Thèse de Doctorat

Qingli NIU

Mémoire présenté en vue de l'obtention
du grade de Docteur d'Oniris
Sous le label de l'Université Nantes Angers Le Mans

Discipline : *Biologie, Médecine, Santé*
Spécialité : *Aspects moléculaires et cellulaires de la biologie*
Laboratoire : UMR1300 INRA Oniris « Biologie, Epidémiologie et Analyses de Risques en santé animale », Oniris, Atlanpole, La Chantrerie, CS 40706, 44307 Nantes cedex 3

Soutenue le 28 avril 2014

École doctorale : Biologie Santé

Comparison of sequence, organisation and expression of the rhoptry-associated-protein-1 genes in *Babesia* isolates responsible for ovine Babesiosis in China

JURY

Rapporteurs : **Isabelle FLORENT**, Professeur au Muséum National d'Histoire Naturelle, UMR 7245 CNRS, MCAM, Paris

Stéphane DELBECQ, Maître de conférences, EA 4558 "Vaccination antiparasitaire" Université de Montpellier 1

Examinateurs : **Marie Laure DARDE,** Professeur, CNR/CRB Toxoplasmose, Université de Limoges

Alain CHAUVIN, Professeur, UMR1300 INRA Oniris BioEpAR, Nantes

Directrice de Thèse : **Laurence MALANDRIN,** Chargée de Recherches, UMR1300 INRA Oniris BioEpAR, Nantes

ACKNOWLEDGEMENTS

Professor Isabelle Florent, Muséum National d'Histoire Naturelle, Paris and professor Stéphane Delbecq, Maître de conférences, EA 4558 "Vaccination antiparasitaire" Université de Montpellier 1, who were willing to judge this work and be rapporteurs and members of the jury.

Professors Marie Laure Darde, CNR/CRB Toxoplasmose, Université de Limoges and Alain Chauvin, Oniris-Ecole Nationale Vétérinaire de Nantes, for agreeing to serve on the jury.

It is a great honor for me to express my sincere gratitude and deep appreciation to my supervisor, Dr Laurence Malandrin, INRA, Biology, Epidemiology and Risk Analysis in Animal Health, LUNAM Université, Oniris-National Veterinary School of Nantes, France, for her constant encouragement, patient guidance and instructive advices including those in the experiment design stages and proposal writing on my thesis, she has spent much time reading through each draft and provided me with inspiring advice. She has walked me through all the course of the writing of this thesis. Without her enlightening instruction, my thesis could not have reached its present form.

My deepest gratitude and sincere to Dr. Emmanuelle Moreau, LUNAM Université, Oniris-National Veterinary School of Nantes, for her supervision of this work, I do appreciate her assistance and patient instructions on my thesis and for her help to solve critical and technical problems as well as her careful reading and correcting in the completion of this manuscript.

I feel grateful to Madam Claire Bonsergent, for her help and encouragement on my life, offering me valuable advices and effective training during my study, helping me work out my problems during the difficult course of the thesis, as well as for her a lot of technical guidance and reading and correcting during preparation of my thesis in the past three years.

I would like to express my special thanks to professors Hong Yin and Jianxun Luo, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, their numerous encouragement and unwavering support have sustained me through frustration and depression and the most difficult time. Without their support, the completion of this thesis would also be impossible.

I greatly appreciate Professor Bernard Carcy, Dr. Blandine Lieubeau, Dr. Christine Fourichon and Dr. Catherine Belloc, for their insightful and excellent comments and very valuable suggestions on each year of my studying.

I deeply appreciate Mr. Jordan Marchand, Ms. Charlotte Valentin and Ms. Aurelie Godin for their valuable assistance during the study of *rap-1* genes and locus organization of *Babesia* sp. Xinjiang and evaluation of *rap-1* gene polymorphism among *B. motasi*-like isolates.

Special thanks should go to members of the UMR1300 Biology, Epidemiology and Risk Analysis in Animal Health, the professor Monique L'Hostis, Dr. Olivier Plantard and Albert Agoulon, Mrs. Maggy Jouglin, Suzanne Bastian, Françoise Armand, Agnès Bouju-Albert, Nathalie de la Cotte, Axelle Hermouet, Nadine Brisseau, Catherine Mangion, Sylvie Martin, Nelly Le Martret and Juliette Bordot. My thanks also go to my colleague of PhD student Elsa Quillery for her kind help and computer administrator Erwann Helleu for his technical support.

I am deeply indebted to all the people who are working at Oniris-National Veterinary School of Nantes, which have one way or another, contributed to this work by providing advices and assistance technique.

I wish to extend my special thanks to Dr. Guiquan Guan and Mrs. Aihong Liu for their help on the preparation and sending of genomic DNA to France, and thanks to Mr. Congshan Yang, for his assistance on the performance of RT-PCR for the transcription analysis of *rap-1* genes from *Babesia* sp. Xinjiang in China, as well as thanks to Mrs. Qiaoyun Ren, for her caring on my own affairs.

I also would like to express my appreciation to Mr. Bingyi Lu in the office of scientific research and management, Dr. Youquan Li, Dr. Zhijie Liu, Dr. Junlong Liu, Mr. Jifei Yang, and Mrs. Miling Ma, Dr. Ze Chen in the laboratory of "Tick and Tick-borne disease". My thanks go to all my colleagues at Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Science, Lanzhou, China.

I would like to extend my gratitude to my Chinese friends in France, including Dr. Jingtao Zhang and his wife, Dr. Yi Sun, Dr. Yan Xu and Dr. Yanhua Wen and Ms. Yuan Liu. Without their generous support and help, my study life would be much hard.

I would like to express sincere gratitude to my beloved parents and sister, for their continuous and unselfish love, supporting and caring for me without a word of complaint on all of my life. In addition, I want to express my sincere appreciation to my husband, Rongzeng Hao, who abandoned work in China and traveled many miles to accompany with me through difficult time, and support and encourage me silently rearward for many years. Finally, I am also pleased to extend my thanks to my lovely daughter, Sophie Hao, who is present with me through wonderful time.

CONTENTS

LIST OF FIGURES 3
LIST OF TABLES 6
LIST OF ANNEXES 7
LIST OF ABBREVIATIONS 8

RESUME FRANCAIS-FRENCH SUMMARY 9

INTRODUCTION-OBJECTIFS 45

LITERATURE REVIEW 53
A. Description and classification of *Babesia* species infecting ruminants 54
1. Taxonomic position and main biological features of *Babesia* 54
2. Phylogenetic relationships of *Babesia* spp. infecting ruminant 55
3. *Babesia* spp. infecting sheep and goat 57
3.1. Discovery of ovine *Babesia* species 58
3.2. Biological characteristics 59
3.3. Phylogenetic positions of the small ruminants *Babesia* 64
3.4. Diagnostic methods developed to detect and identify ovine *Babesia* species 66
3.5. Geographical distribution of ovine babesiosis 68
B. Life cycle of *Babesia* sp. and the red blood cell invasion 71
1. Life cycle of *Babesia* sp. 71
2. The merozoite and apical complex structure 73
2.1. Overview of the merozoite structure 73
2.2. Detailed description of organelles involved in the invasion process 74
3. Detailed description of the invasion process 75
3.1 Attachment 76
3.2 Apical re-orientation 76
3.3 Tight junction formation: irreversible attachment 77
3.4 Release of the contents of the apical complex organelles accompanied with invasion 77
3.5 Internalization 77

4. Roles of the apical complex organelles in the invasion process of erythrocytes ·· 78

5. Proteins involved in the invasion and their function ···· 79

5.1 Surface-coating molecules of extracellular merozoites: the family of variable merozoite surface antigens (VMSAs) ···· 80

5.2 Organelles-secreting molecules ···· 81

C. Control of babesiosis ···· 90

1. Vector control ···· 90

1.1. Chemical acaricides ···· 90

1.2. Biological control of ticks ···· 91

1.3. Vaccines against ticks ···· 91

2. *Babesia* control ···· 92

2.1. Anti-*Babesia* chemotherapeutic drugs ···· 92

2.2. Vaccines ···· 93

PUBLICATIONS ···· 97

Article N° 1: Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 (Lintan) (*B. motasi* phylogenetic group) ···· 99

Article N° 2: Strong conservation of rhoptry-associated-protein-1 (RAP-1) locus organization and sequence among *Babesia* isolates infecting sheep from China (*Babesia motasi*-like phylogenetic group) ···· 125

Article N° 3: Expression of *rap-1* genes in *Babesia* sp. BQ1 (Lintan) (*B. motasi*-like phylogenetic group), a pathogen of sheep in China ···· 179

Article N° 4: Rhoptry-Associated-Protein (*rap-1*) genes in the sheep pathogen *Babesia* sp. Xinjiang: multiple copies differing by 3' end repeated sequences ···· 211

GENERAL DISCUSSION ···· 245

CONCLUSIONS-PERSPECTIVES ···· 261

REFERENCES ···· 265

ANNEXES ···· 293

LIST OF FIGURES

Résumé français – French summary

Figure A. Hypothétique amplification de tandem arrangées tête-queue de gènes avec des amorces conçues dans la région centrale de *rap-1* gène.

Figure B. Alignement de la région centrale de 5 clones avec la taille de 250 pb, trois séquences représentatives ont été sélectionnées pour concevoir des amorces spécifiques (type1: clone 61, type 2: clone 67 et type 3: c: clone 86).

Figure C. Hypothétique *rap-1* locus organisation de *Babesia* sp. BQ1 (Lintan).

Figure D. Alignement protéique de le premier 300 acides aminés séquences de RAP-1 montrant les 4 cystéines conservées ainsi que les séquences conservées utilisées (deux boîte) pour dessiner des amorces dégénérées et amplifier une partie interne du gène.

Figure E: Méthodologie utilisée pour l'amplification des gènes *rap-1*.

Figure F. Alignement de la région de en 3' les *rap-1aα* et *rap-1aβ* gènes, il est indiqué 28 substitutions.

Figure G. L'amplification de la région variable de *rap-1aα* et *rap-1aβ* types de gènes.

Figure H. Alignement de 13 séquences nucléotidiques.

Figure I. Conservation du *rap-1a* régulation de la transcription de la séquence 5'UTR putatif dans les régions intergéniques *rap-1a*.

Figure J. Prédiction d'épitopes des épitopes B putatifs de protéines RAP-1aα2 et RAP-1aβ1 en utilisant la méthode "de prédiction linéaire Bepipred épitopes". Épitopes prédits B ont été indiqués en jaune.

Introduction–Literature review–Discussion

Figure 1. Phylogenetic analysis of *Babesia* spp. of ruminants (domestic and wild) based on the sequences of the 18S rRNA gene.

Figure 2. Morphological characteristics of different species of *Babesia* infecting small ruminants in China, Europe and Iran.

Figure 3. Phylogenetic analysis of *Babesia* species identified of the small ruminant based on the sequences of the 18S rRNA gene.

Figure 4. Geographical distribution of small ruminant *Babesia* spp. in China.

Figure 5. The development life Cycle of *Babesia*.

Figure 6. Cellular structure of a merozoite and major organelles of the apical complex of *Plasmodium falciparum* .

Figure 7. Schematic representation of erythrocyte invasion process: example of *Babesia bovis*

Figure 8. The role of apical complex organelles, secretion of the contents of the these organelles and invasion of erythrocyte by the merozoite of Apicomplexa.

Figure 9. Proteins involved in erythrocyte invasion and characterized in the genus *Babesia*.

Figure 10. Structure of the *rap-1* locus in *B. bovis*.

Figure 11. Structure of the *rap-1* locus in *B. bigemina* .

Figure 12. An alignment of RAP-1a amino acid sequences of several *Babesia* species showing the conserved features of members of this family.

Figure 13. Phylogenetic analysis based on the 18S rRNA gene, including 18s rRNA sequences of ovine *babesia* representatives and other *Babesia* species having *rap-1* sequences.

Figure 14. Phylogenetic trees of several *Babesia* isolates based on their 28S rRNA genes. This tree was calculated using the BI and MP methods.

Figure 15. Phylogenetic tree depicting the relationship among several *B. motasi*-like Chinese isolates and other *Babesia* species based on ITS region sequences.

Figure 16. Conservation of the putative *rap-1* transcription regulation 5'UTR sequence in the *rap-1* intergenic regions (IG1: *rap-1a* to *b*; IG2: *b* to *a*; IG3: *a* to *c*)

LIST OF TABLES

Résumé français – French summary

Table A. Disposition des répétitions dans la région variable de *rap-1a* copies du gène dans *Babesia* sp. Xinjiang.

Table B. Séquence de répétitions de nucléotides et le niveau d'acide aminé.

Introduction–Literature review–Discussion

Table 1. The traditional classification of *Babesia* spp.

Table 2. *Babesia* species of small ruminant in the world and isolates of *Babesia* infecting sheep and goats in China.

Table 2. History of *Babesia* species of sheep and goats in China.

Table 4. Proteins involved in at different invasion stages of *Babesia* species.

Table 5 : Protein of *Babesia* sp as vaccine candidate.

Table 6. The summary of *rap-1* genes, gene numbers, intergenic regions and locus in two different sheep *Babesia* species in China.

Table 7. Repeats in *rap-1* genes from different *Babesia* species.

Table 8. Percentage identity of RAP-1a amino acid sequences of *B. motasi*-like, *Babesia* sp. Xinjiang and *B. ovis* deduced after CLUSTAL W alignment.

LIST OF ANNEXES

Annexe 1. List of publications of Qingli Niu

Annexe 2. Article 1: **Niu Q**, Luo J, Guan G, Ma M, Liu Z, Liu A, Dang Z, Gao J, Ren Q, Li Y, Liu J, Yin H 2009. Detection and differentiation of ovine *Theileria* and *Babesia* by reverse line blotting in China. Parasitol Res. 104:1417-23.

Annexe 3. Article 2: **Niu Q**, Luo J, Guan G, Liu Z, Ma M, Liu A, Gao J, Ren Q, Li Y, Qiu J, Yin H. 2009. Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. Exp Parasitol. 121:64-68.

Annexe 4. Article 3: **Niu Q**, Guan G, Liu Z, Ma M, Li Y, Liu A, Ren Q, Liu J, Luo J, Yin H. 2012. Simultaneous detection of piroplasma infections in field *Haemaphysalis qinghaiensis* ticks by reverse line blotting. Exp Appl Acarol. 56:123-132.

LIST OF ABBREVIATIONS

ADN=DNA	Acide desoxyribonucleique	NO	Nitric oxide
AMA-1	Apical Membrane Antigen-1	NX	Ningxian
ARNr18S=18 S rRNA	Acide ribonucleique ribosomal 18s	OD	Optical density
AA	Amino acid	OIE	World Organization for Animal Health
BboRhop68	*B. bovis* rhoptry protein	pb=bp	base pair
BQ1MA	Merozoite antigens of *Babesia* sp. BQ1 (Lintan)	PCR	Polymerase chain reaction
Bov57	*B. bovis* syntenic gene	PV	parasitophorous vacuole
Bovipain-2	papain-like cysteine protease-2	PVM	parasitophorous vacuole membrane
BYC	Yolk pro-Cathepsin	RAP-1	Rhoptry-associated-protein-1
CAAS	Chinese Academy of Agricultural Science	RBCs	Red Blood Cells
cDNA	Complementary Acide desoxyribonucleique	RLB	Reverse line blot
CFT	Complement Fixation Test	rpm	Round per minute
ELISA	Enzyme Linked Immunosorbent Assay	RRA	RAP-1 related antigen
FBS	Fetal Bovine Serum	RT-PCR	Reverse Transcription PCR
GPI	glycosyl-phosphatidylinositol	SBPs	Spherical Body Proteins
HB	Hebei	SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
IFAT	Inidrect Fluorescent Antibody Test	SmORFs	Small open reading frames
IFNg	Gamma interferon	SUB	Subolesin
IG	Intergenic region	TBS	Tris buffered saline
ITS	Internal Transcribed Spacer	TBST	TBS tween
kDa	Kilo-dalton	TMB	3,3’5,5’ Tetramethylbenzidiene
LAMP	Loop-mediated isothermal amplification	TRAP	Thrombospondin-related anonymous protein
LB	Luria-Bertani	TSP1	thrombospondin type 1
LT	Lintan	TZ	Tianzhu
LVRI	Lanzhou Veterinary Research Institute	VESA	variant erythrocyte surface antigen
MEGA	Molecular Evolutionary Genetics Analysis	VMSA s	Variable merozoite surface antigens
MIC-1	Micronemal proteins	VTDCE	vitellin-degrading cysteine endopeptidase
NC	nitrocellulose	VVBD	Vector and vector Borne disease
nt	nucleotide	vWFA	vonWillebrand factor A domain

RESUME FRANCAIS
FRENCH SUMMARY

Introduction

Les parasites du genre *Babesia* sont des Piroplasmidés et font donc partie du phylum des Apicomplexes. Ces parasites ont deux hôtes: l'hôte vertébré chez qui se déroule une phase de reproduction asexuée, et la tique, hôte définitif, chez qui se déroulent une phase de reproduction sexuée puis asexuée. Ces parasites obligatoires se multiplient chez l'hôte vertébré exclusivement dans les hématies, contrairement à *Plasmodium* dont les premières étapes du développement chez le vertébré ont lieu dans les hépatocytes, et à *Theileria*, avec un développement au niveau des lymphocytes. C'est la multiplication rapide de *Babesia* dans les érythrocytes qui provoque les symptômes bien connus de la babésiose : anémie, fièvre, hémoglobinurie, ictère.

Les babésioses sont décrites chez de très nombreux vertébrés, sauvages et domestiques, et quelques espèces sont zoonotiques. Dans les régions principalement tropicales et subtropicales, la babésiose peut engendrer de graves pertes économiques, notamment au niveau du bétail.

Plusieurs espèces de *Babesia* sont responsables de la babésiose ovine, avec des niveaux de virulence variables et des répartitions géographiques différentes selon les espèces. Parmi les espèces connues de longue date, on peut citer *B. ovis*, *B. motasi* et *B. crassa*. Récemment, en Chine, plusieurs autres parasites pathogènes des ovins ont été décrits. Ils se classent dans deux groupes phylogénétiques. L'un comporte pour l'instant uniquement des parasites récemment décrits dans le monde entier chez les ruminants sauvages (*B. pecorum* chez le cerf en Espagne et plusieurs autres en Afrique du Sud chez la girafe notamment) et chez le mouton en Chine (*Babesia* sp. Xinjiang). Ce groupe phylogénétique forme une clade soeur avec celle d'autres pathogènes des ruminants, *B. ovis* chez les ovins et *B. bovis*, parasite des bovins. L'autre regroupe plusieurs parasites décrits chez le mouton et/ou la chêvre en Chine : *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ninxiang), *Babesia* sp. Hebei, *Babesia* sp. Liaoning, *Babesia* sp. Tianzhu et *Babesia* sp. Madang. Ces parasites forment une clade soeur à celle de *B. motasi* décrit en Europe, clade appelée *B. motasi*-like, en attendant que des critères plus précis permettent de classer ces parasites en tant qu'espèce ou genre, différent ou non de *B. motasi*.

La multiplication de *Babesia* chez son hôte vertébré passe par des cycles successifs de développement et de division asexuée dans le globule rouge qui est finalement lysé, et de

recherche puis d'invasion de nouvelles cellules cibles. Au cours de cette dernière étape, décrite comme très rapide, le parasite est accessible directement au système immunitaire de son hôte vertébré et donc très vulnérable. C'est donc cette étape qui est principalement visée lors de la mise au point de vaccins susceptibles de bloquer la multiplication du parasite. Deux catégories de molécules sont principalement visées dans ces travaux : les protéines de surface du parasite qui servent à son attachement non spécifique à la surface du globule rouge, et les protéines qui lui permettent de pénétrer dans le globule rouge. Chez les Apicomplexes dont *Babesia* fait partie, ces dernières sont localisées au niveau du complexe apical, ensemble d'organelles (micronèmes, rhoptries et granules denses) dont les contenus sont sécrétés séquentiellement au cours du processus d'invasion. Si de nombreuses protéines localisées dans ces différents organelles sont connues maintenant chez *Plasmodium*, les travaux chez *Babesia* sont beaucoup moins avancés.

RAP-1, pour Rhoptry-Associtaed-Protein 1, est l'une des protéines dont le rôle dans l'invasion du globule rouge par *Babesia* est étudié. Cette protéine, caractérisée chez toutes les espèces de *Babesia* étudiées jusqu'ici, fait partie d'une famille multigénique dont les membres partagent des caractéristiques communes: organisation des gènes en tandem tête-queue, 4 cystéines très conservées et quelques motifs protéiques également très conservés.

Mon projet de thèse consiste à étudier cette famille multigénique (séquence, organisation et expression) chez les *Babesia* récemment décrits en Chine comme pathogènes du mouton. L'ensemble de mon doctorat rentre dans les objectifs du projet PIROVAC.

Le projet PIROVAC est un projet européen du 7ème PCRD qui regroupe 16 partenaires et dont l'objectif est d'améliorer la vaccination contre les théleirioses et les babésioses des petits ruminants (mouton et chèvre). Plusieurs voies sont suivies, allant de l'amélioration des vaccins vivants actuels, à la recherche de candidats vaccins pour la production de vaccins recombinants multivalents, en passant par l'étude de la réponse immunitaire des animaux infectés.

Dans ce projet, ma thèse s'inscrit dans la partie «recherche de candidats vaccins pour la production de vaccins recombinants multivalents».

Les travaux de recherche qui sont présentés dans ce document sont organisés sous forme de 4 articles, précédés d'une synthèse bibliographique et suivis d'un discussion/conclusion générales. Le choix a été fait de présenter à la fin de chaque article les données supplémentaires accessibles en ligne uniquement, et d'intégrer les figures dans

le corps du texte pour les articles soumis ou en cours de rédaction, afin de faciliter la lecture du document. Un résumé détaillé en Français de chacun des articles est présenté ci-après.

Article 1

Le locus *rap-1* chez *Babesia* sp. BQ1 (Lintan) : de nombreuses homologies avec celui décrit chez *B. bigemina*, parasite des bovins

Résumé détaillé de l'article

Niu Q, Bonsergent C, Guan G, Yin H, Malandrin L 2013. Sequence and organization of the rhoptry-associated-protein-1 (rap-1) locus for the sheep hemoprotozoan Babesia sp. BQ1 (Lintan) (B. motasi phylogenetic group). Vet Parasitol. 2013 Nov 15;198(1-2):24-38.

Pour ce travail, QN a réalisé les amplifications, clonages et séquençages; CB a formé puis encadré QN au niveau expérimental; GG et HY ont fourni la souche; QN, CB et LM ont rédigé le manuscrit; LM a supervisé l'ensemble des travaux.

Introduction

Babesia sp. BQ1 (Lintan) a été décrit chez le mouton en Chine pour la première fois en 1997 (Yin *et al.*, 1997a), en même temps que plusieurs autres parasites proches. Des études moléculaires à l'aide de différents marqueurs ont placé ces isolats dans le même groupe, phylogénétiquement proche de *B. motasi*, autre parasite du mouton décrit en Europe (Liu *et al.*, 2007; Niu *et al.*, 2009a; Gou *et al.*, 2013). Nous appellerons dans notre étude ce groupe d'isolats chinois pathogènes des petits ruminants *B. motasi*-like.

La transmission vectorielle de *Babesia* sp. BQ1 (Lintan) par les tiques *Haemaphysalis qinghaiensis* et *Haemaphysalis longicornis* a été démontrée récemment (Guan *et al.*, 2010a). *Babesia* sp. BQ1 (Lintan) est peu pathogène pour le mouton et la chêvre (Guan *et al.*, 2002), et ce parasite a été probablement confondu dans les premiers travaux sur les parasites des petits ruminants avec *B. ovis* (Chen, 1982; Zhao *et al.*, 1986). Grâce à la mise en place de la culture in vitro de ce parasite, un test sérologique a été développé à partir d'antigènes de mérozoites (Guan *et al.*, 2010b et c) et utilisé pour étudié la prévalence et la distribution de ce pathogène en Chine. Sa présence possible dans 22 provinces de Chine a ainsi été mise en évidence dans une étude séro-épidémiologique récente (Wang JM *et al.*, 2013). Toutefois, en raison de réactions sérologiques croisées fortes avec *Babesia* sp.

Tianzhu (Guan *et al.*, 2010c), un autre parasite du groupe *B. motasi*-like, on ne peut pas clairement attribuer cette distribution géographique et cette forte prévalence dans certaines provinces de Chine à *Babesia* sp. BQ1 seul.

Grâce à la culture in vitro, la sensibilité des moutons à ce parasite a été étudiée, montrant une sensibilité plus forte des races chinoises en comparaison avec les races françaises (Guan *et al.*, 2010b).

L'objectif de ce travail a été d'identifier les gènes *rap-1* chez *Babesia* sp. BQ1(Lintan).

Méthodologie

Au départ de ce travail, nous ne disposions d'aucunes données sur les gènes *rap-1* chez cette espèce. Je suis donc partie des travaux de Dalrymple *et al.* (1996), qui se sont servis de régions conservées dans les séquences des gènes *rap-1* alors connues pour définir des amorces permettant l'amplification d'une partie centrale d'environ 250 bp du gène. Disposant de séquences supplémentaires par rapport à l'étude de l'époque, nous avons redéfini d'autres amorces en nous basant plus précisément sur les gènes *rap-1* décrits chez les ruminants. L'utilisation de ces amorces a permis l'amplification d'un produit de taille attendue, autour de 250 pb. Une des caractéristiques de la famille rap étant son caractère multigénique, nous pouvions nous attendre à la présence d'un mélange de différents produits d'amplification. L'amplicon a donc été cloné et les différents clones ayant un insert de taille attendue envoyés pour séquençage. Sur les 5 inserts, le séquençage a mis en évidence trois types de séquences. Des amorces spécifiques de chacune des trois séquences ont été définies. Une deuxième caractéristique de la famille *rap* étant la présence de copies multiples en tandem arrangées tête-queue, ces 6 amorces ont été utilisées dans différentes combinaisons pour amplifier les séquences en amont et aval de la région centrale, entre deux gènes potentiellement successifs (voir figure explicative A). Des amplicons de tailles variables, allant jusqu'à environ 5Kb ont été obtenus, taille qui pouvait correspondre aux extrémités 3' et 5' de deux gènes, à deux régions intergéniques et un autre gène intercalé. Les différents amplicons ont été clonés pour les raisons expliquées plus haut, et les inserts de plusieurs clones séquencés dans chaque cas. Les séquences des gènes complets ont ensuite été obtenues après avoir défini de nouvelles amorces au niveau des sites start et

stop putatifs des protéines. Là encore, le polymorphisme de chaque type de gène et de chaque région intergénique a été étudié par séquençage des inserts à partir de multiples clones. Le ratio des différents types de gènes a été évalué grossièrement par une digestion des produits d'amplification de multiples clones, permettant de différencier les différents copies possibles. L'existence en aval du locus *rap* d'une même ORF chez *B. bigemina* et *B. bovis*, nous a permis de séquencer l'extrémité 3' du locus *rap* chez *Babesia* sp. BQ1 (Lintan).

Figure A. Hypothétique amplification de tandem arrangées tête-queue de gènes avec des amorces conçues dans la région centrale de *rap-1* gène. Deux région conservée utilisé pour concevoir des amorces dégénérées ont été indiquées avec des boîtes roses et bleus.

Résultats

<u>Présence chez *Babesia* sp. BQ1 (Lintan) d'au moins trois gènes *rap-1* différents</u>

L'amplification d'une région de 250 pb à l'aide des amorces dégénérées correspondant aux motifs conservés de la protéine a été obtenue. Le séquençage des inserts de cette taille pour 5 clones a révélé 3 séquences différentes (figure B). Sur cette petite portion de gène, deux des séquences (clones 61 et 67) ont été identifiées par analyse BLAST comme correspondantes à deux copies différentes du gène *rap-1a*. La troisième (clone 86) s'est

avérée correspondre à un orthologue putatif du gène *rap-1c*, type n'ayant été jusqu'ici mis en évidence que chez *B. bigemina*.

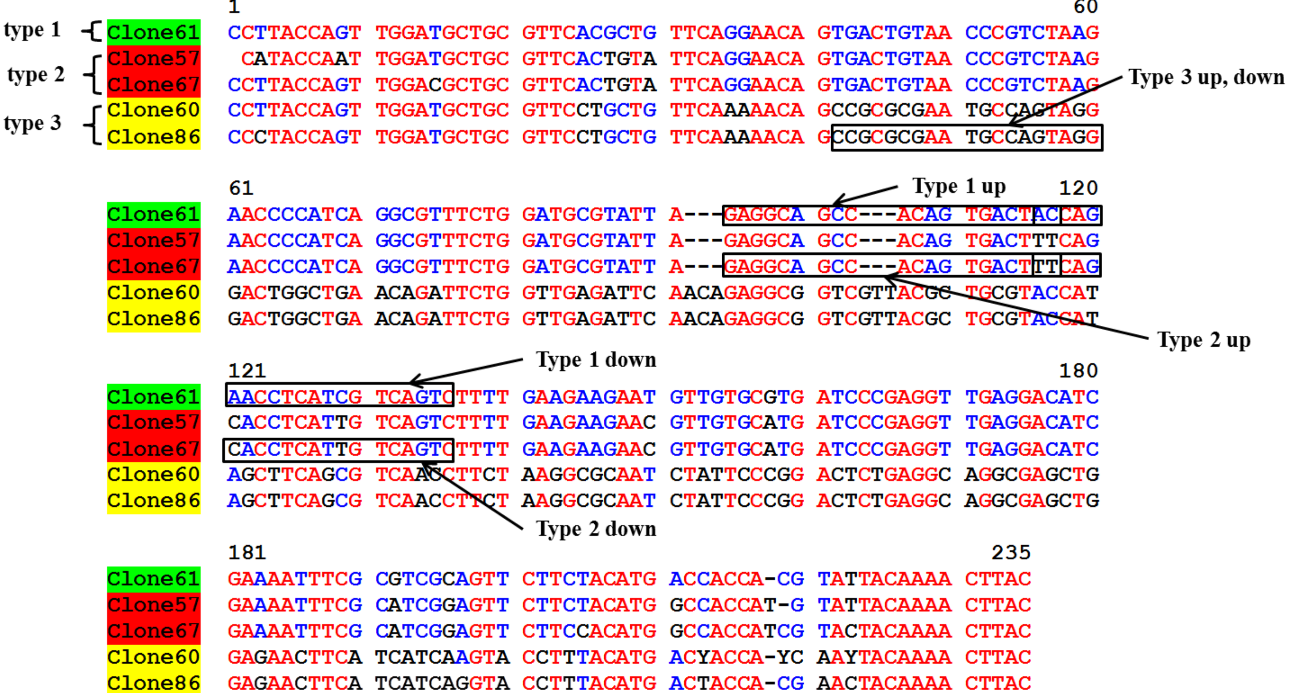


Figure B. Alignement de la région centrale de 5 clones avec la taille de 250 pb, trois séquences représentatives ont été sélectionnées pour concevoir des amorces spécifiques (type1: clone 61, type 2: clone 67 et type 3: c: clone 86). Les trois séquences de cinq clones ont été indiqués par différente couleur, corronsponding à la figure A.

<u>Des orthologues des gènes *rap1a*, *rap-1b* et *rap-1c* décrits chez *B. bigemina* existent aussi chez *Babesia* sp. BQ1 (Lintan)</u>

De nouvelles amorces spécifiques (6) de chacun des gènes ci-dessus ont été définies et utilisées dans différentes combinaisons. Les régions localisées entre les trois gènes caractérisés ci-dessus ont ainsi été étudiées et ont permis de mettre en évidence un autre type de gène chez cette espèce, *rap-1b*, qui comme pour *rap-1c*, n'avait été jusqu'ici mis en évidence que chez *B. bigemina*. Il existe donc chez *Babesia* sp. BQ1(Lintan), trois types de gènes *rap1a*, *rap-1b* et *rap-1c*.

<u>Les gènes *rap-1b* sont intercalés entre les gènes *rap-1a*</u>

La présence des gènes *rap-1b* intercalés entre les gènes *rap-1a* a été démontrée avec deux organisations possibles : 61-b-67 et a-b-61 (l'identité des séquences en 3' des gènes *rap-1a* n'a pas permis de déterminer s'il s'agissait de 61 ou 67). L'amplification/clonage/séquençage du produit b-a-b a mis en évidence l'absence de gènes

de type *rap-1a67* entre deux gènes *rap-1b*. Les gènes *rap-1a* et *rap-1b* sont donc intercalés, dans un ordre qu'il reste à confirmer quant aux gènes *rap-1a*.

Les multiples copies du gène *rap-1b* sont identiques, tandis que celles du gène *rap-1a* sont polymorphes

Les séquences obtenues au cours des étapes préalables ont permis de définir des amorces permettant l'amplification des gènes putatifs *rap-1a* et *rap-1b* complets. L'analyse de leur polymorphisme a donc été entrepris. Les différentes copies du gène *rap-1b* se sont révélées totalement identiques. Par contre, le polymorphisme de *rap-1a*, déjà évident d'après les résultats précédents, a été étudié de façon plus approfondie.

Le gène de type *rap-1a61* existe en fait sous forme de deux séquences différentes au niveau de seulement 3 mutations ponctuelles. Le gène de type *rap-1a67* existe sous forme de deux séquences différentes uniquement au niveau de la région 3'. Le gène ayant la séquence la plus courte (*rap-1a67-2*) est retrouvé uniquement en amont du gène *rap-1c*, tandis que le gène *rap-1a67-1* est présent en amont de *rap-1b*.

Le polymorphisme entre les gènes *rap-1a61* et *rap-1a67* est du type substitutions et est restreint à une région du gène entre 192 et 321nt après le site start putatif. Malgré ce polymorphisme, les 4 cystéines localisées dans cette portion sont conservées.

Le gène *rap-1c* est en copie unique et présent à l'extrémité 3' du locus *rap*

Les différentes approches réalisées permettent de conclure que le gène *rap-1c* est présent en copie unique dans le locus et localisé à l'extrémité 3' du locus:

–aucun polymorphisme entre les différents inserts *rap-1c* séquencés,

–pas de gène *rap-1c* entre 2 gènes *rap-1b* ou entre 2 gènes *rap-1a*,

–aucune amplification jamais obtenue pour c-a ou c-b,

–présence de l'ORF YJR070 en aval de *rap-1c*, comme chez *B. bigemina* et *B. bovis*.

Dans sa région 3', le gène *rap-1c* présente 5 répétitions conservées de 51 à 54 nt. Cette caractéristique est fréquente dans la famille *rap.*

Présence de 3 régions intergéniques différentes

Les différentes amplifications réalisées entre gènes en tandem ont permis de séquencer les régions intergéniques. Elles sont différentes en taille et en séquence entre les différents types de gènes: IG1 (853 pb) est localisée entre *rap-1a* et *rap-1b*, IG2 (1840 pb)

entre *rap-1b* et *rap-1a*, et IG3 (1333 pb) entre *rap-1a67-2* et *rap-1c*.

