

RPS8—a New Informative DNA Marker for Phylogeny of *Babesia* and *Theileria* Parasites in China

Zhan-Cheng Tian*, Guang-Yuan Liu*, Hong Yin, Jian-Xun Luo, Gui-Quan Guan, Jin Luo, Jun-Ren Xie, Hui Shen, Mei-Yuan Tian, Jin-feng Zheng, Xiao-song Yuan, Fang-fang Wang

State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, China

Abstract

Piroplasmosis is a serious debilitating and sometimes fatal disease. Phylogenetic relationships within piroplasmida are complex and remain unclear. We compared the intron–exon structure and DNA sequences of the RPS8 gene from *Babesia* and *Theileria* spp. isolates in China. Similar to 18S rDNA, the 40S ribosomal protein S8 gene, RPS8, including both coding and non-coding regions is a useful and novel genetic marker for defining species boundaries and for inferring phylogenies because it tends to have little intra-specific variation but considerable inter-specific difference. However, more samples are needed to verify the usefulness of the RPS8 (coding and non-coding regions) gene as a marker for the phylogenetic position and detection of most *Babesia* and *Theileria* species, particularly for some closely related species.

Citation: Tian Z-C, Liu G-Y, Yin H, Luo J-X, Guan G-Q, et al. (2013) RPS8—a New Informative DNA Marker for Phylogeny of *Babesia* and *Theileria* Parasites in China. PLoS ONE 8(11): e79860. doi:10.1371/journal.pone.0079860

Editor: Kelly A. Brayton, Washington State University, United States of America

Received: June 26, 2013; **Accepted:** September 25, 2013; **Published:** November 7, 2013

Copyright: © 2013 Tian et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported, in part, by the National Natural Sciences Foundation of China (numbers 31201899 and 30972182), the Natural Science Foundation of Gansu Province (number 096RJZA128), the Science Fund for Creative Research Groups Of Gansu Province (number 1210RJIA006), Key Project of Gansu Province (092NKDA031). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: liuguangyuan2002@sina.com (GYL); tianzhancheng@caas.cn (ZCT)

Introduction

The piroplasms, comprising mainly the genera *Babesia* and *Theileria*, are tick-transmitted protozoa that are highly pathogenic to ruminants, horses, pigs, dogs, cats and cattle, and in some cases, even to humans. In the vertebrate hosts, the infection usually causes fever, anemia and haemoglobinuria, and in severe cases, death [1]. Animals that recover from acute or primary infections naturally remain persistently infected, and act as reservoirs for infecting ticks.

There are some controversial species placements such as *T. equi* and *B. microti* [2–6]. In the most thorough phylogenetic examination to date, incongruencies in the phylogenetic evolution with a taxonomically different dataset were displayed [6–9]. However, in the complex phylogenetic relationships between *Babesia* and *Theileria* spp., previous studies only relied on 18S rDNA gene [4]. The rDNA genes possess both conserved stems and variable loop regions which provide signals for different levels of phylogenetic inference [10], the 18S rDNA sequences support many piroplasm clades, but, being a slow-evolving marker it may fail to provide enough phylogenetic signal to resolve relationships at the species level such as some closely related *Babesia* species in China [11].

Internal transcribed spacer 2 (ITS2) may be a more ideal DNA barcode based on the current database for piroplasma [12], however, complex and unpredictable evolutionary behavior of ITS reduces its utility for phylogenetic analysis [11,13]. Information on mitochondrial DNA (mtDNA) from *Babesia* and *Theileria* spp., is limited, which precludes its use despite its advantage for use as a molecular marker for lower-level phylogeny [11,14,15]. Genome-wide analysis may aid in determining the taxonomy of species such as *B. microti* and *T. equi* but its usefulness as an everyday tool to classify Apicomplexan parasites is limited by availability of complete genome sequences and requirement for a large number of permutations [4–6]. It is therefore highly desirable to test other genes besides the 18S rRNA - to further improve phylogenetic analysis of *Babesia* and *Theileria* species.

Structure and sequence signatures in ribosomal RNA and proteins are defining characteristics of the three domains of life and instrumental in constructing the modern phylogeny [16]. Based on the Ribosomal protein S8 and L4, amino acid sequence alignments of orthologous ribosomal proteins found in Bacteria, Archaea, and Eukaryota display an unusual segment or block structure with major evolutionary implications [17]. A set of 50 informative genes that could be analyzed in a

broader sampling of Piroplasmida taxa to gain a greater understanding of the evolutionary relationships of the piroplasms were revealed in the comparative genomic analysis of *T. equi* [5]. Among them, we selected the 40S ribosomal protein S8 (RPS8) gene locus (including coding and non-coding regions) as phylogenetic marker to evaluate the merits and shortcomings of the phylogeny based on the comparative analysis of the RPS8 and 18S rDNA of *Babesia* and *Theileria* isolates from China for phylogenetic analysis in this study.

Results and Discussion

RPS8 gene sequence information

The complete sequences of the RPS8 (coding and non-coding regions) gene were obtained from twenty-three isolates representing seven *Babesia* and six *Theileria* spp. in China (Table 1). The primers were designed to amplify 561 bp of coding sequence for *Babesia* species and 573 bp of coding sequence for *Theileria* species, which are known to be interrupted by an intron in *Babesia* species and two introns in *Theileria* species. The size of the introns varied from 146 bp in *T. sinensis* to 282 bp in *Babesia* sp. Kashi2 and were specifically located at the conserved position in all the species (Figure 1). The identity of nucleotide sequences of RPS8 between *Babesia* and *Theileria* species varied from 63.5% to 82.9% (interspecific variability), whereas within *Babesia* and *Theileria* species (intraspecific variability) the identity was approximately 89% to 99.7% respectively (data not shown). The nucleotide sequences of the RPS8 (coding and non-coding regions) genes of *Babesia* and *Theileria* species in this study have been deposited in the GenBank database under accession Nos. JN400408 to JN4004028 and Nos. JX402859 to JX402860. The nucleotide sequences of the 18S rDNA genes of several *Theileria* species in this study have been deposited in the GenBank database under accession Nos. KF559355 to KF559357. The nucleotide sequences of the RPS8 (coding and non-coding regions) genes of *B. bovis* USA isolate (NW_001820855), *B. microti* USA isolate (FO082874), *T. annulata* Ankara strain (NC_011099), *T. parva* Muguga strain (NC_007345), *T. orientalis* Shintoku strain (AP011947), *T. equi* Florida strain (ACOU00000000) were drawn from the GenBank database.

