Jpn. J. Infect. Dis., 64, 403-410, 2011

Original Article

Intron Sequences from the *CCT7* Gene Exhibit Diverse Evolutionary Histories among the Four Lineages within the *Babesia microti*-Group, a Genetically Related Species Complex That Includes Human Pathogens

Kohei Fujisawa[1], Rui Nakajima[1,5], Michio Jinnai[1,3], Haruyuki Hirata[1], Aya Zamoto-Niikura[1,2], Takako Kawabuchi-Kurata[1,3], Satoru Arai[4], and Chiaki Ishihara[1]*

[1]*Laboratory Animal Science, School of Veterinary Medicine, Rakuno Gakuen University, Ebetsu 069-8501;*
[2]*Division of Experimental Animal Research, National Institute of Infectious Diseases, Tokyo 208-0011;*
[3]*Osaka Prefectural Institute of Public Health, Osaka 537-0023;*
[4]*Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Tokyo 162-8640, Japan; and*
[5]*Division of Infectious Diseases, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, Republic of China*

(Received May 2, 2011. Accepted July 12, 2011)

SUMMARY: *Babesia microti*, the primary causal agent of human babesiosis in North America, was thought to distribute in Europe in association with ixodid ticks and rodents. Recent analyses of *β-tubulin* and the eta subunit of the chaperonin-containing t-complex protein 1 (*CCT7*) genes revealed discrete clusters (a species-complex comprised of at least 4 taxa for the U.S., Kobe, Munich, and Hobetsu). To further assess the micro-evolutionary history and genetic variability within the taxon, we combined a set of 6 introns from the *CCT7* gene to use as a rapidly evolving DNA marker. Phylogenetic and comparative sequence analyses subdivided the U.S. taxon into 3 geographic subclades—North America, western to central Eurasia, and northeastern Eurasia ($\geq$98% bootstrap supports for each node). The Kobe taxon, which occurs only in a few geographic foci of Japan, could further be subdivided into 2 subgroups (100% support). The Munich and Hobetsu taxa, common to Europe and Japan, respectively, exhibited little or no pairwise sequence divergence among geographically diverse samples, suggesting an extreme population bottleneck during recent history. Despite the small sample size, this study provides a better understanding of the micro-evolutionary relationships and the genetic variability present within each lineage of the *B. microti*-group.

INTRODUCTION

Babesia microti, an erythroparasitic protozoon that maintains a life cycle between the tick *Ixodes dammini* (*I. scapularis*) and the white-footed mouse (*Peromyscus leucopus*), is known to be the primary agent of human babesiosis in North America (1–3). It was once thought that a single species of *B. microti* was the etiologic agent of human babesiosis and was distributed widely throughout temperate zones in the Northern hemisphere. In Asia and Europe, morphologically indistinguishable parasites with similar or identical small subunit rRNA gene sequences (rDNA) also have been identified in association with ixodid ticks, various small wild mammals (Murinae, Arvicolinae, and occasionally Soricidae), and in incidents of human infestation (4–15).

The study of rapidly evolving genes in proteins is used to identify the internal topology of the *Babesia microti*-group. By using strains isolated largely from the North American Continent, Goethert and Telford (6) inferred a *β-tubulin* gene tree with multiple subclusters. Further, Zamoto et al. (14,15) and Tsuji et al. (16) documented substantial genetic diversity among the *β-tubulin* genes, which themselves further subdivided tested samples into 3 separate clades—the U.S., Kobe, and Hobetsu—each having higher node supports (from 90% to 100%). The phylogenetic status recently became more evident following analysis on the 18 samples that comprised 3 zoonotic taxa and a closely related non-zoonotic Munich taxon, which was based on the eta subunit of the chaperonin-containing t-complex protein 1 (*CCT7*, formerly described as *CCT eta*) gene (17). Consequently, the *B. microti*-group is now recognized as a genetically diverse species-complex that consists of at least 4 distinct phylogenetic lineages, i.e., the U.S. taxon broadly distributed in North America, Europe, and Asia (1,3,6,7,14,15,18); the Kobe taxon focally located in a few spotted areas of Japan (Mikura Island in the Izu-Bonin Trench, Awaji Island near Osaka, and Daito-Town in Shimane Prefecture) (9,13,19), the Hobetsu taxon common in Japan (11,14,15,17), and the non-zoonotic Munich taxon enzootic in Europe (20,21).

*Corresponding author: Mailing address: Laboratory Animal Science, School of Veterinary Medicine, Rakuno Gakuen University, 582-1 Bunkyodai Midorimachi, Ebetsu 069-8501, Japan. Tel & Fax: +81-11-388-4774, E-mail: ishihara@rakuno.ac.jp

An understanding of intra-taxon genetic variability and phylogenetic relationships is important since these factors bear on the evolutionary process that explains both geographic spread and the biological mechanisms essential to maintaining host(s)-vector(s) cycle(s). This recognition is necessary to improve lineage-specific risk assessments, to create practical screening tools, and to develop appropriate control strategies. Here, we introduce *CCT7* introns as promising new phylogenetic markers. The *CCT7* nucleotide lengths at relative locations recently have been proven highly conserved by lineage (17). We measured the rapidly evolving DNA to determine its intra-taxon level.