En dehors d'une substitution pour IG1, les différentes copies des régions intergéniques sont conservées.

Organisation générale du locus *rap* chez *Babesia* sp. BQ1 (Lintan)

Les différentes amplifications/séquençage des gènes en tandem ont permis d'aboutir aux conclusion suivantes:

–*rap-1a* et *rap-1b* intercalés,

–*rap-1a67* jamais localisé entre deux *rap-1b*, uniquement *rap-1a61* entre 2 *rap-1b*,

–*rap-1a67-2* en amont de *rap-1c*, et en aval d'un gène *rap-1b,*

–*rap-1c* à l'extrémité 3' du locus.

L'analyse RFLP des inserts de *rap-1a* a permis de déterminer la présence d'un nombre probablement équivalent de *rap-1a67*, de *rap-1a61-1* et de *rap-1a61-2* au niveau du locus. Comme la présence de deux gènes *rap-1a67* minimum est prouvée (*rap-1a67-1* et *rap-1a67-2*), alors 4 gènes *rap-1a61* sont présents potentiellement. Une organisation probable du locus *rap* chez *Babesia* sp. BQ1 (Lintan), remplissant les différentes conditions ci-dessus, est proposée (figure C).

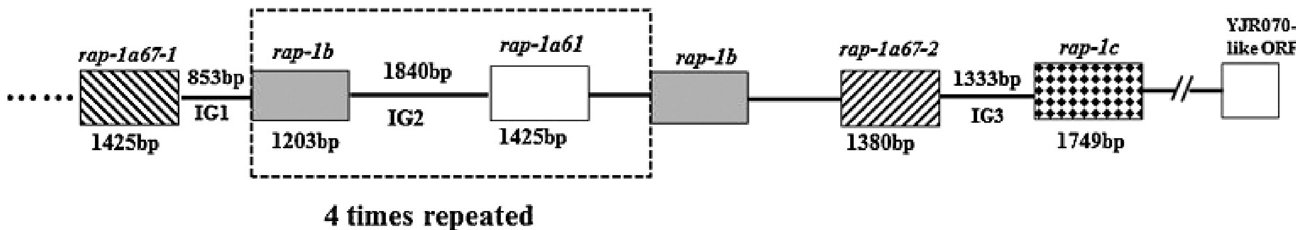


Figure C. Hypothétique *rap-1* locus organisation de *Babesia* sp. BQ1 (Lintan). 4 fois répétées partielles locus sont indiqués avec cadre en pointillés. (figure 8 de article N°1)

Lors de cette partie de mon étude, j'ai séquencé partie par partie un locus dont la taille est d'environ 31 kB. Le séquençage précis de telles régions du génome comportant des séquences répétées identiques est compliqué, et lors de séquençages de génomes complets, ces régions gênent souvent l'assemblage en raison de leur caractère répété.

Il est apparu assez rapidement que chez ce parasite, le locus *rap-1* serait assez complexe, et comparable à celui de *B. bigemina*, puisque des gènes étaient retrouvés (*rap-*

1b et *rap-1c*), jusqu'alors mis en évidence uniquement chez *B. bigemina* (Suarez *et al.*, 2003). L'organisation générale du locus est bien conservée entre ces deux espèces, sauf au niveau du gène précédant *rap-1c*, qui est de type *rap-1a* chez *Babesia* sp. BQ1 (Lintan), et de type *rap-1b* chez *B. bigemina*. Cette différence génère la présence d'une troisième région intergénique (IG3) différente des deux autres, uniquement chez la souche chinoise.

La taille des gènes et des régions intergéniques est également conservée entre ces deux espèces. La conservation des copies du gène *rap-1b* et des régions intergéniques et le polymorphisme (parfois très limité entre certaines copies) uniquement au niveau de *rap-1a,* font également partie des nombreux critères de ressemblance.

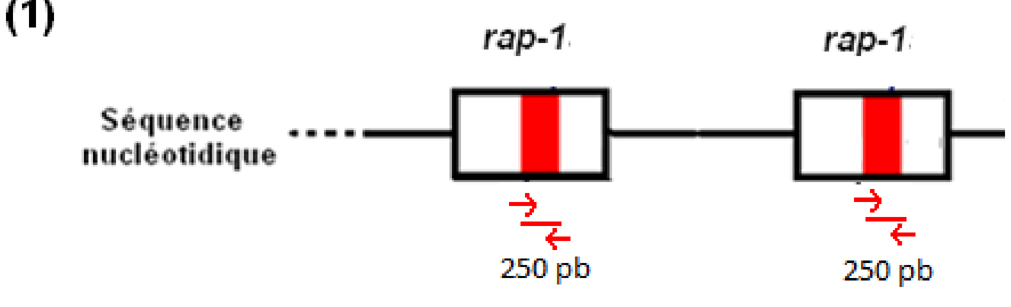
Babesia sp. BQ1 (Lintan) et *B. bigemina* font partie d'un même groupe phylogénétique, qui rassemble aussi *B. crassa*, *B. major*, *B.motasi* et *B. ovata* (Schnittger *et al.*, 2012; Lack *et al.*, 2012). La conservation du locus avec les gènes spécifiques *rap-1b* et *rap-1c* est peut être une caractéristique de ce groupe phylogénétique. Si le rôle de RAP-1a dans l'invasion du globule rouge par le parasite a été démontré au moins chez *B. bovis* (Mosqueda *et al.*, 2002a; Yokoyama *et al.*, 2002), le rôle de RAP-1b et RAP-1c est loin d'être évident. En effet, si la transcription des ces deux gènes a été démontrée, leur traduction n'a jusqu'ici pas été mise en évidence (Suarez *et al.*, 2003). La conservation d'orthologues de ces gènes chez d'autres espèces de *Babesia*, avec une organisation du locus similaire (multicopie de *rap-1b*, *rap-1c* en fin de locus) laissent à penser qu'il ne s'agit pas de pseudogènes, qui auraient disparu ou tout du moins qui n'auraient pas conservé les caractéristiques fondamentales de cette famille (4 cystéines, motifs conservés). Leur absence chez d'autres espèces est surprenante, car le mécanisme et donc les protéines nécessaires au processus d'invasion du globule rouge sont probablement similaires. La protéine RRA, quoique différente des autres RAP, présente les caractéristiques de la famille RAP, et pourrait correspondre à un homologue fonctionnel de RAP-1b (Suarez *et al.*, 2011). Cependant, le gène *rra* n'est pas localisé à proximité de *rap1-a* chez *B. bovis*, contrairement à *rap-1b* chez *B. bigemina.*

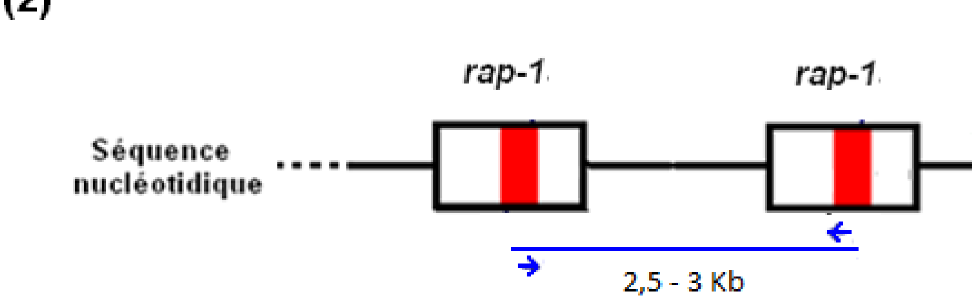
Comme chez *B. bigemina*, seul le gène *rap-1a* est polymorphe chez *Babesia* sp. BQ1 (Lintan). Ce polymorphisme entre *rap-1a61* et *rap-1a67* est dans une région de la protéine située entre les aa 64 et 107, ce qui correspond à la zone de polymorphisme observée entre les copies du gène *rap-1a* chez *B. bigemina* (aa 68 à 87) (Hötzel *et al.*, 1997). C'est étonnamment dans cette région très polymorphe que les 4 cystéines conservées sont présentes. Sur les 26 aa dans la région englobant les 4 cystéines, seuls 11 sont conservés, dont les 4 cystéines. Elles sont supposées avoir un rôle dans la structure de la protéine, en

établissant des ponts disulfures.

Le processus d'invasion du globule par les *Babesia* est loin d'être élucidé, et très peu de gènes et de protéines impliquées sont caractérisés au niveau moléculaire et encore moins au niveau fonctionnel, si on compare avec *Plasmodium* (Lobo *et al*., 2012). Nos travaux futurs vont s'orienter vers l'analyse du polymorphisme de ce locus et des gènes s'y trouvant au niveau des isolats du groupe *B. motasi*-like présents chez les petits ruminants en Chine. En parallèle, l'expression et le rôle des protéines RAP chez *Babesia* sp. BQ1 (Lintan) sera mené. Ces travaux conjoints devraient permettre de cibler dans cette famille le meilleur candidat vaccin fin de protéger les ovins contre les différents membres du groupe *B. motasi*-like.

Figure D: Alignement protéique de le premier 300 acides aminés séquences de RAP-1 montrant les 4 cystéines conservées ainsi que les séquences conservées utilisées (deux boîte) pour dessiner des amorces dégénérées et amplifier une partie interne du gène. *B. bovis* (GenBank accession number: AAB84267), *B. ovis* (GenBank accession number: AAA27805), *B. bigemina* (GenBank accession number: A45614), *B. divergens* (GenBank accession number: CAA89970).(Supplementary figure 1 de article N°1)





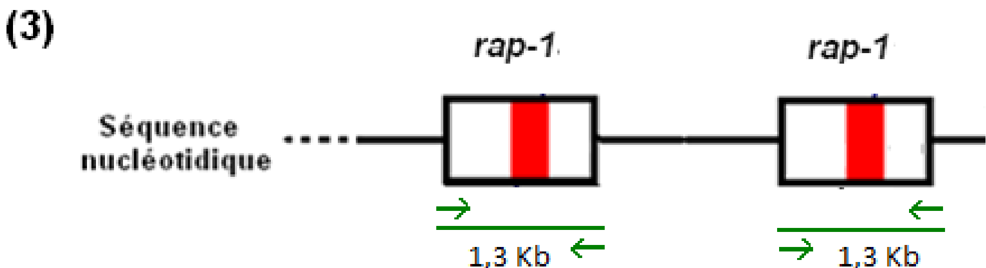


Figure E: Méthodologie utilisée pour l'amplification des gènes *rap-1*. (1) Amplification d'une région interne de 250 pb du gène à l'aide d'amorces dégénérées situées au niveau des motifs conservés décrits figure D. (2) Dessin d'amorces spécifiques à partir des séquences obtenues en 1 et amplification de l'extrémité 3', de la région intergénique et de l'extrémité 5' des gènes arrangés en tandem. (3) Amplification sélective de chaque gène à l'aide d'amorces spécifiques définies selon les séquences obtenues en (2).

Article 2

Caractérisation moléculaire du locus et des gènes *rap* chez trois membres du groupe *B. motasi*-like infectant les moutons en Chine

Résumé détaillé de l'article
Niu Q, Valentin C, Bonsergent C, Malandrin L. Strong conservation of rhoptry-associated-protein-1 (RAP-1) locus organization and sequence among Babesia isolates infecting sheep from China (Babesia motasi-like phylogenetic group).

Soumis à Parasitology Research le 12 Mars 2014.

Pour ce travail, QN et CV ont réalisé les amplifications, clonages et séquençages; CB a formé puis encadré QN et CV au niveau expérimental; QN et LM ont rédigé le manuscrit; LM a supervisé l'ensemble des travaux.

Introduction

A la fin des années 90, plusieurs agents responsables de babésiose chez les petits ruminants en Chine ont été décrits : *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Hebei, *Babesia* sp. Tianzhu, *Babesia* sp. Madang, *Babesia* sp. Liaoning and *Babesia* sp. Xinjiang (Yin *et al*., 1997a). Depuis, ces agents ont été plus ou moins étudiés et caractérisés, au niveau moléculaire et phylogénétique, comme au niveau fonctionnel. Certains ont été isolés, pour d'autres, le vecteur a pu être défini, et enfin différents tests diagnostics, sérologiques comme moléculaires ont été mis au point.

Rapidement, la séparation de ces isolats en deux groupes phylogénétiquement distincts est apparue (Liu *et al*., 2007, Niu *et al*., 2009a). La plupart des agents ont été placé dans un même groupe phylogénétique, proche de celui représenté par *B. motasi*, agent de babésiose ovine en Europe. Nous appellerons dans notre étude ce groupe d'agents pathogènes des petits ruminants en Chine *B. motasi*-like. Phylogénétiquement, seul *Babesia* sp. Xinjiang s'est retrouvé en dehors de cette clade, se trouvant plus proche phylogénétiquement de *B. ovis*, sans pourtant faire partie de cette espèce. Depuis, d'autres parasites ont rejoint ce groupe phylogénétique, provenant de girafe et d'antilope en Afrique du sud (Oosthuizen *et al*., 2009), ou isolés de cerfs en Espagne avec l'espèce *B. pecorum* (Jouglin et al., 2014, in

press).

Même s'il est clairement établi que les 6 agents de babésiose ovine précédemment énumérés font partie du groupe phylogénétique *B. motasi*-like, leur statut en tant qu'espèces distinctes de *B. motasi* et appartenance à une même autre espèce est toujours à définir. En effet, selon les marqueurs phylogénétiques utilisés, ces différents agents se séparent ou non dans différentes clades soeur (Niu *et al.*, 2009a; Gou *et al.*, 2013; Tian *et al.*, 2013a, b et c).

L'objectif de ce travail a été d'étudier le polymorphisme des gènes *rap-1* chez 3 autres membres du groupe *B. motasi*-like (*Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu et *Babesia* sp. Hebei) avec toujours comme objectif à termes de développer un vaccin recombinant contre les babésioses ovines.

Méthodologie

Comme pour Babesia sp. BQ1(Lintan), nous ne disposions, au départ de ce travail, d'aucunes données sur les gènes *rap-1* chez ces 3 agents. Toutefois, la connaissance du locus *rap* chez appartenant au même groupe phylogénétique a été utilisée.

Des amorces basées sur les séquences des gènes *rap-1a*, *rap-1b* et *rap-1c* déterminées pour *Babesia* sp. BQ1 ont été utilisées pour amplifier ces 3 gènes en entier. L'utilisation de ces amorces a permis l'amplification de produits de taille attendue pour ces 3 gènes chez deux des agents étudiés, *Babesia* sp. Tianzhu et *Babesia* sp. BQ1 (Ningxian), mais pas pour *Babesia* sp. Hebei.

Pour *Babesia* sp. Hebei, dont le polymorphisme semblait donc d'ores et déjà plus important, une autre stratégie a été testée. Etant donné le polymorphisme plus important de *rap-1a* au sein d'un même locus, nous avons décidé de commencer par amplifier *rap-1b*. Pour cela des amorces ont été définies dans des régions conservées de ce gènes, que nous avons pu définir grâce à l'alignement des deux séquences de *rap-1b* maintenant disponibles. Une fois la séquence partielle de *rap-1b* obtenue, des amorces ont permis d'amplifier les zones intergéniques entre deux gènes *rap-1b* et de déterminer la séquence du gène *rap-1a* intercalé, comme on pouvait s'y attendre, entre les deux gènes *rap-1b*. Une séquence partielle de *rap-1c* a été obtenue après utilisation d'amorces définies, comme pour *rap-1b*, dans les régions conservées déterminées sur la base de l'alignement des deux séquences de ce gène maintenant disponibles. Si l'obtention de la séquence complète de

rap-1c en 5' n'a pas posé de problèmes, nous n'avons pas réussi à obtenir celle en 3'.

Résultats

Conservation très forte des séquences des gènes *rap-1a*, *rap-1b*, *rap-1c* et des régions intergéniques chez *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) et *Babesia* sp. Tianzhu

Seules les séquences de *rap-1c* présentent un très faible polymorphisme, avec 11 substitutions dans la partie 3' du gène, identiques pour *Babesia* sp. BQ1 (Ningxian) et *Babesia* sp. Tianzhu. La majorité des ces substitutions (8/11) sont localisées au niveau des séquences répétées mises en évidence dans cette région du gène chez *Babesia* sp. BQ1 (Lintan) et retrouvées chez les deux autres membres. Sur ces 11 substitutions, seuls 5 sont non synonymes, 3 d'entre elles étant localisées dans les séquences répétées.

Les deux types de gènes *rap-1a* sont retrouvés, à savoir *rap-1a61* et *rap-1a67*. Un nouveau type de gène *rap-1a61*, *rap-1a61-3*, est défini sur la base des 3 substitutions déjà décrites pour ce gène chez *Babesia* sp. BQ1 (Lintan).

Même les séquences des régions intergéniques sont totalement identiques chez ces trois agents du groupe *B. motasi*-like.

Chez *Babesia* sp. Hebei, les séquences des tous les gènes *rap* et des régions intergéniques diffèrent de celles des 3 autres membres du groupe *B. motasi*-like

Cet agent se différencie très nettement des deux autres étudiés. L'ensemble des séquences du locus diffère de celles de *Babesia* sp. BQ1 (Lintan), avec en moyenne une identité autour de 78%, que ce soit au niveau des gènes ou au niveau des régions intergéniques.

Pour les gènes *rap-1a*, la présence des deux types de gènes bien différents *rap-1a61* et *rap-1a67* n'a pas été mise en évidence. Le gène *rap-1a* chez *Babesia* sp. Hebei est différent de ces deux gènes sur l'ensemble de sa longueur, avec un polymorphisme nettement plus marqué dans la région 3' du gène. C'est d'ailleurs dans cette région que se différencient les trois séquences de *rap-1a* (*rap-1a1*, *rap-1a2* et *rap-1a3*) présentes dans le locus.

Conservation de l'organisation du locus chez les membres du groupe *B. motasi*-like étudiés

Même si *Babesia* sp. Hebei est différent des 3 autres agents étudiés du groupe *B. motasi*-like au niveau des séquences de l'ensemble du locus *rap*, l'organisation générale du

locus est conservée, avec une alternance de gènes *rap-1a* et *rap-1b*, trois régions intergéniques différentes, un gène *rap-1a* en aval du dernier gène du locus, *rap-1c*.

Discussion

Dans cette partie de nos travaux, nous avons démontré la conservation de l'organisation du locus *rap* au niveau du groupe *B. motasi*-like comportant des agents de babésiose ovine en Chine.

Pour 3 des agents, *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) et *Babesia* sp. Tianzhu, l'ensemble de la séquence du locus est pratiquement identique sur une taille totale d'environ 31 Kb. Même les séquences des régions intergéniques, généralement moins conservées, sont identiques. Par contre, le quatrième agent étudié, *Babesia* sp. Hebei, se distingue très clairement des trois autres. En moyenne, sur l'ensemble du locus, l'identité des séquences au niveau nucléotidique baisse très nettement avec une valeur d'environ 78%. Ces résultats sont en accord avec ceux du 18S, qui placent cet agent à part des 3 autres (Lack *et al.*, 2012). D'autres travaux sur la phylogénie des membres de ce groupe, basés sur différents marqueurs chromosomiques et mitochondriaux (ITS, 28S, RPS8, sous-unité 3 de la cytochrome c oxidase et Cytochrome b), donnent des résultats contradictoires (Niu *et al.*, 2009a; Gou *et al.*, 2013; Tian *et al.*, 2013 a, b et c), plaçant souvent *Babesia* sp. BQ1 Ningxian à part des autres. Dans ces travaux, *Babesia* sp. Hebei n'est souvent pas étudié, ce qui ne permet donc pas de comparer nos résultats. Quoi qu'il en soit, il est clair que de plus amples travaux sont nécessaires pour élucider la phylogénie de ce groupe.

Ces quatres agents sont très proches phylogénétiquement puisque leurs séquences ADNr 18S sont identiques à plus de 98%. L'analyse des séquences du locus *rap* sur plus de 30 Kb indique que ce locus semble encore plus conservé que l'ADNr 18S pour des agents très proches. En effet, le locus est identique à 99,9% pour trois des agents dont l'identité au niveau de l'ADNr 18S est d'environ 99%. Par contre, pour *Babesia* sp. Hebei, dont l'identité avec les trois autres agents au niveau de l'ADNr 18S est d'environ 98%, l'identité nucléotidique au niveau du locus *rap* chute à 78% environ. L'analyse des gènes *rap* présente donc un intérêt phylogénétique, qui demeure néanmoins limité au genre *Babesia* pour le gène *rap-1a*, voire au groupe phylogénétique *B. motasi/B. bigemina* pour les gènes *rap-1b* et *rap-1c*.

La présence de copies multiples et conservées de *rap-1b* présente un intérêt pour le

diagnostic moléculaire des membres de ce groupe, permettant d'améliorer le seuil de détection.

Parmi ces différents gènes, seule la traduction de RAP-1a a été démontrée chez *B. bigemina* (Suarez *et al*., 2003). Le niveau élevé de conservation de ce gène au sein d'un locus, mais aussi entre les différents agents du groupe *B. motasi*-like responsables de babésiose ovine, suggère que le développement d'un vaccin recombinant commun à ce groupe est probablement réaliste. Mais de nombreux travaux restent à mener pour atteindre ce but.

Article 3

Analyse de l'expression des gènes *rap-1* chez *Babesia* sp. BQ1 (Lintan) in vitro et in vivo

Résumé détaillé des travaux en cours pour rédaction du 4ème article

Niu Q, Bonsergent C, Guan G, Malandrin L, Moreau E. Expression of rap-1 genes in Babesia sp. BQ1 (Lintan) (Babesia motasi-like phylogenetic group), a pathogen of sheep in China

Données complémentaires à acquérir avant publication

Pour ce travail, les quatre auteurs ont participé aux réalisations expérimentales. QN, LM et EM ont rédigé le manuscrit; LM et EM ont supervisé l'ensemble des travaux.

Introduction

Au cours des travaux précédents réalisés au cours de cette thèse, nous avons montré que chez *Babesia* sp. BQ1 (Lintan) le locus *rap-1* était un locus complexe de 31 Kb environ, comportant 12 gènes *rap-1* appartenant à trois types différents (*rap-1a*, *rap-1b* et *rap-1c*). Nous avons également montré que cette organisation était très conservée au sein du groupe phylogénétique *B. motasi*-like comprenant divers agents de babésiose. Des travaux menés chez *B. bigemina*, chez qui l'organisation du locus est similaire, ont montré que même si la transcription des trois types de gènes se produisait, seule *rap-1a* était traduit en protéine (Suarez *et al.*, 2003).

Dans le cadre du développement de vaccins, il est essentiel de déterminer au sein d'une famille multigénique, quels membres sont traduits afin de concentrer ses efforts de développement sur ces protéines. C'est l'objectif de cette partie de mon travail.

Méthodologie

L'analyse de l'expression des différents gènes du locus *rap-1* chez *Babesia* sp. BQ1 a été réalisé au niveau de la transcription et de la traduction des différents types de gènes mis en évidence au sein du locus.

L'analyse de la transcription a été réalisée à partir de cultures in vitro de parasites par RT-PCR classique à l'aide d'amorces spécifiques de différents gènes décrits : *rap-1a61*, *rap-1a67-1*, *rap-1a67-2*, *rap-1b* et *rap-1c*. Les produits d'amplification ont été clonés et séquencés afin de vérifier l'absence d'introns dans chacun de ces gènes.

Pour l'analyse de la traduction, les épitopes B putatifs ont été recherchés à l'aide de logiciels bio-informatiques. Nous avons ensuite défini des peptides spécifiques de chaque type de protéine (RAP-1a, RAP-1b et RAP-1c, RAP-1a61 et RAP-1a67) que nous avons fait synthétiser. Nous avons ensuite fait produire par la même société (GeneCust) des anticorps polyclonaux de lapin dirigés contre ces peptides (2 lapins/ peptide). La production des différentes protéines par le parasite cultivé in vitro mais également par le parasite lors d'une infection expérimentale de mouton a été testée par western blot en utilisant les anticorps polyclonaux de lapin dirigés spécifiquement contre les différents peptides. Les sérums de moutons collectés lors d'infections expérimentales par *Babesia* sp. BQ1 (Lintan) (Guan *et al.*, 2010c) ont été testés en ELISA contre les différents peptides utilisés pour l'immunisation, fin d'évaluer la présence dans ces sérums d'anticorps anti RAP.

Les protéines recombinantes sont en cours de production afin d'affiner les résultats obtenus et d'apporter les contrôles nécessaires à la validation des résultats.

Résultats

Analyse de la transcription des gènes *rap*

Les trois types de séquences intergéniques (IG1, IG2 et IG3) ont été analysées pour rechercher dans la région 5'-UTR la présence de motifs conservés, suspectés d'être impliqués dans la régulation de la transcription des gènes en aval (Suarez *et al.*, 1998b). Sur les trois motifs décrits, la mRNA box est la moins conservée, en séquence et en position. Par contre les deux autres motifs situés en amont sont très conservés.

L'analyse de la transcription par RT-PCR démontre que les trois types de gènes sont transcrits (*rap-1a*, *rap-1b* et *rap-1c*) lorsque le parasite est cultivé in vitro. Pour *rap-1a*, la transcription de trois des types de gènes a également été démontrée (*rap-1a61*, *rap-1a67-1* et *rap-1a67-2*).

Le séquençage des produits de RT-PCR confirme l'absence d'introns dans ces gènes.

Sélection des peptides et contrôle de la spécificité des sérums obtenus

Les peptides ont été sélectionnés sur la base de leur localisation au niveau d'épitopes B potentiels et de la spécificité au niveau de leurs séquences. L'absence de réactions croisées des sérums dirigés contre chacun des peptides envers les autres peptides a été confirmée par ELISA.

Mise en évidence directe par western blot de la production des protéines RAP par *Babesia* sp. BQ1 (Lintan) cultivé in vitro

Des extraits protéiques de mérozoites ont été produits à partir de culture in vitro de *Babesia* sp. BQ1 (Lintan) et utilisés comme antigène en western blot, L'analyse menée avec les sérums polyclonaux de lapin dirigés contre les différents peptides indique que la protéine RAP-1a est produite in vitro par le parasite. L'analyse avec les sérums spécifiques des deux types de protéines RAP-1 (RAP-1a61 et RAP-1a67) indique que seule RAP-1a-61est détectée, et par un seul des deux sérums produits. La présence de RAP-1b et RAP-1c n'est pas détectée. Des résultats identiques ont été obtenus lorsque des antigènes produits in vivo sont utilisés, en début d'infection au plus fort du pic de parasitémie.

Mise en évidence indirecte, par ELISA, de la production des protéines RAP par *Babesia* sp. BQ1 (Lintan) lors d'une infection expérimentale de moutons

Afin d'évaluer indirectement la production des différentes protéines RAP in vivo, nous avons dosé les anticorps dirigés contre ces protéines présents chez le mouton avant et après infection expérimentale, à différentes dates post-infection. Pour les trois moutons analysés, les résultats indiquent de façon significative que des anticorps sont produits contre les différents peptides spécifiques des différentes protéines RAP. La cinétique de production semble différente selon les protéines. Les anticorps anti RAP-1a apparaissent rapidement et tendent à disparaître, tandis que le taux d'anticorps anti RAP1c augmente encore trois mois après le début de l'infection

Discussion

Dans cette partie, la transcription des différents gènes *rap* a été démontrée. Ce résultat est en accord avec les données publiées chez *B. bigemina* (Suarez *et al.*, 2003).

Les résultats concernant la traduction des différents gènes sont à prendre avec précaution, et de nombreux témoins sont encore à fournir (voir ci-dessous). Toutefois, les résultats concernant la production des protéines RAP in vitro rejoignent ceux décrits chez

B. bigemina (Suarez *et al.*, 2003): pas de traduction apparente de RAP-1b et RAP-1c, et production d'une partie seulement des protéines RAP-1a. Toutefois, l'analyse des sérums produits au cours d'infections expérimentales de moutons indique la présence d'anticorps dirigés contre les différentes protéines, et donc leur production in vivo. Toutefois, la cinétique de production est différente selon les protéines et pourrait correspondre à un mécanisme d'échappement du parasite au système immunitaire.

La formulation de conclusions plus fermes sur cette partie du travail requiert des travaux complémentaires qui seront menés prochainement.

Travaux complémentaires à mener

La production des protéines recombinantes reste encore à réaliser, afin de vérifier la capacité des sérums à reconnaître la protéine correspondante, et pas seulement le peptide. Ce contrôle est nécessaire, car on ne peut pas à l'heure actuelle conclure de façon certaine sur l'absence de production des protéines RAP-1a67, RAP-1b et RAP-1c sans avoir ces témoins. Le fait que sur les deux sérums anti RAP-1a61, un seul donne un résultat positif confirme bien la nécessité d'un tel contrôle.

Après purification, les protéines recombinantes pourront être utilisées afin de tester en western blot les sérums de moutons infectés expérimentalement, et confirmer ainsi les résultats obtenus par ELISA.

L'analyse de l'effet des sérums produits sur la croissance de *Babesia* sp. BQ1 (Lintan) permettra de compléter nos données, en indiquant un rôle possible des protéines reconnues dans le développement du parasite, et aussi l'intérêt ou non de ces protéines dans le cadre d'une stratégie vaccinale visant à bloquer la multiplication parasitaire asexuée.

Les sérums seront utilisés pour vérifier la localisation apicale au niveau des rhoptries des protéines RAP.

Article 4

Les gènes *rap-1* chez *Babesia* sp. Xinjiang : caractérisation et analyse de l'expression d'une famille multigénique conservée avec une partie variable en taille et en séquence, composée de répétitions plus ou moins conservées

Résumé détaillé de l'article

Qingli Niu, Jordan Marchand, Congshan Yang, Claire Bonsergent, Guiquan Guan, Hong Yin, Laurence Malandrin

Rhoptry-Associated-Protein (rap-1) genes in the sheep pathogen Babesia sp. Xinjiang: multiple copies differing by 3' end repeated sequences.

Soumis à PloS One le 28 Février 2014.

Pour ce travail, QN et JM ont réalisé les amplifications, clonages et séquençages; CY a réalisé les analyses transcriptomiques en Chine; CB a encadré QN et JM au niveau expérimental; GG et HY ont encadré CY en Chine; QN, CY et LM ont rédigé le manuscrit; LM a supervisé l'ensemble des travaux en France.

Introduction

Babesia sp. Xinjiang est un isolat parasite du mouton en Chine. Il a été découverte récemment (Guan *et al*., 2001). Sa transmission est assurée par la tique *Hyalomma anatolicum anatolicum* (Guan *et al*., 2009), et une étude séro-épidémiologique a montré sa très forte prévalence chez les petits ruminants en Chine ainsi que sa très large distribution géopgraphique dans ce pays (Guan *et al*., 2012a).

Babesia sp. Xinjiang se place dans un groupe phylogénétique à part de celui des autres espèces responsables de babésiose ovine, i.e. le groupe *B. motasi*, et les espèces *B. crassa* et *B. ovis* (Guan *et al*., 2009). La spécificité de ce parasite pour le mouton ne semble pas stricte, car les études in vitro ont montré sa capacité à se développer en hématies de bovin et de Cervidés (Guan et al., 2012b). *Babesia* sp. Xinjiang se place dans le même groupe monophylétique que l'espèce *B. pecorum* nouvellement décrite chez le Cerf en Espagne (Jouglin et al., 2014 in press), et que plusieurs parasites caractérisés au niveau moléculaire

chez la girafe et l'antilope de Roan en Afrique du Sud (Oothuizen *et al.*, 2009).

L'objectif de ce travail a été d'identifier les gènes *rap-1* chez cet isolat. Devant la multiplicité des gènes séquencés, une analyse de leur transcription a été également menée. Les résultats de ce travail en termes de diagnostic, de vaccin recombinant, de fonction et d'évolution des ces gènes sont discutés.

Méthodologie

Comme dans le cas de *Babesia* sp. BQ1 (Lintan), nous ne disposions d'aucune donnée sur les gènes *rap-1* chez cette espèce. J'ai donc utilisé la même méthodologie que celle utilisée pour *Babesia* sp. BQ1 (figures D et E). Notre expérience chez cette dernière espèce a montré que les amorces dégénérées définies pour amplifier la région centrale des gènes *rap-1* ne permettaient pas l'amplification des gènes de type *rap-1b*. De nouvelles amorces ont donc été définies sur la base des deux séquences maintenant connues de ce gène (*B. bigemina* et *Babesia* sp. BQ1).

Comme pour *Babesia* sp. BQ1, j'ai eu recours de façon systématique au clonage des produits d'amplification avant de procéder au séquençage en raison du caractère multigénique de cette famille et du polymorphisme éventuel entre les différentes copies.

Plusieurs copies du gène *rap-1a* ayant été mises en évidence, avec une région 3' de taille variable, j'ai défini de nouvelles amorces pour amplifier de façon sélective cette région afin de l'analyser de façon plus approfondie (visualisation des amplicons de taille différente possible sur gel, séquençage d'un plus grand nombre de clones, PCR avec 30 cycles d'amplification pour limiter l'apparition de chimères). Ces amorces ont permis également d'analyser la transcription des différentes copies du gène mises en évidence par RT-PCR classique.

Résultats

<u>Absence apparente des gènes *rap-1b* et *rap-1c*</u>

Un produit de taille attendue (environ 250 pb) a été obtenu après amplification à l'aide des amorces dégénérées localisées au niveau des motifs conservés de la protéine RAP-1. Après clonage, le séquençage de plusieurs inserts n'a mis en évidence qu'un seul type de séquences, qui correspond à un gène de type *rap-1a*. L'amplification de la partie centrale

du gène *rap-1b* n'a donné aucun amplicon de taille attendue, et le séquençage des amplicons obtenus n'a pas permis d'identifier ce gène.

Chez *Babesia* sp. Xinjiang, seul le gène *rap-1a* semble donc être présent.

rap-1a: plusieurs copies du gène avec un polymorphisme dans la région 3'

A partir du fragment séquencé dans la région centrale du gène *rap-1a*, de nouvelles amorces ont été définies afin d'amplifier les copies potentielles de ce gène, organisées en tandem. Les amorces ont donc été utilisées pour amplifier la région 3' d'une première copie, la région intergénique et la région 5' de la copie suivante (figure E).

Un amplicon de taille attendue (environ 2,6 kb) a été obtenu, cloné et les séquences des inserts de plusieurs clones bactériens ont été comparées. L'analyse des séquences en amont de la région centrale (partie 5' du gène et région intergénique) a montré une absence complète de polymorphisme de cette région. Par contre, les séquences obtenues pour les différents clones dans la partie du gène situé en 3' de cette région connue se sont avérées différentes en tailles et en séquences, et cette région a été analysée de façon plus approfondie.

Les sites start et stop putatifs des gènes ont été définis en traduisant les séquences obtenues. La localisation des motifs conservés DAAF et YKTYL, et des cystéines conservées permet de confirmer le site start. La longueur conservée moyenne des séquences connues permet de contrôler le positionnement des sites stop.