Interspecific Genetic Distance

The ribosome, with its conserved central role in protein synthesis, has long constituted a prime subject for phylogenetic analysis [17]. The coding sequence of the RPS8 gene displays moderate conservation as a constitutive component of the ribosome. Compared with the moderate sequence identity seen in RPS8 coding regions from *Babesia* and *Theileria* species, the RPS8 non-coding regions are highly divergent. Using the Tamura–Nei model of sequence evolution, d values (Genetic Distance) were calculated independently across coding and non-coding region sequences (Data not shown). Across *Babesia* and *Theileria* species, d ranges were 0.07903–0.47956 (coding) and 0.12511–1.14484 (non-coding regions). The d-values for non-coding regions are on average 2-fold greater than those for coding regions. Furthermore, the degree

of variation in the RPS8 (coding and non-coding regions) dataset was considerable when observing pairwise differences between sequences. Among the eight *Babesia* and seven *Theileria* species sequenced, the average p-distances were 0.45, while for the 18S rDNA the average p-distance is 0.083, indicating the RPS8 (coding and non-coding regions) being more variable than the 18S rDNA (Data not shown).

Comparison of the RPS8 (coding and non-coding regions) gene-based phylogenetic trees and the 18S rDNA gene-based phylogenetic trees

Phylogenetic trees based on the 18S rDNA and RPS8 (coding and non-coding regions) genes were constructed by the best-fit model of Bayesian and Maximum likelihood (ML) analysis (Figure 2 and 3). The Bayesian trees were congruent with those obtained under the ML criterion. Nodes receiving $\geq 85\%$ bootstrap (BP) support in the Maximum likelihood analysis and/or ≥ 0.98 posterior probability (PP) in the Bayesian analysis were considered strongly statistically supported, and nodes receiving < 60 BP and/or < 0.90 PP were considered poorly supported, revealing significant levels of genetic diversity. In the RPS8 tree, the Bayesian and ML analyses returned nearly identical topologies, however, for *Theileria* species, there was less resolution in the ML analysis given that the RPS8 gene is highly variable. The datasets with significant heterogeneity result in the long-branch attraction and poor resolution at relatively deeper nodes when using the ML approach [18].

Both *Babesia* and *Theileria* were polyphyletic with four major clades being defined (Figure 3). *Theileria* species were clustered within clade I except *T. equi*, this clade was strongly supported by both ML and Bayesian analyses (PP > 0.95, BP > 85). Clade II consisted of only three representatives of a single species, *T. equi/B. equi* (PP = 1.00, BP = 98–100), for which the phylogenetic status is controversial [2,3]. In our analysis, this clade was strongly supported as a separated taxon to other *Theileria* spp. in the 18S rDNA tree (PP = 1.00), furthermore, there are poorly supported as a sister taxon to other *Theileria* spp. with ML (BP = 16) analysis of RPS8. However, there was relatively moderate statistical support for *T. equi* as a sister taxon to other *Theileria* spp. in the Bayesian analysis (PP = 0.91). Its position relative to Clade I and Clade III comprising the majority of *Babesia* sp. was unresolved, indicating the unclear evolutionary position of *T. equi* [4,14]. In both the trees, *B. microti* was placed at the root of piroplasms, thereby, separating it from the *Babesia* and *Theileria* clades (Figure 2, 3), which is consistent with Genome-wide phylogenetic analyses suggesting a new genus for the *B. microti* group of strains [6].

Overall topology of the two trees showed similar major branching orders, beginning with the outgroup *Toxoplasma gondii* followed by polyphyletic *Theileria* and *Babesia* species. However, phylogenetic reconstructions show that, although it is the least variable, the 18S rDNA tree is more resolved than the RPS8 tree in the nodes of some clades. For example, *B. motasi* isolates were clearly separated from the branch composed of *B. ovata* and *B. bigemina* with statistical support (PP=0.97, BP=64) in the 18S rDNA tree as well as there was

Table 1. The host, location, vector and RPS8 and 18S rDNA gene sequences for *Babesia* and *Theileria* species used in this study, * indicated that the sequence were drawn from the database.

parasite	Host	Location	Tick vector	Intron location	Genbank Accession No.	
					RPS8	18S rDNA
<i>Babesia bovis</i>	Cattle	Shaanxian	<i>Rhipicephalus microplus</i>	212-458	JN400408	AY603398*
<i>B. bovis</i>	Cattle	Lushi	<i>R. microplus</i>	212-458	JN400409	JX495403*
<i>B. bovis</i>	Cattle	USA	<i>R. microplus</i>	212-415	NW_001820855*	-
<i>B. bigemina</i>	Cattle	Kunming	<i>R. microplus</i>	212-487	JN400410	AY603402*
<i>B. bigemina</i>	Cattle	Lushi	<i>R. microplus</i>	212-487	JN400411	JX495402*
<i>B. major</i>	Cattle	Yili	<i>Haemaphysalis punctata</i>	212-485	JN400412	AY603399*
<i>B. ovata</i>	Cattle	Lushi	<i>H. longicornis</i>	212-486	JN400413	AY603401*
<i>B. ovata</i>	Cattle	Wenchuan	<i>H. longicornis</i>	212-484	JN400414	AY603403*
<i>B. ovata</i>	Cattle	Zhangjiachuan	<i>H. longicornis</i>	212-487	JN400415	AY603400*
<i>Babesia</i> sp. Kashi2	Cattle	Kashi	<i>Hyalomma</i> spp.	212-493	JN400416	AY726557*
<i>B. motasi</i>	sheep	Lintan	<i>H. qinghaiensis</i>	212-493	JN400417	AY260181*
<i>B. motasi</i>	sheep	Ningxian	<i>H. longicornis</i>	212-491	JX402860	AY260180*
<i>B. motasi</i>	sheep	Tianzhu	<i>H. qinghaiensis</i>	212-493	JX402859	DQ159072*
<i>Babesia</i> sp. Xinjiang-2005	sheep	Kashi	<i>Hyalomma anatolicum</i>	212-403	JN400418	DQ159073*
<i>B. microti</i>	human	USA	<i>Ixodes scapularis</i>	212-362	FO082874* (2706040.2706769)	AF231348*
<i>Theileria annulata</i>	Cattle	Shanmenxia	<i>H. detritum</i>	212-328, 482-512	JN400419	KF559356
<i>T. annulata</i>	Cattle	Xinjiang	<i>H. scupense</i>	212-328, 482-512	JN400420	EU073963*
<i>T. annulata</i>	Cattle	Ningxia	<i>H. detritum</i>	212-328, 482-512	JN400428	EU083800*
<i>T. annulata</i>	Cattle	Ankara	<i>H. detritum</i>	212-328, 482-512	NC_011099*	-
<i>T. parva</i>	Cattle	Muguga	<i>Rhipicephalus appendiculatus</i>	212-328, 482-512	NC_007345*	HQ895968*
<i>T. sergenti</i>	Cattle	Lushi	<i>H. longicornis</i>	212-331, 485-516	JN400421	AF081137*
<i>T. orientalis</i>	Cattle	Shintoku (Japan)	<i>H. longicornis</i>	212-331, 485-516	AP011947*	-
<i>T. sinensis</i>	Yak	Lintan	<i>H. qinghaiensis</i>	212-328, 482-510	JN400422	EU274472*
<i>T. sinensis</i>	Yak	Weiyuan	<i>H. qinghaiensis</i>	212-328, 482-510	JN400423	EU277003*
<i>T. sinensis</i>	Cattle	Lintao	<i>H. qinghaiensis</i>	212-328, 482-510	JN400427	KF559355
<i>T. luwenshuni</i>	sheep	Ningxian	<i>H. qinghaiensis</i>	212-333, 487-515	JN400424	JF719833*
<i>T. uilenbergi</i>	sheep	Longde	<i>H. qinghaiensis</i>	212-329, 483-511	JN400425	JF719835*
<i>T. equi</i>	horse	Zhaoyuan	<i>Demacntor</i> spp.	212-339, 493-528	JN400426	KF559357
<i>T. equi</i>	horse	Florida (USA)	<i>D. variabilis</i> ; <i>Amblyomma cajennense</i> ; <i>R. microplus</i>	212-339, 493-528	ACOU00000000*	-