MATERIALS AND METHODS

Parasite strains: The parasite strains investigated in this study are listed in Table 1. The individual collections have been classified previously on the basis of the *small subunit rRNA* gene (rDNA) sequences (9,11,13), *β-tubulin* (14,15) gene sequences, and eta subunit of *CCT7* gene sequences (17), which were placed into 4 separate groups: U.S. (9 strains), Kobe (3 strains), Munich (5 strains), and Hobetsu (7 strains) taxa (Table 1). In this study, 2 Kobe (Hatahiyodori and Mikura) and 4 Hobetsu strains (Tsukiyo, Takanosu, Akkeshi, and Awaji) were propagated in Syrian hamsters or NOD/shi-scid mice, from which erythrocytes were substituted with rat erythrocytes (Mikura strain) as described previously (11). Animal experimentation was conducted according to the Laboratory Animal Control Guidelines of Rakuno-Gakuen University. The strains in this study originated from diverse geographical areas of Japan and were used to elucidate taxon-specific evolutionary history, geographic spread, and genetic structure. The extracted DNA from parasite-containing erythrocytes was subjected to polymerase chain reaction (PCR) amplification and sequencing.

Sequencing of *CCT7* introns: The intron sequences from the 18 strains of the *B. microti*-group and 2 additional outgroup control sequences from raccoon and *B. rodhaini* (Table 1) were obtained from GenBank. To determine the intron sequences of 6 other strains, PCR was performed on genomic DNA by targeting the regions that contained full length genes. Genomic DNA was extracted from infected erythrocytes using a whole-blood DNA extraction kit (DNA Extractor WB Kit; Wako Pure Chemical Industries, Osaka, Japan) according to the manufacturer's instructions. The primer combinations used for the Hatahiyodori and Mikura strains in the Kobe taxon and the Tsukiyo, Takanosu, Akkeshi, and Awaji strains in the Hobetsu taxon were UScct-1 (5′-GGTGTGCCAGGCAAATGCCATTTTAAG-3′)/KoCCT3′R1 (5′-AATCCTGCATTTCTCATTGTTAAG-3′), and BmCCT5′A (5′-AACCCAATATATCTGGCACRTA-3′)/HoCCT3′R1 (5′-ACCCGTGGCATTTTAGAGTTATCG-3′), respectively (Table 1). The PCR mixtures were prepared using a high-fidelity polymerase (Takara La Taq polymerase, Takara Bio Inc., Otsu, Japan) according to the manufacture's instructions and then denatured for 2 min at 94°C. Thirty thermal cycles were applied; each cycle contained 30 s at 94°C, 40 s at 50°C, and 90 s at 72°C, all using a PC320-program temperature control system (ASTEC Co., Fukuoka, Japan). The PCR products were purified from a 1% agarose gel with the MinElute Gel Extraction Kit (QIAGEN, Tokyo, Japan). The nucleotide sequences were determined next by direct sequencing of the PCR fragment using the CEQ8000 automated sequencer (Beckman Coulter, Brea, Calif., USA) with the DTCS DNA Sequence kit (Beckman Coulter). The exon-intron structures were verified by aligning the complementary DNA (cDNA) sequences from each parasite strain in the Kobe and Hobetsu taxa (the Kobe524 and Ho234, accession nos. AB366755 and AB366750, respectively) with the genomic DNA sequences from the 6 strains.

Alignment of nucleotide sequences: We used the set of 6 conserved introns in the *CCT7* gene (intron locations and sizes are shown in Table 1), which presented in single copy in the *B. microti*-group parasite genomes (22). The individual intron at each location (1st, 2nd, 4th, 6th, 10th, and 12th) was aligned with sequences from the same-location counterpart within the lineage. Alignment was performed with the Clustal W Alignment program (23) in MacVector 10.5 and adjusted manually according to a method reported previously (17). For both multiple and pairwise alignments, only the introns in the exact same codon phase (position of the intron within a codon) were considered as identical in location. The nucleotide sequences of all 6 introns were aligned and compared for similarity with the sequence from other strains in the same lineage.

Comparison of evolution rates: The evolution rates of *CCT7* introns were compared to the short term rates (rapidity of sequence change) of other DNA markers used for inferring genetic variability and evolutionary relationships among strains of the *B. microti*-group. The evolutionary rates of the nucleotide sequences were derived from present introns and from the rates derived from *CCT7* coding sequences (CDS) for the Gray and Xinjiang samples of the U.S. taxon, and also from the rates in *18S rDNA* sequences available exclusively from public databases. The database profiles were obtained from DDBJ [Gray, accession no. AY693840; Xinjiang1637, accession no. AB083375].

Phylogenetic analysis: Phylogenetic trees were constructed by the neighbor-joining method (24), and the support for resulting tree nodes was calculated with 1,000 bootstrap replicates by the bootstrap tree algorithm. A bootstrap value of $\geq 95\%$ was considered statistically significant and values $\geq 99\%$ were considered highly significant.

RESULTS

New sequences: Six new sequences, 2 from the Hatahiyodori and Mikura strains in the Kobe taxon and 4 from the Tsukiyo, Takanosu, Akkeshi, and Awaji strains in the Hobetsu taxon, were determined in this study and submitted to GenBank (accession number shown in Table 1). Three sequences in the Hobetsu group (Tsukiyo, Takanosu, and Akkeshi) were identical to each other and to the Ho234 sequence reported previously (17) (accession number shown in Table 1).

Sequence alignment and genetic analysis: The *CCT7* genes from the 24 samples distributed among the 4 separate taxa each contained all 6 introns (Table 1). All 6 introns had standard 5′-GT...AG-3′ boundaries and