Deux types de gènes rap-1 définis sur la base du polymorphisme en 3'

Deux types de séquences ont été obtenues en 3' de la région variable. Elles se distinguent sur la base de 28 substitutions. Ces deux types de séquences permettent de distinguer deux types de gènes que j'ai appelé α et β (figure F).

```
         1                                                          60
rap-1aα  TTCGCTCCCG GTGGCAGGAG ATTCCTGCGC ATCAAGTCCC CTGTATCGGA TGCAGAAACG
rap-1aβ  GTATTGCCCA GAGCTAAAGA GGTTTTGGGT ATTAACTCTC CTGCGTCTGA TGCAGAAACG

         61                                                        120
rap-1aα  GTTGCAGAAG AGTCAACTGC TACCATCGTA GAAGAGCCTA AGATCGATTC CGAAATTACC
rap-1aβ  GTTGCAGAAG AGTCAACTGC TACCATCGTA GAAGAGCCTA AGATCGATTC CGAAACCACC

         121                     147
rap-1aα  ACGGAGGATG AGGTATCTCA AGAATAA
rap-1aβ  ACGGAGGATG AGGTATCTCA GGAATAA
```

Figure F. Alignement de la région de en 3 ' les *rap-1aα* et *rap-1aβ* gènes, il est indiqué 28 substitutions.

Une région variable constituée d'un nombre variable de séquences répétées proches

Afin d'analyser plus précisément la région variable mise en évidence, des amorces ont été définies afin d'encadrer cette région pour l'amplifier. La taille des fragments ainsi générés permet leur séparation sur gel d'électrophorèse, et ainsi d'analyser visuellement la répétabilité des amplifications moléculaires réalisées. En effet, en raison de la présence de gènes multiples de séquences très proches, et de l'utilisation d'amorces conservées pour amplifier ces différentes séquences, les risques de création de chimères sont élevés. Pour ces raisons, le nombre de cycles de PCR a également été réduit à 30. L'amplification a également été répétée 3 fois avec deux jeux d'amorces différentes. Les résultats ont été identiques à chaque fois (figure G), avec l'écart de taille attendu lié à la localisation des amorces.

Les différents amplicons ont été clonés et les inserts de 97 clones ont été séquencés. Les alignements ont été réalisés manuellement, la ressemblance des séquences couplée à des différences de taille rendant difficiles les alignements automatiques.

Des séquences de taille variant de 72 à 360 nucléotides ont été retrouvées, correspondant bien à la taille des amplicons visualisés sur gel. La région variable est en fait constituée d'une succession de répétitions de 36 nucléotides, allant de 2 à 10 répétitions. La présence de répétitions de longueur correspondant à un multiple de 3 permet de ne pas décaler le cadre de lecture, ce qui est particulièrement important dans une séquence codante, comme cela est le cas ici. La comparaison des séquences des répétitions présentes dans les différentes régions variables permet de distinguer 13 séquences nucléotidiques différentes (figure H). Leur agencement permet de définir 7 séquences différentes de la région variable, 5 qui appartiennent au type α décrit précédemment, et 2 au type β (table

A). Certaines répétitions présentes dans les séquences de type *β* (4, 5, 6, 8 et 9) sont uniques à ce type.

Le nombre de substitutions nucléotidiques entre les répétitions varie entre 1 et 15, mais beaucoup d'entre elles se situent au niveau du 3ème codon (7/12), n'engendrant ainsi que très rarement une modification de l'acide aminé codé. Ainsi, 6 séquences nucléotidiques vont coder pour la même séquence peptidique (table B).

Il en résulte que les gènes *rap-1aα3* et *α4* codent une protéine putative identique.

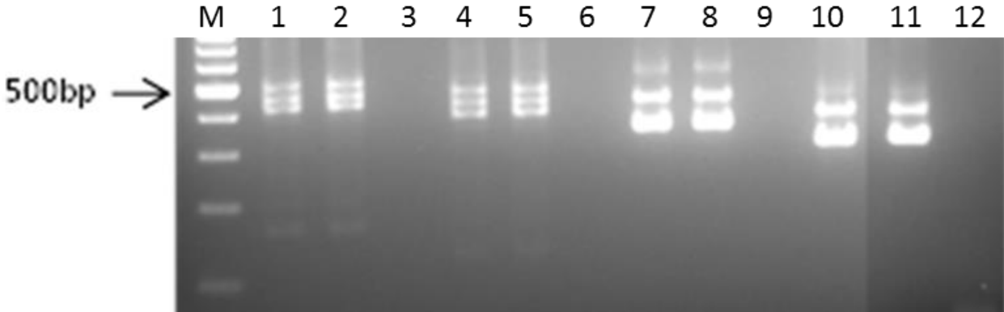


Figure G. L'amplification de la région variable de *rap-1aα* et *rap-1aβ* types de gènes. Voies 1-6: *rap-1aα*; voies 7-12: *rap-1aβ*. Voies 1 et 2: deux amplifications séparées en utilisant des amorces rap1a-F-875/rap1aα-R, voie 3: contrôle à blanc; Voies 4 et 5: dupliquer amplifications utilisant des amorces rap1a-F-915/rap1aα-R, voie 6: contrôle à blanc; voies 7 et 8: dupliquer amplifications utilisant des amorces rap1a-F-875/rap1aβ-R, voie 9: contrôle à blanc; voies 10 et 11: dupliquer amplifications utilisant des amorces rap1a-F-915/rap1aβ-R, voie 12: contrôle à blanc. M: 100 pb marqueur moléculaire.(figure 2 de article 4).

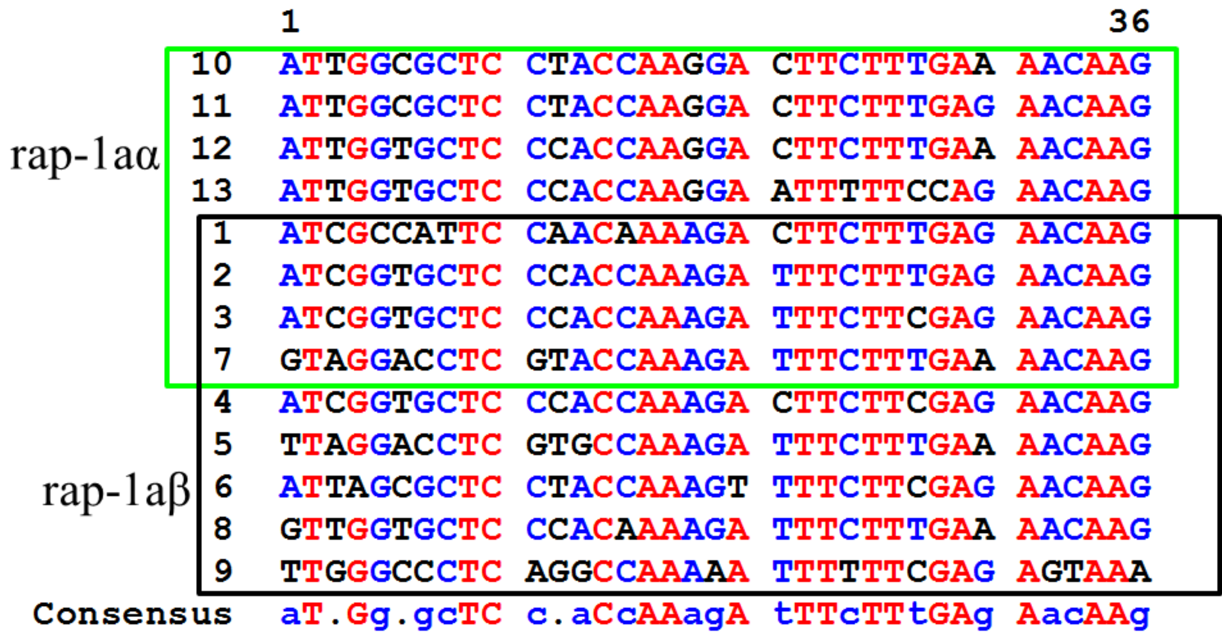


Figure H. Alignement de 13 séquences nucléotidiques. La séquence de deux *rap-1a* types répétitions indiqué par la couleur différente, *rap-1aα* avec la boîte verte et le *rap-1aβ* avec la boîte noire.

Table A. Disposition des répétitions dans la région variable de *rap-1a* copies du gène dans *Babesia* sp. Xinjiang. Les couleurs identiques utilisés pour des nombres différents de répétition identiques indiquent des séquences d'acides aminés. (Table 3 de Article 4)

Répétition type et la disposition										**numéro de répétition**	**Nom de *rap-1a* copie**	**Longueur de gène**
1	2	3	3	7	10	11	12	11	13	10	*rap1-aα1*	1443 pb
1	2	3	3	3	7	10	11	13		9	*rap1-aα2*	1407 pb
1	2	3	3	7	10	11	13			8	*rap1-aα3*	1371 pb
1	3	3	3	7	10	11	13			8	*rap1-aα4*	1371 pb
1	13									2	*rap1-aα5*	1155 pb
1	2	3	4	5	6	7	8	9		9	*rap1-aβ1*	1407 pb
1	4	5	6	7	8	9				7	*rap1-aβ2*	1335 pb

Table B. Séquence de répétitions de nucléotides et le niveau d'acide aminé. La couleur de la séquence repetidic différents sont correspondait aux couleurs du tableau A.

Nombre de répétition	Séquence nucléotidique	Séquence peptidique	Copie de *rap-1a* type
1	ATCGCCATTCCAACAAAAGACTTCTTTGAGAACAAG	**IAIPTKDFFENK**	α et β
2	ATCGGTGCTCCCACCAAAGATTTCTTTGAGAACAAG	**IGAPTKDFFENK**	α et β
3	ATCGGTGCTCCCACCAAAGATTTCTTCGAGAACAAG	**IGAPTKDFFENK**	α et β
4	ATCGGTGCTCCCACCAAAGACTTCTTCGAGAACAAG	**IGAPTKDFFENK**	β
5	TTAGGACCTCGTGCCAAAGATTTCTTTGAAAACAAG	**LGPRAKDFFENK**	β
6	ATTAGCGCTCCTACCAAAGTTTTCTTCGAGAACAAG	**ISAPTKVFFENK**	β
7	GTAGGACCTCGTACCAAAGATTTCTTTGAAAACAAG	**VGPRTKDFFENK**	α et β
8	GTTGGTGCTCCCACAAAAGATTTCTTTGAAAACAAG	**VGAPTKDFFENK**	β
9	TTGGGCCCTCAGGCCAAAAATTTTTTCGAGAGTAAA	**LGPQAKNFFESK**	β dernier
10	ATTGGCGCTC CTACCAAGGACTTCTTTGAAAACAAG	**IGAPTKDFFENK**	α
11	ATTGGCGCTC CTACCAAGGACTTCTTTGAGAACAAG	**IGAPTKDFFENK**	α
12	ATTGGTGCTC CCACCAAGGACTTCTTTGAAAACAAG	**IGAPTKDFFENK**	α
13	ATTGGTGCTC CCACCAAGGAATTTTTCCAGAACAAG	**IGAPTKEFFQNK**	α dernier

Les copies du gène rap-1a mises en évidence chez Babesia sp. Xinjiang sont donc au nombre de 7, avec une taille allant de 1155 à 1443 nt. Elles sont constituées d'une région 5' conservée de 936 nucléotides, suivie d'une région variable en taille et en séquence constituée de séquences répétées qui définie les 7 copies, et d'une région 3' de 147 nt dont la séquence définit deux types de gènes rap-1, α et β.

<u>Organisation en tandem des gènes *α* et *β* sur le même locus</u>

Afin de vérifier la présence des différents types de gènes *rap-1a* sur le même locus, et d'étudier leur arrangement en tandem, des amplifications ont été réalisées à l'aide des amorces spécifiques de alpha et beta définies pour amplifier la région variable. Cette fois-ci, ces amorces ont été utilisées pour vérifier l'existence de différents arrangements possibles des gènes (α-α, β-β, α-β et β-α), et donc amplifier les extrémités des gènes et les régions intergéniques correspondantes. Des amplicons de taille attendue ont été obtenus pour toutes les combinaison, sauf β-β, malgré des essais répétés. Les deux types de gènes, α et β, sont donc localisés sur un même locus et arrangés en tandem tête à queue. Les amplicons ont été clonés et séquencés, et les arrangements suivants ont été validés (α-β1, β-α2, et α-α1). Le type du gène en amont ne peut être connu par cette méthode car la séquence obtenue dans cette région est non spécifique. Leur arrangement en tandem sur plusieurs loci différents ne peut toutefois pas être écarté.

<u>Analyse des régions intergéniques</u>

Les régions intergéniques localisées entre les gènes *rap-1a* ont été amplifiées et séquencées à partir de plusieurs types d'amplicons. Leur longueur est constante (1632 pb). Très peu de variations de séquences ont été mises en évidence, avec seulement 3 substitutions (311, 822 et 1006 pb après le site stop putatif) dont les arrangements sont au nombre de 5 (TAA, TAC, CAC, CAA et TCA, l'arrangement TCA n'ayant été retrouvé qu'après un gène de type *rap-1aβ*.

<u>Analyse de la transcription</u>

Analyse de la région 5' UTR des gènes *rap-1a*

La région 5' UTR localisée en amont des gènes *rap-1a* a été analysée, et les 3 blocs de séquences conservées déjà décrits dans cette région pour les gènes *rap-1* par Suarez et al.

en 1998b ont été retrouvés (figure I). Ces blocs sont supposés avoir un rôle dans la régulation de la transcription des gènes *rap-1*. Leur présence est un indicateur de la transcription effective de ces différents gènes.

	-59 box	-36 box	mRNA box
1	TCGCACTTAGCTGCA	GAGGTGCAG	TATAGCAGTGCTATATA
2	TCGCACTTNTTTGCA	AAGGTGCAC	TATAGCAGTGCTATATA
3	TCGCACTCGCTTGCA	GAGGTGCAG	TATAGCAGTGCTATATA

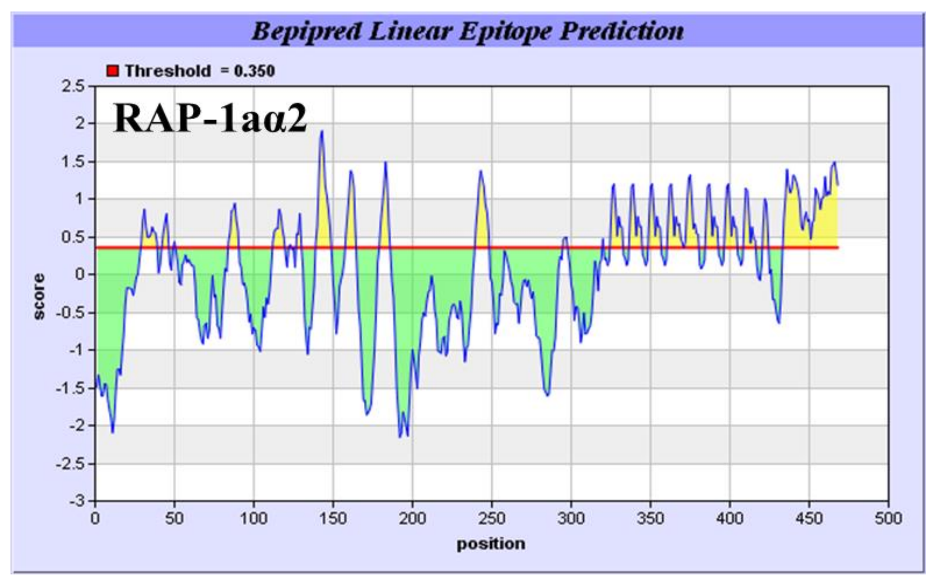
Figure I. Conservation du *rap-1a* régulation de la transcription de la séquence 5'UTR putatif dans les régions intergéniques *rap-1a*. La position des boîtes -59, -36 et mARN sont désignés par Suarez et al. en 1998. 1: *Babesia* sp. Xinjiang; 2: Consensus from Suarez et al. en 1998; 3: *B. ovis*. Le *Babesia* sp. Xinjiang a souligné séquences correspondent à la séquence identique au consensus. (figure 5 de Article 4)

Les gènes de type α et β sont transcrits

Afin de faire une analyse rapide de la transcription des différentes copies des gènes *rap-1a*, les amorces utilisées pour amplifier la région variable ont été utilisées dans une analyse RT-PCR classique. L'ARN a été préparé à partir de culture in vitro de *Babesia* sp. Xinjiang. De la même façon que pour l'analyse PCR, de multiples amplicons sont obtenus, indiquant une transcription active d'au moins 3 types de gènes *rap-1aα*, et des deux gènes *rap-1aβ*. Seul le gène *rap-1aα5* ne semble pas être transcrits, et la preuve de la transcription des deux copies *rap-1aα3* et *α4* (région variable de même taille) ne peut être apportée dans cette analyse.

La région 3' variable et constituée de répétitions contient la majorité des épitopes B putatifs

Les outils bioinformatiques (http://tools.immuneepitope.org/tools/bcell/iedb_input) ont été utilisés pour prédire la localisation des épitopes B putatifs des protéine RAP-1a chez *Babesia* sp. Xinjiang. Le résultat le plus frappant est la concentration de ces épitopes dans la région 3' de la protéine, et plus particulièrement la présence d'épitopes répétés correspondants aux répétitions de la région variable (figure J).



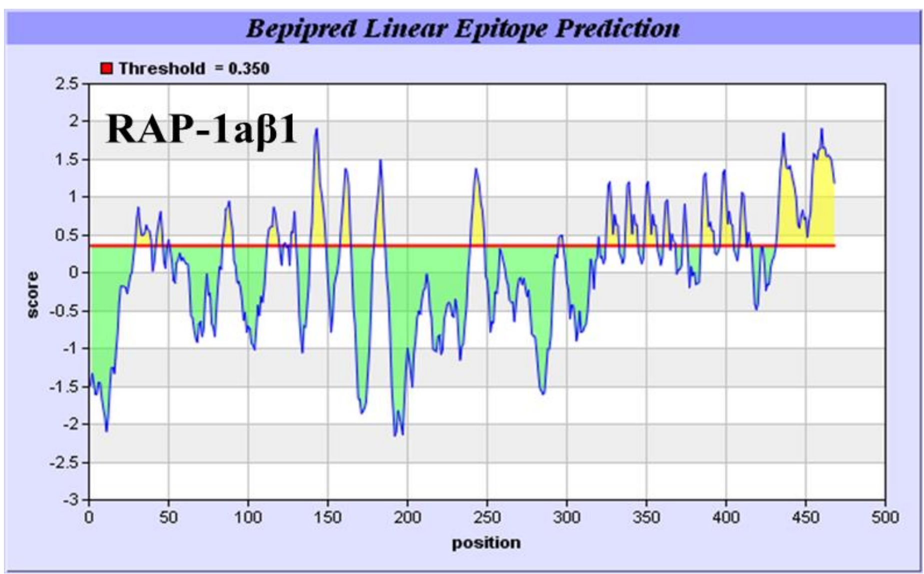


Figure J. Prédiction d'épitopes des épitopes B putatifs de protéines RAP-1aα2 et RAP-1aβ1 en utilisant la méthode "de prédiction linéaire Bepipred épitopes". Épitopes prédits B ont été indiqués en jaune.

Discussion

Dans cette partie de mon travail, j'ai montré la présence d'au moins 7 gènes *rap-1a* différents chez *Babesia* sp. Xinjiang. Les gènes de type *rap-1b* et *rap-1c* n'ont pas été trouvés, ce qui semble limiter leur présence au groupe phylogénétique comprenant *B. motasi* et apparentés, ainsi que *B. bigemina*.

Une très forte conservation des séquences des 7 gènes a été démontrée, notamment dans la partie 5'. La plus forte conservation de cette partie des séquences *rap-1*, avec notamment la présence des motifs très conservés DAAF et YKTYL et des 4 cystéines, est une caractéristique de cette famille (Dalrymple *et al*., 1996; Skuce *et al*., 1996).

Le caractère multigénique de cette famille pour *Babesia* sp. Xinjiang a également été démontré. En dehors du groupe phylogénétique *B. motasi* (avec plusieurs isolats étudiés dans ma thèse)/*B.bigemina*, le caractère multigénique de la famille est rarement clairement démontré. Le nombre de copies de *rap-1a* est faible chez *B. bovis* (2 copies identiques), estimé à 4 ou 5 copies chez *B. ovis*, et supposé multiple chez *B. divergens* et *B. canis* (Dalrymple *et al*., 1996). La présence de copies multiples peut être supposée également au regard des résultats chez *B. caballi* : amplification PCR avec produits multiples (Bhoora *et al*., 2010), immunoblots avec bandes multiples (Awinda *et al*., 2013). Même s'il s'agit d'une famille multigénique, le nombre de copies de *rap-1a* reste faible au regard d'autres familles multigéniques décrites chez *B. bovis*, telles que la famille *ves* (environ 150 copies) (Brayton *et al*., 2007) ou la famille *smorf* (44 copies) (Ferreri *et al*., 2012).

La création des familles multigéniques se fait par duplication de gènes ancestraux qui sont ensuite maintenus si leur présence multiple représente un avantage pour la survie de l'organisme ou du parasite qui les possède : évolution adaptative (Innan, 2009). Dans le cas de cette famille multigénique, nous avons montré que les séquences intergéniques sont également dupliquées. Ces séquences intergéniques étant impliquées dans la régulation de la transcription et de la traduction des gènes situés en aval (Suarez *et al*., 2004), la conservation de copies multiples dans le but d'augmenter la quantité des protéines correspondantes produites est très probable (Innan, 2009). Ce besoin important en protéines *rap-1* peut correspondre chez ces parasites à un stade précis du développement parasitaire, plus spécifiquement une étape du processus d'invasion. De plus, il a été démontré que l'organisation des gènes en tandem tête-queue augmente de façon significative la quantité de protéines produites (Suarez *et al*., 2004), et nous avons

démontré cette organisation pour les gènes *rap-1a* de *Babesia* sp. Xinjiang. Nous avons également montré que au moins 5 des gènes *rap-1a* chez *Babesia* sp. Xinjiang étaient transcrits. Par contre, nous n'avons pas mené les analyses pour déterminer si les différentes copies étaient traduites. Le degré de similitude entre les séquences n'aurait permis de différencier que les protéines α et *β*, et donc de n'obtenir qu'une réponse partielle sur la traduction des différentes copies de *rap-1a*.

Le polymorphisme observé entre les différentes copies du gène *rap-1a* chez *Babesia* sp. Xinjiang est limité à la région 3' du gène. La partie 3' terminale (147 nt) diffère par la présence de 28 mutations ponctuelles qui se traduisent par le changement de 11 acides aminés, et plus en amont par la présence d'un nombre variable de séquences répétées de 36 nt plus ou moins dégénérées. La variation de séquence 3' terminale, concentrée dans les 15 acides aminés suivants la dernière séquence répétée (changement de 2/3 des acides aminés), permet de définir deux types de gènes, α et *β*. Les séquences des répétitions sont très conservées, notamment dans les différentes copies α, avec au total 6 répétitions nucléotidiques qui codent la même séquence peptidique. Un tel degré de conservation entre les différentes copies d'un gène suggère une évolution concertée des différentes copies, avec l'intervention de mécanismes de recombinaison génétique entre copies tels que la conversion génique. Ce mécanisme permet en effet d'homogénéiser les séquences par un mécanisme de copier-coller, et donc de réduire ainsi le taux effectif de mutations (Innan, 2009). L'accumulation de mutations entre les gènes de types α et *β*, en 3' mais aussi au niveau des répétitions, suggère qu'à un moment donné ces copies ont échappé au mécanisme de conversion génique, accumulant alors plus de mutations. Une telle différence est également visible chez B. ovis, pour qui 3-4 des copies ne diffèrent que par le nombre de répétitions, alors que la copie la plus en aval du locus ne présente qu'une homologie faible avec les autres copies (72% pour la partie 5'), et se démarque par l'absence de séquences répétées. Là aussi, cette copie semble avoir été « exclue » du mécanisme de conversion génique et avoir accumulée un grand nombre de mutations ponctuelles. Sa présence en copie unique, contrairement aux deux copies de type *β* présentes chez *Babesia* sp. Xinjiang, n'aurait pas permis au processus de conversion génique de limiter le taux effectif de mutations, comme cela a pu se produire entre les 2 copies *β*.

L'ensemble du processus d'invasion est mal/peu caractérisé pour les parasites du genre

Babesia, et les protéines impliquées sont encore très mal connues. Le rôle des protéines RAP-1 dans le processus d'invasion est suspecté chez *B. bovis* depuis une douzaine d'années (Yokoyama *et al.*, 2002, Mosqueda *et al.*, 2002a), mais aucune donnée complémentaire sur leur fonction précise n'a été obtenue à ce jour. La présence de séquences répétées dans les gènes *rap-1a* de *Babesia* sp. Xinjiang a probablement un lien avec la fonction de ces protéines. En effet, une étude récente sur la comparaison de la fréquence et de la nature des séquences répétées dans les génomes de Protozoaires parasites indique que globalement, les protéines ayant ces séquences ont un rôle soit dans l'invasion des cellules cibles, soit dans les mécanismes d'échappement (Mendes *et al.*, 2013). La nature conservée des séquences répétées chez les gènes décrits dans mon travail indiquerait une fonction plutôt en lien avec les interactions protéine-protéine, et donc dans l'invasion. L'attachement des protéines RAP-1 à la surface des érythrocytes et leur rôle dans l'adhésion du parasite aux globules rouges ont été démontrés (Yokoyama *et al*, 2002, Mosqueda *et al.*, 2002a). Le caractère conservé de ces séquences répétées dépasse l'échelle de l'espèce parasitaire, puisque la même séquence répétée IGAPTKDFFENK est retrouvée chez une autre espèce infectant le mouton, *B. ovis*, alors que la partie 5' normalement plus conservée n'est similaire qu'à 57% (Dalrymple *et al.*, 1993a). Le fait que la majorité des épitopes B putatifs soit localisée au niveau de la partie 3', incluant les séquences répétées, pourrait faire penser à un rôle de ces séquences dans les mécanismes d'échappement immunitaire. Ces épitopes pourraient agir comme un écran de fumée pour distraire le système immunitaire et protéger la partie N-terminale conservée et fonctionnelle de la protéine. La forte conservation de ces séquence répétées ne semble cependant pas aller dans ce sens.

Des travaux complémentaires sont nécessaires fin d'élucider le rôle de ces protéines dans l'invasion, et de définir leur intérêt vaccinal. Ce dernier point sera abordé de façon plus poussée dans la discussion générale.

INTRODUCTION
OBJECTIFS

Introduction

Babesiosis is a haemoparasitic disease, caused by the intraerythrocytic multiplication of protozoa of the genus *Babesia* (Apicomplexa phylum, class Aconoidasida, order Piroplasmida, family Babesiidae) and transmitted by ticks. This disease is also called piroplasmosis, including Theileriosis due to an other Apicomplexa of the genus *Theileria*. Babesiosis is a frequent infection of domestic and wild animals worldwide, and is considered as the second most common blood-borne parasites of mammals after trypanosomes (Levine, 1985; Telford et al., 1993; Bush et al., 2001). The clinical symptoms of babesiosis are due to the repeated asexual rounds of multiplication of parasites inside of host erythrocytes, and are usually characterized by fever, depression, hemolytic anemia, hemoglobinuria, icterus, finally resulting in death in severe cases if not treated. The disease can be responsible of great direct and indirect economic losses, due to the death of the animals, a reduction in the production or restriction in animal movements (McCosker, 1981). The prevalence of babesiosis correlates with the geographic distribution and activity of vector ticks. Furthermore environmental conditions changing, especially global warming, favors tick survival and reproduction, which correlate with a significant increase in the abundance of ticks (Slenning, 2010). This could explain the increase and the emergence of tick-borne disease, in particular piroplasmosis.

These parasites were first discovered in erythrocytes of cattle in Romania in the late 19th century by Victor Babes, a Romanian microbiologist, who also found similar organisms in red blood cells of sheep later. One year later in the USA, Theobald Smith described the presence of an intraerythrocytic parasite of cattle, which is considered as the agent of Texas Cattle Fever (babesiosis). New species responsible for Texas cattle fever disease were discovered, which they called *Pyrosoma bigeminum* (*Babesia bigemina*) in later four years. In 1893, these described parasites were named as *Babesia bovis*, *Babesia ovis* and *Babesia bigemina*, respectively by Starcovici (Mihalca, 2010). Newly recognized babesial pathogens parasitizing the blood of other domestic and wild animals were then discovered and described from all around the world, stretching from the polar circle to the equator, overlapping their tick vector distribution. Up to now, more than 100 *Babesia* species infecting domestic and wild animals, cattle, sheep/goats, horse, pigs and dogs and humans are described in Asia, Europe, Africa, Australia and America and the number of discovered and described species is increasing (Kuttler 1988a; Levine, 1988; Uilenberg, 2006; Gray and Weiss, 2008; Lack et al., 2012; Schnittger et al., 2012). Six *Babesia*

species responsible for ruminant babesiosis in cattle (*Babesia bovis*, *B. bigemina*, *B. divergens*, *B. major*, *B. ovata* and *B. orientalis*) and three in sheep/goat (*B. ovis*, *B. motasi* and *B. crassa*) have been mainly described in the world. *B. bigemina* and *B. bovis* in cattle, *B. motasi* and *B. ovis* in sheep/goat are considered as the more pathogenic (Levine, 1985; Kuttler, 1988a; Friedhoff, 1988, 1997; Uilenberg, 2006).

In China, the babesiosis is caused by various protozoan parasites of the genus *Babesia*, in which numbers of agent are highly pathogenic for ruminant cattle, buffalo, yak, sheep, goat (Yin et al., 1997a; Liu ZL et al., 1997). The cattle and sheep/goat grazing areas are mainly distributed in the north of China and the *Babesia* species of ruminants are therefore common parasites in this area. Among the 109 species of ticks identified to date in China, 45 infest sheep and goats. Babesiosis of ruminant in China is considered to be one of the most frequent tick-borne diseases with an important economic impact. The economic losses caused by babesiosis and other related tick-borne disease as Theileriosis and Anaplasmosis, are enormous, in the range of several millions USD per annum of mutton production, goat wool and animal meat output (Yin and Luo, 2007). Bovine babesiosis was first described in the fifties in China (Wang, 1993). At present the reported causative agents of bovine babesiosis are *B. bigemina*, *B. bovis*, *B. ovata*, *B. major*, *B. orientalis* and the recently described *Babesia* U sp. Kashi derived from *Hyalomma anatolicum anatolicum* collected from the field (Lu et al., 1992; Yin et al., 1997a; Bai and Liu, 1990; Luo et al., 2005a, b). The *B. bigemina* infection normally was characterized by a low parasitemia, in contrast with the infection by *B. bovis* which resulted in high parasitemia and more severe disease than *B. bigemina*. Both species are transmitted by the genus *Rhipicephalus* (*Boophilus*) and infected cattle and buffalo. These two species were considered to have the more economic impact on cattle health. *B. major* and *B. ovata* are transmitted by *Haemaphysalis punctata* and *H. longicornis* respectively, while *B. orientalis* is transmitted by *Rhipicephalus* and highly pathogenic to only buffalo (Luo et al., 2005a, b; Yin et al., 1996; He et al., 2012). *B. divergens* was detected in anemic patients but not isolated in cattle in China (Qi et al., 2011).

Ovine babesiosis was first recorded in China in 1982, and the parasite was suspected to be *B. ovis* due to its morphological characteristics observed on blood smears, the clinical manifestations and its relatively high pathogenicity (Chen, 1982). In 1985, an outbreak of babesiosis in sheep occurred in Heilongjiang province, and the causative agent was again identified as *B. ovis*. However, there were neither related *B. ovis* parasite to be isolated nor other cases to be reported in these areas since (Zhao et al., 1986). Ten years later, in 1996,

many sheep death due to an unexpected explosion of piroplasmosis infection were found in Ningxian county of Gansu Province and it caused serious harm for the sheep industry (Lian et al., 1997). Only a large sheep *Babesia* was observed in blood smears from three Small-Tailed Han sheep (short-tail) at that time. However, in 1997, two morphologically different *Babesia* parasites were collected from naturally infected sheep in this region, one classified as a large form and the other one as a small form. These two *Babesia* parasites isolated by the laboratory of the Veterinary Institute of Lanzhou (Gansu, China) were initially considered as *B. ovis* (small size) and *B. motasi*/unknown *Babesia* (called *Babesia* sp. BQ1 (Ningxian)) (Lian et al., 1997; Liu GY et al., 1997, 1999; Yin et al., 1997b; Bai et al., 2002). However, further unpublished studies using the isolate considered as *B. ovis* have shown that this small *Babesia* isolated from sheep was probably confused with parasites of the genus *Theileria* (personnal discussion with Yin and Bai). In 2002, a second *Babesia* parasite was isolated from splenectomized sheep experimentally infected with *H. qinghaiensis*, collected in pastures Lintan County of Gansu Province (Guan et al., 2002). This isolate was named *Babesia* sp. BQ1 (Lintan). In 2007, four other ovine *Babesia* parasites were further isolated from different geographical parts of China, by inoculating blood from field asymptomatic sheep to splenectomised sheep. They were named *Babesia* sp. Tianzhu, *Babesia* sp. Madang, *Babesia* sp. Hebei and *Babesia* sp. Liaoning, according to their geographical origin, plus Ningxian and Lintan isolates, that these six isolates were defined *B. motasi*-like, since the similar morphological size, tick vectors and 18S rRNA sequence (about 93%) with *B. motasi* European isolate were found (Liu et al., 2007). In 2001, another B*abesia* parasite was isolated from sheep experimentally infected by *Rhipicephalus sanguineus* and *Hy. a. anatolicum* collected from sheep in Kashi region, Xinjiang Uygur Autonomous Region in northwestern China , and later named *Babesia* sp. Xinjiang (Guan et al., 2001, 2002). So, at least 7 isolates from different regions in China have been isolated from sheep in the last 15 years, and further description of their respective virulence, vector, geographical distribution and taxonomy will be given in details in the litterature review.

At present, effective and actual control methods of babesiosis are still based on the chemotherapy/chemoprevention targeting ticks and/or *Babesia*: elimination of the vector in endemic areas, chemotherapy during acute babesiosis, prevention of infection by tick populations,. However, due to the drawbacks of chemotherapy, such as drug-resistance and drug residues in animals' products, more effective ways against babesiosis have been searched (de Waal and Combrink, 2006; Mosqueda et al., 2012). Vaccine application is an

alternative to prevent babesiosis. Some live vaccines have been studied and they could control babesiosis to a certain degree. There is still no commercial vaccine against small ruminants babesiosis currently available. The efforts are now mainly directed towards development of recombinant vaccines against a number of *Babesia* species targeting antigens which would block the parasite multiplication within the vertebrate host or the parasite transmission between ticks and vertebral hosts. A combination of tick vaccine research and vaccine development based on asexual growth cycle of blood *Babesia* stage were thought practical and ideal control strategies of babesiosis, particularly against virulent *Babesia* infection on the long-term (Gohil et al., 2013).