doi: 10.1371/journal.pone.0079860.t001

moderate statistical support (PP=0.92, BP=78) in the RPS8 tree. These data indicate that 18S rDNA and RPS8 reliably distinguishes the deeper branches among some *Babesia* species. Within *B. ovata* isolates, the phylogenetic relationship between the *B. ovata* Zhangjiachuan isolate and the Wenchuan isolate is reliable in the RPS8 tree (PP=0.99, BP=84) as showed in the 18S rDNA tree (PP=1.00, BP=83). The RPS8 gene seems to be equal to the 18S rDNA in recognizing close lineages among some *Babesia* and *Theileria* species.

In the 18S rDNA tree, the phylogenetic relationship between *B. bovis* and other two *Babesia* species (*Babesia* sp. Xinjiang-2005 and *Babesia* sp. Kashi2) were poorly statistically supported (PP=0.87, BP=37) but the latter two *Babesia* species are closely related with high statistical support (PP=1.00, BP=95), which was not reconciled with morphological and biological data in previous studies (Figure 2) [19-24]. On the contrary, the phylogeny generated using the RPS8 gene did provide extremely strong support for the close sister-taxon relationship between *B. bovis* and the other two *Babesia*

species (*Babesia* sp. Xinjiang-2005 and *Babesia* sp. Kashi2) (PP=1.00, BP=92) and unreliable sister-taxon relationship between *Babesia* sp. Xinjiang-2005 and *Babesia* sp. Kashi2 (PP=0.69, BP=50) (Figure 3). This is probably due to the use of a small dataset and two markers with different modes of evolution. Thus, for more accurate definition of certain *Babesia* isolates, it may be necessary to incorporate more markers that have differing evolutionary rates.

For *Theileria* species, clade I in the 18S rDNA tree was composed of two subclades, one subclade comprising *T. annulata* isolates and *T. parva*, with the other subclade consisting of *T. uilenbergi*, *T. luwenshuni*, *T. sinensis* isolates and *T. sergenti*. The phylogenetic status of two subclades resolved with strong statistical support (PP=1.00, BP=96). Furthermore, *T. annulata* and *T. parva* were closely related with a strong statistical support (PP=1.00, BP=100). *T. luwenshuni* formed a separate branch that is distantly related to *T. sergenti* and *T. sinensis* with statistical support (PP=0.76, BP=84) (Figure 3). *T. uilenbergi* formed a separate branch with