Table 1. *CCT7* genes of *B. microti*-group parasite strains used

Taxon	Strain[1]	Origin	Geographical origin	Gene size[2] (CDS)[3]	Accession no.	Locations and sizes of *CCT7* introns											
						1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
U.S.	Gray	human	U.S.A.	2009 (1629)	AB362586	21[5]	22	—[6]	20	—	49	—	—	—	247	—	21
	GI	human	U.S.A.	2009 (1629)	AB362581	21	22	—	20	—	49	—	—	—	247	—	21
	NM69	vole	Hokkaido, Japan	2010 (1629)	AB362582	21	22	—	20	—	49	—	—	—	248	—	21
	AK2273	mouse	Hokkaido, Japan	2010 (1629)	AB362587	21	22	—	20	—	49	—	—	—	248	—	21
	Vladivostok38	vole	Vladivostok, Russia	2010 (1629)	AB362584	21	22	—	20	—	49	—	—	—	248	—	21
	Irkutsk16	vole	Irkutsk, Russia	2010 (1629)	AB366748	21	22	—	20	—	49	—	—	—	248	—	21
	Korea8	mouse	Korea	2010 (1629)	AB362583	21	22	—	20	—	49	—	—	—	248	—	21
	Xinjiang1637	vole	Xinjiang, China	2009 (1629)	AB362585	21	22	—	20	—	49	—	—	—	247	—	21
	HK	vole	Hannover, Germany	2008 (1629)	AB366747	21	22	—	20	—	48	—	—	—	247	—	21
Kobe	Kobe524	human	Kobe, Japan	1888 (1629)	AB366755	20	22	—	20	—	45	—	—	—	132	—	20
	Hatahiyodori*	mouse	Shimane, Japan	1888 (1629)	AB583230	20	22	—	20	—	45	—	—	—	132	—	20
	Mikura*	rat	Mikura Island, Japan	1916 (1629)	AB583231	20	22	—	20	—	72[7]	—	—	—	133	—	20
Munich	Munich	mouse	Germany	2003 (1629)	AB362588	21	22	—	20	—	49	—	—	—	241	—	21
	Poland5	tick	Poland	2003 (1629)	identical[4]	21	22	—	20	—	49	—	—	—	241	—	21
	Poland6	tick	Poland	2003 (1629)	identical[4]	21	22	—	20	—	49	—	—	—	241	—	21
	UK5	vole	England	2003 (1629)	identical[4]	21	22	—	20	—	49	—	—	—	241	—	21
	UK6	vole	England	2003 (1629)	identical[4]	21	22	—	20	—	49	—	—	—	241	—	21
Hobetsu	Ho234	mouse	Hokkaido, Japan	1813 (1629)	AB366750	21	22	—	20	—	46	—	—	—	54	—	21
	Tsukiyo*	mouse	Tokushima, Japan	1813 (1629)	AB583229	21	22	—	20	—	46	—	—	—	54	—	21
	Takanosu*	mouse	Akita, Japan	1813 (1629)	AB583228	21	22	—	20	—	46	—	—	—	54	—	21
	Akkeshi*	vole	Hokkaido, Japan	1813 (1629)	AB583227	21	22	—	20	—	46	—	—	—	54	—	21
	Ot1	mouse	Chiba, Japan	1813 (1629)	AB366754	21	22	—	20	—	46	—	—	—	54	—	21
	Awaji*	mouse	Hyogo, Japan	1813 (1629)	AB583226	21	22	—	20	—	46	—	—	—	54	—	21
	Da116	mouse	Shimane, Japan	1813 (1629)	AB366753	21	22	—	20	—	46	—	—	—	54	—	21
Parasties related distantly to *B. microti*-group																	
Parasite of raccoon		raccoon	Hokkaido, Japan	2199 (1629)	AB366751	21	22	—	20	—	254	—	—	—	231	—	23
B. rodhaini		vole	Africa ?	1752 (1629)	AB366752	21	—	—	21	—	19	—	—	—	43	—	20

[1]: Asterisk indicates sequencses determined newly in this study. Other sequences were reported previously (17).
[2]: Nucleotide length (nt) from 5′ end to 3′ end of *CCT7* gene including coding sequences and introns.
[3]: Nucleotide length (nt) of the complete coding sequence (CDS) of *CCT7* gene.
[4]: The nucleotide sequence was 100% identical to that of Munich strain.
[5]: Number indicates intron size (nt).
[6]: — indicates luck of intron at the location.
[7]: A single 27-nt insertion (CGCCT ATATA TATAT ATATA TGTTA TA) was found after 10th position of this intron.