Objective of the thesis

My PhD thesis is included in the large European Union cooperation project "PIROVAC" that aims in developing vaccines against ovine theileriosis and babesiosis. The research work of "PIROVAC" project for babesiosis of small ruminants contains mainly 7 parts of work: 1). Establishment of parasite cultures and preparation of infective material; 2). Determination of strain diversity of *Babesia* parasites; 3). Identification of potential attenuation markers; 4). In vivo assessment of attenuation of *Babesia*; 5). Genetics of host response to infection and vaccination; 6). Comparative genomics of ovine *Babesia* species; 7). Identification and characterization of proteins involved in invasion of host cells by *Babesia* species of small ruminants for vaccine development.

My project is included in the seventh part of the PIROVAC project and is focused on ovine *Babesia* from China (5 different isolates representative of the main *Babesia* responsible of ovine babesiosis in China) and on one potential vaccine candidate, RAP-1 (Rhoptry- Associated-Protein-1), as a protein with a role in erythrocyte invasion by *Babesia*. The objective is to determine if this protein could be a good vaccine candidate against ovine babesiosis and could be added to a subunit vaccine in addition to other candidates studied by other groups in this European project.

Due to the complete lack of knowledge on the proteins involved in the erythrocyte invasion process in the chinese ovine parasites, my project has first consisted in the characterization of *rap-1* genes. It has been divided in three main parts:

1- determination of the *rap-1* gene sequences, gene copy numbers as well as gene locus organization in two phylogenetically very distant ovine parasites, namely *Babesia* sp. BQ1 (Lintan) and *Babesia* sp. Xinjiang (articles 1 and 3).

2- evaluation of *rap-1* genetic polymorphism among phylogenetically related (*B. motasi*-like group) but geographically distinct isolates in China. Polymorphism information on the selected *rap-1* genes is a key information to select conserved region in the RAP-1 protein as a common vaccine candidate construction (article 2).

3- evaluation of the expression of the different characterized *rap-1* genes, in order to target the expressed copy of this multigene family and to study the function of the *Babesia* sp. BQ1 (Lintan) RAP-1 proteins in the invasion process of red blood cells (articles 3 and 4).

Thanks to a previous collaboration between the Lanzhou Veterinary Research Institute in China and the UMR BioEpAR in France (G. Guan PhD thesis), the in vitro culture of *Babesia* sp. BQ1 (Lintan) was already established and sheep experimental infections had already been performed, from which the collected sera were available to use in the *rap-1* expression studies. In vitro culture of *Babesia* sp. Xingjang has been set up in LVRI in China. However, restrictions have been imposed by the Chinese institute, preventing living material from leaving their Institute. The gene locus organization will then be studied with genomic DNA produced in China and sent to France. The expression studies (RT-PCR) will be performed in collaboration with this Institute.

The manuscript is organized in 9 main parts : an extensive french résumé of my thesis, one general review of the litterature on the subjects related to my study area, 4 chapters corresponding to a published paper, two submitted papers and a manuscript in preparation, a general discussion, a conclusion and perspectives and finally annexes.

The supplementary data of each published or submitted paper are presented at the end of each respective paper, and other tables and figures are inserted in the text to facilitate reading and comprehension.

LITERATURE REVIEW

A. Description and classification of *Babesia* species infecting ruminants

Babesiosis is a tick-borne disease, caused by the infection and the subsequent intraerythrocytic multiplication of protozoa of the genus *Babesia*. This genus contains many species of major veterinary and economic importance (Levine, 1985; Bush et al., 2001). At the end of 19th century, these parasites were first discovered in erythrocytes of cattle in Romania. Up to now, more than 100 different *Babesia* species were identified in the world, attributed to present-day rapid development of molecular biology technique.

1. Taxonomic position and main biological features of *Babesia*

Babesia is an unicellular eukaryote, that belong to the Protista or Protozoa. Together with *Theileria*, they belong to the Piroplasmida (table 1). With the Haemosporida, including the famous genus *Plasmodium*, they belong to the Aconoidasida, since they do not have a conoid at the top of the apical complex. Members of these three genera are vertebrate blood parasites transmitted by vectors, ticks in the case of *Babesia* and *Theileria*, mosquitoes in the case of *Plasmodium*. *Babesia* characteristic biological feature in the vertebrate host is to target only red blood cells, while *Theileria* has a first developmental stage in the lymphocytes and *Plasmodium* in the hepatocytes. In the tick vector, only *Babesia* is able to be transmitted trans-ovarially from the femelle to the eggs and subsequently larvae.

Inside the *Babesia* genus, species were traditionally divided into 2 groups based on the merozoite morphology, or more precisely its size: the "large" *Babesia* with the length of the merozoite > erythrocyte radius, and the "small" *Babesia*, with the length of the merozoite < erythrocyte radius.

If phylogenetically closely related parasites (based on molecular new phylogeny) belong to the "large" *Babesia* group (figure 1), this morphological separation is not so clear for example for the phylogenetic group that includes *Babesia* sp. Xinjiang (described as a large form) and *B. bovis*/B. *ovis* (described as small forms).

Table 1. The traditional classification of *Babesia* (**Levine, 1971; Levine et al., 1980**)

Kingdom	Protozoan		
Phylum	Apicomplexa		Apical complex organelles (micronemes, rhoptries, dense granules)
Class	Aconoidasida (Sporozoa)	Subclass Piroplasmea	No conoid at the merozoite apex
Order	Piroplasmida	Suborder Piroplasmorina	Sexual reproduction in the tick vector
Family	Babesiidae		Absence of schizogony
Genus	*Babesia*		Asexual reproduction by binary fission, lack of hemozoin

2. Phylogenetic relationships of *Babesia* spp. infecting ruminant

The traditional taxonomy of species within the *Babesia* genus was entirely based on the biological and morphology criteria: morphology and ultrastructure of the merozoite under optical and electronic microscope, pathogenicity, antigenicity, host specificity, transmitting tick vectors, and epidemiological data (Uilenberg, 2006; Lack et al., 2012). These descriptions have progressively been completed and now often replaced by molecular criteria that allow phylogenetic analyses of parasites that have not even been isolated (molecular data from parasites in ticks for example).

Molecular biology and the use of 18S rDNA sequences allow a new approach of phylogenetic classification, since it contains highly conserved regions, variable and hypervariable regions. Therefore, 18S rDNA sequences are widely used as a genetic marker for taxonomic and phylogenetic studies and led to a more precise classification of some unknown species (Ludwig and Schleifer, 1994; Woese, 1987; Ahmed et al., 2006; Altay et al., 2007; Schnittger et al., 2012; Lack et al., 2012). Studies of molecular criteria, associated with morphological and biological data, could be useful tools to redefine the relationships between various *Babesia* species (Schnitter et al., 2012).

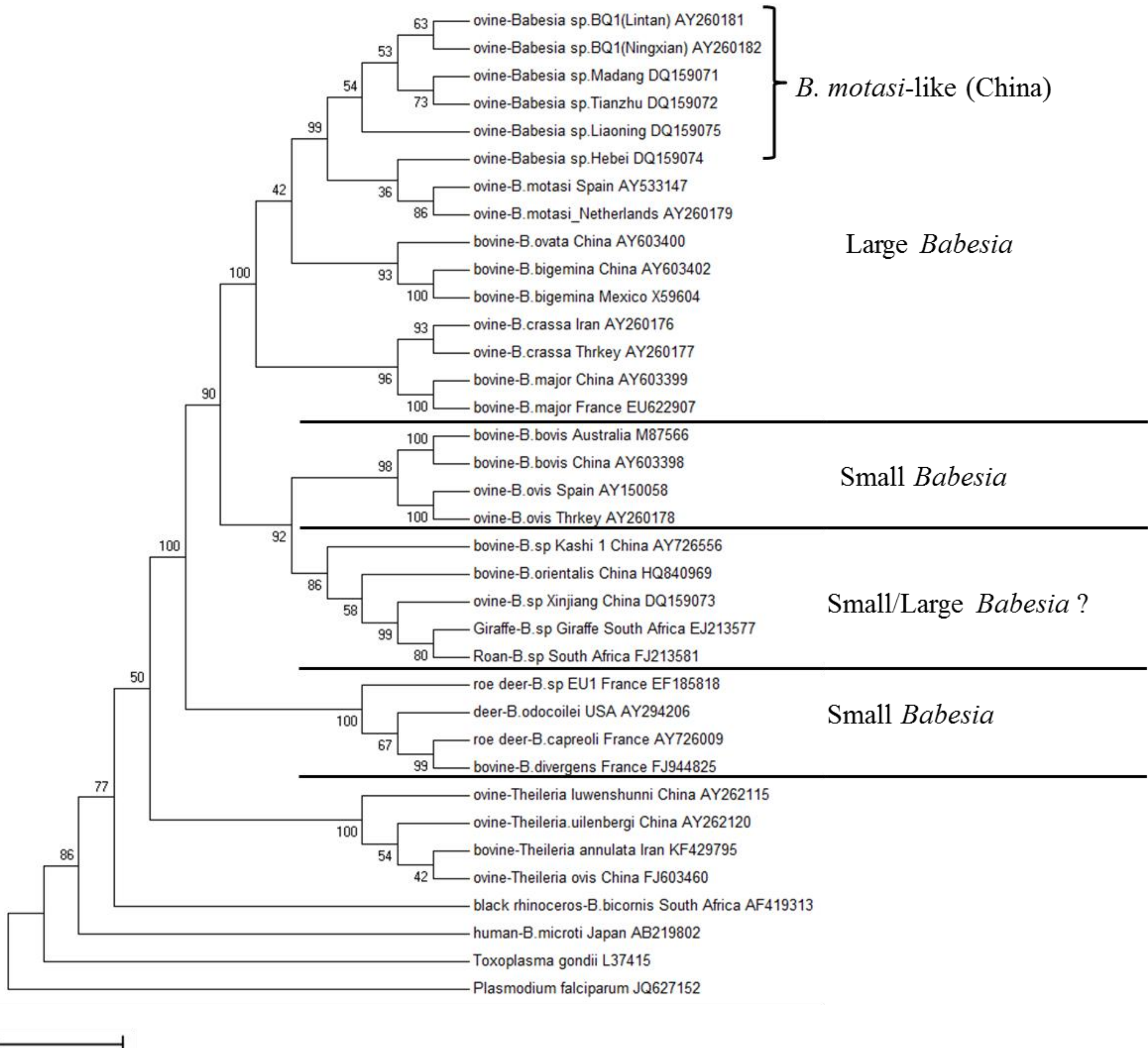


Figure 1. Phylogenetic analysis of *Babesia* spp. of ruminants (domestic and wild) based on the sequences of the 18S rRNA gene. The tree was inferred using the neighbor joining method of MEGA 5.2, bootstrap values are shown at each branch point. Numbers above the branch demonstrate bootstrap support from 1000 replications. The model of nucleotide substitution used corresponds to the method of "maximum composite likelihood", The scale used to represent the branch lengths is expressed in the same unit as that used to calculate the distances between sequences. All sites of the alignment containing insertions-deletions, missing data were eliminated from the analysis (option "complete deletion"). *Theileria* spp. of cattle: *Theileria annulata*, *T. ovis* and of sheep: *T. uilenbergi* and *T. luwenshunni*, *Toxoplasma gondii* and *Plasmodium falciparum* were used as outliers.

More than 111 valid *Babesia* species have been described so far, including six *Babesia* species or groups responsible of cattle babesiosis (*B. bigemina*, *B. bovis*, *B. ovata, B. divergens*, *B. major* and *B. orientalis)*, five in sheep/goat (*B. ovis*, *B. motasi*, the *B. motasi*-like group in China, *B. crassa*, *Babesia* sp. Xinjiang) and several identified or unidentified *Babesia* species of wild ruminants (*B. capreoli* in roe deer, *B. odocoilei* in deer) (Ristic, 1988; Uilenberg, 2006; Malandrin et al., 2010; Holman et al., 2003; Schnitter et al., 2012; Guan et al., 2001; Liu et al., 2007). A phylogenetic tree based on the 18s rDNA gene of *Babesia* species of ruminants (domestic and wild) (figure 1) highlights that *Babesia* species infecting ruminants belong to several different clades, and that ovine, bovine or cervid infecting species are often present as sister clades (*B. major*/*B. crassa*, *B. bigemina*/*B. motasi*, *B. bovis*/*B. ovis*, *B. divergens*/*B. capreoli*) (Lack et al., 2012; Schnittger et al., 2012). *Babesia* species infecting sheep are scattered in 4 different clades as indicated on figure 1. The chinese ovine *Babesia* isolates (*Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu, *Babesia* sp. Madang, *Babesia* sp. Liaoning and *Babesia* sp. Hebei) mainly belong to a phylogenetic group closely related to *B. motasi* sheep parasites described in Europe. Since species or genus status of members of this group is not clearly defined, they will be referred to as *B. motasi*-like. Only *Babesia* sp. Xinjiang clusters in another phylogenetic clade together with recently molecularly described parasites from ruminants in South Africa (Oosthuizen et al., 2009).

3. *Babesia* spp. infecting sheep and goat

Ovine babesiosis is one of the most important tick-borne haemoparasitic disease of small ruminants in tropical, subtropical and temperate regions. Compared with babesiosis of cattle, this disease in sheep and goat seems to be somewhat neglected. However, the economic losses in sheep and goat production due to babesiosis are significant in tropical and subtropical areas and even higher than babesiosis of cattle in some countries as China and Iran, where it therefore increasingly attracts more attention (Mehlhorn and Schein, 1984; Luo and Yin 1997; Ahmed et al., 2002; Bai et al., 2002; Yin et al., 2002a, 2004).

3.1. Discovery of ovine *Babesia* species

Five *Babesia* isolate, species or phylogenetic group responsible of small ruminant babesiosis have been described around the world.

B. ovis was one of the first *Babesia* species to be discovered (1888). *B. motasi* discovery and description in 1926 was then confirmed by many studies. Two others parasites have also been described: *B. taylori (Piroplasma taylori)* in goat in India in 1935 and *B. foliate* in 1941 in sheep in India also and, but the validity of these two species is doubtful and further research has not been published by now (Ristic, 1988; Yin et al., 1997a; Guan et al., 2001; Bai et al., 2002; Schnittger et al., 2003; Uilenberg, 2006; Liu et al., 2007). About one hundred years after *B. ovis* first description, *B. crassa* as a new species was reported from Iran (Hashemi-Fesharki and Uilenberg, 1981).

Reports of ovine babesiosis in China and identification of the etiological agents are recent (table 2). *B. ovis* was first reported but never identified again in later reports in China, casting doubts on the species identification. Several sheep parasites were isolated, mainly by sheep experimental infection with blood from naturally infected sheep. They cluster in two different phylogenetic groups and their status as two or more new species remains to be clarified (see below, part 3.3).

Table 2. History of *Babesia* species of sheep and goats in China.

Years	Species	Location (Province)	Reference
1982	*B. ovis*	Sichuan	Chen, 1982
1985		Heilongjiang	Zhao et al., 1986
1997	*Babesia* sp. BQ1 (Ningxian)	Gansu	Lian et al., 1997; Yin et al., 1997b,
2001	*Babesia* sp. Xinjiang	Xinjiang Uygur Autonomous Region	Guan et al., 2001
2002	*Babesia* sp. BQ1 (Lintan)	Gansu	Guan et al., 2002
2007	*Babesia* sp. Tianzhu, Madang, Hebei, Liaoning	Gansu, Hebei and Liaoning	Liu et al., 2007

3.2. Biological characteristics

B. ovis **(Babes, 1888; Starcovici, 1893)**

B. ovis is considered as a small *Babesia* according to its morphology (figure 2). It exhibits a high degree of polymorphism: a single pyriform, round ring, double pyriform, rod-shaped, oval-shaped and cross-shaped. Round ring parasites are mostly located in the edge of red blood cells. Pyriform parasite length is from 1.0 to 2.0 x 0.5 microns.

Sometimes, binary fission of the parasite in red blood cells can be seen. *B. ovis* has been isolated from blood of infected sheep. This species is considered as highly pathogenic for sheep and causes severe infections. The rate of mortality in sheep is 30 to 50% in field infections in Iran (Hashemi-Fesharki, 1997). However, it is rarely pathogenic for goats.
The description of *B. ovis* as a mixture of small and large forms of *Babesia* has been described recently in Iran (Shayan et al., 2008). In this study, the authors demonstrated the confusion between *B. ovis* large forms and *B. motasi* as well as a difference in pathogenicity with milder clinical symptoms when infections were performed with the large form of *B. ovis* compared with the small one. A subsequent molecular survey on asymptomatic sheep revealed that *B. ovis* is the main cause of ovine babesiosis in this country (Ranjbar-Bahadori et al., 2012).Three species of *Rhipicephalus* are able to transmit *B. ovis*: *Rhipicephalus bursa*, *R. turanicus* and *R. sanguineus*. *B. ovis* has been described in Asia, Europe and Africa (Ristic and Kreier, 1981; Friedhoff, 1997; Yeruham et al., 1998; Razmi et al., 2002; Uilenberg, 2006; Shayan et al., 2008; Altay et al., 2008; Ekici et al., 2012).

B. motasi **(Wenyon, 1926)**

B. motasi belongs to the large *Babesia,* with a typical double pyriform parasites size; measuring 2.5 to 4.0 x 2.0 microns (figure 2). It is described as double pyriform, single pyriform, round ring, oval, three-leaf-shaped and irregularly shaped. The characteristic form is three-leaf-shaped. Some published data indicate that *B. motasi* is not or poorly pathogenic for intact sheep and is not infective for goats (Alani et al., 1988a, b; Uilenberg et al., 1980), while some describe that *B. motasi* shows high pathogenicity with death in sheep (Jagannath et al., 1974; Razmi et al., 2002). It is thus possible that the pathogenicity of *B. motasi* from different geographical regions differ in their virulence. A recent report described the pathogenicity of *B. motasi* in detail and the *B. motasi* species might be consisted of two species or subspecies according to their pathogenicity, infectivity to sheep

and goat, serology and morphology: one has a low pathogenicity in northern Europe, while one has high pathogenicity in southern Europe and the Mediterranean basin (Uilenberg et al., 1980; Lewis et al., 1981). Finally, *Haemaphysalis punctata* is the transmission vector of *B. motasi* (Friedhoff, 1997; Uilenberg, 2006).

***B. crassa* (Hashemi-Fesharki and Uilenberg, 1981)**

B. crassa is a large parasite but smaller than *B. motasi*. The length of the pyriform body was measured 2.2 to 2.5 (± 0.43) microns (figure 2). Its morphology is mainly pear-shaped and one of its characteristics is the frequent occurrence of four merozoites in one erythrocyte caused by quadruple division or two successive binary divisions. It was initially isolated from sheep in Iran. This parasite is low or no pathogenic for sheep and goats. The tick vector and the mode of transmission remain unknown (Hashemi-Fesharki and Uilenberg, 1981; Friedhoff, 1997; Uilenberg, 2006).

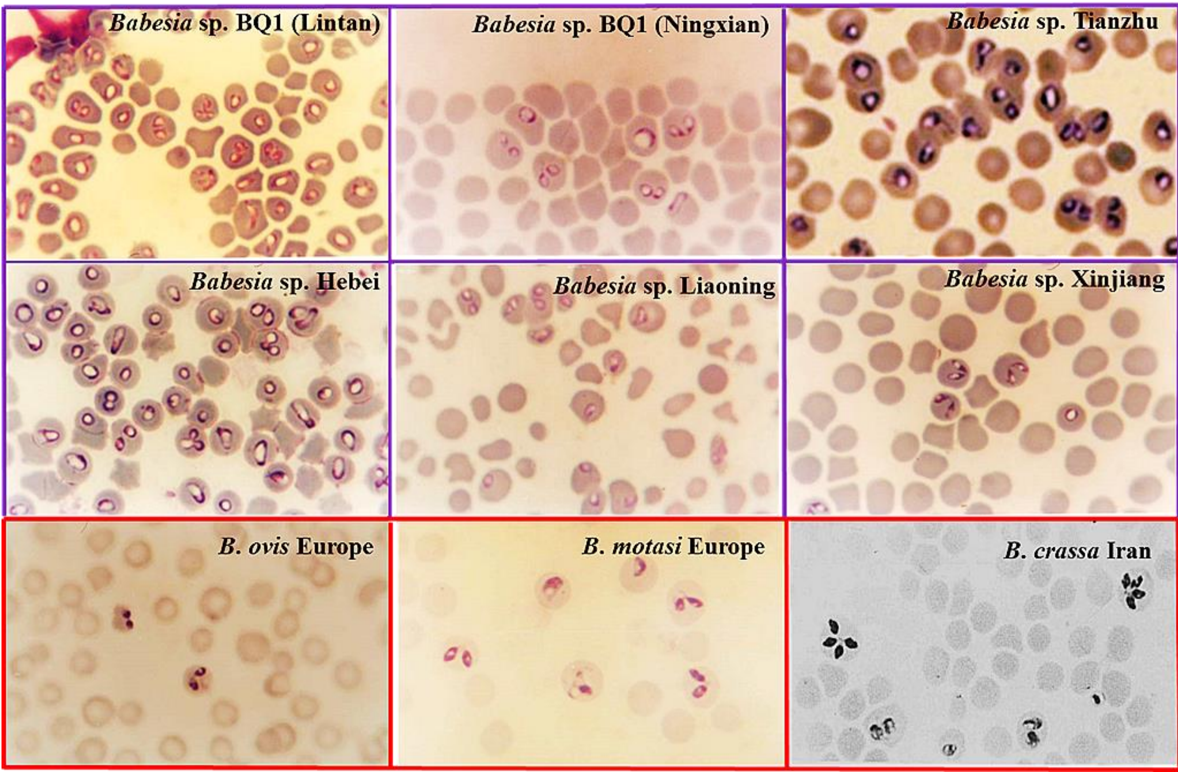


Figure 2. Morphological characteristics of different species of *Babesia* infecting small ruminants in China, Europe and Iran. The blood smears of China's strains were provided by the laboratory VVBD, LVRI, CAAS, China. The stained slides of European species were provided by the Faculty of Veterinary Medicine, Utrecht University (Netherlands). The strained slides of B. crassa was cited from paper published by Hashemi-Fesharki and Uilenberg in 1981.

B. motasi-like (In China) (Yin et al., 1997a; Lian et al., 1997; Bai et al., 2002; Guan et al., 2002; Liu et al., 2007)

The 6 isolated parasites of the group *B. motasi*-like showed a similar morphology with European *B. motasi* isolates.

Babesia sp. BQ1 (Ningxian) was isolated from diseased sheep bloods, and was characterized by a similar morphology with *B. motasi* and a strong virulence in sheep and goats. *Haemaphysalis longicornis* is clarified as its transmitted vector (Yin et al., 1997b; Lian et al., 1997; Bai et al., 2002). *Babesia* sp. BQ1 (Ningxian) isolate is strongly pathogenic associated with a parasitaemia as high as 23.7% and with a high mortality rate for sheep and goats. *Haemaphysalis longicornis* has been indicated as its transmitted vector (Bai et al., 2002). *Babesia* sp. BQ1 (Ningxian) isolate is a large *Babesia*, slightly smaller than *B. motasi* in Europe, with a wide variety of morphological forms in the infected erythrocytes (figure 2). Double pyriform is the typical forms, with the size of 1.8 to 2.5 × 0.9 to1.8 microns, with mean dimensions of 2.21 (±0.12) ×1.17 (±0.18) microns and with a frequency of about 36.1%. The others forms include ring form, with a ratio of 5.4%; single pyriform, with a frequency around 16.9%. three-leafed shaped formed from original double pyriform parasite, with a proportion about 11.5%; rod-shaped showing the same diameter as the red blood cells with about 6.0% in the total number of parasites, irregularly shaped which is large and dominates the entire erythrocyte with the proportion of about 17.4%, and rare comma-shaped and budding forms (Bai et al., 2002).

Babesia sp. BQ1 (Lintan) was isolated from splenectomized sheep experimentally infected with *H. qinghaiensis* collected in pastures. The parasites show moderate virulence for sheep and goats (this isolate was confused with *B. ovis* in the publication of Guan et al., 2002). *H. qinghaiensis* and *H. longicornis* are potential transmitting vector ticks (Guan et al., 2010b). *Babesia* sp. BQ1 (Lintan) isolate has similar size than *Babesia sp*. BQ1 (Ningxian) isolate, with the most common forms of single pyriform (25% in total parasite), round and oval (27%), and some various forms with double pyriform (24%), budding form (7%) and elongated shape (17%) (table 3). *H. qinghaiensis* and *H. longicornis* have been recently described as the vector of *Babesia sp*. BQ1 (Lintan) isolate (Guan et al., 2010a).

Babesia sp. Tianzhu, Madang, Liaoning and Hebei were isolated by inoculating the blood of the asymptomatic sheep from the field where *Haemaphysalis* are the dominant tick species, to the splenectomized sheep experimentally (Liu et al., 2007). Their transmitted vector remains unknown.

***Babesia* sp Xinjiang (In China) (Guan et al., 2001)**

Babesia sp. Xinjiang has been recently isolated from sheep experimentally infected with *R. sanguineus* and *Hy. a. anatolicum* collected from pastures in China (Guan et al, 2001). The pathogenic and morphological characteristics of this species are totally different from the other *Babesia spp*. of small ruminant previously reported in the world and in China. It is a large parasite with variously morphological forms in infected erythrocytes: single and paired piriform are the most common shapes, ring form, three-leafed-shaped, rod shaped, budding forms and oval forms. The size of single piriform was measured with average being of 2.64 (±0.37) × 1.43 (±0.21) microns and the paired piriform with the mean dimensions of 2.42 (±0.35) × 1.61 (±0.22) microns. It is weakly virulent for sheep since no detectable clinical symptoms have been observed in spleen intact sheep. The demonstrated vector is *Hy. a. anatolicum*.

Table 3. *Babesia* species of small ruminant in the world and isolates of *Babesia* infecting sheep and goats in China (Yin *et al*., 1997a; Lian *et al*., 1997; Bai *et al*., 2002; Guan *et al*., 2001, 2002, 2009, 2010a; Liu *et al*., 2007; Uilenberg, 2006; Shayan *et al*., 2007; Razmi and Nouroozi., 2010; Hashemi-Fesharki and Uilenberg, 1981)

Species	Morphological characteristics	Isolation	Transmitted vector ticks	Virulence	Geographical distribution
B. motasi	Large (2.5-4.0×2.0 μm), double pyriform, single pyriform, three leafed.	Blood of infected sheep	*Haemaphysalis punctata* *Rhipicephalus bursa* ?	Variable	Asia, Europe, Africa
B. ovis	Small (1-2×0.5 μm), double pyriform, single pyriform, three leafed and irregular.	Blood of infected sheep	*Rhipicephalus bursa,* *R. turanicus* *R. sanguineus*	Variable	North of Africa, East of Asia, Southeast of Europe
B. crassa	Large (2.2～2.5±0.43 μm),, four parasites in a RBC, oval or round after division.	Blood of infected sheep	unknown	Low	Asia
B. motasi-like	Large, slightly small than B. motasi European isolate (2.21 (±0.12) × 1.17(±0.18) μm), single pyriform, round and oval, double pyriform, budding form.	Blood of infected sheep	Ningxian: *H. longicornis* Lintan: *H. qinghaiensis*, *H. longicornis* Others: Unknown	Ningxian: Strong Lintan: Low Others:Unknown	Gansu, Hebei and Liaoning provinces in China
Babesia sp. Xinjiang	Large, double pyriform (2.42 (±0.35) × 1.61 (±0.22)), single syriform (2.64 (±0.37) ×1.43 (±0.21))	Adult semi-engorged of *Hyalomma anatolicum anatolicum* in pasture	*Hy. a. anatolicum*	Low	Kashi city, Xinjiang Autonomous Region in China

3.3. Phylogenetic positions of the small ruminants *Babesia*

Phylogenetic analysis using 18S rDNA gene reveals that the five above-mentionned species or group form distinct phylogenetic groups. *B. ovis* strains from all over the world and *B. crassa* each form homogeneous clades, with strong bootstrap values (figure 3).

For the isolates from China, the morphology of *Babesia* sp. Xinjiang was different from that of the other chinese *Babesia* isolates (Liu et al., 2007; Guan et al., 2009). Molecular taxonomy and phylogenetic reconstruction based on 18S rDNA gene and ITS rDNA confirmed this biological difference, placing *Babesia* sp. Xinjiang in a clade clearly separated from all the other small ruminants *Babesia* described so far (Liu et al., 2007; Niu et al 2009a). *Babesia* sp. Xinjiang is phylogenetically closely related to other *Babesia* recently described from wild ruminants in South Africa and with *B. pecorum* isolated from red deer in Spain (Guan et al., 2001; Oosthuizen et al., 2009; Jouglin et al., 2014 in press).

B. motasi can be divided with strong statistical support in two clades, one with the European isolates, and the second with the recently described Chinese isolates (Yin et al., 1997a; Lian et al., 1997; Bai et al., 2002; Guan et al., 2002; Liu et al., 2007). The naming of members of this group as *B. motasi*-like reflects their relatedness to this species, but they form a phylogenetic group rather distinct from *B. motasi.*

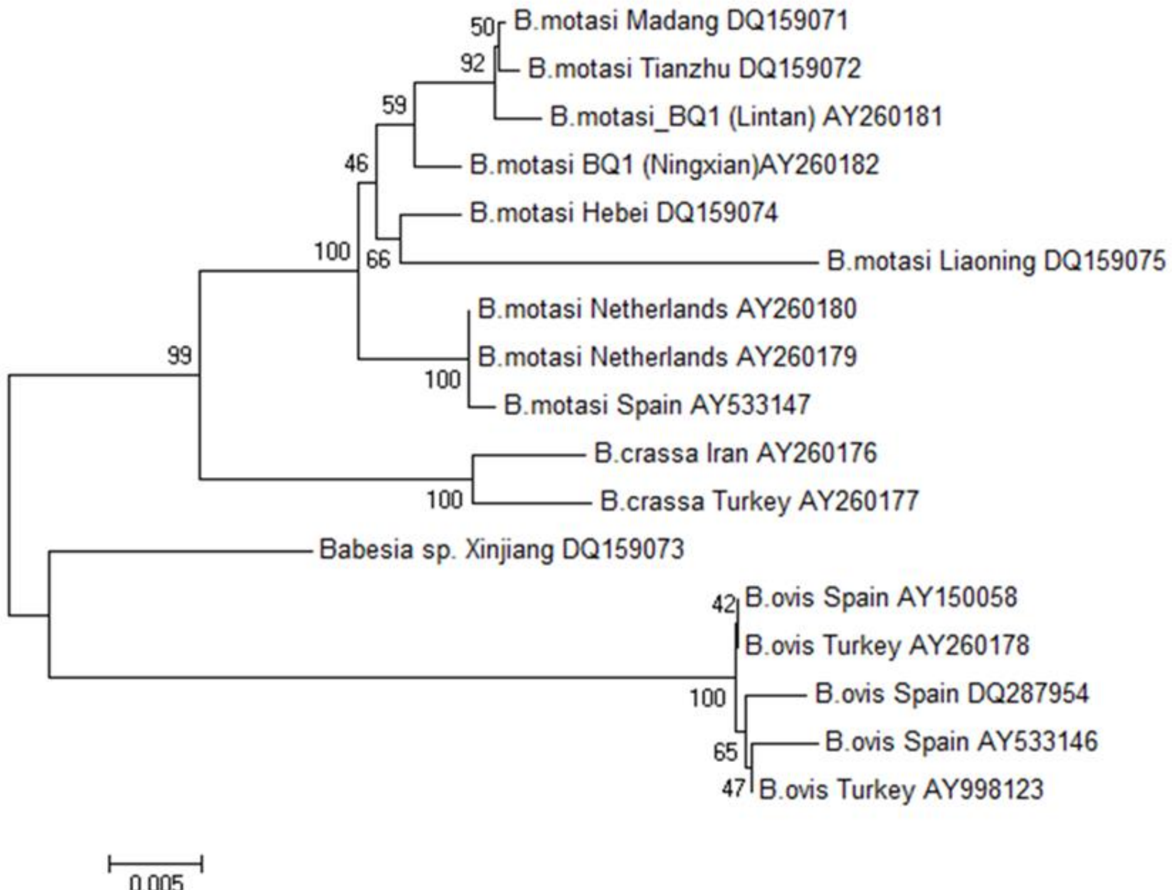


Figure 3. Phylogenetic analysis of *Babesia* species identified of the small ruminant based on the sequences of the 18S rRNA gene. The tree was inferred using the neighbor joining method of MEGA 5.2, bootstrap values are shown at each branch point. Numbers above the branch demonstrate bootstrap support from 1000 replications. The model of nucleotide substitution used corresponds to the method of "maximum composite likelihood", The scale used to represent the branch lengths is expressed in the same unit as that used to calculate the distances between sequences. All sites of the alignment containing insertions-deletions, missing data were eliminated from the analysis (option "complete deletion ").

Recently, some new molecular markers including mitochondrial ones, as well as more refined phylogenetic tree reconstruction were used to clarify the phylogeny of Piroplasma species or isolates of ruminants in China. Several selected genes or spacers (18S rDNA, 28S rDNA, ITS and COI) have been recently studied as candidates for a DNA barcode to identify species of the order Piroplasmea. ITS2 has 98% and 92% identification efficiency at the genus and species level respectively. The ITS2 could therefore be an ideal DNA barcode for the identification of piroplasma (Gou et al., 2012).

Within the *B. motasi*-like phylogenetic group, the relationships between the different isolates is far from being clear and the status of the members of the *B. motasi*-like group as one species is even doubtfull. The topology of the 18S rDNA based tree within this group is not strongly supported, but two isolates, *Babesia* sp. Hebei and *Babesia* sp. Liaoning seem to diverge. The 28S rDNA phylogenetic reconstruction also groups *Babesia* sp. BQ1 (Lintan) with *Babesia* sp. Tianzhu isolates in the same clade, but separates *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Hebei from this group (Gou et al., 2013). More markers were also used, RPS8 gene (Tian et al., 2013a), cytochrome b gene (Tian et al., 2013b) and the cytochrome C oxidase subunit III (Tian et al., 2013c) (but *Babesia* sp. Hebei was not included in these three studies) and ITS (Niu et al., 2009a) (*Babesia* sp. Liaoning was not included in this study). Using these markers, no clear pattern of isolates grouping could be highlighted, except the close relatedness of *Babesia* sp. BQ1 (Lintan) and *Babesia* sp. Tianzhu.