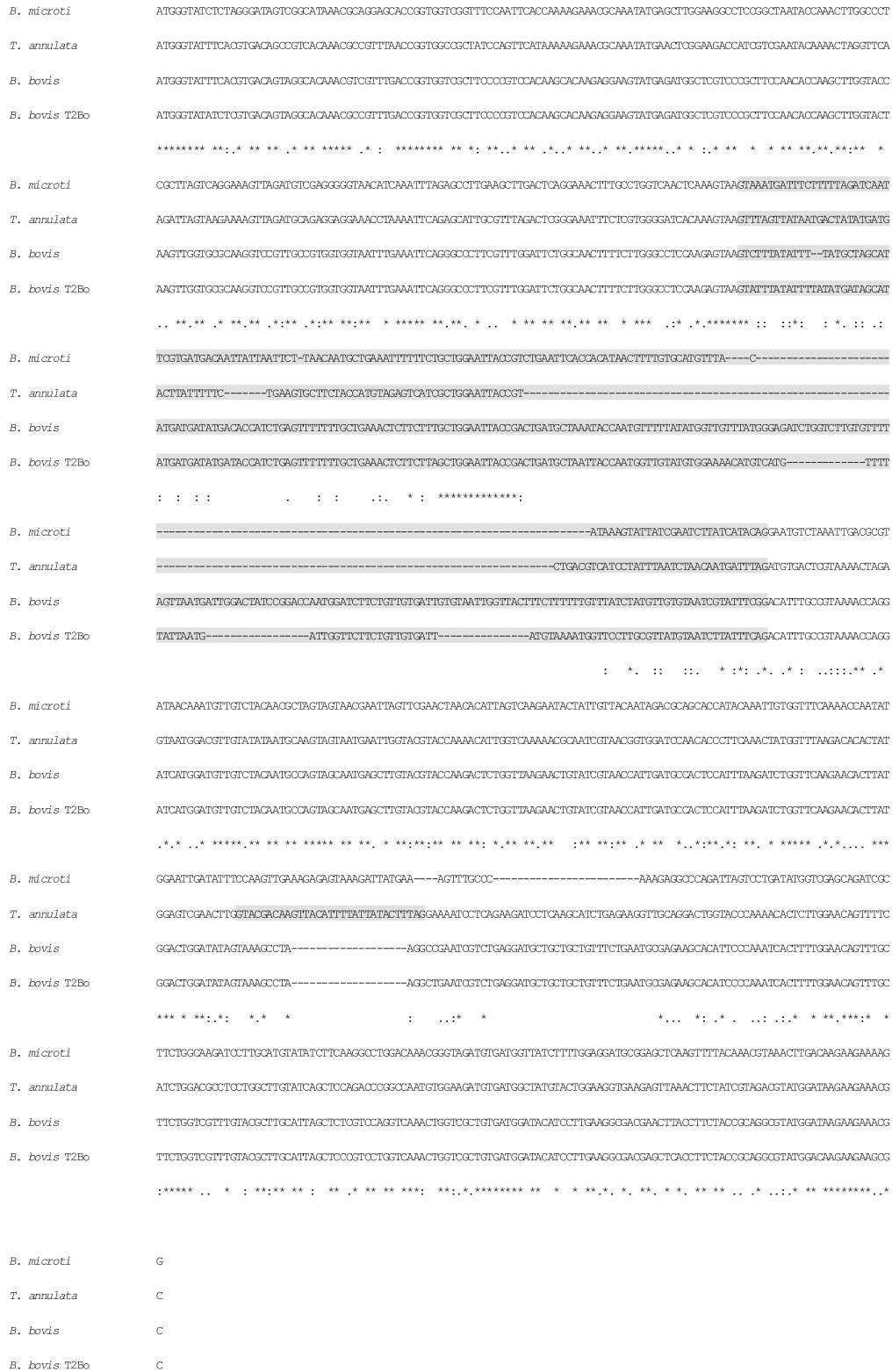


Figure 1. The alignment of RPS8 (coding and non-coding regions) sequences from *B. bovis* and *Theileria* species. *B. bovis* (GenBank accession no. JN400409); *B. microti* (GenBank accession no. FO082874); *T. annulata* (GenBank accession no. NC_011099); *B. bovis* (GenBank accession no. NW_001820855) were aligned using the ClustalW2 program. The non-coding region is marked with gray shading.

doi: 10.1371/journal.pone.0079860.g001

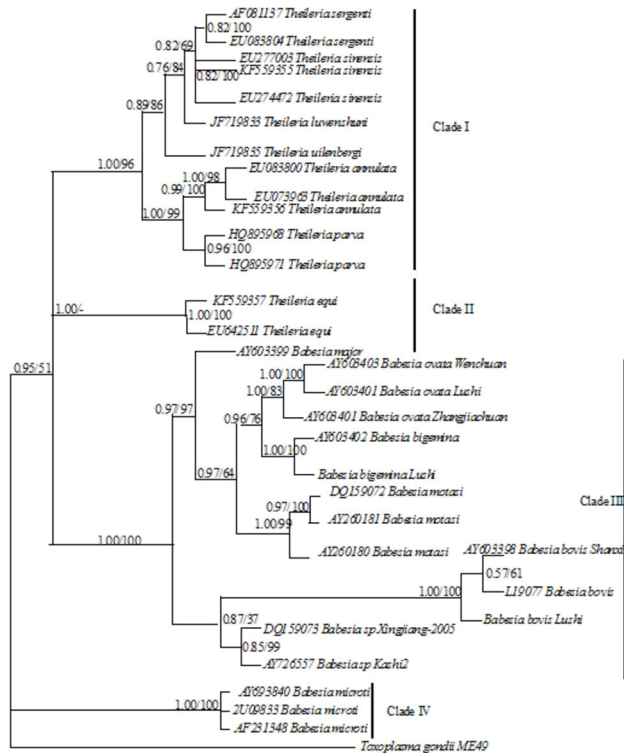


Figure 2. Inferred phylogenetic relationship among representative *Babesia* and *Theileria* species based on 18S rDNA sequences. The 18S rDNA sequences were analyzed utilizing Bayesian analysis (Bayes) and maximum likelihood (ML), using *Toxoplasma gondii* as outgroup. The numbers along branches indicate posterior probability (PP) and bootstrap probability (BP) values resulting from different analyses in the order: Bayes/ ML. Highly statistically supported nodes were BP \geq 85; PP \geq 0.98; while poorly statistically supported nodes were BP $<$ 60; PP $<$ 0.90.

doi: 10.1371/journal.pone.0079860.g002

statistical support (PP=0.89, BP=86). In the RPS8 tree, *T. luwenshuni* formed a separate branch with strong statistical support (PP=0.98, BP=76) [25,26].

In short, 18S rDNA as a slow-evolving marker more reliably distinguishes deeper branches among some *Babesia* species than the RPS8 gene. However, the RPS8 gene seems to be equal to the 18S rDNA in recognizing lineages among some closely related *Babesia* and *Theileria* species.

Combined analysis

A phylogenetic tree was reconstructed based on the combined dataset of the two markers. The resulting tree is presented in Figure 4. It is mostly reconciled with the phylogenetic trees constructed based on the RPS8 gene and 18S rDNA apart from the phylogenetic relationship of *T. equi* to other *Theileria* spp. The clade of *T. equi* was strongly supported as sister taxon to other *Theileria* spp. with ML (BP=97) analysis of the combined dataset of the two markers. However, the phylogenetic tree constructed based on the