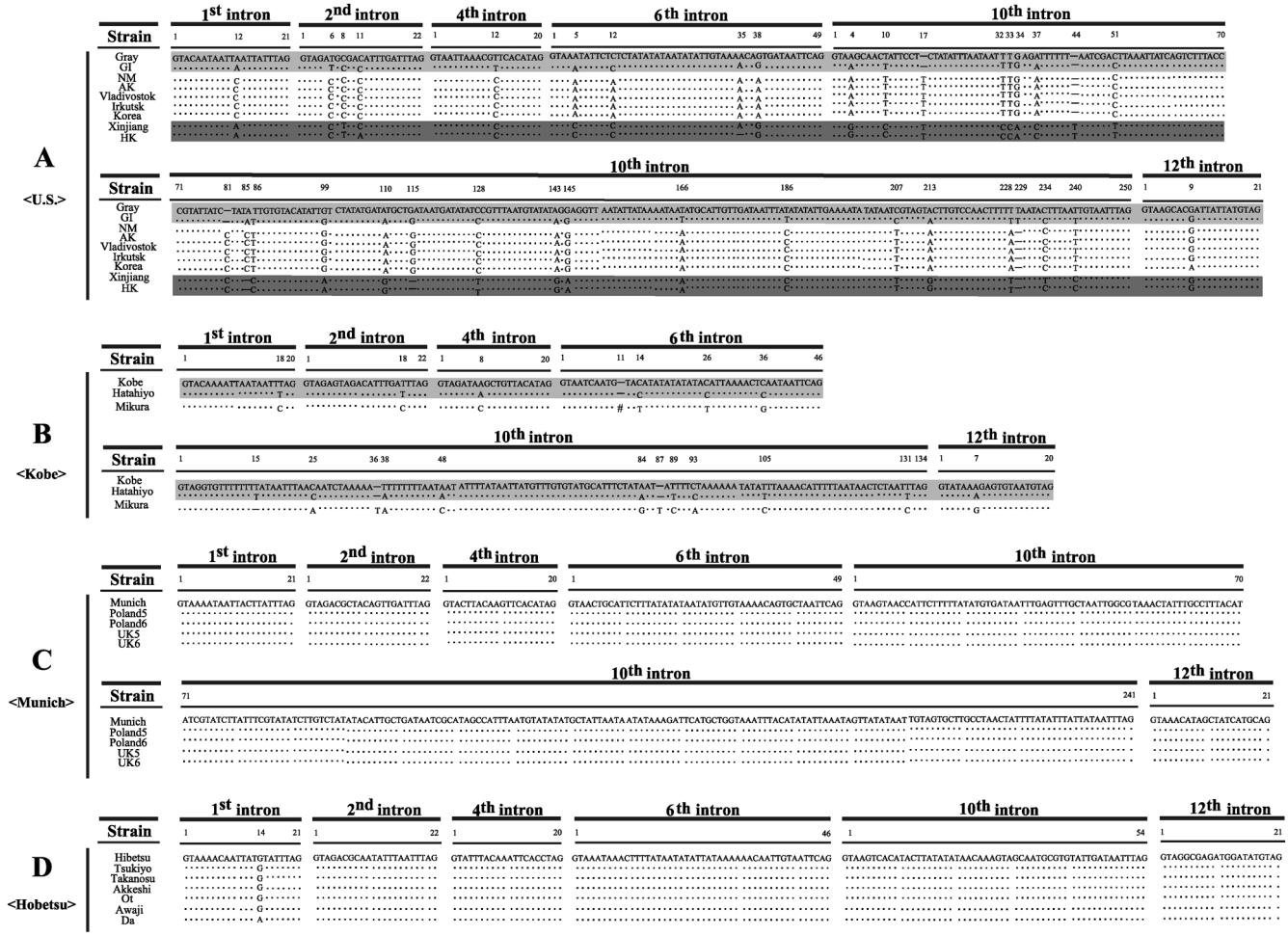

Fig. 1. Alignment of nucleotide sequences of *CCT7* introns from strains of the *B. microti*-group. The individual intron at each location (1st, 2nd, 4th, 6th, 10th, and 12th) was aligned by sequences with its same-location counterparts within lineage. Dots indicate identity with the consensus sequence. Line of letters with gray (North America), dark-gray (western to central Eurasia), or white background (northeastern Eurasia) highlights subgroups that appeared to be further divided within a lineage. #), A 27-nt insertion (CGCCT ATATA TATAT ATATA TGTTA TA) occurs after 10th position of the 6th intron in Mikura strain in the Kobe taxon.

the CDS aligned confidently and exhibited no size heterogeneity at the exon-intron junctions. The introns ranged in size from 19 to 254 nucleotides (nt) depending on location and strain (Table 1). Introns shorter than 25 nt (1st, 2nd, 4th, and 12th locations) retained their size consistently, regardless of taxon, with exception of the 1st and 12th introns of the Kobe taxon (Table 1). The 2 largest introns (6th and 10th locations) varied in size, but at the intra-taxon level, the relative length was fairly well conserved, i.e., size divergence did not exceed 1 base, except for a single 27-nt insertion (CGCCT ATATA TATAT ATATA TGTTA TA) that occurred after 10th position of the 6th intron in the Mikura strain of the Kobe taxon (Table 1 and Fig. 1B).

Pairwise sequence comparisons were performed independently for each lineage of the *B. microti*-group. For the U.S. lineage, the intra-taxon sequence divergence values ranged from none (identical) to around 8%; the latter applied to the Gray and Xinjiang strains (7.6%; 23 point-mutations [substitutions] and 6 insertion/deletions [indels] for 383 nt) or to Gray and HK strain; (7.8%; 23 substitutions and 7 indels for 383 nt) (Fig. 1). Sequence alignment separated these 9 strains into 3 subgroups, (i) North America (GI and Gray), (ii) western to central Eurasia (Xinjiang1637 and HK), and (iii) northeastern Eurasia, including Japan (Vladivostok38, Irkutsk16, Korea8, NM69, and AK2273), by setting nucleotide changes at 36 positions as the sum of substitutions and indels (Fig. 1A). Four other nucleotide changes were determined based on intra-subgroup sequence differences that occurred either within the western to central Eurasia subgroup (2 substitutions at the 11th position of the 2nd intron and the 234th position of the 10th intron, respectively, and 1 indel change after the 34th position of the 6th intron) or in the northeastern Eurasia subgroup (1 substitution at the 9th position of the 12th intron of the Korea strain) (Fig. 1A). Among the 3 Kobe samples, 2 strains (Kobe and Hatahiyodori) exhibited very similar sequences with only a single substitution difference at the 38th position of the 10th intron (Fig. 1B). However, the remaining strain isolated from a wild rat in Mikura Island, the southern-most island of Izu Seven Islands about 200 km south of Tokyo, showed a 7.3% sequence differences (19/262; a 27-nt concatenated insertion sequence found in the 6th intron of the Mikura strain was treated expediently as 1 incident; Fig. 1B and Table 2). For the Munich taxon, all the 5 samples from England, Germa-

Table 2. Frequency and pattern of mutations

Taxon	Type of mutation	Mutation frequency for each intron location					
		1st	2nd	4th	6th	10th	12th
U.S.	Single-base substitution	1(4.8)[1)]	3(13.6)	1(5.0)	3(6.1)	20(8.0)	1(4.8)
	Insertion/Deletion[2)]	0	0	0	1(2.0)	6(2.4)	0
Kobe	Single-base substitution	1(5.0)	1(4.5)	1(5.0)	3(6.5)	8(6.0)	1(5.0)
	Insertion/Deletion[2)]	0	0	0	1(2.2)[3)]	3(2.2)	0
Munich	Single-base substitution	0	0	0	0	0	0
	Insertion/Deletion	0	0	0	0	0	0
Hobetsu	Single-base substitution	1(4.8)	0	0	0	0	0
	Insertion/Deletion	0	0	0	0	0	0