3.4. Diagnostic methods developed to detect and identify ovine *Babesia* species

Clinical observation can support a provisional diagnosis of the disease. Thin or thick smears stained with Giemsa can confirm the presence of *Babesia*, especially from peripheral blood, and the diagnosis is based on the morphological examination. Usefull to confirm acute cases, this method has limited value for chronic cases, where only low numbers of piroplasms exist. Moreover, even with a significant degree of expertise, the species differentiation remains challenging (refer to *B. ovis* co-existence of small and large forms that were first identified as *B. motasi*).

Serological diagnostic tests such as indirect haemagglutination, complement fixation tests and ELISA have been established for a number of *Babesia* species, with the drawback

of serological cross-reactions between closely related species (Bakheit et al., 2007). The first serological diagnosis test of *B. ovis* is IFAT (Ozkuc, 1979). It requires few reagents and moderate skill. This method is also currently and widely being used in some countries to detect *B. ovis* (Ferrer et al., 1998a, b, c; Sevinc et al., 2007, 2013a; Ekici et al., 2012). The specificity of IFAT method often seems not satisfactory on the species level of same genus (Papadopoulos et al., 1996a), except in few reported studies (Hasherni-Fesharki and Uilenberg, 1981; Habela et al., 1990). For example, the IFAT serological detection of *B. crassa* (12.6% of the sheep sera) and *B. motasi* (10.5% of the sheep sera) in the Macedonia region of Greece was attributed by the authors to cross-reactions with *B. ovis* (Papadopoulos et al., 1996a and b) and the same problem was also reported in earlier studies (Lewis et al., 1981; Kjemtrup et al., 1995).

The development of ELISA tests requires preliminary knowledge on parasite polymorphism, and/or in vitro culture achievement to produce antigens, and/or prior genetic knowledge on the targeted antigen to select specific peptidic sequences as antigen. ELISA tests has been developed for the detection of *B. ovis* using a recombinant *B. bovis* derived antigen (Duzgun et al., 1991, Sevinc et al., 1996; Emre et al., 2001; Cicek et al., 2004), and research for new specific antigens is being performed (Sevinc et al., 2013b; Guan et al., 2012c). ELISA tests specific of *Babesia* sp. BQ1 (Lintan) and *Babesia* sp. Xinjiang were developed using in vitro produced merozoite extracts (Guan et al., 2010c, 2012a). The first test also revealed cross-reactions with another chinese ovine parasite, *Babesia* sp. Tianzhu (Guan et al., 2010c).

In contrast to these methods, the application of molecular genetic tests allows direct, specific and sensitive detection of parasites as well as rapid, simultaneous detection and differentiation of co-infecting piroplasms in a given animal. Different PCR-based molecular diagnostic tools have been established and the reference gene for these applications is the well-characterized 18S rDNA (Criado-Fornelio A. 2007). PCR is a highly specific and sensitive method for the diagnosis of *Babesia* infections and widely used to detect and identify *B. ovis* and *B. motasi* (Aktaş et al., 2005; Theodoropoulos et al., 2006; Shayan et al., 2007; Razmi et al., 2013). In China, a nest-PCR was successfully developed to detect *B. motasi*-like from other Piroplasma with same transmitted vector ticks (Sun et al., 2008). However, there are often false negative results due to the presence of inhibitor of DNA polymerase in the sample. The LAMP (loop-mediated isothermal amplification) technic has been recently used for detection and identification different *Babesia* or *Theileria* species in China (Liu et al., 2012, 2013; Xie et al., 2013), since

compared to PCR, it has advantages of high specificity, sensitivity and efficiency, to diagnose disease quickly and accurately. This method has been established to detect *Babesia* spp. from sheep and goats in China and the result demonstrated that *B. motasi*-like can be distinguished from *Babesia* sp. Xinjiang (Guan et al., 2008). Recently, PCR-based reverse line blotting (RLB) has been used for detection and discrimination all the species of *Babesia* and *Theileria* species (Gubbels et al., 1999; Schnittger et al., 2004). This method was also developed to detect sheep *Babesia* species: *B. motasi*-like in China (Niu et al., 2009b), and *B. ovis* in Iran (Ranjbar-Bahadori et al., 2012; Iqbal et al., 2013). The advantage of PCR-RLB is that multiple samples can be analyzed against multiple probes to enable the simultaneous detection of several species from each other (Rijpkema et al., 1995; Schouls et al., 1999).

All these tests allowed the identification and survey of ovine babesiosis all around the world.

3.5. Geographical distribution of ovine babesiosis

B. ovis as well as *B. motasi* have been described in Asia, Europe and Africa (Uilenberg, 2006). However due to the recent description of *B. ovis* as also having isolates with large forms (Shayan et al., 2008), the molecular confirmation of this distribution is needed, due to potential morphological confusion between *B. motasi* and *B. ovis*.

B. crassa was initially isolated from sheep in Iran (Hashemi-Fesharki and Uilenberg, 1981), and the analysis of a *Babesia* sp. sequence from Turkey revealed homology with *B. crassa* (Schnittger et al. 2004). The presence of *B. crassa* in sheep in Iran and Turkey was not confirmed in recent studies that analyzed by Reverse Line Blot the blood of asymptomatic sheep (Altay et al., 2012; Ranjbar-Bahadori et al., 2012). But the RLB analysis of ticks collected from humans in Turkey confirmed the presence of this species (Aktaş, 2014).

The cases of ovine babesiosis in China were reported in many provinces, such as Inner Mongolia, Henan, Shannxi, Yunnan, Sichuan, Heilongjiang and Qinghai (Yang et al, 2003; Li et al, 2006; Zhang et al, 2010) (figure 4). The isolates of *B. motasi*-like have been first isolated and described in the provinces of Gansu, Hebei and Liaoning provinces. *Babesia* sp. Xinjiang has originally been described only in the province of Xinjiang (figure 4) (Guan et al, 2009).

Sero-epidemiological investigations by ELISA for the infections of small ruminants by *Babesia* species have indicated infection rate average of 66.84% for *Babesia* sp. BQ1 (Lintan) in 974 sera samples from Gansu province (Guan et al., 2010a), and 31.66% of *Babesia* sp. Xinjiang in 22 province from 3857 sera samples (Guan et al., 2012a). Another recent large-scale sero-epidemiology survey of *Babesia* sp. BQ1 (Lintan) in small ruminants from these 22 province of China by ELISA assay indicated that infection by this isolate were found in all investigated 22 provinces with significantly different positive rates from 1.6% (Jilin), for *Babesia* sp. Xinjiang, no positive sera were detected in this province to 91.0 % (Inner Mongolia), with and average prevalence of 43.5 % (Wang et al., 2013). However, this prevalence not only includes *Babesia* sp. BQ1 (Lintan) but certainly also other related parasites, due to the known serological cross-reactions with *Babesia* sp. Tianzhu, and certainly with other not yet identified isolates. In the same two provinces (Jilin and Inner Mongolia), *Babesia* sp. Xinjiang prevalence was 0 % and 43.07% respectively (Guan et al., 2012a). Therefore, these two phylogenetically unrelated parasites co-circulate in China, with a rather important prevalence in China.

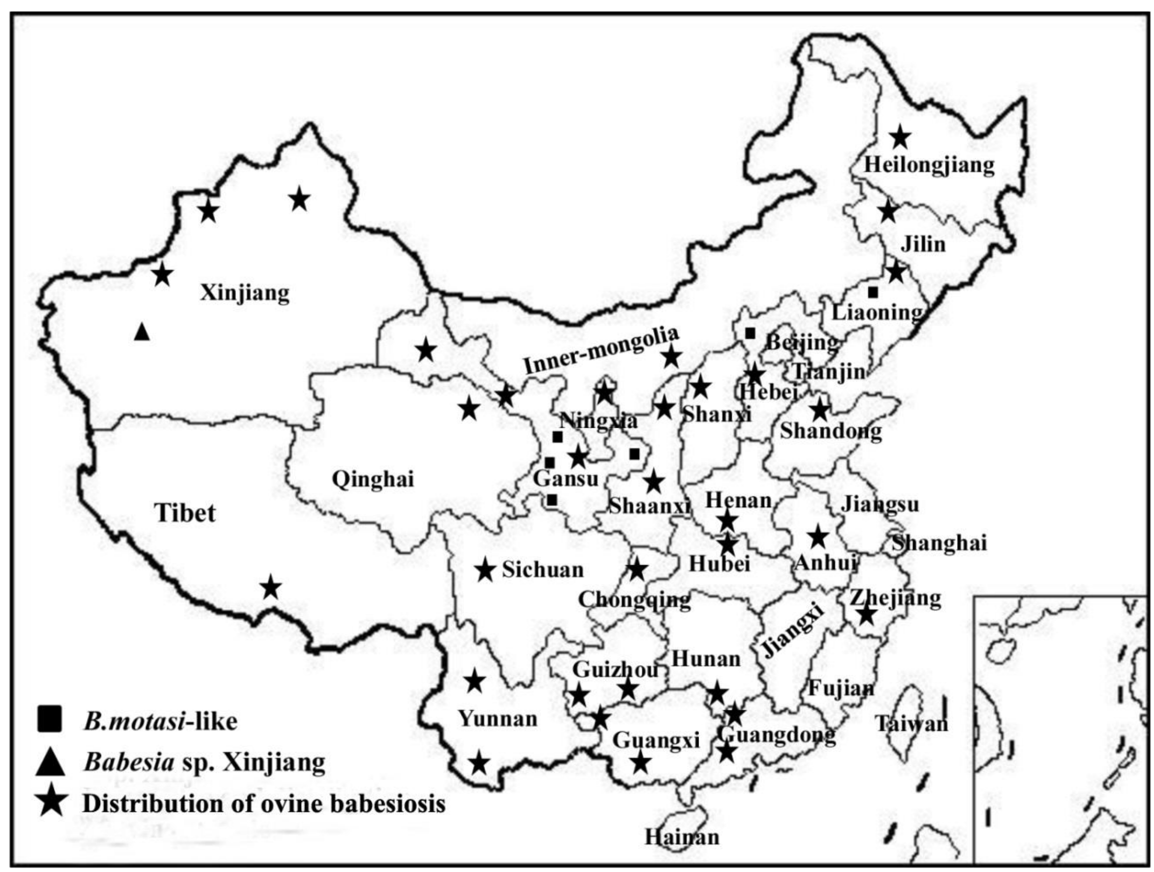


Figure 4. Geographical distribution of small ruminant *Babesia* spp. in China. (Chen, 1982; Zhao *et al*., 1986; Guan *et al.*, 2001, 2002, 2012a; Bai *et al*., 2002; Yang *et al*., 2003; Li *et al*., 2006; Liu *et al*., 2007; Zhang *et al*., 2010; Wang *et al*., 2013).

B. Life cycle of *Babesia* sp. and the red blood cell invasion

1. Life cycle of *Babesia* sp.

The life cycle of *Babesia* parasite is rather complex and requires two hosts, an invertebrate tick vector and a vertebrate host (figure 5). In the vertebrate host, the parasite multiplies asexually in the erythrocytes (merogony). In its vector, *Ixodidae* tick, sexual reproduction occurs (gametogony or gamogony) as well as asexual multiplication in different organes including the salivary glands (sporogony). During their blood meal, the ticks introduce the sporozoites present in their salivary glands into the vertebrate host. Sporozoites directly enter in erythrocytes in the blood, a distinctive feature of this genus. Other Apicomplexa either develop first in the hepatocytes (*Plasmodium*) or in the lymphocytes (*Theileria*). Within the erythrocyte, the parasite develops into a trophozoite that divides to give 2 to 8 merozoites liberated in the blood stream by the red blood cell rupture. At a low frequency, some merozoites develop into gametocytes, the precursor forms of gametes. These gametocytes do not develop further until they are taken up by the ticks, where they develop into gametes in the midgut. Sexual reproduction occurs in the midgut, the gametocytes have changed environment during the passage from vertebrate host blood stream to the midgut of vector ticks to further develop into gametes, which fuse to form a zygote (spherical cell). The zygote then infects digestive tick cell to develop into kinetes. These kinetes invade different organs of the tick: ovary infection leads to the trans-ovarian transmission of the parasite while salivary gland infection leads to the transmission to the vertebrate host (Mehlhorn and Shein, 1984).

The pathological changes caused by *Babesia* species are similar to those seen in *Plasmodium falciparum* (Wright et al., 1988). Parasites in the genus *Babesia* are transmitted when sporozoites are released with the saliva during tick feeding (figure 5). Clinical symptoms of the disease appear when the merozoites invade and replicate within host erythrocytes, and reach high parasitemia (Wright and Goodger, 1988). Sporozoites and merozoites are two asexual forms that invade the erythrocytes of the vertebrate host.

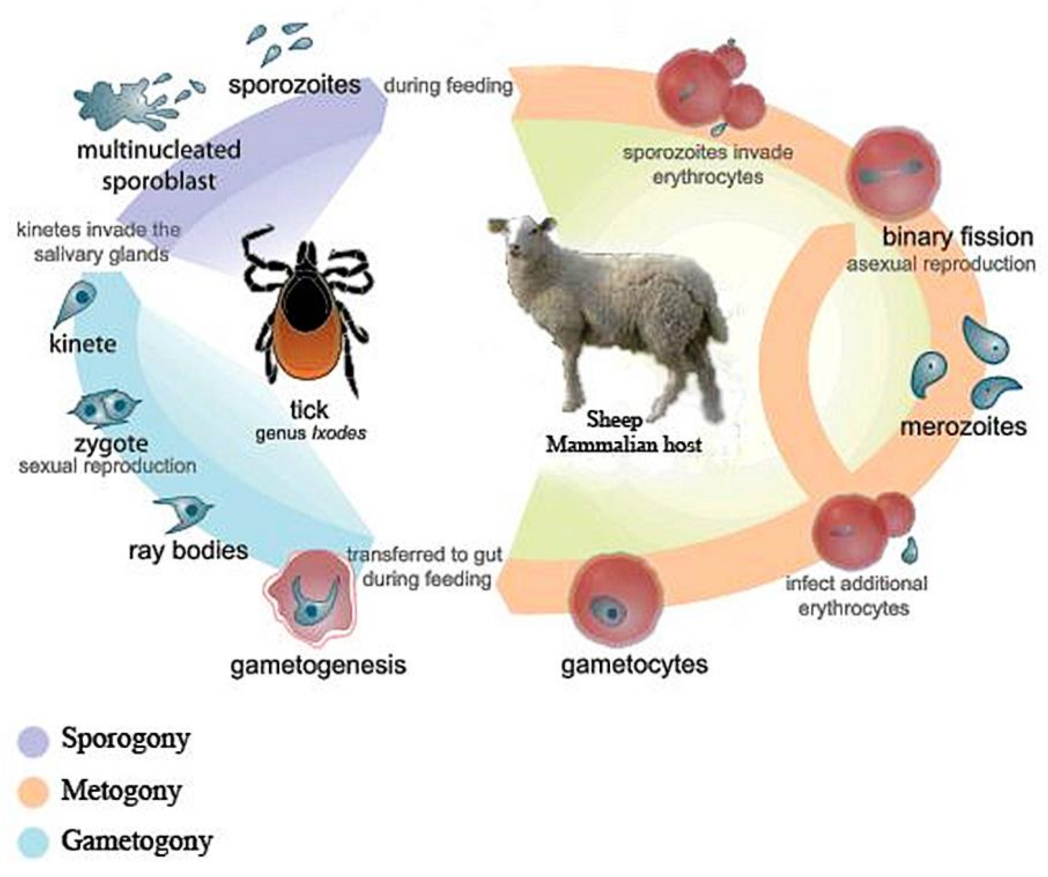


Figure 5. The development life Cycle of *Babesia* (modified from Wikipedia, the free encyclopedia. Author: LadyofHats. http://en.wikipedia.org/wiki/Babesia#cite_note-Ristic-4) (Adapted from Mehlhorn and Shein, 1984).

Protective immune response against *Babesia* sp. is mainly mediated by antibodies which block the invasion process of the extracellular merozoites and the carbon monoxide (NO) produced by IFNg-activated macrophages, which is toxic for the intracellular merozoites (Brown et al, 2006a). Once the parasites are inside the red blood cells, they are very well protected and cannot be reached by the circulating antibodies (Yokoyama et al., 2002). To prevent its elimination from the blood stream, Apicomplexan parasites have developed an efficient machinery to rapidly invade a new cell, escaping as fast as possible from the immune system.

Even if the invasion process of merozoites is rapid, it represents the only period of the life cycle when the parasite is directly exposed to the host immune system. A solution of the vaccine development based on the invasion factors is thus to target the molecules on the extracellular merozoite or those involved in the rapid process of invasion (this stage is exposed to the host humoral immune system), but early enough in the process for these

antigens to be in contact with the immune system and to induce production of IFNg and neutralizing antibodies. A number of different antigens involved in invasion process have proven to induce significant protection upon challenge infection in *Plasmodium* species (Rojas-Caraballo et al., 2009) and several *Babesia* species (McElwain et al., 1991; Brown et al., 1993; Hines et al., 1995a; Rodriguez et al., 1996; Machado et al., 1999; Hadj-Kaddour et al., 2007; Fish et al., 2008).

2. The merozoite and apical complex structure

2.1. Overview of the merozoite structure

The morphology and ultrastructure of merozoites from Apicomplexan share similar features (figure 6). They contain all organelles and molecules necessary to invade and establish itself in a new RBC. The free merozoite is very small, 1.2 µm long and is pear shaped with a pointed apical pole. Merozoites are surrounded by fibrillar material, a surface a coat, an outer cell membrane and an inner osmiophilic membrane. The coat that covers the merozoite is thick and adhesive. Apicomplexan parasites are characterized by the presence at their apical pole of a complex of organelles, called the apical complex. It is composed by three sets of secretory vesicles and a polar ring: the micronemes are small and numerous organelles, the 2 to 8 rhoptries are twin flask-shaped, and the small rounded dense granules. Despite the usual nucleus, mitochondria and microtubules, the cytoplasm contains a plastid, another specific feature of Apicomplexa (figure 6) (Sam-Yellowe, 1996; Bannister and Mitchell, 2003; Kat et al., 2008; Rodriguez et al., 2008).

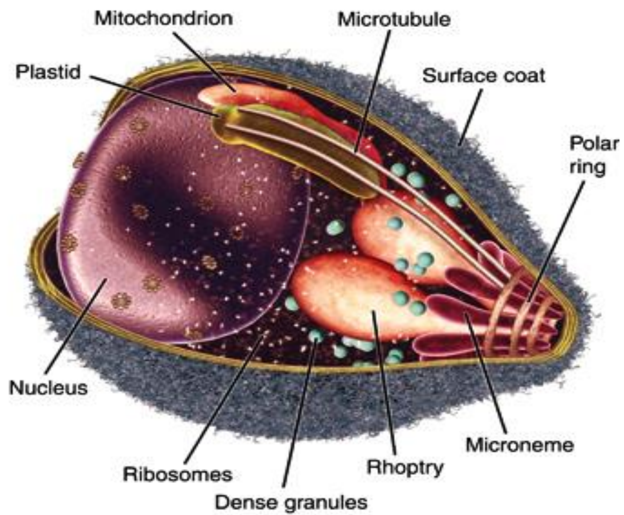


Figure 6. Cellular structure of a merozoite and major organelles of the apical complex of *Plasmodium falciparum* (Cowman and Crabb, 2006).

2.2. Detailed description of organelles involved in the invasion process

Micronemes

Micronemes vary in shape and numbers according to the parasite species. They are generally diminutive, form fusiform sacs which are very elongated about 120 nm×40 nm and attached to the rhoptry canal. They disappear in the course of invasion, due to triggering calcium ion discharge and the release of the contents of the micronemes, which then mediate parasite attachment. Their contents are then translocated towards the posterior end of the parasite and shed. Micronemes are thought to be formed by budding from the Golgi apparatus (Bannister et al., 2000; Preiser et al, 2000; Kats et al., 2008).

Rhoptries

Rhoptries are membrane-bounded, twin club-shaped organelles with a number varying according to the parasite species (2 to 8) and stage of life cycle. They are the largest organelles of the apical complex with typical bilayer structure of their membrane. The size of rhoptries is variable from diminutive flask-shaped in *Babesia* (*B. caballi*) to the very elongated rhoptries of about 650 nm × 300 nm in *Plasmodium*. They consist of two

distinct parts: a basal bulb which is composed of homogeneously packed granular materials and a narrowing rhoptry duct which consists of fibrillar, granular and some membranous ingredients. When their contents are discharged after adhesion of the merozoite to the host cell surface during invasion process, rhoptries change their conformation to become irregular, with the most of complex lamellar, or accompanied with some membranous vesicles, tubules and vesicular structures. The duct parts combine with each other as well as with the plasma membrane of merozoite. Finally, the packed materials of rhoptries are converted from granular to lamellar contents that will form the parasitophorous vacuole (PV) containing the merozoite, internalized within red blood cell (Blackman and Bannister, 2001; Preiser et al, 2000).

Spherical bodies (Dense granules in *Plasmodium*)

These organelles are rounded membranous vesicles. They are located between the rhotries bulbs and the nucleus, freely lying within the cytoplasm of the merozoite. The size of the dense granules is smaller than microneme and rhoptries with about 80 nm diameter. They contain similar contents (packed granules) than rhoptries. They move to the surface of merozoite after its internalization in the host red blood cell, and release their contents by exocytosis from the parasite's plasma membrane into the parasitophorous vacuole (PV). According to the parasite genus, they are either called spherical bodies (*Babesia*) or dense granules (*Plasmodium*), in relation with their density in electron microscopy (Bannister et al., 2000; Preiser et al, 2000).

3. Detailed description of the invasion process

The process of erythrocyte invasion by *Babesia* parasites (Igarashi et al, 1988; Yokoyama et al., 2006; Lobo et al., 2012) is considered to be similar to *Plasmodium*'s which has been well studied and reviewed (Bannister and Mitchell, 2003; Preiser et al., 2000). Several steps in this process have been described, namely attachment (figure 7, 1), re-orientation (figure 7, 2), tight junction formation (figure 7, 3), invasion (figure 7, 4), and internalization (figure 7, 5) (Yokoyama et al., 2002, 2003, 2006; Kumar et al., 2004; Rodriguez et al., 2008). One major difference between *Plasmodium* and *Babesia* is the maintenance (*Plasmodium*) or disappearance (*Babesia*) of the parasitophorous membrane once the parasite is inside the erythrocyte (dotted line on figure 7) (Yokoyama et al., 2006).

However, the destruction of the PV in the genus *Babesia* still needs to be confirmed using more advanced imaging techniques before the complete absence of a PV membrane (PVM) in infected RBCs could be concluded (Gohil et al., 2010).

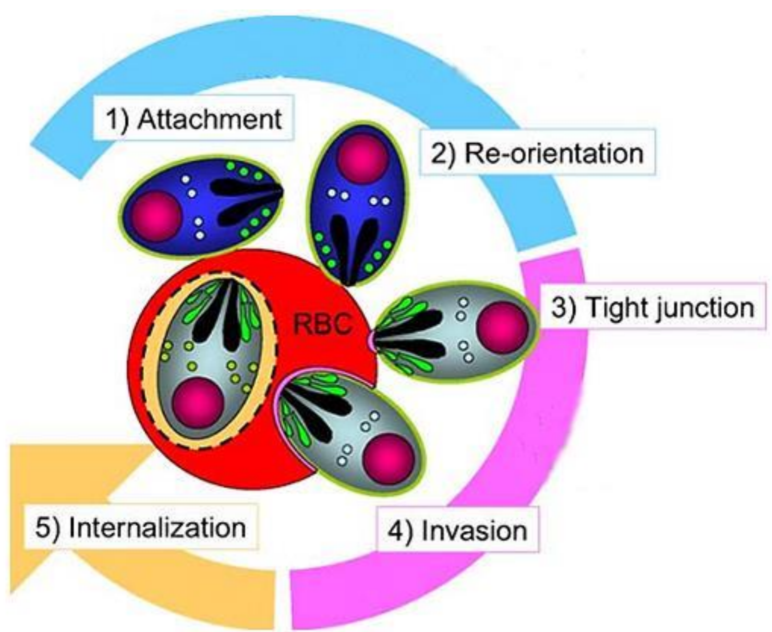


Figure 7. Schematic representation of erythrocyte invasion process: example of *Babesia bovis* (Yokoyama et al., 2006).

3.1 Attachment

The merozoites are covered with a protein surface coat and roll over the surface of red blood cell. One of these proteins belongs to the Variable Merozoite Surface Antigen family (VMSA), which are shed in the course of the invasion process and released in the surrounding medium (As depicted in figure 7, 1 showing the disappearance of the green surface coat). These VMSAs are responsible for the initial recognition and attachment with the erythrocyte surface in any orientation. This first attachment is nonspecific and reversible, mediated apparently by weak electrostatic interactions (Yokoyama et al., 2006).

3.2 Apical re-orientation

To achieve invasion, the merozoite must re-orientate and become perpendicular to the red blood cell surface, so that its apical end comes in contact with the surface of red blood

cell. The reorientation process is maintained with low-affinity contacts, once the apical end is adjacent to the erythrocyte. This proximity will allow the creation of the tight junction between the two membranes of the erythrocyte surface and the parasite apical part (Yokoyama et al., 2006).

3.3 Tight junction formation: irreversible attachment

The microneme organelles exocytosis first occurs, then the apical part binds to the host cell and the moving junction, which is a ring-like proteic structure, is formed involving high-affinity ligand-receptor interactions between parasite and host cell. Once the tight junction is formed, the merozoite is committed to the active invasion process. The two membranes become associated and the junction moves backward from the anterior to the posterior end of the parasite. This moving junction resulted in its encapsulation in a protective vacuole within the host cell. The assembly of a functional moving junction is thus crucial to parasite survival (Suss-Toby et al., 1996; Yokoyama et al., 2006).

3.4 Release of the contents of the apical complex organelles accompanied with invasion

The contact of the parasite apical complex with the red cell surface triggers the sequential secretion of the contents of the micronemes (AMA-1 and TRAP) and rhoptries (RAP-1). Rhoptries are then exocytosed after micronemes, while the moving junction glides backward. Rhoptry material is integrated in the parasitophorous vacuole membrane (Yokoyama et al., 2006).

3.5 Internalization

When the tight junction reaches the posterior pole of the parasite powered by the parasite's actin-myosin motor, the surface coat is completely shed by serine protease or sheddase and the parasite is eventually enclosed within the parasitophorous vacuole (PV) consisting mainly in red cell membrane components. During the final phase of invasion, spherical bodies (dense granules) are exocytosed in the vacuolar space and the content of the spherical bodies is secreted into the parasitophorous vacuole, causing further expansion

of the membrane lining the vacuole. The merozoite has finally entered the erythrocyte and the erythrocyte membrane reseals. The whole invasion process is followed by a second wave of erythrocyte destruction (Yokoyama et al., 2006).

Though, the invasion process of erythrocytes is complex (Dvorak et al., 1975). The entire invasion process of parasites was initially proved and completed in approximately 30s (Aikawa et al, 1978; Sun et al., 2011), and under in vitro conditions merozoites generally recognize new target RBCs within 1 min of their release from their host RBC for *Plasmodium*. (Gilson and Crabb, 2009). Studying the kinetics of invasion revealed that 80% of invasion events occur within 10 min of mixing merozoites and RBCs in *Plasmodium* (Boylea et al., 2010). For *B. divergens*, 70% of the invasion-competent parasites (merozoites) invaded the RBC in less than 45 s and all living merozoites achieved invasion within 10 min of contact (Sun et al., 2011).

4. Roles of the apical complex organelles in the invasion process of erythrocytes

During the invasion process, the merozoite is structurally modified. Two major modifications occur: the shedding of the proteins from the adhesive coat and the disappearance of the apical organelles whose contents are discharged. The proteins from these organelles play crucial roles in the molecular events occurring during invasion (Aikawa et al., 1978; Bannister et al., 1986, 1989, 1990; Sam-Yellowe, 1996). The invasion is a rapid process (few seconds), and during this short time window, the sequential protein secretion from micronemes, rhoptries and then dense granules is orchestrated. This leads to host cell invasion and parasitophorous vacuole membrane formation. The maintenance of the latter varies between genus Apicomplexan (Preiser et al., 2000, Yokoyama et al., 2006; Gohil et al., 2010).

The detailed function of each organelle during invasion process is depicted in figure 8. (1) Contact of the parasite with a host cell surface; and transduction of a putative signal to the apex; (2) Re-orientation, apical binding to the host cell and micronemes exocytosis resulting in formation of the moving junction; (3) Rhoptries exocytosis and progressive gliding of the moving junction backward; (4) Expansion of the vacuole with integration of parasite proteins into the lipidic bilayer of the internalized host cell membrane, sealing of

the vacuole. When the moving junction reaches the posterior end of the parasite, dense granules exocytose in the vacuolar space (Dubremetz et al., 1998).

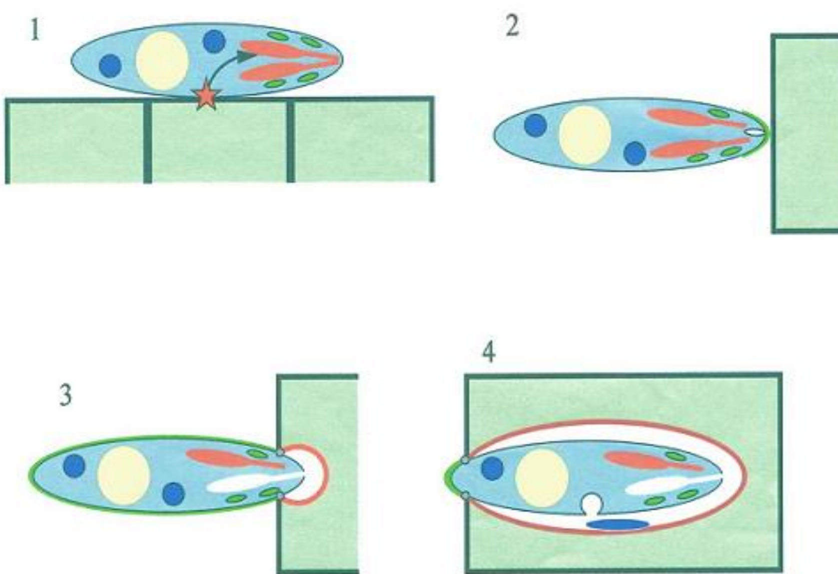


Figure 8. The role of apical complex organelles, secretion of the contents of the these organelles and invasion of erythrocyte by the merozoite of Apicomplexa (Dubremetz *et al*., 1998).

In the invasion process, the secreted proteins are responsible of very diverse molecular interactions from simple attachment, to modification of the erythrocyte plasma membrane and propelling (Blackman and Bannister, 2001). While more than 50 different proteins involved in the various stages of invasion have been characterized in *Plasmodium* (Rodriguez et al., 2008), this number remains below 10 for *Babesia.*

5. Proteins involved in the invasion and their function

Proteins involved in attachment and invasion are localized respectively on the merozoite surface or inside the apical organelles and they are secreted during invasion. Knowledge on these proteins for *Babesia* are limited the VMSAs (Variable Merozoite Surface Antigens), Rhoptry-Associated-Protein-1 (RAP-1), Associated Membrane Antigen-1 (AMA-1), Thrombospondin-related anonymous protein (TRAP) and Spherical Body Proteins (SBPs) (Yokoyama et al., 2006) (figure 9). These different proteins, with the

exception of SBP which are secreted too late in the invasion process, represent potential vaccine targets (Brown et al., 2006b; Carcy et al., 2006; Suarez and Noh, 2011).

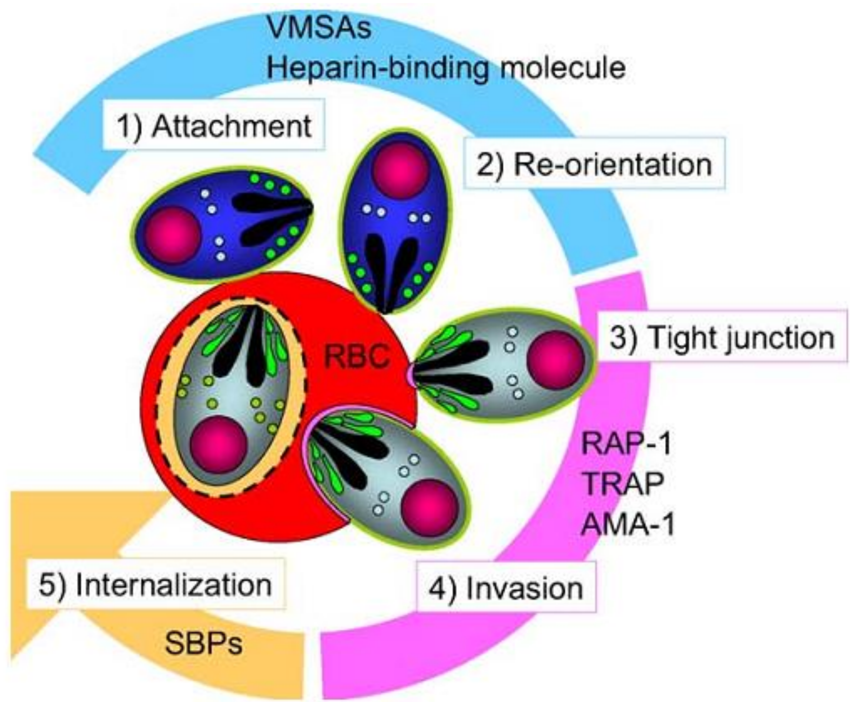


Figure 9. Proteins involved in erythrocyte invasion and characterized in the genus *Babesia* (Yokoyama et al., 2006)

5.1 Surface-coating molecules of extracellular merozoites: the family of variable merozoite surface antigens (VMSAs)

The variable merozoite surface antigens family gathers proteins localized on the surface of merozoites that play a crucial role in the non-specific binding to the erythrocyte and that are shed during the invasion process. They are surface-exposed proteins, often glycosylated, anchored in the cell membrane by glycosyl-phosphatidylinositol (GPI) moieties, and expressed on the surface of both merozoites and sporozoites. In *B. bovis*, the merozoite surface bears at least five proteins (MSA-1, MSA-2a1, a2, b and c) of molecular weights ranging from 42 to 44 kDa (Florin-Christensen et al., 2002; Mosqueda et al., 2002a, b; Sivakumar et al., 2013; Tattiyapong et al., 2014). The *msa-1* gene is a single copy gene (Suarez et al., 2000), while *msa-2* consists of four genes tandemly arranged in the genome of *B. bovis* (Florin-Christensen et al., 2002). Antibodies against MSA proteins

inhibit parasite invasion, and they are therefore considered as vaccine candidate antigens (Hines et al., 1992; Brown and Palmer, 1999; Palmer and McElwain, 1995; Dominguez et al., 2010). Members of this family have been described in other species and named *gp45/gp55* in *B. bigemina* (McElwain et al, 1991; Fisher et al, 2001), *Bc28* in *B. canis* (Carcy et al., 2006; Yang et al., 2012), *Bd37* in *B. divergens* (Carcy et al., 1995; Précigout et al., 2004; Delbecq et al., 2006, 2008; Hadj-Kaddour et al., 2007) and *Bgp32*, *45*, *47* and *50* in *B. gibsoni* (Goo et al., 2012). Nothing is known about these variable merozoite surface antigens for the sheep *Babesia* species.