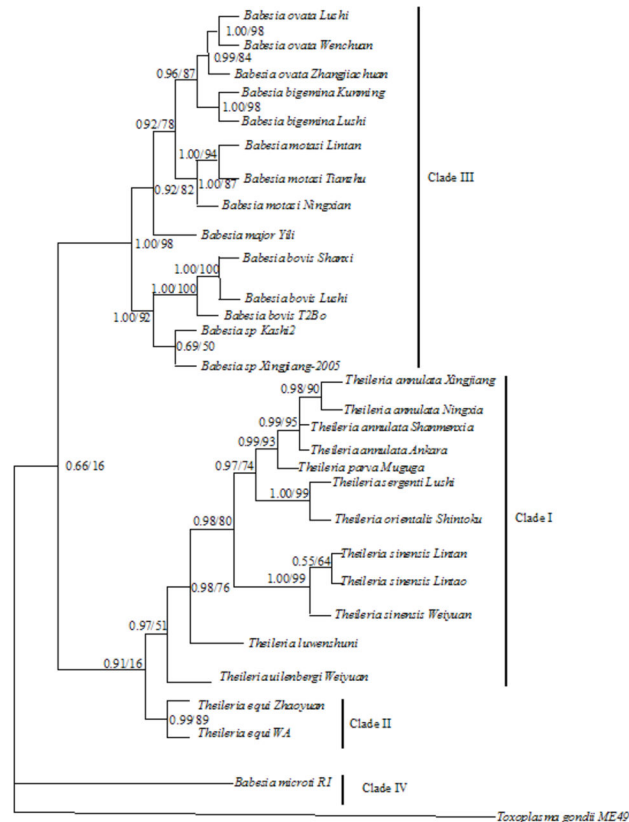


Figure 3. Inferred phylogenetic relationship among representative *Babesia* and *Theileria* species based on RPS8 (coding and non-coding regions) gene sequences. The RPS8 (coding and non-coding regions) gene sequences were analyzed utilizing Bayesian analysis (Bayes) and maximum likelihood (ML), using *Toxoplasma gondii* as outgroup. The numbers along branches indicate posterior probability (PP) and bootstrap probability (BP) values resulting from different analyses in the order: Bayes/ ML. The accession numbers of the isolates used in the phylogenetic tree were listed in Table 1. High statistically supported nodes had BP \geq 85; PP \geq 0.98; while poorly statistically supported nodes had BP $<$ 60; PP $<$ 0.90.

doi: 10.1371/journal.pone.0079860.g003

combined dataset of the two markers is not reliable with poor statistical support (Data not shown).

Materials and Methods

Ethics statement

This study was approved by the Animal Ethics Committee of the Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences. All sheep and calves, were handled in accordance with good animal practices required by the Animal Ethics Procedures and Guidelines of the People's Republic of China.

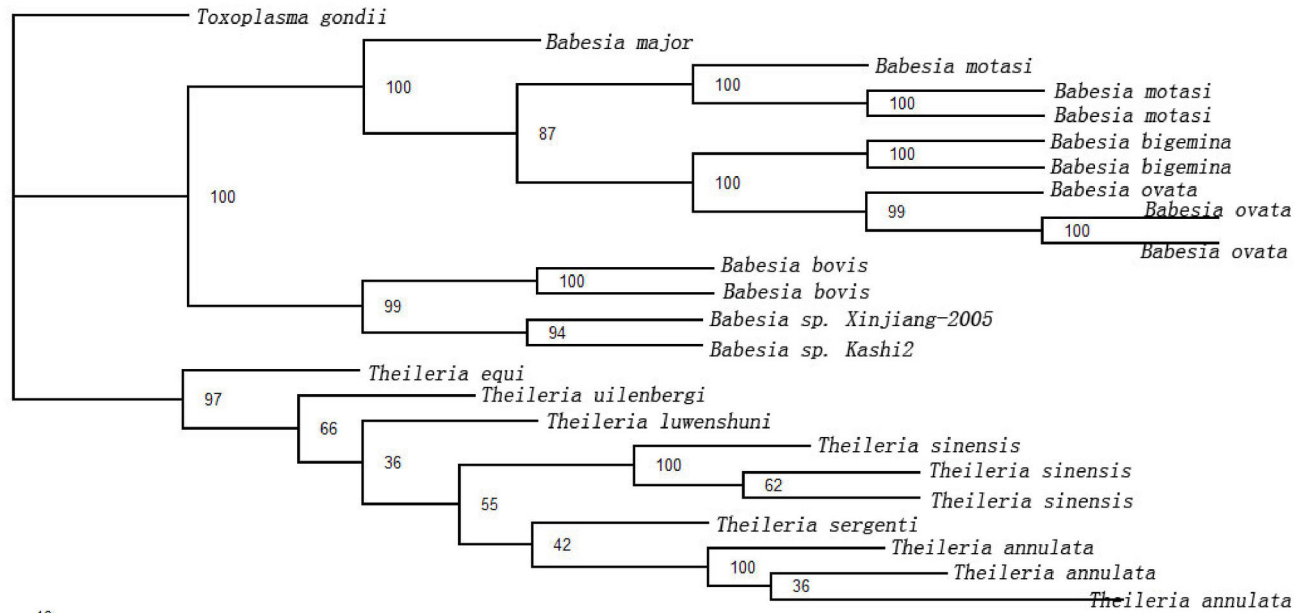


Figure 4. Inferred phylogenetic relationship among representative *Babesia* and *Theileria* species based on the combined data of 18S rDNA and RPS8 (coding and non-coding regions) gene sequences. The combined data of 18S rDNA and RPS8 (coding and non-coding regions) gene sequences were analyzed utilizing maximum likelihood (ML), using *Toxoplasma gondii* as outgroup. The numbers along branches indicate bootstrap probability (BP) values. The accession numbers of the isolates used in the phylogenetic tree were listed in Table 1. High statistically supported nodes had BP \geq 85; while poorly statistically supported nodes had BP $<$ 60.

doi: 10.1371/journal.pone.0079860.g004

Animals

All sheep and calves were purchased from a *Babesia* and *Theileria*-free area and maintained in an isolated stable. All sheep and calves were splenectomized and treated with antibiotics to promote the healing of wounds one month before the study. During this time, blood films were taken weekly from the ears of the sheep and calves to be examined by Giemsa stain for the presence of haemoprotozoan parasites. The experimental animals were tested by PCR with the universal primers for *Babesia* and *Theileria* species based on the 18S rDNA sequences prior to use to ensure that they were free of hemoparasites. After collecting blood containing parasites, the experimental animals were treated with anti-piroplasmosis drug and they were rehabilitated.

Parasite species

The isolates used in this study are listed in Table 1. *Babesia bovis* (Shanxian and Lushi) [27], *B. bigemina* (Kunming and Lushi) [28], *B. major* (Yili) [29], *B. ovata* (Wenchuan and Lushi and Zhangjiachuan) [30,31,32], *Babesia* sp. Kashi2 (Kashi) [24], *B. motasi* (Ningxian and Tianzhu and Lintan) [33,34], *Babesia* sp Xingjiang-2005 (Kashi) [21], *Theileria annulata* (Xingjiang and Ningxia and Shanmenxia) [30,35], *T. sergenti* (Lushi) [36], *T. sinensis* (Weiyuan and Lintan and Lintao) [37,38], *T. uilenbergi* (Longde) [26], *T. luwenshuni* (Ningxian) [26], *T. equi* (Zhaoyuan) [39]. The reference parasite species

are as follows: *T. annulata* (Ankara strain) [40], *T. orientalis* (Shintoku strain) [41], *T. equi* (USDA strain) [5], *T. parva* (Muguga strain) [42], *B. microti* (RI isolate) [6], *Toxoplasma gondii* (ME49 strain) [43], *B. bovis* (T2Bo strain) [44].