[1)]: Mutation frequency. Numbers in parenthesis indicate the normalized frequency per 100 bases as per cent (%).
[2)]: All of 11 insertion/deletion mutations are identified as single nucleotide indels except one case found in the 6th intron of the Mikura strain.
[3)]: To determine the normalized frequency of mutation, a single 27-nt mutation found in the 6th intron of the Mikura strain is counted as a single occurrence at the location.

ny, and Poland (Table 1) shared identical sequences (Fig. 1C). The Hobetsu sample sequences were identical, except for the Da strain, which exhibited a single base substitution (G→A) at the 14th position of the first intron (Fig. 1D).

The evolutionary rate of the *CCT7* introns was determined and compared to the rates of other gene sequences available exclusively from public databases. The total evolutionary rate of 7.6% of the intron sequence diversity (29 nt changes, 23 substitutions, and 6 indels from the 383 nt total length) demonstrated here for the Gray and Xinjiang samples in the U.S. taxon was 63 times higher than the rate determined for the *18S rDNA* gene sequences (0.12%; 2 mutations in 1665 nt) and 2.6 times higher than the rate for the *CCT7* CDS (2.9%; 47 alterations in 1629 nt).

Lineage-specific patterns of sequence variation: Substitutions and indels were observed predominantly in the U.S. and Kobe taxa, but rarely in the Munich and Hobetsu taxa (Fig. 1 and Table 2). From the 383 nt total of the combined sequences in the U.S. taxon, 29 single-nt substitutions and 7 indels (of which 26 substitutions and 6 indels were identified as inter-subgroup differences) and the remaining 3 substitutions (at the 11th, 234th, and 9th positions in the 2nd, 10th, and 12th introns, respectively) and 1 indel change (at the 35th position in the 6th intron), were identified as intra-subgroup variations (Fig. 1A and Table 2). In the Kobe taxon, 15 substitutions and 4 length variations were detected from a total of 262 nt of the combined sequences (Fig. 1B and Table 2). The 2 strains, Kobe and Hatahiyodori, in the Kobe taxon were almost identical in sequence with only a single base difference at the 38th position of the 10th intron, but the Mikura strain differed notably from the other 2 strains because it showed 18 or 19 nt differences (14 or 15 substitutions and 4 indels; Fig. 1B and Table 2). In the 5 strains of the Munich group, originating from diverse geographic areas of Europe, all the sequences grouped together with 100% identity (Fig. 1C and Table 2). Likewise in the Hobetsu taxon, the sequences from the 7 samples isolated from various small mammals (including rodents and shrews) across widely dispersed geographic regions of Japan were so crowded together they appeared as 1 common pattern, with the exception of the Da strain from Daito Town, Shimane Prefecture (southwestern portion of the Japanese mainland) which showed a single-base substitution occurring in the 1st intron (Fig. 1D and Table 2). The overall transition and transversion substitutions obtained for each pair of sequences in the present data-set were 32 (71%) and 13 (29%), respectively (Fig. 1).

Upon comparison of normalized frequencies (%) per 100 bases in the U.S. and Kobe taxa, point mutations occurred with similar frequency in every intron (4.5% to 6.5%), although a somewhat elevated frequency was observed in 2 introns from each type, an extremely short (13.6% for the 2nd intron) and a normal (8.0% for the 10th intron) length intron (Table 2, in the U.S. taxon). Contrary to point-mutations, the indel-mutations occurred exclusively in the normal size introns at the 6th and 10th location (Fig. 1 and Table 2). In the U.S. taxon, the 6th intron exhibited 1 indel mutation and the 10th intron had 6; the Kobe group exhibited 1 in the 6th intron and 3 in the 10th intron (Fig. 1 and Table 2). All but 1 of the indel mutations were limited to a 1-nt gap, but in the 6th intron of the Mikura strain (Kobe taxon), from which an AT repeat was present; the mutation involved a unique insertion measuring 27 nt in length (CGCCT ATATA TATAT ATATA TGTTA TA).

Phylogenetic analyses: From the phylogenetic tree constructed on the basis of the combined sequences of the 6 *CCT7* introns (Fig. 2), the 9 strains of the U.S. taxon were separated clearly into 3 independent geographic clusters (or subclades), i.e., (i) North America (GI and Gray), (ii) western to central Eurasia (Xinjiang1637 and HK), and (iii) northeastern Eurasia, including Japan (Vladivostok38, Irkutsk16, Korea8, NM69, and AK2273). The nodes of the individual subclades were supported by high bootstrap values (98% to 100%) as shown in Fig. 2A. Despite the very small sample (since parasites of this lineage are rare and difficult to be detected), data analysis revealed that the Kobe taxon splits into 2 subclades with high bootstrap supports (100%) when a 27-nt insertion is included (data not shown) or deleted (100%, Fig. 2B). In contrast, the 5 Munich strains originated from multiple geographically