5.2 Organelles-secreting molecules

Apical membrane antigen-1 (AMA-1)

AMA-1 protein is highly conserved among apicomplexan parasites (Chesne-Seck et al., 2005) and has been characterized in several species of *Babesia*: *B. bovis* (Gaffar et al., 2004a), *B. gibsoni* (Zhou et al., 2006a), *B. bigemina* (Torina et al., 2010) and *B. divergens* (Montero et al., 2009; Tonkin et al., 2013). It is stored in the microneme organelles and is transported to the parasite surface immediately after its secretion during host cell invasion. This protein is involved in the tight interaction between the host cell and the parasite surface, called the "moving junction". AMA-1 plays also an important role in apical reorientation. It is a type I integral membrane protein with three characteristic structures: (i) an N-terminal, cysteine-rich ectodomain, (ii) a single transmembrane domain, and (iii) a C-terminal cytoplasmic tail (Gaffar et al, 2004a). In *Plasmodium*, the full-length AMA-1 (83 kDa) is cleaved in the N-terminal ectodomain and transported to the merozoite surface membrane as a 66-kDa protein. AMA-1 is then further processed to 44- and 48-kDa soluble fragments during invasion of merozoites (Howell et al., 2001, 2003). The cleavage of *B. divergens* AMA-1 has also been demonstrated (Montero et al., 2009). Anti-AMA-1 antibody could inhibit erythrocyte invasion and the immunization with *Plasmodium* AMA-1 has been demonstrated to induce protective immunity against blood-stage of a homologous parasite challenge in many malarial models (Cortes et al., 2003; Dautu et al., 2007; Remarque et al., 2008). In *B. bovis*, the use of specific antibodies directed against certain synthetic peptides (Gaffar et al., 2004a) and conserved region (Salama et al., 2013) of the AMA-1 protein could inhibit the invasion of parasite and could serve as a vaccine candidate against *B. bovis* infection. Nothing is known about AMA-1of the sheep *Babesia*

species. The characterization and production of anti-AMA-1 antibodies to be tested in an invasion inhibition test is being carried out by another group in the PIROVAC project.

Thrombospondin-related anonymous protein (TRAP)

TRAP has been described primarily in parasites of the genus *Plasmodium* (Robson et al., 1990; Ménard, 2000). TRAP is also stored within the microneme organelles. It is released onto the cell surface early in the invasion process and translocated along the parasite surface during penetration up to the posterior pole. The role of TRAP in gliding motility and host cell invasion was established by gene-targeting studies. A role of TRAP in mosquito salivary gland invasion has also been demonstrated (Sultan et al., 1997; Wengelnik et al., 1999; Pihlajamaa et al., 2013). TRAP has been investigated as a vaccine candidate (Ogwang et al., 2013; Bauza et al., 2014). TRAP homologs have been identified in *B. bovis* (Gaffar et al., 2004b; Terkawi et al., 2013) and *B. gibsoni* (Zhou et al., 2006b) by the presence of two adhesive conserved domains responsible for host cell recognition and invasion: a von Willebrand factor A domain (vWFA) and one or more thrombospondin type 1 (TSP 1) repeats. Localization to the merozoite apical complex and role of TRAP in invasion was confirmed in *Babesia* with anti-TRAP specific antibodies (Gaffar et al., 2004b; Zhou et al., 2006b).

Rhoptry-associated-protein-1 (RAP-1)

Rhoptry contents were initially identified and purified by using monoclonal antibodies in *Plasmodium yoelii*, which have detected a 235-kDa protein localized at the apical region of merozoites as a dot like immunofluorescence pattern (Freeman et al., 1980). It was further confirmed that the 235-kDa antigen localized in the rhoptries as well as on the merozoite surface of the erythrocytic stage of *P. yoelii* (Oka et al., 1984). These proteins were later characterized by monoclonal antibodies in rhoptries of *P. falciparum* (Holder et al., 1985). Even if they have the same name, *Plasmodium* and *Babesia* RAP-1 do not share characteristic features. They were named according to their localization in the rhoptries, but are not necessarily orthologues. The conserved features described below for the *Babesia rap-1* family are not found in *Plasmodium rap-1*.

In the genus *Babesia*, RAP was initially described in *B. bigemina* and designated p58 at that time (McElwain et al., 1987), and was then characterized in all examined *Babesia* species: *B. bovis* (Suarez et al., 1991b), *B. divergens* (Skuce et al., 1996), *B. gibsoni* (Zhou

et al., 2007), *B. caballi* (Kappmeyer et al., 1999), *B. ovis* and *B. canis* (Dalrymple et al., 1993a). The full sequencing of the gene was published in *B. bovis* (named pBv60) and then in *B. bigemina* (*p58*, later named as *rap-1a*) (Suarez et al., 1991b; Mishra et al., 1991). The name of rhoptry-associated-protein-1 was first then proposed to designate *B. bovis* pBv60 in 1993 (Suarez et al., 1993). A *rra* gene (Rap-1-Related Antigen) encoded by a single gene copy in the genome of *B. bovis* was recently described and contains all the features of RAP-1 family of *Babesia* species (Suarez et al., 2011).

The *rap-1* gene family in all *Babesia* species studied is characterized by the presence of multiple gene copies arranged in tandem head to tail, from the simplest arrangement of two almost identical copies in *B. bovis* separated by an intergenic region (figure 10) (Suarez et al, 1998a).

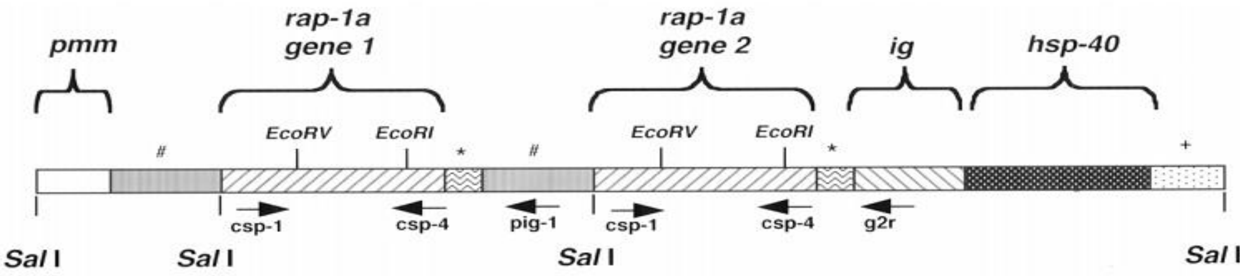


Figure 10. Structure of the *rap-1* locus in *B. bovis* (Suarez *et al*., 1998a)

The same studies performed in *B. bigemina* revealed a totally different organization and complexity. Five variable *rap-1a* interspaced with five *rap-1b* identical gene copies and one 3-terminal *rap-1c,* separated by intergenic regions have been described at this locus (figure 11). This complex multigene *rap-1* organization occurs in several examined strains (Suarez et al., 2003). But even if all *rap-1* genes but *rap-1aβ3* (Mishra et al., 1992) are transcribed, the transcription levels are variable, and both *rap-1b* and *rap-1c* are not translated into proteins in vitro or in vivo, while *rap-1a* is expressed in the merozoite rhoptries. *Rap-1a*, *rap-1b* and *rap-1c* share very low homologies (< 30% at the nucleotide level) and are considered as different gene types, and not only duplicated genes as the variable *rap-1a* or conserved *rap-1b* copies.

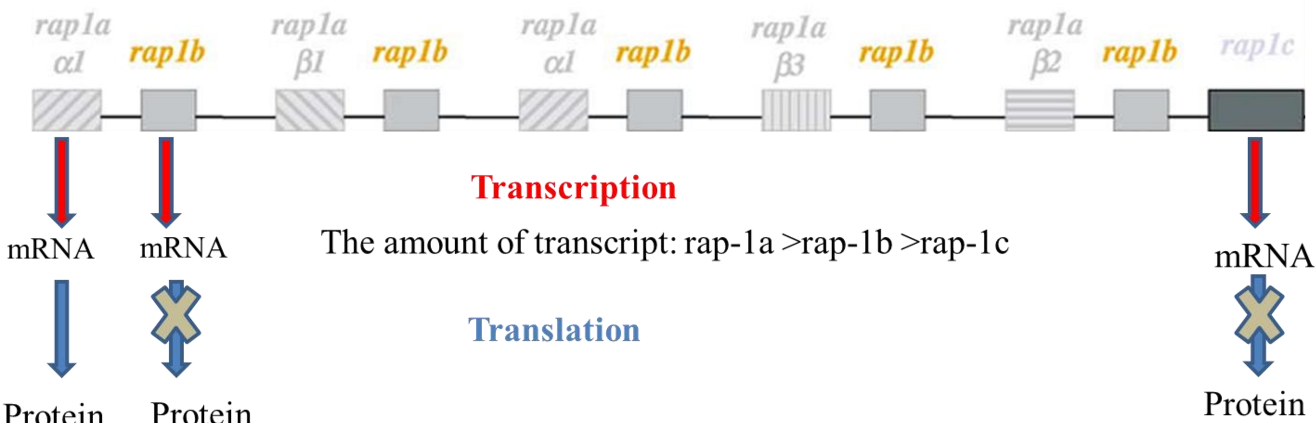


Figure 11. Structure of the *rap-1* locus in *B. bigemina* (Suarez *et al*., 2003)

The genetic organization of the *rap-1* locus in most babesial species is complex. The number of *rap-1* gene copies is variable: 2 arranged and identical genes in *B. bovis* and *B. canis*, 3 genes in *B. divergens* and *B. gibsoni*, 5 in *B. ovis* and 11 genes in *B. bigemina* (Suarez et al., 1991b; Dalrymple et al., 1993a; Skuce et al., 1996; Terkawi et al., 2009; Suarez et al., 2003). *B. ovis* locus organization has not yet been carefully studied and even the gene copy number is not surely attested (Dalrymple et al, 1993a).

The RAP-1 proteins in all studied *Babesia* spp. share some common features, such as: strict conservation of 4 cysteine residues, a highly conserved 14 amino acids motifs and several shorter conserved oligopeptide motifs in the first N-terminal 300 amino acids. They otherwise share overall sequence identity ranging from 30 to 45% (Brown et al., 1999; Suarez et al., 1994). The RAP-1 proteins are detectable not only in extracellular merozoites but also in sporozoite stage. This (these) protein(s) seemed to be involved in the RBC invasion process. Indeed, the RAP-1 specific antisera from immunized rabbit were effectively neutralizing the binding of sporozoites of *B. bovis* to erythrocytes (Mosqueda et al., 2002a), and RAP-1 specific monoclonal antibody blocked the binding of RAP-1 to merozoites and inhibited *B. bovis* growth in vitro (Yokoyama et al., 2002).

The N-terminal regions of RAP-1 sequences are more conserved among babesial parasites than the C-terminal region and several C-terminal sequences encode conserved or degenerate repetitive motifs in many *Babesia* species (Dalrymple et al., 1993a; Terkawi et al., 2009; Bhoora et al., 2010). The conservation of the N-terminal region and in particular the position of the 4 cysteine as well as the 14 aa conserved motif suggests that this region mediated critical function in merozoite invasion of erythrocytes in different *Babesia*

species and parasite survival and that antibodies against critical epitopes of this region could inhibit parasite multiplication and growth (Suraez et al., 1991a, 1994; Hötzel et al., 1996; Norimine et al., 2003). The repeated motifs were rather conserved in *B. ovis* (Dalrymple et al., 1993a), and at a lesser extent in *B. bovis* (Suarez et al., 1991b), *B. caballi* (Bhoora et al., 2010), *B. orientalis* (unpublished), *B. canis* (Dalrymple et al., 1993a) and *B. gibsoni* (Terkawi et al., 2009). In general, amino acid repeats are thought to mediate functions in parasite survival and immune evasion (Ramasamy, 1998; Mendes et al., 2013).

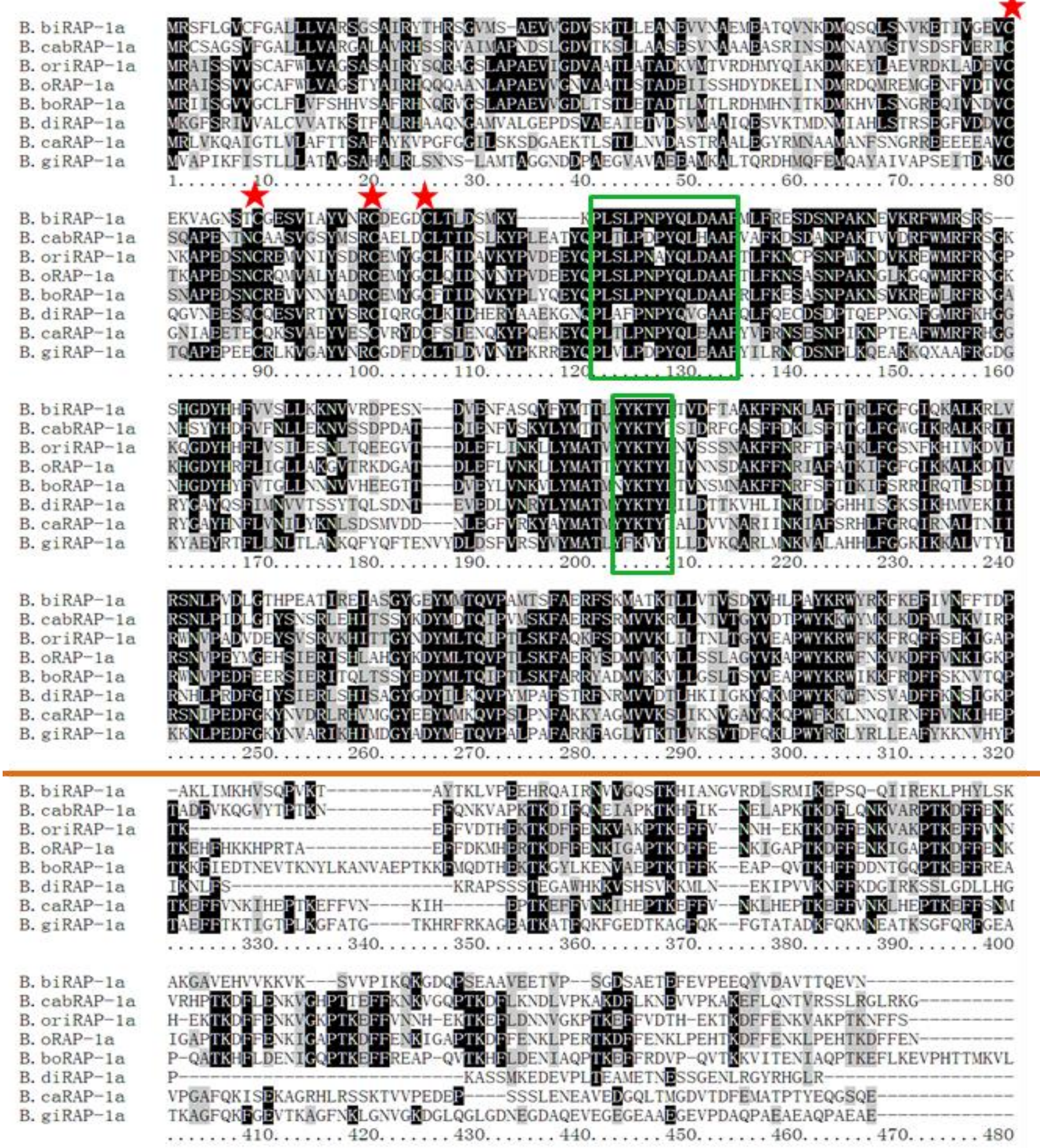


Figure 12. Alignment of RAP-1a amino acid sequences of several *Babesia* species showing the conserved features of members of this family. The strictly conserved 4 cysteine residues are indicated with red asterisks. Two highly conserved motifs are indicated with green boxes. The first relatively conserved 300 amino acids sequence are underlined (dark yellow). B.biRAP-1a: *B. bigemina* RAP-1aα1 (GenBank A45614); B. cabRAP-1a: *B. caballi* RAP-1a (GenBank ADF30791); B. oriRAP-1a: *B. orientalis* RAP-1a (GenBank AFC60006) ; B.oRAP-1a : *B. ovis* RAP-1a (partial sequences from GenBank AAA27811, AAA27804 and AAA27812 linked together); B.boRAP-1a: *B. bovis* RAP-1a (GenBank AACM44006); B. diRAP-1a: *B. divergens* RAP-1a (GenBank CAA89970); B. caRAP-1a: *B. canis* RAP-1a (GenBank CAA01285); B. giRAP-1a: *B. gibsoni* RAP-1a (GenBank ABD52000).

Spherical Body Proteins (SBPs)

Spherical bodies act later in the invasion process than micronemes and rhoptries, when the parasite is fully engulfed in the erythrocyte. Spherical body proteins of *Plasmodium* sp. play a role in the formation of a functional parasitophorous vacuole in order to allow intra-cellular parasite growth and viability (Dowling et al., 1996; Ruef et al., 2000; Bannister and Mitchell, 2003). The role of the spherical body proteins in *Babesia* is undoubtedly different from the one described for *Plasmodium*, if one takes into account that a differential characteristic between these parasites is the disappearance of this vacuole in members of the genus *Babesia* (Yokoyama et al., 2006; Gohil et al., 2010).

In *B. bovis,* four proteins have been characterized: BbSBP-1 (77-80kDa) was isolated from cDNA expression library screened with naturally infected bovine sera (Hines et al., 1995a), BbSBP-2 (225-kDa) (Jasmer et al., 1992) and BbSBP-3 (135-kDa) were isolated using monoclonal antibodies raised against spherical bodies (Ruef et al., 2000). BbSBP-4 (41-kDa) recently described is secreted into the cytoplasm of RBC at a later stage of infection and its production gradually increases until the rupture of RBCs. Thus this molecule is more accessible to host immune cells (Terkawi et al., 2011a, b). This feature was considered very important for subunit vaccine candidate and diagnostic antigen (Yokoyama et al., 2006). A report showed that BbSBPs were conserved among different geographically distant isolates and potentially able to induce significant protection against bovine babesiosis (Goodger et al., 1992).

Table 4. Proteins involved in at different invasion stages of *Babesia* species.

Localization	Protein	*Babesia* species	Putative function and features	References
Sprozoites and merozoite surface	VMSAs	*B. bovis*: MSA-1 (42-kDa), MSA-2, a, b, c (44-kDa); *B. bigemina*: gp45/55; *B. divergens*: Bd37; *B. canis*: Bc28; *B. gibsoni*: Bgp32, 45, 47,50	Expressed on the merozoite and sporozoite surface; mediated initial attachment and facilitated of sporozoites and merozoite to the host RBCs	*B. bovis*: Mosqueda et al., 2002a,b; *B. bigemina*: Fisher et al., 2001; *B. divergens*: Delbecq et al., 2006, 2008; *B. canis*: Yang et al.,2012; *B. gibsoni*: Goo et al., 2012.
Micronemes	AMA-1	*B. bovis* (82-kDa), *B. bigemina* (66-kDa), *B. divergens* (68-kDa), *B. gibsoni* (66-kDa)	Highly conserved; involved in the moving junction formation, mediated early stage of invasion (?)	*B. bovis*: Gaffar et al., 2004a, *B. bigemina*: Torina et al., 2010; *B. divergens*: Montero et al., 2009; *B. gibosni*: Zhou et al., 2006a
	TRAP	*B. bovis* (75-kDa); *B. gebsoni* (80-kDa)	Conserved, expressed on micronemes (?), recognition, possible attachement and invasion of host RBCs.	*B. bovis*: Gaffer et al., 2004b; *B. gebsoni*: Zhou et al., 2006b
	MIC-1	*B. bovis* (21-kDa)	Highly conserved, expressed on micronemes (?), contained a sialic acid (mediated erythrocyte invasion)-binding micronemal adhesive repeat (MAR) domain; cell recogniton, attachment and invasion (?)	Silva et al., 2010
Rhoptries	RAP-1	*B. bovis*, *B. divergens*, *B. ovis*, *B. canis*, *B. gibsoni*, *B. caballi*: RAP-1a (~60-kDa); *B. bigemina*: RAP-1a, RAP-1b (~38-kDa) and RAP-1c(~57-kDa)	Expressed on rhoptries, Erythrocyte binding and invasion of host RBCs on sporozoite and merozoite stages.	*B. bovis*: Suarez et al., 1998a; *B. divergens*, B. ovis, B. canis: Dalrymple et al., 1993a; *B. gibsoni*: Zhou et al., 2007; *B. caballi*: Kappmeyer et al., 1999; *B. bigemina*: Suarez et al., 2003
	BboRhop68	*B. bovis* (68-kDa)	Expressed on rhoptries, transmembrane domains, post-invasion, egression (?)	Baravalle et al., 2010
	RRA	*B. bovis* (~42-kDa)	Low expressed on rhoptries (?) on merozoite stage	Suarez et al., 2011

Localization	Protein	*Babesia* species	Putative function and features	References
Spherical bodies	SBP-1,2,3,4	*B. bovis*: SBP-1 (77-80-kDa); SBP-2 (225-kDa); SBP-3 (135-kDa); SBP-4 (41-kDa)	Expressed on spherical body, proteins deployed to erythrocyte cytoplasm, post-invasion and associate with the RBC membrane	SBP-1: Hines et al., 1995a; SBP-2: Jasmer et al., 1992; SBP-3: Ruef et al., 2000; SBP-4: Terkawi et al., 2011a, b
Infected erythrocyte surface	VESA and smorfs	*B. bovis*	Post-invasion (?), persistent infection, evasion, infected RBC sequestration, infected RBC (vesa), unknown (smorfs)	Al-Khedery and Allerd, 2006; Brayton et al., 2007; Ferreri et al., 2012
Merozoite and kinete (?)	Bov57	*B. bovis* (57.5-kDa)	Expressed in merozoite (in mammalian host) and kinetes (in vector ticks), may use as a target blocking multiplication in the mammalian host and blocing transmission by the tick vector.	Brayton et al., 2007, Freeman et al., 2010
?	Bbo-6cys A, B, C, D, E, F	*B. bovis* (60-kDa)	In blood stages	Silva et al., 2011
?	Bovipain-2	*B. bovis* (42-kDa), *B. bigemina*?	In blood stages	Mesplet et al., 2010

C. Control of babesiosis

A completely effective and satisfactory control method has not yet been developed against babesiosis, due to: (1) the lack of complete comprehension of the biology of *Babesia* and the complex immune mechanism for the *Babesia* infection (Gohil et al., 2013); (2) the global distribution of *Babesia* species corresponding with the global distribution of their ixodid vectors (Zintl et al., 2003); (3) the increasing development of resistance to acaricides by ticks, safety or residue problems of drugs in the vertebrate host and the many drawbacks (non-safety, low-efficiency) of the current live vaccines (Andreotti et al., 2011; Suarez and Noh, 2011; Mosqueda et al., 2012). All these points together result in the slow development of improved control tools for babesiosis worldwide.

Today, control of babesiosis is achieved by preventive methods (elimination of the vector in endemic areas), chemotherapeutic drugs and vaccines targeting ticks and/or *Babesia* (Suarez and Noh, 2011; Mosqueda et al., 2012).

1. Vector control

Ticks and the many pathogens they transmit result in serious diseases and even death of the host animals worldwide. Therefore, ticks have been becoming a major economic problem and subsequently related control strategy has begun to be exploited since the last Century (Graf et al., 2004). There are up to now mainly three methods to control ticks.

1.1. Chemical acaricides

Arsenic derivatives were the first chemical used for tick control. Subsequently, insecticides such as synthetic pyrethroids, organophosphates, amitraz and especially permethrin were widely used for significant tick control in the tropical and subtropical countries (Graf et al., 2004; Roma et al., 2010). Dipping and spraying of ivermectin, coumaphos, deltamethrin, at regular intervals, are currently used to eliminate ticks (Yin et al., 1997a). However, the application of these drugs was accompanied with a series of serious drawbacks, including resistance developed by ticks and low efficacy in reducing tick infestations, significant environmental pollution and milk and meat products contaminated with drug residues (Graf et al. 2004; Chevillon et al., 2007). Some new

products were reported with no resistance so far in some countries, for example an acarine growth regulator based on a benzoyl-phenyl-urea and a slow-release injectable macrocyclic lactone (Graf et al., 2004). Alternative tick control methods were then searched, as the use of biological control methods using natural predators.

1.2. Biological control of ticks

Biological control of tick population using pathogens and predators of ticks is a promising alternative developed recently, due to its low risk for human and animal health and its environmental safety (Samish and Rehacek, 1999). Many pathogens and predators of ticks have been known in nature, as bacteria, fungi, spiders, ants and other living things. They have a potential role toward limiting tick populations, even if only few biological control programs were used. Among these pathogens, entomopathogenic fungi as *Beauveria bassiana* or *Metarhizium anisopliae* can be used against several tick species, including *Rhipicephalus* (*Boophilus*) *microplus*, *R. sanguineus*, *Anocentor nitens*, *Amblyomma cajennense* (Fernandes and Bittencourt, 2008), *Haemaphysalis qinghaiensis* (Ren et al., 2011) and *Hyalomma anatolicum anatolicum* (Sun et al., 2011a, b). Entomopathogenic fungi can penetrate the ticks and infect all stages of tick (eggs, larvae, and adults) (Samish and Rehacek 1999; Gindin et al., 2002; Ren et al., 2011). However, drawbacks associated with biological control method are observable, including inconsistent levels of control, relatively slow speed to kill ticks and the stability of these microbial agents in the field remains unanswered (Polar et al., 2005a, b; Ren et al., 2011).

1.3. Vaccines against ticks

Compared to chemical drugs, the use of vaccine is a non toxic, low-contaminating, cost-effective strategy to tick control and prevents the repeated acaricidal applications resulting in drug-resistant ticks (Guerrero et al., 2012). An effective vaccine against arthropods (*Hypoderma lineatum*) was first developed and used for the protection of animals against hypodermosis in the 1980 (Pruett et al., 1987). Many attempts have been recently made to develop effective vaccine against ticks. Two commercial recombinant vaccines using the midgut glycoprotein antigen Bm86 against ticks have been produced in Australia (TickGARD) (Willadsen et al., 1995) and Cuba (Gavac™) (Canales et al., 1997).

They can decrease the vector reproductive capacity or affect transmission of pathogens by reducing the number of ticks and their weight (Hajdusek et al., 2010). The efficacy of these both vaccines was considerably variable and only the Gavac^{TM} vaccine is still marketed so far (Guerrero et al., 2012). Furthermore, the action of Bm86-based vaccines is very specific to the cattle *Rhipicephalus* (*Boophilus*) spp. and no efficacy against other tick species was observed (de la Fuente and Kocan, 2006; de Vos et al., 2001; Sossai et al., 2005; Canales et al., 2009). Several novel vaccine candidates are being studied, such as glutathione S-transferase from *Haemaphysalis longicornis* (Parizi et al., 2012), subolesin (SUB) (Almazán et al., 2010, 2012), vitellin-degrading cysteine endopeptidase (VTDCE) (Parizi et al.,2012), Yolk pro-Cathepsin (BYC) (Leal et al., 2006,) and ferritin-like protein (Hajdusek et al., 2010) from *Rhipicephalus microplus*. Vaccines to others species of ticks should be given more care in the future.

2. *Babesia* control

To control *Babesia* infection in the vertebrate host, chemotherapy is still an important tool since no effective vaccine against Babesiosis is currently available (Zintl et al, 2003; Mosqueda et al., 2012).

2.1. Anti-*Babesia* chemotherapeutic drugs

Many chemical drugs have been reported to be effective for the treatment of babesiosis, such as suinuronium sulfate, amicarbalide, trypan blue, diminazene aceturate and imidocarb diproprionate (Kuttler, 1988b; Vial and Gorenflot, 2006). However, suinuronium sulfate and amicarbalide have been withdrawn because of safety issues. Imidocarb and diminazene aceturate are at present two widely used chemotherapeutic drugs.

Imidocarb is the only chemotherapeutic drug that provides any significant long-term protection from clinical disease for at least 3 to 6 weeks through subcutaneous or intramuscular injection. Disadvantages of imidocarb is its high toxicity for animal (it could cause death in a few minutes) as well as its residues (Zintl et al., 2003; Mosqueda et al., 2012). Several new anti-*Babesia* drugs (e.g. Triclosan; Nerolidol; Artesunate; Epoxomicin; Gossypol and Atovaquone) more recently developed are used to control and treat

babesiosis. The mechanism of action of triclosan is to inhibit the synthesis of new membrane by the parasites preventing its growth and therefore disease progress without toxicity for host cells (**Bork et** al., 2003; Lund et al., 2005). The efficiency of triclosan was reported at concentration of 100μg/ml that can inhibit growth of *B. bovis* and *B. bigemina* (Mosqueda et al., 2012). Nerolidol can inhibit the growth of *Babesia* parasites through interfering with isoprenoid biosynthesis of the parasite (AbouLaila et al., 2010). The mode of action of artesunate is argued, however, it has shown effect in growth inhibition for several *Babesia* species: *B. bovis*, *B. bigemina*, *B. ovate*, *B .caballi* and *B. microti* (Goo et al., 2010). Epoxomicin is a potent proteasome inhibitor and proposed as anti-protozoa drug by blocking the catalytically active proteasomal subunits, thereby to inhibit the growth of the parasites (Mordmüller et al., 2006). The gossypol was mainly reported to inhibit the growth of in vitro *B. bovis* cultures but with toxic effects on animals. The mode of action of atovaquone is by inhibiting the rate of oxygen consumption by the parasite, and it inhibits the growth of *B. divergens* infecting human erythrocytes in vitro (Vial and Gorenflot, 2006).

However, due to the drawbacks of drug treatment, such as drug-resistance and drug residues in milk and meat, high toxicity for animal, caused by the possible misuse or prolonged use of these drugs, searches for more effective ways against babesiosis have prompted (De Waal and Combrink, 2006, Mosqueda et al., 2012).

2.2. Vaccines

Vaccine application is the best control option to prevent babesiosis. Some live vaccines have been studied and they could control babesiosis to a certain degree.

Live vaccines

In the late 19th century, researchers found that the animals could recover from naturally acquired infection with *Babesia*. Several attempts to produce vaccine were therefore made using blood collected from recovered animals (de Waal and Combrink, 2006). The first live *B. bovis* vaccine was then produced in Australia using virulent *B. bovis* strains attenuated by multiple rapid passages that apparently reduced parasite virulence (Callow et al., 1979; Bock et al., 1992; Pipano, 1995). The current Australian chilled tick fever vaccine and three live attenuated vaccines using *B. bovis* strains were

produced and offered protection for *B. bovis* and *B. bigemina* (Gohil et al., 2013). However, the potential risk of these live attenuated parasites is their reversion to virulence. In addition, the short shelf-life, cold storage conditions, potential contamination with other blood parasites, viruses or bacteria are limiting their broad application (Jenkins, 2001; Hope et al., 2005).

Recombinant and subunit vaccines

The efforts are now mainly being directed towards the development of recombinant or subunit *Babesia* vaccines against a number of *Babesia* species. These vaccines aim at blocking either the parasite transmission between ticks and vertebrate hosts or the parasite multiplication within the vertebrate host. The second approach has been favored for *Babesia* vaccines, since the transmission of *Babesia* to ticks remains poorly studied. Several vaccine candidates acting on parasite asexual multiplication in the vertebrate host have been described (table 5).

Table 5. Protein of *Babesia* sp. as vaccine candidate.

Protein	Blood stage	Tick stage	References
VMSA, VMSA-2 (*B. bovis*) Gp45/55 (*B. bigemina*)	+	+	Goff *et al*., 1988; Hines *et al*., 1995b; Suarez *et al*., 2000; Florin-Christensen *et al*.,2002; LeRoith *et al*., 2006; Fisher *et al*., 2001
AMA-1	+		Gaffar *et al*., 2004a; Torina *et al*., 2010; Salama *et al*., 2013
TRAP	+		Gaffar *et al*., 2004b; Goo *et al*., 2013
RAP-1	+		Dalrymple *et al*., 1993a; Suarez *et al*., 1993, 2003; Kappmeyer *et al*., 1999; Zhou *et al*., 2007
SBP-1, 2, 3, 4			Hines *et al*., 1995b; Dowling *et al*., 1996; Ruef *et al*., 2000; Terkawi *et al*., 2011a, b
RAP-1 related antigen (RRA)			Suarez *et al*., 2011
Micronemal proteins (MIC-1) (*B. bovis)*	+		Silva *et al*., 2010
Bbo-6cys A. B. C, D, E, F	+		Silva *et al*., 2011
Bovipain-2	+		Mesplet *et al*., 2010

Early attempts to develop a vaccine against *Babesia* focused on the merozoite surface coat proteins, and more specifically on the glycosyl phosphatidyl (GPI) anchored proteins (VMSAs), which are sloughed off during the process of host cell invasion (Carcy et al., 1995). However, the antigenic diversity of these proteins resulting in immune evasion could pose challenges to vaccine development and efficiency in the field in the long term application (Palmer et al., 1991; Shkap et al., 1994; Carcy et al., 2006). Other protective proteins involved in erythrocyte invasion process, including Rhoptry Associated Protein-1 (RAP-1), Apical Membrane Antigen-1 (AMA-1), Thrombospondin-related anonymous protein (TRAP) and Spherical Body Proteins (SBP-1, 2, 3,4) that have been described in *B. bigemina* and *B. bovis*, may form the basis of a recombinant vaccine against *Babesia* infections (Dalrymple, 1993b; Suarez and Noh, 2011). Attenuated and subunit vaccines based on the ability of these proteins to induce protective immunity are currently used to aid in control of babesiosis in endemic areas (Suarez and Noh, 2011). Among the limited number of known *Babesia* antigens, Rhoptry-Associated-Proteins may be considered as a strong target in recombinant vaccine development (Zintl et al, 2003; Santangelo et al, 2007). Interestingly, it has been reported that RAP-1 also is expressed in *B. bovis* sporozoites. Inhibition of the initial invasion by sporozoites could be a very effective approach to control babesiosis (Mosqueda et al, 2002a).

Finally, to control the complex tick-borne disease (herein the babesiosis) in the world, there is no adequat single method. Combining the application of anti-tick chemical acaricides, tick appropriate vaccine with the anti-*Babesia* drugs and vaccines development based on parasite asexual growth cycle of blood stage were thought practical and ideal control strategies of babesiosis (Gohil et al., 2013).