DNA extraction

Sheep and calves were infected intravenously with 15 ml of cryopreserved infected blood stock of these *Babesia* and *Theileria* isolates. Daily post-infection rectal temperatures were measured to monitor for disease and blood smears were examined to monitor for presence of piroplasms. When the parasitemia reached more than 5% of whole blood, blood was collected into heparinised tubes. Parasite DNA was isolated using a genomic DNA Purification Kit (Gentra, USA) according to the manufacturer's instructions. The amount of DNA isolated was assessed spectro-photometrically. Control DNA was isolated from venous blood of uninfected sheep and calves.

PCR, cloning, and DNA sequencing

The partial RPS8 gene was amplified from the genomic DNA extracted from each sample except *Babesia microti* by conventional PCR using forward primer 5'- ATGGGTATT(A/C)TCA(G/T)C(A)GT(C/G)GAC(T)AG-3' and reverse primer 5'- GCGTTTCTTCTTA(G)TCCATACG -3'. The reaction mixture consisted of 10 \times PCR buffer, 1.5 mM MgCl₂, 200 mM each deoxynucleotide triphosphate, 40 pmol each primer, 1.5 U of

Taq polymerase (Takara) and approximately 10 ng of DNA, in a final volume of 50 µl. Each PCR consisted of 35 cycles of denaturation at 94°C for 45 s, annealing at 52°C for 60 s, and extension at 72°C for 60 s; an initial denaturation step consisting of incubation at 94°C for 5 min and a final extension step consisting of incubation at 72°C for 10 min was also included. After PCR amplification, the PCR fragment was cloned into the pGEM-T Easy vector (Promega) according to the manufacturer's recommendations. At least two positive clones from each sample were sequenced by using ABI PRISM™ 377XL DNA sequencer (TaKaRa). All new data has been deposited in GenBank.

Bioinformatic study of genetic distances and phylogenetic analysis

Both ML and Bayesian approaches were used to evaluate each of the individual loci separately. The outgroup taxon *Toxoplasma gondii* was used for all loci. ML analyses were conducted with PhyML [45] as implemented within Geneious [46], using the best-fit models of nucleotide substitution detected by jModelTest. Support for nodes was estimated by analyzing 1000 bootstrap pseudoreplicates for each locus. Bayesian phylogenetic analysis was conducted on the total dataset using MrBayes v3.1.2 [47]. Akaike information criterion was used to identify the most appropriate model of nucleotide substitution for the Bayesian analysis in the program MrModeltest v2.2, and the best-fit model determined by jModelTest for 18S and RPS8 was GTR with both a proportion of invariable sites (I) and variation among sites (G). The Bayesian analysis was run for 15,000,000 generations with phylogenies sampled every 1,000 generations, and values for the substitution model parameters were not defined a priori, but were treated as unknown variables with uniform priors. Resulting burn-in values (the point at which the model parameters and tree scores reached stationarity) were determined empirically by evaluating likelihood scores. The nucleotide distance matrices were created under a ML

correction in MEGA 4.0. The extent of sequence disparity between specimens was calculated by averaging pairwise comparisons of sequence differences across all specimens [48].

The sequence data were also examined using Maximum likelihood analyses in PhyML [45]. The combined data of 18S rDNA and RPS8 (coding and non-coding regions) gene sequences from *Toxoplasma gondii* was used as an outlier group. The phylogenetic trees were constructed based on the combined data of 18S rDNA and RPS8 (coding and non-coding regions) gene sequences from *Theileria* and *Babesia* species determined in our laboratory or obtained from the GenBank database.

The program Genescan for predicting the locations and exon-intron structures of genes in genomic sequences from a variety of organisms (<http://genes.mit.edu/GENSCAN.html>). The RPS8 deduced amino acid sequence was analyzed with the Expert Protein Analysis System (<http://us.expasy.org/>). Theoretical molecular weights and isoelectric point (pI) were determined using Peptide-Mass (<http://us.expasy.org/tools/peptide-mass.html>), and prediction of potential phosphorylation sites was carried out at the NetPhos website. Identification of other motifs was conducted using the Motif scan program (<http://hits.isb-sib.ch/cgi-bin/PFSCAN>).

Acknowledgements

We thank five anonymous reviewers and editor Kelly A. Brayton for their valuable comments and suggestions. The authors also thank Sueyakant D. Waghela (Texas A&M University) for improving the manuscript.

Author Contributions

Conceived and designed the experiments: ZCT GYL. Performed the experiments: ZCT. Analyzed the data: ZCT. Contributed reagents/materials/analysis tools: GQG JRX HY JXL JL MYT HS XSY FFW JFZ. Wrote the manuscript: ZCT.