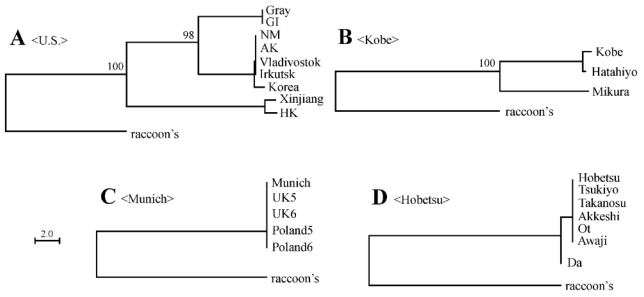

Fig. 2. Phylogenetic trees based on the *CCT7* intron sequences showing diverse evolutionary histories of the four individual lineages of the *B. microti* group. Trees are constructed based on the combined sequence data of all the 6 *CCT7* introns from strains within lineage. Raccoon's sample was used as an out-group. In all trees, the numerals at the internodes are bootstrap values based on 1,000 replicates. Scale bar represents 2.0% of the genetic distance. A 27-nt concatenated insertion sequence found in the 6th intron of the Mikura strain was treated expediently as one incidence. (When *B. rodhaini* was used as an out-group, despite it missed 2nd intron, all the phylogenetic trees obtained were almost identical in topology and very similar in support. Furthermore, trees generated from both neighbor-joining and maximum likelihood quartet-puzzling analyses showed similar topologies with high nordal supports and hence only the neighbor-joining tree is shown here.)

widespread regions form 1 phylogenetic cluster of the tree (Fig. 2C). The sequences from the 7 samples of the Hobetsu taxon were closely related to one another in phylogeny, and together they formed a single clade with strong support (Fig. 2D).

DISCUSSION

The results of this study support our hypothesis that intronic nuclear-DNA loci in the *CCT7* gene qualify as DNA markers for analysis of of the 4 taxa in the *B. microti*-group by study of micro-evolutionary relationships and genetic variability within each taxon. Here, we report on the diverse evolutionary histories and genetic structures of the 4 lineages; the U.S. and Kobe taxa, which distribute either globally across Eurasia and North America or focally in a few areas of Japan. These taxa exhibit substantial intra-taxon sequence variation among samples and can be subdivided into 3 (U.S. taxon) or 2 (Kobe taxon) phylogenetic subclusters, respectively, with high statistical supports ($\geq$98%). The Munich and Hobetsu lineages (common to Europe and Japan, respectively) showed very little to no pairwise sequence divergence ($\leq$0.5%) among geographically diverse samples and formed a single clade, which is suggestive of extreme genetic bottlenecks in the recent past. In comparison, the *CCT7* introns (analyzed with existing DNA markers) exhibited relative sequence divergence between the Gray and Xinjian strains of the U.S. taxa—63 and 2.6 times higher than *18S rDNA* and *CCT7* CDS, respectively. By recruiting more sequences from additional samples and possibly from other genes, intron sequences in future should allow more accurate inferences regarding the population history of the *B. microti*-group parasites which may affect the dynamics and importance of such factors as lineage spreading, selection pressure due to bottlenecking, founder effects, and maintenance of the host-vector-cycle.

Our previous study with the single-copy gene *CCT7* showed that the intron location and surrounding exon boundaries were fully conserved among members of the *B. microti*-group—even among evolutionary distant species such as raccoons and *B. rodohaini* (17) (Table 1). In addition to providing location stability, *CCT7* introns were shown to exhibit size uniformity within each taxon at the respective intron locations (17) (Table 1). However, the *B. microti*-group is unusual because a size difference was observed between 2 different types of introns, i.e., the normal-size and the 19–23 nt (smallest known) introns (17)—introns of less than 25 nt have never been recorded in any eukaryotes except for 3 other protists: *Paramecium tetraurelia* (20–33 nt) (25), *Bigelowiella natans* (18–21 nt) (26), and *Nyctotherus ovalis* (21–29 nt) (27). Mutual sequence alignment of all 6 introns at a specific location would thereby be able to occur in each taxon, and aligned sequences may prove more useful for detecting micro-evolutionary changes than *CCT7* CDS, which is the most sensitive phylogenetic marker available at present (17). To test this hypothesis, we performed a preliminary trial to test whether the 6 introns in the *CCT7* gene shared common evolutionary features among the 4 *B. micros*-group lineages despite intron size differences. The 4 smallest known introns (the 1st, 2nd, 4th, and 12th) were able to be aligned independently with good compatibility across the entire data-set and the 4 clusters determined by sequence similarity and also by the phylogenetic trees (data not shown), which were congruent with the 4 major taxa reported previously from the *CCT7* CDS (17). This finding indicated that the introns and CDS from the *CCT7* gene had undergone similar evolutionary processes during the evolution of these 4 phylogenetic lineages. The 2 largest (6th and 10th) introns also exhibited high-fidelity alignment within each lineage, and the phylogenetics based on the intronic sequences were generally consistent with the inference drawn from the *CCT7* CDS phylogeny (data not shown) shown in our previous report (17). Regardless of size, these introns exhibited similar levels of point mutation frequency ranging from 4.5 to 6.5 changes per 100 bases (Table 2). No disparity was observed in the rate of mutation acceleration between the 2 types, i.e., the 2 accelerated cases from the U.S. taxon were observed in the smallest known intron (13.6%, 3 changes from a 22 nt total length for the 2nd intron) and in the longest intron (8.0%, 20 changes from a total length for the 10th intron) (Table 2). Consequently, extremely short and normal *CCT7* introns of the *B. microti*-group are presumed to have shared evolutionary history along with *CCT7* CDS. Thus, we used a combination of all 6 introns in the present study because the *CCT7* introns appeared to be too short to resolve individually the evolutionary questions for the 4 individual lineages of the *B. microti* group.

