like species in dogs. *Vet. Parasitol.* **89**: 241–248.
43. Zamoto, A., Tsuji, M., Kawabuchi, T., Wei, Q., Asakawa, M. and Ishihara, C. 2004. U.S.-type *Babesia microti* isolated from small wild mammals in Eastern Hokkaido, Japan. *J. Vet. Med. Sci.* **66**: 919–926.
44. Zamoto, A., Tsuji, M., Wei, Q., Cho, S. H., Shin, E. H., Kim, T. S., Leonova, G. N., Hagiwara, K., Asakawa, M., Kariwa, H., Takashima, I. and Ishihara, C. 2004. Epizootiologic survey for *Babesia microti* among small wild mammals in northeastern Eurasia and a geographic diversity in the β-tubulin gene sequences. *J. Vet. Med. Sci.* **66**: 785–792.