References

- Kuttler KL (1988) World-wide impact of Babesiosis, M Ristic, Babesiosis of domestic animals and man. Boca Raton, FL: CRC Press. pp. 1–22. Available: <http://scholar.google.com/hk/scholar>
- Mehlhorn H, Shein E (1984) The piroplasms: life cycle and sexual stages. *Adv Parasitol* 23: 37–103. Available: <http://www.ncbi.nlm.nih.gov/pubmed/6442536>. PubMed: 6442536.
- Schein E, Rehbein G, Voigt WP, Zweggarth E (1981) *Babesia equi* (Laveran 1901) -Development in horses and in lymphocyte culture. *Tropenmed Parasitol* 32: 223–227. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7345686>. PubMed: 7345686.
- Lack JB, Reichard MV, Van Den Bussche RA (2012) Phylogeny and evolution of the Piroplasmida as inferred from 18S rRNA sequences. *Int J Parasitol* 42: 353–363. doi:10.1016/j.ijpara.2012.02.005. PubMed: 22429769; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22429769>. Accessed: 13 March 2012.
- Kappmeyer LS, Thiagarajan M, Herndon DR, Ramsay JD, Caler E et al. (2012) Comparative genomic analysis and phylogenetic position of *Theileria equi*. *BMC Genomics* 13: 603. doi: 10.1186/1471-2164-13-603. PubMed: 23137308; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/23137308>. Accessed: 9 November 2012.
- Cornillot E, Hadj-Kaddour K, Dassouli A, Noel B, Ranwez V et al. (2012) Sequencing of the smallest Apicomplexan genome from the human pathogen *Babesia microti*. *Nucleic Acids Res* 40: 9102–9114. doi:10.1093/nar/gks700. PubMed: 22833609; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22833609>. Accessed: 24 July 2012.
- Criado-Fornelio A, Martinez-Marcos A, Buling-Saraña A, Barba-Carretero JC (2003) Molecular studies on *Babesia*, *Theileria* and *Hepatozoon* in southern Europe part II. Phylogenetic analysis and evolutionary history. *Vet Parasitol* 114: 173–194. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12788253> Accessed: 11 June 2003. doi: 10.1016/S0304-4017(03)00141-9. PubMed: 12788253.
- Allsopp MTE, Allsopp BA (2006) Molecular sequence evidence for the reclassification of some *Babesia* species. *Ann NY Acad Sci* 1081: 509–517. Available: <http://www.ncbi.nlm.nih.gov/pubmed/17135560> Accessed: 2006 Oct
- Morrison DA (2009) Evolution of the Apicomplexa: where are we now? *Trends Parasitol* 25: 375–382. doi:10.1016/j.pt.2009.05.010. PubMed: 17135560. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19635681> Accessed: 25 July 2009.
- Voigt O, Erpenbeck D, Wörheide G (2008) Molecular evolution of rDNA in early diverging Metazoa: first comparative analysis and phylogenetic application of complete SSU rRNA secondary structures in *Porifera*. *BMC Evol Biol* 8: 69. doi:10.1186/1471-2148-8-69. PubMed: 18304338; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/18304338>. Accessed: 27 February 2008.