Combined intron sequence data analysis was used to further divided the 2 known taxa, U.S. and Kobe, into 3 and 2 separate clusters, respectively. This method was based on the sequence similarities (Fig. 1) and phylogenies (Fig. 2). Nucleotide changes occurred mostly ($\geq$90%) at the intra-taxon subgroup level and less frequently among individuals within a subgroup (Fig. 1), indicating that the extent of *CCT7* intron variability is

largely proportional to the micro-evolutionary distance and not to allele or genotype frequencies in a population. In contrast, the Munich and Hobetsu taxa, which exhibited very slight to no sequence variation among geographically diverse samples (Fig. 1 and Table 2), distributed independently in Europe (England, Germany, Poland, and Russia) (17,20,21) or in Japan (11,14,15, 19). The very low genetic diversity within Munich and Hobetsu taxa (Figs. 1, 2 and Table 2) may be due to the natural selection tendency toward fragmentation and clonal proliferation, i.e., the founder effect or a bottleneck event. The regional distribution of the Munich (Europe) and Hobetsu taxa (Japan) and their protracted evolutionary duration (represented by branch length of the CDS phylogeny) (17), indicate a bottleneck event as the most plausible scenario for the clonal population structure of the 2 extant taxa. The bottleneck catastrophe would have had to occur directly within the parasites themselves or indirectly as host and vector. Indeed, the Munich and Hobetsu lineages may have faced a major risk of extinction if clonal recovery in recent evolutionary history did not take place.

In contrast to 2 of these regional taxa, the Kobe taxa have exhibited focal distribution, e.g., in wild mouse and rat reservoirs, and in just a few narrowly defined areas of Japan—the Mikura Island in the Izu-Ogasawara (Izu-Bonin) Trench, Awaji Island near Osaka and Hatahiyodori in Daito-Town in Shimane Prefecture on the western tip of Honshu, mainland Japan (10,13,17,19). Assuming that the Kobe lineage has arisen from a population bottleneck, a plausible explanation for its evolutionary history may be that the lineage, once widely distributed, had become virtually extinct save for the few remnants with pronounced genetic diversity that are surviving currently in remote focal areas of Japan. Another scenario envisions a founder effect in which a few parasites would have dispersed and evolved quickly in niches and without any competition on an isolated Pacific island. In fact, the sequence changes have occurred almost solely in the Mikura strain, i.e., all but 1 of the 19 sequence differences, including a 27-nt insertion, were identified from a Mikura island sample (Fig. 1B). Analyses on more intron data is difficult due to the sample size limitations resulting from scarce distribution, but would allow more accurate inferences to be drawn.

Introns received little to no attention as candidate tools for Piroplasma systematics until 2009, when the presence or absence of introns from the *CCT7* gene were charted and for the first time used in a phylogenetic study on the *B. microti*-group (17). In that particular study, the matrix containing information on intron presence/absence as well as the presence of the smallest known introns proved instrumental in helping to identify the *B. microti*-group, which is suspected to have diverged concurrently from other Piroplasma genera, *Babesia* and *Theileria*, approximately 350 million years ago (28).

The present study demonstrates the utility of introns as a rapidly evolving DNA marker for strengthening the research arsenal. Our analyses have resulted in further demarcation of the U.S. and Kobe taxa into 3 and 2 geographic subgroups, respectively, and has inferred the existence of 2 late bottleneck events associated with the Munich and Hobetsu taxa distributed regionally across Europe and Japan, respectively. These meaningful results improve our knowledge and fundamental insights into genetic structure as it relates to biological and ecological traits. A practical example for its use is in prevention of transfusion-acquired human babesiosis, which often results in an asymptomatic latent-carrier state (13,29). Currently, serological screening is the most sensitive test to detect highly elevated titers of IgG antibody beyond a year (13). However, since many different antigenic types exist, the assay system requires more complexity since the tests are fairly type-specific (11,30). In Japan, 3 distinct lineages exist (9,10,13–15, 17,19), and it is possible that even more diverse and larger archives of antigens exist as well, which makes the standardization of serological tests daunting task. Our present results on the recent bottleneck in the Hobetsu taxon may provide new insights into the assay testing solution in Japan. Serological testing may be useful to detect asymptomatic human carriers; e.g., the occurrence of population bottlenecks with Hobetsu infections in Japan can impact the degree of polymorphism at antigen-encoding loci.

Further, the Hobetsu and U.S. taxa in Japan may differ in specificity for the vector tick species and perhaps for the Kobe taxon (A. Zamoto-Niikura, personal communication). Likewise, the knowledge of lineage-specific genetic variability and evolutionary relationships should aid in the development of disease management strategies.

The stark difference observed between the 2 largest and smallest introns in the occurrence of indels may reflect robust restraint on, or conservation of, the nucleotide length of extremely short introns. Intron size has been linked with genetic factors like recombination rate (31), genome compaction (26), and splicing efficiency (25). It remains unclear why extremely small introns occur exclusively and frequently in the *B. microti*-group from piroplasms (order Piroplasmida) or what factor(s) governs intron size in the organism. Genomic-scale comparison of sequence- and structure-based methods between separate genera of the order Piroplasmida may provide new insights into the genetic basis for the evolution and role of these unique introns. If precise dating is provided for individual sequences, indel mutations (and possibly substitution mutations), may become a promising tool; intron presence/absence matrix data have been shown to function as a very powerful tool for deciphering early evolution piroplasms (17) and for analyzing genetic diversity and evolutionary relationships within the *B. microti*-group parasites at various points in their evolution.

Acknowledgments This work was supported in part by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

We acknowledge the late Professor Masayoshi Tsuji for his seminal contributions to this research effort and especially for his many years of helping with the sample collections and gene sequencing. Dr. Tsuji, who died December 19, 2006, during the early years of this research provided the guidance that has helped in the demarcation of the "*B. microti*-group" and in the ongoing search for improved methodology. 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. We are especially grateful to Dr. Tatsufumi

Takahashi and Dr. Katsuro Hagiwara (Rakuno Gakuen University, School of Veterinary Medicine) for their generous time and energy providing special technical assistance and insightful discussions from multidisciplinary perspectives related to the study. With heartfelt appreciation, we also thank Dr. Danny H.-Kauffmann Jokl, Department of Ophthalmology, Columbia University Medical Center, New York, New York, U.S.A, and Professor Nell L. Kennedy (Rakuno Gakuen University, School of Veterinary Medicine) for critical discussion and help with the manuscript.