PUBLICATIONS

Article N° 1: Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 (Lintan) (*B. motasi* phylogenetic group)

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/authorsrights

Author's personal copy

Veterinary Parasitology 198 (2013) 24–38

Contents lists available at ScienceDirect

Veterinary Parasitology

journal homepage: www.elsevier.com/locate/vetpar

Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group)

Qingli Niu[a,b], Claire Bonsergent[a,b], Guiquan Guan[c], Hong Yin[c], Laurence Malandrin[a,b,*]

[a] INRA, UMR1300 Biology, Epidemiology and Risk Analysis in Animal Health, BP 40706, F-44307 Nantes, France
[b] LUNAM Université, Oniris, UMR BioEpAR, F-44307 Nantes, France
[c] State Key Laboratory of Veterinary Etiological Biology, LVRI, Lanzhou, China

ARTICLE INFO

Article history:
Received 23 May 2013
Received in revised form 27 August 2013
Accepted 27 August 2013

Keywords:
Babesia sp. BQ1 Lintan
Rhoptry-associated-protein-1
Multigenic *rap-1* locus

ABSTRACT

Babesiosis is a frequent infection of animals worldwide by tick borne pathogen *Babesia*, and several species are responsible for ovine babesiosis. Recently, several *Babesia motasi*-like isolates were described in sheep in China. In this study, we sequenced the multigenic *rap-1* gene locus of one of these isolates, *Babesia* sp. BQ1 Lintan. The RAP-1 proteins are involved in the process of red blood cells invasion and thus represent a potential target for vaccine development.

A complex composition and organization of the *rap-1* locus was discovered with: (1) the presence of 3 different types of *rap-1* sequences (*rap-1a*, *rap-1b* and *rap-1c*); (2) the presence of multiple copies of *rap-1a* and *rap-1b*; (3) polymorphism among the *rap-1a* copies, with two classes (named *rap-1a61* and *rap-1a67*) having a similarity of 95.7%, each class represented by two close variants; (4) polymorphism between *rap-1a61-1* and *rap-1a61-2* limited to three nucleotide positions; (5) a difference of eight nucleotides between *rap-1a67-1* and *rap-1a67-2* from position 1270 to the putative stop site of *rap-1a67-1* which might produce two putative proteins of slightly different sizes; (6) the ratio of *rap-1a* copies corresponding to one *rap-1a67*, one *rap-1a61-1* and one *rap-1a61-2*; (7) the presence of three different intergenic regions separating *rap-1a*, *rap-1b* and *rap-1c*; (8) interspacing of the *rap-1a* copies with *rap-1b* copies; and (9) the terminal position of *rap-1c* in the locus. A 31 kb locus composed of 6 *rap-1a* sequences interspaced with 5 *rap-1b* sequences and with a terminal *rap-1c* copy was hypothesized.

A strikingly similar sequence composition (*rap-1a*, *rap-1b* and *rap-1c*), as well as strong gene identities and similar locus organization with *B. bigemina* were found and highlight the conservation of synteny at this locus in this phylogenetic clade.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Babesiosis is caused by multiplication in the red blood cells of protozoa from the genus *Babesia* that belongs to the phylum Apicomplexa. Ovine babesiosis is one of the most important tick-borne hemoparasitic diseases of sheep and goats in tropical and subtropical countries but also in

* Corresponding author at: Oniris, UMR INRA/Oniris 1300, Biology, Epidemiology and Risk Analysis in animal health, Route de Gachet, La Chantrerie, BP 40706, F-44307 Nantes, France. Tel.: +33 02 40 68 78 57; fax: +33 02 40 68 77 51.
E-mail addresses: laurence.malandrin@oniris-nantes.fr, philippe.maestri@gmail.com (L. Malandrin).

0304-4017/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.vetpar.2013.08.025

Author's personal copy

temperate regions. It is caused mainly by *Babesia ovis*, *B. motasi* and *B. crassa* (Kuttler, 1988; Friedhoff, 1997; Uilenberg, 2006). In China, several new *Babesia* spp. naturally infective to sheep, have been recently described (Yin et al., 1997; Bai et al., 2002; Guan et al., 2002; Liu et al., 2007). Most of these newly described *Babesia* belong to the *B. motasi* phylogenetic group and they form a sister clade with the European isolates of *B. motasi*, that includes *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Hebei, *Babesia* sp. Tianzhu, and *Babesia* sp. Liaoning. On the other hand, *Babesia* sp. Xinjiang forms a separate clade together with other *Babesia* recently described in wild ruminants in South Africa (Liu et al., 2007; Niu et al., 2009; Oosthuizen et al., 2009).

Vaccines could be developed against babesiosis by targeting parasite multiplication within the host, and more precisely the erythrocyte invasion step via the proteins involved in this process. Apicomplexa protozoa are characterized by their apical complex, that contains organelles (rhoptries, micronemes and spherical bodies or dense granules) from which proteins are secreted or exocytosed to achieve erythrocyte invasion (Yokoyama et al., 2006; Lobo et al., 2012). Among these proteins, the rhoptry protein RAP-1 (rhoptry-associated-protein-1) has been shown to be a prime candidate for the development of recombinant vaccines against babesiosis (Brown et al., 1998, 1999; Brown and Palmer, 1999). The precise function of RAP-1 has not yet been elucidated, but the inhibition of *Babesia* growth in vitro, of erythrocyte adhesion or invasion using RAP-1 directed antibodies, points toward a relevant role in invasion (Figueroa and Buening, 1991; Mosqueda et al., 2002; Yokoyama et al., 2002).

RAP-1 is a protein of 40–60 kDa that has been characterized in all *Babesia* species examined so far (*Babesia bovis*, *B. bigemina*, *B. divergens*, *B. canis*, *B. ovis* and *B. gibsoni*) (Suarez et al., 1991a; Dalrymple et al., 1993; Skuce et al., 1996; Zhou et al., 2007). The organization, sequences, transcriptional analysis, genetic variation, gene expression (Hötzel et al., 1997; Suarez et al., 1998a, 2003; Mosqueda et al., 2002), and immune response to rhoptry proteins in *B. bovis* and *B. bigemina* have been reported (Suarez et al., 1993; Ushe et al., 1994; Boonchit et al., 2002). Some features are conserved among the RAP-1 proteins that characterize this family: the 4 cysteine residues at the N-terminus are strictly conserved and a 14 amino acids motif as well as several shorter oligopeptide motifs are also conserved (Suarez et al., 1991a,b, 1994; Dalrymple et al., 1996). The *rap-1* gene family in all *Babesia* species is characterized by the presence of multiple gene copies arranged in head to tail tandem, from a simple arrangement in *B. bovis* (only two copies of the *rap-1a* group) (Suarez et al., 1998b) to an extremely complex locus arrangement in *B. bigemina* (5 *rap-1a*, 5 *rap-1b* and 1 *rap-1c*) (Suarez et al., 2003). *rap-1a*, *rap-1b* and *rap-1c* genes have signature motifs typical of *rap-1* family but belong to divergent groups of genes due to their overall divergence in nucleotide sequences. Up to now, *rap-1b* and *rap-1c* gene groups have been described only in *B. bigemina*, and they are very conserved within each strain (for the *rap-1b* copies) and among strains (Suarez et al., 2003). Although present in all *Babesia* species studied so far, the *rap-1a* gene copies share an overall sequence identity ranging from 30 to 45%. In *B. bovis*, two identical or very similar copies of *rap-1a* are present in the locus and their polymorphism is very limited among different strains, with a maximum of 9 amino acid substitutions (Suarez et al., 1998b).

In the present study, we describe the different *rap-1* copies present in *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group), and the organization of these sequences.

2. Materials and methods

2.1. Origin of Babesia sp. BQ1 Lintan

Babesia sp. BQ1 Lintan was isolated from a sheep infested with adult *Haemaphysalis qinghaiensis* ticks from Lintan, Gansu Province in China (Guan et al., 2002). A clonal line of *Babesia* sp. BQ1 from sheep blood samples was isolated using in vitro cultures as described previously (Malandrin et al., 2004) and cryopreserved in liquid nitrogen at Oniris and at the Lanzhou Veterinary Research Institute (LVRI).

2.2. Babesia sp. BQ1 Lintan in vitro culture

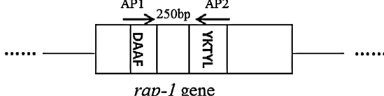
Blood samples cryopreserved in liquid nitrogen were thawed in a water bath at 37 °C, washed with RPMI 1640 (Lonza, Belgium) and centrifuged (1200 × *g*, 10 min). Primary cultures were prepared in 24-well culture plates containing 50 µl of the packed donor sheep red blood cells in 2 ml RPMI 1640 supplemented with heat inactivated Fetal Bovine Serum (FBS 20%) (DE14-801F, Lonza), Amphotericin B (0.5 µg/ml) and Gentamicin (50 µg/ml) to avoid fungal, bacterial and trypanosomal development. The cultures were produced at 37 °C in a humidified 6% CO_2 atmosphere. They were fed every 2–3 days by removing 1.5 ml of the supernatant medium overlying the erythrocyte layer without disturbing the RBC layer and replacing it with an equal volume of fresh medium (RPMI 1640 + FBS 20%). Meanwhile, the parasitemia of parasites developing in the cultures was monitored by microscopic examination of thin blood smears stained with May–Grünwald–Giemsa (MGG) (Diff Quick, Cooper). A subculture was prepared when parasitemia reached 2% by transferring 10 µl of infected RBC into another well containing 2 ml of fresh medium and 50 µl of packed donor sheep RBC. The entire contents of the culture well were then transferred to 25 cm^2 flasks and then 75 cm^2 flasks in 10 ml and then 40 ml of fresh medium, respectively, to expand the culture.

2.3. Genomic DNA extraction

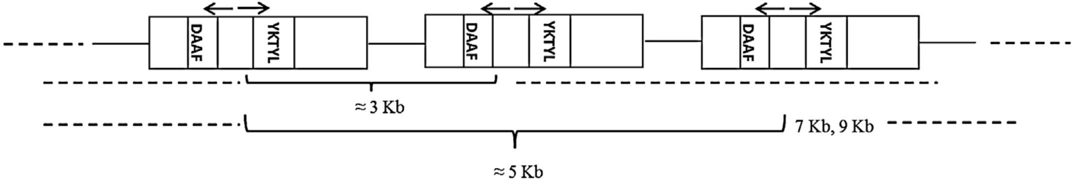
When the cultures in the flask reached a parasitemia of about 10%, the RBC were collected by centrifugation at 1200 × *g* for 10 min and washed twice with PBS. The erythrocyte pellet was diluted 1:1 with PBS and frozen at −20 °C for genomic DNA preparation using the Promega Wizard genomic DNA purification kit (Promega, France).

Author's personal copy

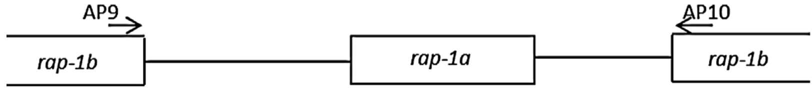
Step 1: Amplification of a central region of *rap-1* gene with degenerated primers



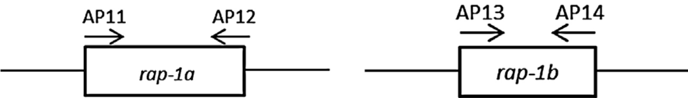
Step 2: Hypothetical amplification of tandemly repeated genes with primers designed in the central region of *rap-1* gene



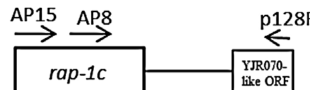
Step 3: Confirmation of *rap-1a* gene location between *rap-1b* genes



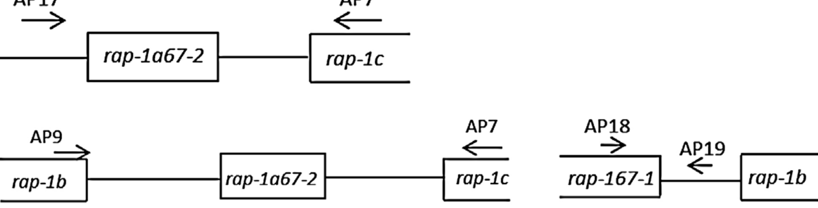
Step 4: *rap-1* full gene amplification



Step 5: Amplification of *rap-1c* gene and of the 3' end region of *rap-1* locus



Step 6: Confirmation of *rap-1a67* genes location before *rap-1b* gene or *rap-1c* gene



Step 7: Amplification of the intergenic regions using *rap-1* specific primers

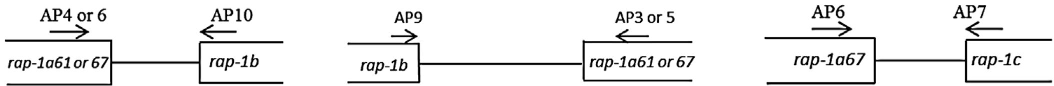


Fig. 1. *rap-1* locus amplification strategy and primers location.

2.4. *Sequencing the rap-1 genes and intergenic regions of Babesia sp. BQ1 Lintan*

The rationale for the experimental strategy used to decipher the *rap-1* gene locus composition is summarized in Fig. 1, with a description of the successive steps and primers used. Primer sequences are given in the supplementary Table 1. To obtain *Babesia* sp. BQ1 *rap-1* sequences, we used two features of this gene family. First, two conserved motifs (DAAF and YKTYL) described for these genes were successfully used to design degenerated primers (AP1 and AP2) in order to amplify the central region of unknown *rap-1* genes from *B. ovis* and *B. canis* (Dalrymple et al., 1993) (Fig. 1, step 1). Second, the tandem head to tail gene arrangement characteristic of this gene family allowed the amplification of gene extremities, intergenic regions and potentially other intercalated genes, by using primers from the central region directed outwards this region. To this purpose, a set of 6 primers (AP3 to AP8) designed from the obtained central region sequences were used in different combinations (6 different combinations that could correspond to 6 different arrangements of the 3 putative genes

Author's personal copy

Table 1
Primer combinations, amplification products that were cloned and sequenced and the corresponding primer locations and gene organization.

Primer pairs	Size of the amplicon(s)	Amplicon cloned and sequenced (number of clones sequenced)	Organization
AP1–AP2	0.26 kb, 0.32 kb, 0.8 kb	226 bp (5)	DAAF 226bp YKTYL
AP6–AP7	3 kb	2710 bp (1)	932bp 67 1333bp 445bp 86
AP4–AP5	5 kb	5360 bp (1)	977bp 61 853bp 1203bp b 1840bp 487bp 67
AP6–AP3	5 kb	5360 bp (1)	977bp 67 853bp 1203bp b 1840bp 487bp 61
AP8–AP3	0.7 kb, 2 kb, 3 kb	711 bp (1)[a]	[a]
AP8–AP5	2 kb, 3 kb	1660 bp (1)[a]	[a]
AP9–AP10	4 kb	4219 bp (11)	30bp b 1840bp 1425bp 61 853bp 71bp b
AP11–AP12	1.4 kb	1425 bp (61) (8) 1380 bp (67) (7)	61 or 67
AP13–AP14	1.3 kb	1203 bp (10)	b
AP8–p128R	2.5 kb, 3.5 kb	3671 bp (2)	AP8 c p128R YJR070-like ORF
AP15–p128R	3 kb, 4.5 kb	4129 bp (2)	AP15 c p128R YJR070-like ORF
AP4–AP10	2 kb	1901 bp (29)	977bp 61 853bp 71bp b
AP9–AP3 or 5	2.5 kb	2337 bp (61) (10) 2337 bp (67) (10)	30bp b 1840bp 487bp 61 or 67
AP6–AP7	3 kb	2710 bp (9)	932bp 67-2 1333bp 445bp c
AP17–AP7	3 kb	2280 bp (4)	AP17 67-2 AP7 c
AP9–AP7	5 kb	5028 bp (1)	30bp b 1840bp 1380bp 67-2 1333bp 445bp c
AP18–AP19	1.5 kb	1553 bp (4)	67-1 b

[a] Sequences were unrelated to each *rap-1* gene.

that were discovered) (Fig. 1, step 2 and Table 1). The 5′ and 3′ putative gene sequences were obtained this way, and primers were designed from these partial sequences to amplify and sequence the full putative genes (Fig. 1, step 4, primers AP11 to AP14). The arrangements in the locus of the different putative gene sequences were further studied using supplementary specific primers designed from the generated sequences (AP9–AP10, step 3 and AP17 to AP19, step 6). Sequences of the intergenic regions between the different putative genes sequences that were discovered

Author's personal copy

were also confirmed (Fig. 1, step 7). The amplification of 3′ end of the locus was performed using the published p128R primer, located downstream the *rap-1* locus in *B. bovis* and *B. bigemina* (Suarez et al., 1998b) (Fig. 1, step 5).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetpar.2013.08.025.

2.4.1. Partial amplification and sequencing of rap-1 genes

According to the defining molecular features of *rap-1a* genes from *Babesia* species infecting ruminants, degenerated primers (Supplementary Table 1, primers AP1 and AP2) were designed from the consensus sequence in two conserved amino acid regions (DAAF and YKTYL) (Supplementary Fig. 1). The amplification of a *rap-1* central region of about 250 bp was expected (Fig. 1, step 1). Suitable amplification conditions were determined by gradient annealing temperature. The amplification conditions were: an initial denaturation of 5 min at 94 °C, followed by 35 cycles of 1 min 30 s at 94 °C, 1 min 30 s at 50 °C/52 °C/54 °C or 56 °C, 1 min 30 s at 72 °C, and final extension for 10 min at 72 °C. The size of the produced amplicons was analyzed on 1.5% agarose gels with ethidium bromide staining. The crude PCR products were then cloned into TOPO TA vector (TOPO TA Cloning Kit for Sequencing, Invitrogen) and transformed into TOP10 *Escherichia coli* cells (according to the instructions of the manufacturer), in order to sequence the potentially different fragments. One hundred colonies were analyzed after the cloning reaction and the size of the inserted fragment was determined by PCR using the vector primers. Plasmids were then extracted (NucleoSpin plasmid extraction, Macherey-Nagel) from the selected colonies and the inserts sequenced. Clones with inserts of about 250 bp were obtained and three different sequences, all of them blasting with *rap-1* sequences, were discovered. They were named *61*, *67* and *86* according to the clone number.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetpar.2013.08.025.

2.4.2. Amplification and sequencing of tandemly repeated rap-1 region (5′ and 3′ extremities and intergenic regions)

Specific primers were designed from these three different internal sequences, and used outwards from the central sequence to amplify upstream and downstream *rap-1* gene regions as well as intergenic regions (Fig. 1, step 2). The presence of three different central regions indicated the existence of at least 3 *rap-1* sequences, and thus 6 possible arrangements. The specific primers were therefore used in different combinations (AP4–AP5; AP4–AP7; AP6–AP3; AP6–AP7; AP8–AP3; AP8–AP5) (Supplementary Table 1 and Table 1). The mean lengths of the published *rap-1* genes and intergenic regions indicated that theoretically amplicon sizes of about 3 kb (with no intercalating gene), 5 kb (with another intercalating gene) or more (depending on the number of *rap-1* genes in the locus) could be expected (Fig. 1, step 2). PCR products with sizes ranging from 0.7 to 5 kb were obtained in the different amplification groups (Table 1). Crude PCR products were cloned into the pGEM-T Easy vector according to the manufacturer's instructions (Promega). Colonies were selected by direct colony PCR using vector primers. Plasmids were then extracted (NucleoSpin plasmid extraction, Macherey-Nagel) and the inserts sequenced. The long inserts were sequenced by successively designing internal primers (Supplementary Table 1).

2.5. Amplification and identification of the rap-1 copies

We could conclude, from the sequences obtained with the above experiments, that three types of sequences were present, *rap-1a* (*61* and *67* types), *rap-1b* and *rap-1c*.

2.5.1. rap-1a

rap-1a entire sequences were determined from amplicons generated by PCR using *rap-1a* specific primers (Fig. 1, step 4, supplementary Table 1 and Table 1, primers AP11 and AP12) designed from the 5′ and 3′ ends of the *rap-1a* sequences determined previously. PCR was performed with Phusion High-Fidelity DNA polymerase (Thermo Scientific), amplicons were cloned into the Zero-Blunt end vector (Invitrogen) and ten plasmids were extracted and their inserts sequenced using the vector primers.

A digestion scheme was developed on the basis of this first set of sequencing results to roughly evaluate the presence and ratio of the different copies of *rap-1a* sequences (see results section). The different *rap-1a* copies were amplified using primers AP11 and AP12, localized in regions conserved in all the sequenced copies, and producing amplicons of the same size for the different copies, to limit ratio biases due to amplification and cloning. After cloning, the different cloned copies of *rap-1a* were amplified using *rap-1a* specific primers (AP11–AP12), and the PCR product was digested with *PflFI* to differentiate *67* type copies from the *61* type copies (Table 2). The *61* sequence variants were differentiated by separately applying the enzymes *HinP1*I and *Aci*I (see results section for a more detailed explanation of the choice of these enzymes). Digestions were performed according to the manufacturer's recommendations (New England BioLabs). The proportion of each copy was estimated by applying this scheme to 27 clones.

2.5.2. rap-1b

The sequence variability among the *rap-1b* copies was then evaluated by cloning crude PCR products (primer combination AP13–AP14) obtained with the Phusion High-Fidelity DNA polymerase into the Zero-Blunt end vector. Several plasmids were extracted and their inserts sequenced using the vector primer M13F/R (Fig. 1, step 4).

2.5.3. rap-1c

The 5′ extremity of *rap-1c* was sequenced from the 3 kb amplicon generated using primer combination AP6–AP7. Amplification of the 3′ end of this gene was attempted using forward AP8 *rap-1c* primer with reverse primers specific for *rap-1a* (AP3 and AP5), *rap-1b* (AP10) or *rap-1c* (AP7), in case these copies were located downstream of *rap-1c*. Primer p128R located downstream of *rap-1* locus and conserved in *B. bovis* and *B. bigemina* YJR070-like *orf* (AF026272)

Author's personal copy

Table 2
Enzymatic scheme used roughly to estimate the ratio of the different *rap-1a* copies based on the point mutations located at positions 143, 683 and 1248.

Gene	Point mutations	Enzyme used *Pfl*FI	*Hinp*1I	*Aci*I	Number of clones
rap-1a67	ATC	297 bp 1069 bp	–	–	8/27
rap-1a61-1	ACC	215 bp 297 bp 852 bp	305 bp 522 bp	224 bp 647 bp	9/27
rap-1a61-2	CTT	215 bp 297 bp 852 bp	424 bp 522 bp	129 bp 647 bp	10/27

(Suarez et al., 2003) was used for *rap-1c* 3′ end amplification (primer combinations AP8–p128R, AP15–p128R) using Ex *Taq* Hot-Start DNA polymerase (TaKaRa). Amplicons were cloned into the pGEM-T Easy vector and 2 clones from each amplification were sequenced (Fig. 1, step 5 and Table 1).

2.6. Bioinformatics analysis

All *rap-1* sequences obtained in this study were subjected to blast searching on the NCBI website http://www.ncbi.nlm.gov/blast/cgi, using the program BLASTn. Different free software tools for signal peptide prediction were used to search for the presence of putative signal peptides (SIG-Pred, http://bmbpcu36.leeds.ac.uk/prot_analysis/Signal.html; SignalP, http://www.cbs.dtu.dk/services/SignalP/; PrediSi, http://www.predisi.de/home.html). Phylogenetic analysis was performed using the software MEGA 5.

3. Results

3.1. Babesia sp. BQ1 genome contains at least three rap-1 sequence types

The amplification of putative *rap-1* genes using degenerated primers (AP1 and AP2, Table 1) produced three different amplicons (230, 320 and 800 bp). The size of the smaller amplicon could correspond to the central region of *rap-1*. Five different clones containing inserts of the expected size were obtained. Three 226 bp different sequences were discovered (named *61*, *67* and *86* from the clone number) (Fig. 2), all of them blasting with *rap-1* genes. *rap-1* sequences *61* and *67* were found to be very similar (91.6% identity) and clearly different from *rap-1* sequence *86* (63.7% and 61.0% identity, respectively). Sequences *61* and 67 blasted with the central region of *B. bigemina rap-1a*, while sequence *86* blasted with the central region of *B. bigemina rap-1c*, and will be named as *rap-1c* later in the manuscript. This first result already indicated the existence of at least three distinct *rap-1* copies in *Babesia* sp. BQ1 *rap-1* locus.

3.2. Babesia sp. BQ1 rap-1 locus contains orthologs of B. bigemina rap-1a, rap-1b and rap-1c genes

PCR amplification with the primer combination AP6–AP7 (Table 1) produced a 3 kb amplicon corresponding to the 3′ end of a *rap-1* copy, an intergenic region and the 5′ end of the following *rap-1* copy (Table 1). After cloning, the entire insert was sequenced. The 5′ region (445 bp) of an ortholog of the *B. bigemina rap-1c* gene (GenBank accession number: AF026272), with a nucleotide identity of 73%, was identified by BLAST analysis. The sequence of the 3′ region (932 bp) was obtained and blasted with the *B. bigemina rap-1a* gene (GenBank accession number: AB594817) (75% identity).

PCR amplification with the primer combinations AP4–AP5 and AP6–AP3 (Table 1) yielded, in both cases, DNA fragments of about 5 kb that were successfully cloned into pGEM-T Easy Vector. For each primer combination, one whole 5 kb insert was sequenced using successively designed internal primers (Supplementary Table 1). The sequencing results indicated that the *rap-1* 3′ extremities were similar for *61* (977 bp, primer combination AP4–AP5) and *67* (977 bp, primer combination AP6–AP3) and blasted with an identity of about 76% with several *B. bigemina rap-1a* gene sequences (*alpha-1*, *beta-2* and *beta-3*) over a length of about 450 bp. Sequences for the 5′ ends of *61* (487 bp, primer combination AP4–AP5) and *67* (487 bp, primer combination AP6–AP3) were different and blasted over a length of about 470 bp with several *B. bigemina rap-1a* gene sequences deposited in GenBank (identities of 72% for the *61* type and 69–70% for the *67* type). We could therefore conclude that the *Babesia* sp. BQ1 *rap-1* locus contained at least two *rap-1a* copies (named *rap-1a61* and *rap-1a67*).

Sequencing of the whole 5 kb inserts (obtained with primer combinations AP4–AP5 and AP6–AP3) revealed in both cases the presence of an intercalating *rap-1* copy, identical in both cases. The strong identity (75%) with *B. bigemina rap-1b* (GenBank accession number: AY146980) over a length of about 880 bp led us to conclude that *Babesia* sp. BQ1 *rap-1* locus contains orthologs of the *B. bigemina rap-1b* gene.

3.3. Babesia sp. BQ1 rap-1a sequences are interspaced with rap-1b sequences

Sequencing of the 5 kb DNA fragment obtained with the primer combinations AP4–AP5 and AP6–AP3 revealed the presence of *rap-1b* sequences between the *rap-1a* sequences (Table 1). However, due to the sequence identity between the 3′ ends of the *61* and *67 rap-1a* sequence types, no conclusion could be drawn about the presence of

Author's personal copy

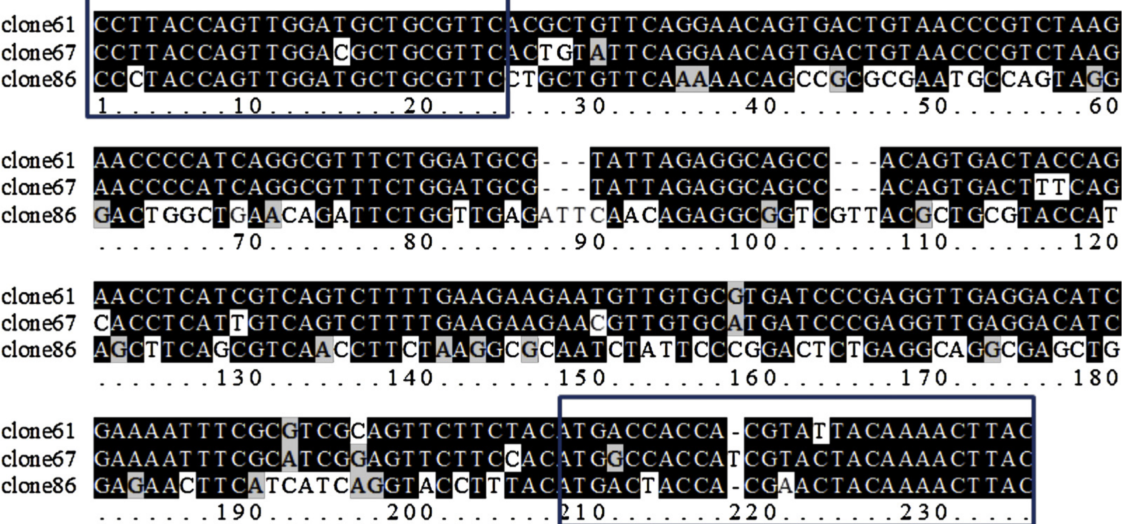


Fig. 2. Alignment of three nucleotide sequences of the *rap-1* 250 bp central regions obtained from amplification, with the degenerated primers AP1–AP2, cloning and sequencing. The forward and reverse degenerated primers, sequences used to amplify *rap-1* central regions are indicated with boxes.

either *61* or *67* type sequences upstream of a *rap-1b* gene. The presence of different *rap-1a61* sequence types between the *rap-1b* genes was confirmed by sequencing the 4 kb amplicons obtained using *rap-1b* outward primers (primer combination AP9–AP10, Fig. 1, step 3; Table 1). Sequences of *rap-1a67* were never found between two *rap-1b* copies. We could therefore conclude that *rap-1b* copies could be located upstream or downstream of *61* or *67 rap-1a* types, but never between two *rap-1a67* types.

3.4. The Babesia sp. BQ1 rap-1 locus contains at least two different rap-1a types of sequences (61 and 67), each of them with variants within the locus

3.4.1. rap-1a61 and rap-1a67 types of sequence

The existence of different types of *rap-1a* sequences within the *Babesia* sp. BQ1 *rap-1* locus was hypothesized following the first experiments to sequence the 226 bp central regions of the *rap-1* copies. Two different copies of *rap-1a* type (*61* and *67*) could already be differentiated. This was further confirmed by subsequent amplifications and sequencing with several specific primer combinations (AP9–AP10; AP11–AP12).

The putative sizes of the different *rap-1a* sequences were different: 1425 bp for *rap-1a61* and *rap-1a67-1* encoding a putative 474 amino acids protein, and 1380 bp for *rap-1a67-2* encoding a putative 459 amino acids protein (Fig. 3). The molecular features of the *rap-1* family members were conserved, i.e., the 4 cysteines and the patches of conserved motifs including PLTLPNPYQL**DAAF** and Y**YKTYL**TVD. These two *rap-1a* types (*61* and *67*) showed identities of 94.9% and 93.5%, respectively, at the nucleotide (60 substitutions) and amino acid (27 aa modifications with 43 codons affected by nucleotide substitutions) levels. The differences in the protein sequences were mostly limited to a region between aa 64 and 107, with few modifications between aa 108 and aa 201. The sequences were conserved at the 5′ end (aa 1–63 with only one modification at aa 48) and perfectly conserved at the 3′ end (aa 202–456). Among the substitutions, the transitions were slightly more frequent (32/60) with a majority of A/G changes (22/32). The 4 different transversions were in roughly equivalent amounts (6 A/T, 5 A/C, 7 G/C and 6 G/T). About half of the substitutions affected only one nucleotide in the codon (32/60), 17 of them were located on the third position of the codon, 4 on the second and 10 on the first. Most of these single substitutions at the third position were synonymous (16/17), whereas only one substitution at the first position was synonymous. Double substitutions affected 11 codons (all of them resulting in aa changes), while triple substitutions affected 2 codons (non-synonymous). Identities of 73% and 72% were calculated between the *B. bigemina rap-1a* (GenBank accession number: M60878) and *rap-1a61* and *rap-1a67*, respectively, at the nucleotide level (57% for both at the protein level, GenBank accession number: AAA65583).

3.4.2. rap-1a61 copies are polymorphic

Precise sequencing of the 5′ and 3′ regions of *rap-1a61* for 20 different DNA inserts (Fig. 1, step 7) demonstrated the presence of point mutations at three different positions (one in the 5′ region at position 143 from the putative start site with either A or C, and two in the 3′ region at positions 683 and 1248 from the putative start site with either C or T at both positions). The combination of these three point mutations was then determined by amplifying the full *rap-1a* (primers AP11–AP12) and sequencing 8 clones with Phusion High-Fidelity DNA polymerase. Two combinations of these substitutions were evidenced (ACC and CTT) coding for two different putative proteins (RAP-1a61-1 and RAP-1a61-2), since two of these mutations were not synonymous (position 143 nt A/C – aa E/A; position 683 nt C/T – aa A/V; position 1248 nt C/T – G/G) resulting in the AA combinations EAG or AVG at these three positions.