11. Tian Z, Luo J, Zheng J, Xie J, Shen H et al. (2013a) Phylogenetic analysis of *Babesia* species in China based on cytochrome b (COB) gene. *Infect Genet Evol* 13: 36–40. doi:10.1016/j.meegid.2012.09.001. PubMed: 23041715; Available: <http://www.ncbi.nlm.nih.gov/pubmed/23041715>. Accessed: 4 October 2012.
12. Gou H, Guan G, Liu A, Ma M, Xu Z et al. (2012) A DNA barcode for Piroplasma. *Acta Trop* 124: 92–97. doi:10.1016/j.actatropica.2012.07.001. PubMed: 22796448; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22796448>. Accessed: 14 July 2012.
13. Alvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol* 29: 417–434. Available: <http://www.ncbi.nlm.nih.gov/pubmed/14615184>. Accessed: December 2003.
14. Hikosaka K, Watanabe Y, Tsuji N, Kita K, Kishine H, et al. (2010) Divergence of the mitochondrial genome structure in the apicomplexan parasites, *Babesia* and *Theileria*. *Mol Biol Evol* 27: 1107–1116. doi:10.1093/molbev/msp320. PubMed: 14615184. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20034997>. Accessed: 24 December 2009.
15. Tian Z, Liu G, Yin H, Luo J, Guan G et al. (2013b) Cytochrome c oxidase subunit III (COX3) gene, an informative marker for phylogenetic analysis and differentiation of *Babesia* species in China. *Infect Genet Evol* 18: 13–17. doi:10.1016/j.meegid.2013.04.002. PubMed: 23619098; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/23619098>. Accessed: 22 April 2013.
16. Chen K, Eargle J, Sarkar K, Gruebele M, Luthy-Schulten Z (2010) Functional role of ribosomal signatures. *Biophys J* 99: 3930–3940. doi:10.1016/j.bpj.2010.09.062. PubMed: 21156135; Available: <http://www.ncbi.nlm.nih.gov/pubmed/21156135>. Accessed: 15 December 2010.
17. Vishwanath P, Favaretto P, Hartman H, Mohr SC, Smith TF (2004) Ribosomal protein-sequence block structure suggests complex prokaryotic evolution with implications for the origin of eukaryotes. *Mol Phylogenet Evol* 33: 615–625. Available: <http://www.ncbi.nlm.nih.gov/pubmed/15522791>. Accessed: December 2004
18. Felsenstein J (2004) *Inferring Phylogenies*. Sunderland, MA, USA: Sinauer Associates, Inc.. Available: <http://scholar.google.com/hk/scholar>. doi:10.1016/j.ympve.2004.07.003. PubMed: 15522791?
19. Guan G, Yin H, Luo JX, Lu WS, Zhang QC et al. (2001) Isolation of a large ovine *Babesia* sp. in Xinjiang, China. *Chin J Vet Sci Technol* 31: 35–36. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
20. Guan GQ, Yin H, Luo JX, Lu WS, Zhang QC et al. (2002) Transmission of *Babesia* sp. to sheep with field-collected *Haemaphysalis qinghaiensis*. *Parasitol Res* 88: S22–S24. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12051601>. Accessed: May 2002
21. Guan G, Ma M, Moreau E, Liu J, Lu B, et al. (2009) A new ovine *Babesia* species transmitted by *Hyalomma anatolicum anatolicum*. *Exp Parasitol* 122: 261–267. doi:10.1016/j.exppara.2009.05.001. PubMed: 12051601. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19460377>. Accessed: 19 May 2009.
22. Luo J, Chen F, Lu W, Guan G, Ma M et al. (2003) Experimental transmission of an unnamed bovine *Babesia* by *Hyalomma* spp., *Haemaphysalis longicornis* and *Boophilus microplus*. *Vet Parasitol* 116: 115–124. Available: <http://www.ncbi.nlm.nih.gov/pubmed/14519316>. Accessed: 8 October 2003. doi:10.1016/S0304-4017(03)00262-0. PubMed: 14519316.
23. Luo J, Yin H, Liu Z, Yang D, Guan G et al. (2005a) Molecular phylogenetic studies on an unnamed bovine *Babesia* sp. based on small subunit ribosomal RNA gene sequences. *Vet Parasitol* 133: 1–6. Available: <http://www.ncbi.nlm.nih.gov/pubmed/15982821>. Accessed: 10 October 2005, doi:10.1016/j.vetpar.2005.02.014. PubMed: 15982821.
24. Luo J, Yin H, Guan G, Yang D, Liu A et al. (2005b) A comparison of small-subunit ribosomal RNA gene sequences of bovine *Babesia* species transmitted by *Haemaphysalis* spp. in China. *Parasitol Res* 95: 145–149. Available: <http://www.ncbi.nlm.nih.gov/pubmed/15609061>. Accessed: 18 December 2004, doi:10.1007/s00436-004-1268-4.
25. Schnittger L, Yin H, Gubbels MJ, Beyer D, Niemann S et al. (2003) Phylogeny of sheep and goat *Theileria* and *Babesia* parasites. *Parasitol Res* 91: 398–406. Available: <http://www.ncbi.nlm.nih.gov/pubmed/14505044>. Accessed: 18 September 2003. doi:10.1007/s00436-003-0979-2. PubMed: 14505044.
26. Yin H, Schnittger L, Luo J, Seitzer U, Ahmed JS (2007) Ovine theileriosis in China: a new look at an old story. *Parasitol Res* 101: S191–S195. Available: <http://www.ncbi.nlm.nih.gov/pubmed/17823827>. Accessed: September 2007.
27. Bai Q, Liu GY, Zhou JY (1991) Isolation of single species of *Babesia* bovis. *Chin J Vet Sci Technol* 1: 20–22. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.. doi:10.1007/s00436-007-0689-2 PubMed: 17823827
28. Bai Q, Yin SX, Chen ZH (1987) Isolation of single species of *Babesia bigemina*. *Chin J Vet Sci Technol* 9: 25–27. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
29. Yin H, Lu W, Luo J, Zhang Q, Lu W et al. (1996) Experiments on the transmission of *Babesia major* and *Babesia bigemina* by *Haemaphysalis punctata*. *Vet Parasitol* 67: 89–98. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9011018>. doi:10.1016/S0304-4017(96)01022-9. PubMed: 9011018.
30. Luo JX, Yin H, Liu GY, Guan GQ, Liu ZJ et al. (2006) Collection and identification of Piroplasma infected to cattle and sheep in China. *Chin J Parasitol Parasit Dis* 24: S48–S53. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October.
31. Bai Q, Liu GY, Zhang L (1990) Discovery and isolation of *Babesia ovata* in China. *Chin J Vet Sci* 16: 2–4. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
32. Bai Q, Liu GY, Han GF, Hui Y (1994) Isolation and complementary transmission of *Babesia ovata* in Zhangjiachuan of Gansu Province. *Chin J Vet Sci Technol* 24: 9–10. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
33. Yin H, Luo JX, Lu WS, Lu WX, Zhang QC et al. (1997) Isolation and Morphology survey of the *Babesia motasi* and *Babesia ovis* in China. *Chin J Vet Sci Technol* 27: 7–9. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
34. Yin H, Lu W, Luo JX (1997) Babesiosis in China. *Trop Anim Health Prod* 29: 11–15. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9512738>. doi:10.1007/BF02632908.
35. Luo JX, Lu WS (1997) Cattle theileriosis in China. *Trop Anim Health Prod* 29: 4–7. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9512736>. doi:10.1007/BF02632906.
36. Liu A, Guan G, Liu Z, Liu J, Leblanc N et al. (2010) Detecting and differentiating *Theileria sergenti* and *Theileria sinensis* in cattle and yaks by PCR based on major piroplasm surface protein (MPSF). *Exp Parasitol* 126: 476–481. doi:10.1016/j.exppara.2010.05.024. PubMed: 20685208; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20685208>. Accessed: 1 June 2010.
37. Bai Q, Liu GY, Yin H, Zhao QZ, Liu DK et al. (2002) *Theileria Sinensis* sp nov: A New Species of Bovine *Theileria* - Classical Taxonomic Studies. *Acta Veterinaria et Zootechnica Sinica* 33: 73–77. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
38. Yin H, Luo J, Guan G, Gao Y, Lu B et al. (2002) Transmission of an unidentified *Theileria* species to small ruminants by *Haemaphysalis qinghaiensis* ticks collected in the field. *Parasitol Res* 88: S25–S27. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12051602>. Accessed: May 2002, PubMed: 12051602.
39. Luo J, Liu GY, Tian ZC, Xie JR, Zhang P et al. (2011) Taxonomic Status Assays of *Theileria equi* Based on Hsp70 Gene. *Sci Agricultura Sinica* 44: 3694–3700. (in Chinese) Available: <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CJFQ>. Accessed: 22 October 2013.
40. Hayashida K, Hara Y, Abe T, Yamasaki C, Toyoda A et al. (2012) Comparative genome analysis of three eukaryotic parasites with differing abilities to transform leukocytes reveals key mediators of *Theileria*-induced leukocyte transformation. *mBio* 3: e00204-12. doi:10.1128/mBio.00204-12. PubMed: 22951932; Available: <http://www.ncbi.nlm.nih.gov/pubmed/22951932>. Accessed: 4 September 2012.
41. Pain A, Renaud H, Berriman M, Murphy L, Yeats CA et al. (2005) Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*. *Science* 309: 131–133. Available: <http://www.ncbi.nlm.nih.gov/pubmed/15994557>. Accessed: 1 July 2005. doi:10.1126/science.1110418. PubMed: 15994557.
42. Gardner MJ, Bishop R, Shah T, de Villiers EP, Carlton JM et al. (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* 309: 134–137. Available: <http://www.ncbi.nlm.nih.gov/pubmed/15994558>. Accessed: 1 July 2005, doi:10.1126/science.1110439. PubMed: 15994558.
43. Bontell IL, Hall N, Ashelford KE, Dubey JP, Boyle JP et al. (2009) Whole genome sequencing of a natural recombinant *Toxoplasma gondii* strain reveals chromosome sorting and local allelic variants. *Genome Biol* 10: R53. doi:10.1186/gb-2009-10-5-r53. PubMed: 19457243; Ncbi website. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19457243>. Accessed: 20 May 2009.

44. Brayton KA, Lau AO, Herndon DR, Hannick L, Kappmeyer LS et al. (2007) Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. *PLOS Pathog* 3: 1401–1413. Available: <http://www.ncbi.nlm.nih.gov/pubmed/17953480> Accessed: 19 October 2007, PubMed: 17953480.
45. Guindon S, Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704. Available: <http://www.ncbi.nlm.nih.gov/pubmed/14530136> Accessed: October 2003, doi:10.1080/10635150390235520. PubMed: 14530136.
46. Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M et al. (2009) Geneious v4.7. Available: <http://www.geneious.com>.
47. Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11524383> Accessed: August 2001, doi: 10.1093/bioinformatics/17.8.754. PubMed: 11524383.
48. Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17: 368–376. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7288891>. doi:10.1007/BF01734359. PubMed: 7288891.