Conflict of interest None to declare.

REFERENCES

1. Homer, M.J., Aguilar-Delfin, I., Telford, S.R., 3rd, et al. (2000): Babesiosis. Clin. Microbiol. Rev., 13, 451–469.
2. Kjemtrup, A.M. and Conrad, P.A. (2000): Human babesiosis: an emerging tick-borne disease. Int. J. Parasitol., 30, 1323–1337.
3. Telford, S.R., 3rd, Gorenflot, A., Brasseur, P., et al. (1993): Babesial infections in humans and wildlife. p. 1–47. *In* J.P. Kreiner (ed.), Parasitic Protozoa. 2nd ed. vol. 5. Academic Press, New York.
4. Duh, D., Petrovec, M. and Avsic-Zupanc, T. (2001): Diversity of *Babesia* infecting European sheep ticks (*Ixodes ricinus*). J. Clin. Microbiol., 39, 3395–3397.
5. Duh, D., Petrovec, M., Trilar, T., et al. (2003): The molecular evidence of *Babesia microti* infection in small mammals collected in Slovenia. Parasitology, 126, 113–117.
6. Goethert, H.K. and Telford, S.R., 3rd. (2003): What is *Babesia microti*? Parasitology, 127, 301–309.
7. Hildebrandt, A., Hunfeld, K.P., Baier, M., et al. (2007): First confirmed autochthonous case of human *Babesia microti* infection in Europe. Eur. J. Clin. Microbiol. Infect. Dis., 26, 595–601.
8. Hunfeld, K.P., Hildebrandt, A. and Gray, J.S. (2008): Babesiosis: recent insights into an ancient disease. Int. J. Parasitol., 38, 1219–1237.
9. Saito-Ito, A., Tsuji, M., Wei, Q., et al. (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.
10. Saito-Ito, A., Yano, Y., Dantrakool, A., et al. (2004): Survey of rodents and ticks in human babesiosis emergence area in Japan: first detection of *Babesia microti*-like parasites in *Ixodes ovatus*. J. Clin. Microbiol., 42, 2268–2270.
11. Tsuji, M., Wei, Q., Zamoto, A., et al. (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.
12. Walter, G. (1981): [Isolation of *Babesia microti* (Franca 1912) from free-living nymphs of *Ixodes ricinus* (Linnaeus 1758) (author's transl)]. Acta Trop., 38, 187–188.
13. Wei, Q., Tsuji, M., Zamoto, A., et al. (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.
14. Zamoto, A., Tsuji, M., Kawabuchi, T., et al. (2004): U.S.-type *Babesia microti* isolated from small wild mammals in Eastern Hokkaido, Japan. J. Vet. Med. Sci., 66, 919–926.
15. Zamoto, A., Tsuji, M., Wei, Q., et al. (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.
16. Tsuji, M., Zamoto, A., Kawabuchi, T., et al. (2006): *Babesia microti*-like parasites detected in Eurasian red squirrels (*Sciurus vulgaris orientis*) in Hokkaido, Japan. J. Vet. Med. Sci., 68, 643–646.
17. Nakajima, R., Tsuji, M., Oda, K., et al. (2009): *Babesia microti*-group parasites compared phylogenetically by complete sequencing of the CCTη gene in 36 isolates. J. Vet. Med. Sci., 71, 55–68.
18. Meer-Scherrer, L., Adelson, M., Mordechai, E., et al. (2004): *Babesia microti* infection in Europe. Curr. Microbiol., 48, 435–437.
19. Tabara, K., Arai, S., Kawabuchi, T., et al. (2007): Molecular survey of *Babesia microti*, *Ehrlichia* species and *Candidatus* Neoehrlichia mikurensis in wild rodents from Shimane Prefecture, Japan. Microbiol. Immunol., 51, 359–367.
20. Rar, V.A., Epikhina, T.I., Livanova, N.N., et al. (2011): Genetic diversity of *Babesia* in *Ixodes persulcatus* and small mammals from North Ural and West Siberia, Russia. Parasitology, 138, 175–182.
21. Sinski, E., Bajer, A., Welc, R., et al. (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 (Suppl. 40), 137–143.
22. Nishisaka, M., Yokoyama, N., Xuan, X., et al. (2001): Characterisation of the gene encoding a protective antigen from *Babesia microti* identified it as eta subunit of chaperonin containing T-complex protein 1. Int. J. Parasitol., 31, 1673–1679.
23. 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.
24. Saitou, N. and Nei, M. (1987): The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4, 406–425.
25. 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.
26. Gilson, P.R., Su, V., Slamovits, C.H., et al. (2006): Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. Proc. Natl. Acad. Sci. USA, 103, 9566–9571.
27. Ricard, G., de Graaf, R.M., Dutilh, B.E., et al. (2008): Macronuclear genome structure of the ciliate *Nyctotherus ovalis*: single-gene chromosomes and tiny introns. BMC Genomics, 9, 587.
28. Nguyen, H.D., Yoshihama, M. and Kenmochi, N. (2007): The evolution of spliceosomal introns in alveolates. Mol. Biol. Evol., 24, 1093–1096.
29. Leiby, D.A., Chung, A.P., Gill, J.E., et al. (2005): Demonstrable parasitemia among Connecticut blood donors with antibodies to *Babesia microti*. Transfusion, 45, 1804–1810.
30. Arai, S., Tsuji, M., Kaiho, I., et al. (2003): Retrospective seroepidemiological survey for human babesiosis in an area in Japan where a tick-borne disease is endemic. J. Vet. Med. Sci., 65, 335–340.
31. Carvalho, A.B. and Clark, A.G. (1999): Genetic recombination: intron size and natural selection. Nature, 401, 344.