3.4.3. rap-1a67, one copy or several identical copies

The whole *rap-1a67* localized upstream *rap-1c* was amplified using Phusion High-Fidelity DNA polymerase

Author's personal copy

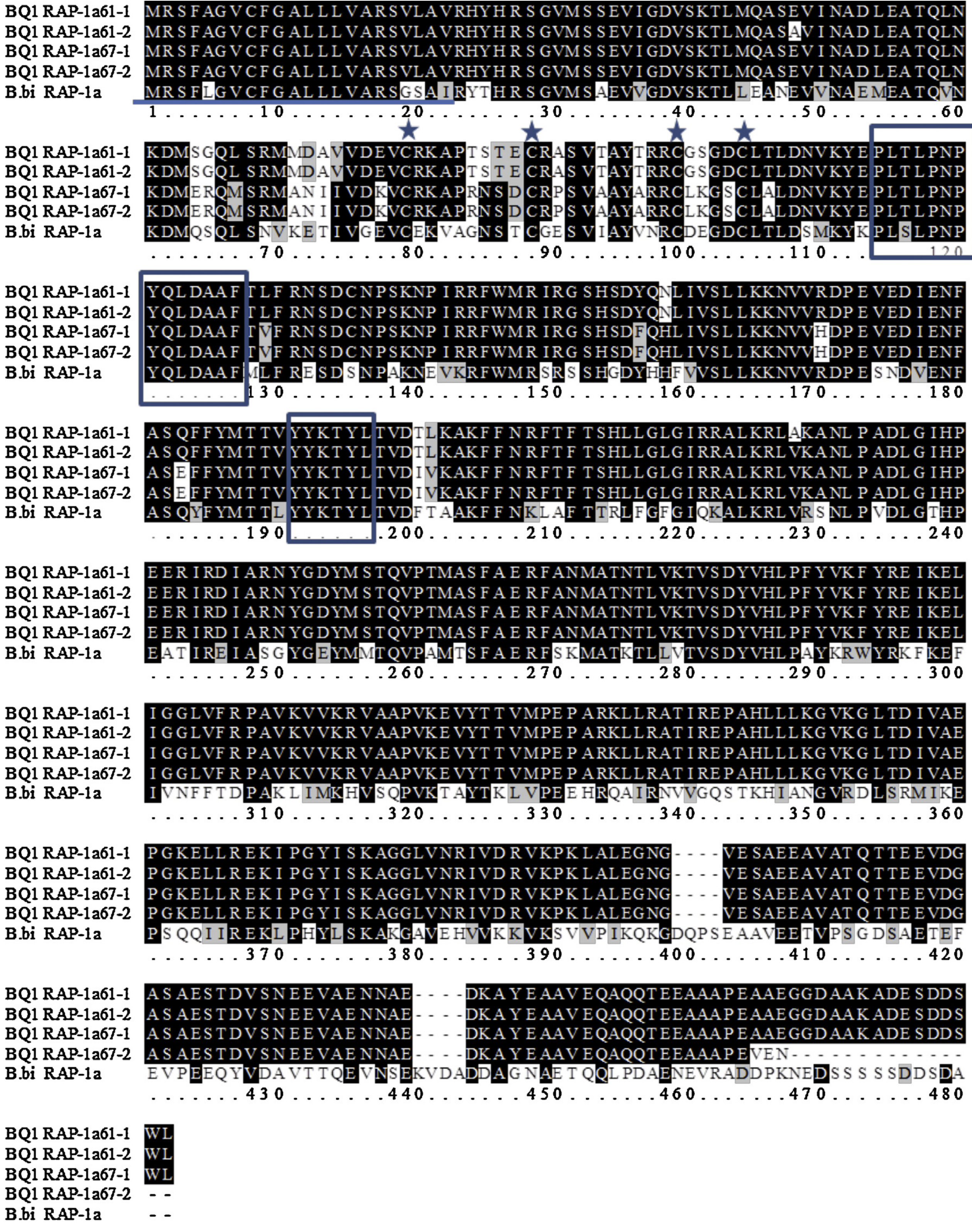


Fig. 3. Alignment of predicted amino acid sequences of *Babesia* sp. BQ1 *rap-1a* copies (61 and 67) with *B. bigemina rap-1a α1* gene (GenBank accession number: AAA65583). The 4 conserved cysteine residues are marked with asterisks, the two, conserved amino acid sequences used to design degenerated primers indicated with boxes, and predicted signal peptide sequences are underlined.

and primer combinations AP17–AP7 (4 clones) and AP9–AP7 (1 clone) (Fig. 1, step 6) with reverse primers specific of the *rap-1c* sequence. All obtained sequences were of 1380 bp and were identical. The whole *rap-1a67* localized upstream *rap-1b* was also amplified with the primer combinations AP18–AP19 (4 clones) and AP9–AP5 (10 clones) (Fig. 1, steps 6 and 7), the reverse primers being specific to the *rap-1b* sequences. All obtained sequences

Author's personal copy

Fig. 4. Alignment of predicted amino acid sequences of *Babesia* sp. BQ1 *rap-1b* with *B. bigemina rap-1b* gene (GenBank accession number: AAN84516) and with *B. bovis* RRA sequence (NCBI reference sequence: XP_001610950). The 4, conserved cysteine residues are marked with asterisks, the two conserved amino acid sequences used to design degenerated, primers indicated with boxes, and predicted signal peptide sequences are underlined.

were of 1425 bp and were identical. Taken together, these results show that there are two *rap-1a67* (*rap-1a67-1*, 1425 bp located before *rap-1b* and *rap-1a67-2*, 1380 bp located before *rap-1c*). These two sequences differ only at the 3′ end where nucleotide substitutions created an earlier stop site in the case of *rap-1a67-2*.

3.4.4. Estimation of the relative abundance of the rap-1a copies

A *rap-1a* restriction enzyme digestion scheme was set up to roughly estimate the ratio between the different *rap-1a* copies (type *67* and types *61*). After amplification of the different *rap-1a* sequences using AP11–AP12 *rap-1a* primers and cloning of the different amplicons of the same size, inserts from plasmids extracted from 29 different colonies were amplified using the same primers and then digested with *Pfl*FI to first differentiate the *67* type from the *61* type. *rap-1a* amplicons corresponding to the *61* type were then digested separately with the enzymes *Aci*I and *Hinp1*I to differentiate the different *61* variants (Table 1).

The expected digestion patterns were obtained for the 27 cloned amplicons. Eight corresponded to type *67*, nine to type *61-1* (ACC) and ten to type *61-2* (CTT). This ratio indicates the presence of roughly an equal number of each of these copies. Two clones gave unexpected digestion patterns. They had nucleotide sequences corresponding to the *67* sequence in the 5′ end (up to position 466 for clone *67/61-1* and to position 228 for clone *67/61-2*) and to the *61* sequence in the 3′ end (Supplementary Fig. 2). These sequences were most probably generated during PCR by shuffling between amplicons of the 3 closely related genes, as explained in the discussion section.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetpar.2013.08.025.

3.5. Babesia sp. BQ1 rap-1b copies are conserved

Primers allowing specific amplification of the full *rap-1b* (AP13–AP14) were designed. The length of *rap-1b* was 1203 bp with putative start and stop sites encoding a 400 amino acid protein (Fig. 4). Identity with the *B. bigemina rap-1b* gene (GenBank accession number: AY146980 and AAN84516), at the nucleotide and the protein levels, was 70.3% and 63.7%, respectively.

Since multiple *rap-1b* copies were present within the *rap-1* locus, the genetic diversity of these copies was investigated by cloning and sequencing 10 different clones (primers AP13–AP14, amplification with high fidelity Phusion *Taq*). Ten identical sequences were obtained.

Author's personal copy

Fig. 5. Alignment of predicted amino acid sequences of *Babesia* sp. BQ1 *rap-1c* with *B. bigemina rap-1c* gene (GenBank accession number: AAN84521). The 4 conserved cysteine residues are marked with asterisks, the two conserved amino acid sequences used to design degenerated primers indicated with boxes, and predicted signal peptide sequences are underlined.

3.6. *Babesia sp. BQ1 rap-1c*

Partial sequences (618 bp) of the 5′ region of the *rap-1c* gene were obtained (primers AP6–AP7, high fidelity Phusion *Taq*, Zero Blunt end cloning), and were sequenced using the vector primers. Sequences from 9 clones were found to be identical.

Many attempts were made to amplify the 3′ end of *rap-1c* using reverse primers specific to the *rap-1a*, *rap-1b* and *rap-1c* sequences, in case these genes were located downstream of *rap-1c*. Depending on the DNA polymerase used, either no amplification occurred, or the fragments amplified did not contain any sequences which were related to any of the *rap-1* genes. The attempt to amplify the full *rap-1c* was successful, with both AP8/p128R and AP15/p128R primer combinations (4 clones sequenced). Taken together, *rap-1c* full length was 1749 bp with putative start and stop sites encoding a 582 amino acid protein. Sequence identities with *B. bigemina rap-1c* (GenBank accession numbers: AF026272 and AAN84521), at the nucleotide and protein levels, were 62.8% and 52.9%, respectively (Fig. 5). Interestingly, five relatively conserved blocks of 17–18 aa were found in the 3′ region of *rap-1c* gene (from nucleotide position 1258–1521) (Fig. 6). These repeated sequences were not identified in *B. bigemina rap1-c* sequences and this probably explains the lower sequence identity compared with *rap-1a* and *rap-1b* identities between the two species.

3.7. *Presence of signal peptides*

A search to detect the presence of putative signal peptide sequences was carried out with free bioinformatics software (SignalP4-1, SIG-Pred and PrediSi). All three software detected putative signal peptides in the case of *rap-1a* (21 amino acids) and *rap-1c* (18 amino acids) but none could identify a putative signal peptide in the case of *rap-1b* (Supplementary Fig. 3).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetpar.2013.08.025.

3.8. *Three different intergenic regions exist in the Babesia sp. BQ1 rap-1 locus*

Sequencing of the cloned fragments obtained after amplification with different primer pairs (AP4–AP5, AP6–AP3, AP9–AP10, AP9–AP3, AP9–AP5, AP4–AP10)

Author's personal copy

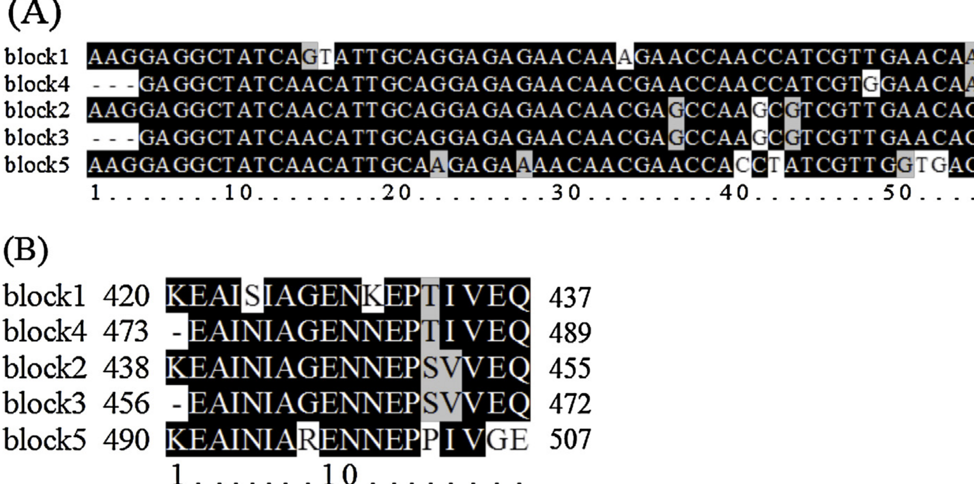


Fig. 6. Alignment of repeated nucleotide (A) and amino acid (B) blocks in the 3′ region of *Babesia* sp. BQ1 *rap-1c* sequence.

revealed intergenic regions of 853 bp (IG1) between *rap-1a* and *rap-1b*, and of 1840 bp (IG2) between *rap-1b* and *rap-1a* (*61* or *67*) (Supplementary Figs. 4 and 5). Sequencing of the 1840 bp IG region for 20 different DNA inserts demonstrated the absence of sequence variability (amplification with primer pairs AP9–AP3 and AP9–AP5, Phusion High-Fidelity DNA polymerase, insert sequencing for 10 clones of each amplification) (Fig. 1, step 7 and Table 1).

The same strategy (amplification with Phusion High-Fidelity DNA polymerase before cloning) was used for the 853 bp IG region located between *rap-1a61* (forward primer AP4) and *rap-1b* (reverse primer AP10). Sequencing of the 853 bp IG region was obtained for 10 different DNA inserts. One reproducible point mutation C (IG located after *rap-1a61-1* in 6 clones) or T (IG located after *rap-1a61-2* in 4 clones) was located 62 bp after the *rap-1a61* putative stop site. Taken together, we named IG1 as IG1.1 and IG1.2, downstream *rap-1a 61-1* and *rap-1a 61-2*, respectively (Supplementary Fig. 4).

The intergenic region between *rap-1a67-2* and *rap-1c* (IG3) differed from the two other intergenic regions and had a length of 1333 bp between the *rap-1a67-2* putative stop site and the *rap-1c* putative start site (Supplementary Fig. 6). Identical sequences of this 1333 bp IG were obtained for the 9 inserts analyzed (primer combination AP6–AP7, Phusion High-Fidelity DNA polymerase). Overall, the A + T residue content of IG1 (54.6%) is similar compared to the A + T content (52.5%) in the *rap-1b* sequence. In contrast, the A + T content of intergenic regions IG-2 and IG-3 (53.5 and 62.5%, respectively) are distinct compared to the contents of the putative corresponding coding regions (*rap-1a* with 46% and *rap-1c* with 53%). None of these three IG regions contains significant long open reading frames. The intergenic regions IG1 (from *rap-1a* to *rap-1b*) and IG2 (from *rap-1b* to *rap-1a*) in *Babesia* sp. BQ1 Lintan *rap-1* locus showed 27.6% and 22.3% identities with IG1 and IG2 in *B. bigemina rap-1* locus at the nucleotide level. For IG1 and IG2, significant homologies with *B. bigemina* corresponding intergenic regions were found only in the 5′ upstream regions of the following putative genes corresponding to putative transcription regulation sequences. No sequence with significant homology with IG3 was found in the databases.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetpar.2013.08.025.

3.9. Phylogenetic analysis

A phylogenetic tree was constructed by neighbor-joining method using the program MEGA5.1. All babesial RAP-1 and RAP-1 Related Antigen (RRA) sequences deposited in GenBank were used. *Theileria annulata* and *Plasmodium falciparum* RAP-1 sequences were used as outgroups (Fig. 7). As expected from the similarities already described above, all the RAP-1a sequences from *Babesia* sp. BQ1 formed a sister clade with the *B. bigemina* RAP-1a sequences, and clustered with all the babesial RAP-1a type sequences (see discussion about *B. gibsoni rap-1* gene naming). The RAP-1b and RAP-1c sequences of *B. bigemina* and *Babesia* sp. BQ1 each formed separate clades. The recently described RAP1 related antigen (RRA) gene formed a separate branch in this tree.

3.10. Putative gene number and arrangement in the Babesia sp. BQ1 rap-1 locus

We could conclude from the obtained amplification and sequencing data that the *rap-1* locus in *Babesia* sp. BQ1 contains orthologs of the *B. bigemina rap-1a*, *rap-1b* and *rap-1c* genes (supplementary Fig. 3). The presence of four different *rap-1a* sequences (*61-1*, *61-2* and *67-1*, *67-2*) was demonstrated. The amplification (primer pairs AP4–AP5, AP6–AP3, AP9–AP10 and AP18–AP19) and subsequent sequencing results allow us to conclude that the *rap-1b* sequences are interspaced between the *rap-1a61* sequences. *rap-1a67-1* is located upstream a *rap-1b* and *rap-1a67-2* downstream a *rap-1b*. The presence of a *rap-1a67-2* type upstream from the *rap-1c* sequence was also demonstrated (primer pairs AP17–AP7 and AP6–AP7). The position of *rap-1c* at the end of the locus was also demonstrated, followed by the YJR070-like *orf*, also found

Author's personal copy

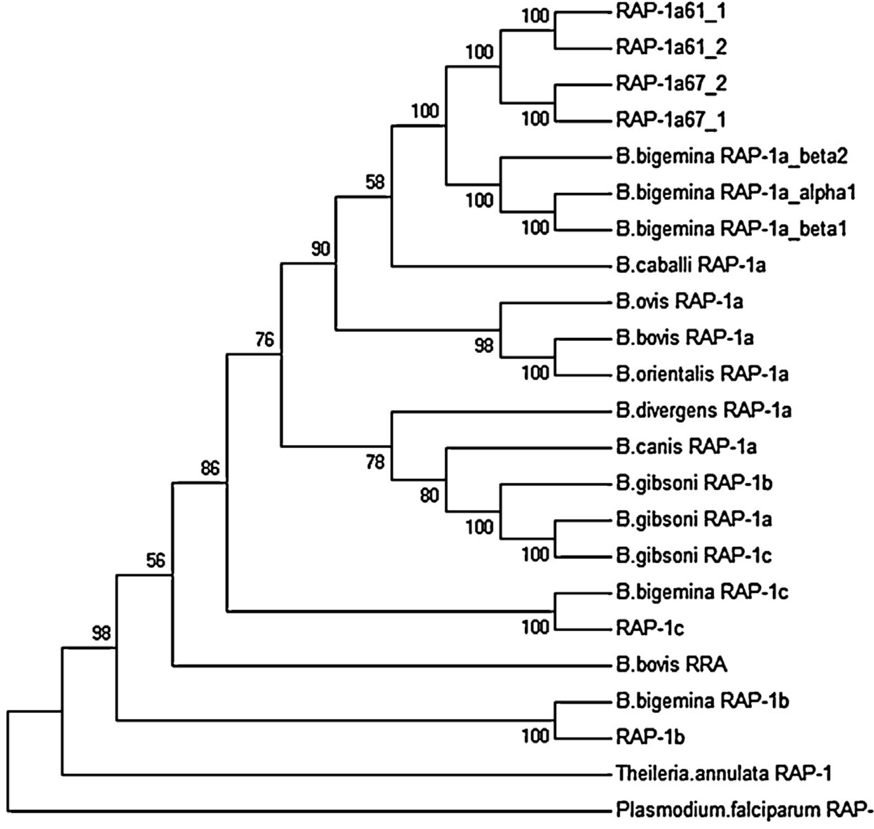


Fig. 7. Phylogenetic tree of the amino acid sequences of *Babesia* sp. BQ1 RAP-1 and of all known members of, the RAP-1 family in *Babesia*: *B. bigemina* RAP-1aα1(GenBank accession number: AAA65583), *B. bigemina*, RAP-1aβ1(D45561), *B. bigemina* RAP-1aβ2(C45561), *B. bigemina* RAP-1b (AAN84516), *B. bigemina* RAP-1c (AAN84521), *B. bovis* RAP-1 (AAB84267), *B. ovis* RAP-1 (AAA27805), *B. divergens* RAP-1 (CAA89970), *B. gibsoni* RAP-1a (ABD52000), *B. gibsoni* RAP-1b (BAH96476), *B. gibsoni* RAP-1c (BAH96477), *B. canis* RAP-1 (CAA01285), *B. caballi* RAP-1 (BAA83725), *B. orientalis* RAP-1 (AGC60006), *B. bovis* RRA (NCBI reference sequence: XP_001610950). The tree was inferred using the neighbor joining method, bootstrap values are shown at each branch point. The *Theileria annulata* (NCBI reference sequence: XP_954078) and *Plasmodium falciparum* (ADH84092) amino acid sequences were used as outliers.

at this position in the case of *B. bovis* and *B. bigemina rap-1* locus (Suarez et al., 1998b). We also concluded, from the restriction enzyme digestion scheme used to evaluate the relative abundance of the *rap-1a* genes, that the locus could comprise one *67-1* gene upstream of *rap-1b* and one *67-2* located upstream of *rap-1c* and downstream of a *rap-1b* gene, plus two *61-1* and two *61-2* copies types. These various data lead us to propose the locus organization indicated in Fig. 8. The relative position and numbers of the 4 *rap-1a61* genes could not be ascertained. The mutation in the IG region between *rap-1a61* and *rap-1b* (62 bp after the *61* putative stop site) is related to the type *61* gene which is upstream (C with *61-1* and T with *61-2*). Based on the estimate number and ratio of the different sequences obtained, the expected size of the *rap-1* locus is about 31 kb.

3.11. Nucleotide accession numbers

All sequences have been deposited in GenBank with the following accession numbers: KC953701 (*rap-1a61-1* and IG1.1), KC953700 (*rap-1a61-2* and IG1.2), KC953702 (*rap-1b* and IG2), KC953703 (*rap-1a67-2* and IG3), KF039723 (*rap-1a67-1*) and KF039724 (*rap-1c*).

4. Discussion

The rhoptry-associated-protein 1 (RAP-1) is thought to have an essential role in the process of erythrocyte invasion by Apicomplexan parasites of the genus *Babesia* (Yokoyama et al., 2006), and thus represents a potential vaccine candidate (McElwain et al., 1987; Wright et al., 1992; Dalrymple, 1993; Brown et al., 1998; Brown and

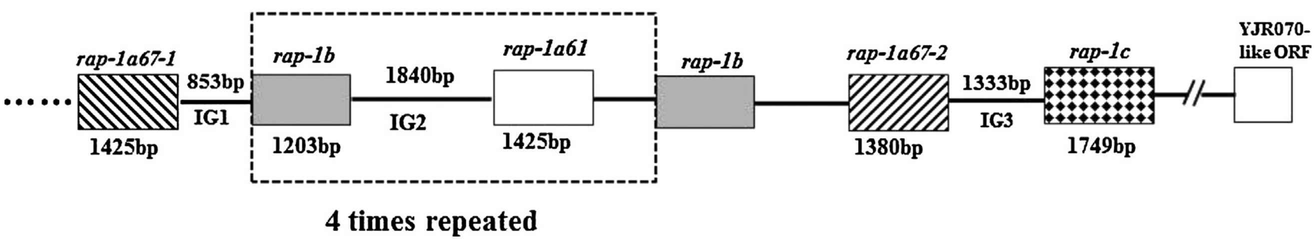


Fig. 8. Hypothetic *rap-1* locus organization of *Babesia* sp. BQ1. 4 times repeated partial locus are indicated with dotted line box.

Author's personal copy

Palmer, 1999). This protein was initially described in *B. bigemina* and designated p58 (McElwain et al., 1987), and was then characterized in all known Babesia species subsequently examined (*B. bovis*, *B. ovis*, *B. divergens*, *B. caballi*, *B. canis*) (Goff et al., 1988; Wright et al., 1992; Dalrymple et al., 1993; Skuce et al., 1996; Suarez et al., 1998b; Kappmeyer et al., 1999; Suarez et al., 2003). The full sequencing of the gene including the description of a signal peptide was published the same year first in *B. bovis* (named pBv60) and then in *B. bigemina* (p58 named later as *rap-1a*) (Suarez et al., 1991a; Mishra et al., 1991). However, the RAP-1 name (rhoptry-associated-protein 1) was proposed later to designate *B. bovis* pBv60 (Suarez et al., 1993). Despite their rather low sequence identity, conserved features for members of this family were described: four conserved cysteines, several short conserved motifs at the N-terminal region and a signal peptide essential for export of the protein to the rhoptry organelle (Suarez et al., 1991b, 1994; Dalrymple et al., 1993; Skuce et al., 1996), as well as their occurrence as multicopy genes (Dalrymple et al., 1993; Skuce et al., 1996; Suarez et al., 1998a,b, 2003; Terkawi et al., 2009). The potential of these proteins as candidates for subunit vaccine development has been postulated (McElwain et al., 1987; Wright et al., 1992).

In this study, we characterized 6 different *rap-1* copies in the *Babesia* sp. BQ1 genome, that share the characteristic features of the family. Based on their sequence identities, these 6 sequences could be separated into 3 different types. Due to the strong identities between the sequences described in our study and those described for *B. bigemina*, we named the *rap-1* sequences of *Babesia* sp. BQ1 according to the gene naming system adopted by Suarez et al. in 2003: *rap-1a* (with 4 different sequences named *rap-1a61-1*, *rap-1a61-2*, *rap-1a67-1* and *rap-1a67-2*), *rap-1b* and *rap-1c*. As reported for *B. bigemina*, the identities between these three types of *rap-1* are very low and they cluster with their *B. bigemina rap-1* gene orthologs in three separate clades.

Recently, *rap-1a*, *rap-1b* and *rap-1c* genes have also been described for *B. gibsoni* (Terkawi et al., 2009). However, they all cluster in the *rap-1a* clade in the phylogenetic tree presented, and *rap-1b* and *rap-1c* from *B. gibsoni* in fact are *rap-1a* polymorphic gene copies and should not be confused with the *rap-1b* and *rap-1c* clade members. Greater care needs to be taken when naming sequences to avoid this kind of confusion. The study that recently described *rra* gene (Rap-1-Related Antigen) in *B. bovis* indicated that RRA is phylogenetically close to the RAP-1b of *B. bigemina*. Despite conservation of the four cysteines, the patches of conserved motifs are very limited (Fig. 4) and the overall homologies with other *rap-1* genes are low (Suarez et al., 2011). In our study, the phylogenetic tree constructed from all known RAP-1 genes of *Babesia*, indicated that RRA branched separately on the *rap-1* phylogenetic tree (Fig. 7).

Full gene amplification, cloning and sequencing of *rap-1a* led us to identify two sequences (*rap-1a67/61-1* and *rap-1a67/61-2*) which corresponded to two different combinations of the 5′ end of *rap-1a67* and the 3′ end of *rap-1a61* (Supplementary Fig. 2). These sequences were most probably chimeras that were produced in the course of PCR. Use of PCR to amplify a template of multiple homologous genes (4 different *rap-1a* copies in our case) with the same set of primers, can result in the appearance of recombinant or "shuffled" molecules (Meyerhans et al., 1990), created either by annealing of an incompletely extended primer or by template switching during DNA synthesis (Kanagawa, 2003). The frequency of shuffled amplicons is usually rather low (<5%), as in the present study. These PCR chimeras were therefore not taken into account and only those sequences with a high number of represented clones were considered as valid sequences (Table 2).

The four different *rap-1a* gene copies found in *Babesia* sp. BQ1 Lintan genome are relatively well conserved, with identities of about 94%, and variability mostly limited to a 5′ region of 43 amino acids (aa 64–107). Interestingly, this region encompasses the NT variable region described in *B. bigemina rap-1a* genes (aa 68–87) (Hötzel et al., 1997). The polymorphism in the C-terminal end of the *rap-1a* gene is not as pronounced for *Babesia* sp. BQ1, as compared with *B. bigemina*, despite a shorter *67-2* gene due to 8 point mutations and the creation of a putative stop site. The extent of polymorphism between the RAP-1a copies differs considerably according to the *Babesia* species. For example, the two copies are identical in the case of *B. bovis* (Suarez et al., 1998b), with an average identity of 96% in the case of the four RAP-1a copies of *B. bigemina* (Mishra et al., 1992) and between 33.7 and 57% for the three RAP-1a copies of *B. gibsoni* (Terkawi et al., 2009).

From many points of view, the features of *B. bigemina* and *Babesia* sp. BQ1 *rap-1* are very similar. The locus is composed of three different types of *rap-1* (*rap-1a*, *rap-1b* and *rap-1c*), and up to now, the last two types have been described for only these two species. The *rap-1a* genes share an average sequence identity, at the nucleotide and protein levels, of 72.5% and 57%, respectively. For *rap-1b*, the sequence identities were as high as 70.3% and 63.7%, while the values for the *rap-1c* genes were 62.8% and 52.9%. The level of *rap-1a* copy polymorphism is within the same range (94 and 99%) and localized in the same 5′ region of the gene. As described for *B. bigemina*, *rap-1b* is also present as multiple copies, all of them identical in sequence. The only unique gene copy type is *rap-1c* and the presence of an YJR070-like *orf* in the 3′ end of *rap-1* locus in *B.* sp. BQ1 Lintan, homologous to YJR070-like *orf* from *B. bigemina*, demonstrated that an YJR070-like *orf* was located the same position in these two *Babesia* species genome (Suarez et al., 2003). The locus sizes are comparable, 27 kb for *B. bigemina*, about 31 kb for *Babesia* sp. BQ1, as well as the general locus organization, i.e., an alternation of *rap-1a* and *rap-1b* genes, the locus starting with a *rap-1a* gene and finishing with a *rap-1c* gene. The only difference is the presence of a *rap-1a* gene upstream of *rap-1c* in *Babesia* sp. BQ1, where a *rap-1b* gene is found in the case of *B. bigemina*. Due to this difference, there are three different intergenic regions at the *Babesia* sp. BQ1 *rap-1* locus, while only two have been described for *B. bigemina*. These features of locus structure and gene composition might be common to members of sister clades including *B. motasi*, *B. motasi*-like (from China), *B. ovata* and *B. bigemina* parasites, based on the 18S rRNA phylogeny (Paparini et al., 2012).

Author's personal copy

The precise function of the *Babesia* RAP-1 remains unknown. A role in erythrocyte invasion is strongly suspected for members of the *rap-1a* clade, as RAP-1a specific antibodies block invasion or adhesion of the protein to the erythrocyte surface (Mosqueda et al., 2002). However, such a role is difficult to attribute to RAP-1b and RAP-1c, since these proteins could not be detected in *B. bigemina*, either in vivo or in vitro (Suarez et al., 2003). Not even a putative signal peptide could be found for *Babesia* sp. BQ1 *rap-1b*, while such a signal peptide has been reported for *B. bigemina rap-1b*, as well as for all *rap-1* genes described so far. Whether *rap-1b* and *rap-1c* are transcribed pseudogenes remains to be clarified.

Understanding the genetic variation of a gene is essential when designing a vaccine. Conservation of the gene sequence between different copies of *rap-1a* would be an advantage for a vaccine candidate. Most of the non-synonymous point mutations are limited to a 5′ variable region of 43 aa. The interest of RAP-1a as a vaccine candidate is based on many aspects. These include the conservation of these proteins at the species level, which clearly offers another advantage when a vaccine is being developed to target geographically distant strains (Suarez et al., 1998b; Hötzel et al., 1997); the expression of RAP-1 protein not only in the merozoite but also in the sporozoite stage of *B. bovis* (Mosqueda et al., 2002); and the immunization of cattle with purified native *B. bigemina* RAP-1 (McElwain et al., 1991; Rodriguez et al., 1996) and recombinant *B. bovis* RAP-1, which resulted in partial protection upon challenge (Fish et al., 2008). It could therefore be included in a subunit recombinant vaccine.

Many studies still need to be performed on expression of the different copies of *rap-1* sequences in *Babesia* sp. BQ1, on their genetic variability and their respective role in the invasion process to determine their potential interest as a vaccine candidate.

Acknowledgements

This study was supported financially by a grant from the European Union cooperation project PIROVAC (KBBE-3-245145), the « Région Pays de la Loire » (France), and INRA.

References

Bai, Q., Liu, G.Y., Liu, D.K., Ren, J.X., Li, X., 2002. Isolation and preliminary characterization of a large *Babesia* sp. from sheep and goats in the eastern part of Gansu Province, China. Parasitol. Res. 88, S16–S21.

Boonchit, S., Xuan, X.N., Yokoyama, N., Goff, W.L., Wagner, G., Igarashi, I., 2002. Evaluation of an enzyme-linked immunosorbent assay with recombinant rhoptry-associated protein 1 antigen against *Babesia bovis* for the detection of specific antibodies in cattle. J. Clin. Microbiol. 40, 3771–3775.

Brown, W.C., McElwain, T.F., Hötzel, I., Suarez, C.E., Palmer, G.H., 1998. Helper T-cell epitopes encoded by the *Babesia bigemina rap-1* gene family in the constant and variant domains are conserved among parasite strains. Infect. Immun. 66, 1561–1569.

Brown, W.C., McElwain, T.F., Palmer, G.H., Chantler, S.E., Estes, D.M., 1999. Bovine CD4(+) T-lymphocyte clones specific for rhoptry-associated protein 1 of *Babesia bigemina* stimulate enhanced immunoglobulin G1 (IgG1) and IgG2 synthesis. Infect. Immun. 67, 155–164.

Brown, W.C., Palmer, G.H., 1999. Designing blood-stage vaccines against *Babesia bovis* and *B. bigemina*. Parasitol. Today 15, 275–281.

Dalrymple, B.P., 1993. Molecular variation and diversity in candidate vaccine antigens from *Babesia*. Acta Trop. 53, 227–238.

Dalrymple, B.P., Casu, R.E., Peters, J.M., Dimmock, C.M., Gale, K.R., Boese, R., Wright, I.G., 1993. Characterization of a family of multi-copy genes encoding rhoptry protein homologues in *Babesia bovis*, *Babesia ovis* and *Babesia canis*. Mol. Biochem. Parasitol. 57, 181–192.

Dalrymple, B.P., Peters, J.M., Böse, R., Wright, I.G., 1996. A polymerase chain reaction method for the identification of genes encoding members of the Bv60/p58 family of rhoptry protein homologues in the genus *Babesia*. Exp. Parasitol. 84, 96–100.

Figueroa, J.V., Buening, G.M., 1991. In vitro inhibition of multiplication of *Babesia bigemina* by using monoclonal antibodies. J. Clin. Microbiol. 29, 997–1003.

Fish, L., Leibovich, B., Krigel, Y., McElwain, T., Shkap, V., 2008. Vaccination of cattle against *B. bovis* infection with live attenuated parasites and non-viable immunogens. Vaccine 26, 29–33.

Friedhoff, K.T., 1997. Tick-borne disease of sheep and goats caused by *Babesia*, *Theileria* or *Anaplasma* spp. Parassitologia 39, 99–109.

Goff, W.L., Davis, W.C., Palmer, G.H., McElwain, T.F., Johnson, W.C., Bailey, J.F., McGuire, T.C., 1988. Identification of *Babesia bovis* merozoite surface antigens by using immune bovine sera and monoclonal antibodies. Infect. Immun. 56, 2363–2368.

Guan, G.Q., Yin, H., Luo, J.X., Lu, W.S., Zhang, Q.C., Gao, Y.L., Lu, B.Y., 2002. Transmission of *Babesia* sp. to sheep with field-collected *Haemaphysalis qinghaiensis*. Parasitol. Res. 88, S22–S24.

Hötzel, I., Suarez, C.E., McElwain, T.F., Palmer, G.H., 1997. Genetic variation in the dimorphic regions of *rap-1* genes and *rap-1* loci of *Babesia bigemina*. Mol. Biochem. Parasitol. 90, 479–489.

Kanagawa, T., 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). J. Biosci. Bioeng. 96, 317–323.

Kappmeyer, L.S., Perryman, L.E., Hines, S.A., Baszler, T.V., Katz, J.B., Hennager, S.G., Knowles, D.P., 1999. Detection of equine antibodies to *Babesia caballi* by recombinant *B. caballi* rhoptry-associated protein 1 in a competitive-inhibition enzyme-linked immunosorbent assay. J. Clin. Microbiol. 37, 2285–2290.

Kuttler, K.L., 1988. World-wide impact of babesiosis. In: Ristic, M. (Ed.), Babesiosis of Domestic Animals and Man. CRC Press, Boca Raton, FL, pp. 1–22.

Liu, A.H., Yin, H., Guan, G.Q., Schnittger, L., Liu, Z.J., Ma, M.L., Dang, Z.S., Liu, J.L., Ren, Q.Y., Bai, Q., Ahmed, J.S., Luo, J.X., 2007. At least two genetically distinct large *Babesia* species infective to sheep and goats in China. Vet. Parasitol. 147, 246–251.

Lobo, C.A., Rodriguez, M., Cursino-Santos, J.R., 2012. *Babesia* and red cell invasion. Curr. Opin. Hematol. 19, 170–175.

Malandrin, L., L'Hostis, M., Chauvin, A., 2004. Isolation of *Babesia divergens* from carrier cattle blood using in vitro culture. Vet. Res. 35, 131–139.

McElwain, T.F., Perryman, L.E., Davis, W.C., McGuire, T.C., 1987. Antibodies define multiple proteins with epitopes exposed on the surface of live *Babesia bigemina* merozoites. J. Immunol. 138, 2298–2304.

McElwain, T.F., Perryman, L.E., Musoke, A.J., McGuire, T.C., 1991. Molecular characterization and immunogenicity of neutralization-sensitive *Babesia bigemina* merozoite surface proteins. Mol. Biochem. Parasitol. 47, 213–222.

Meyerhans, A., Vartanian, J.P., Wain-Hobson, S., 1990. DNA recombination during PCR. Nucleic Acids Res. 18, 1687–1691.

Mishra, V.S., Stephens, E.B., Dame, J.B., Perryman, L.E., McGuire, T.C., McElwain, T.F., 1991. Immunogenicity and sequence analysis of recombinant p58: a neutralization-sensitive, antigenically conserved *Babesia bigemina* merozoite surface protein. Mol. Biochem. Parasitol. 47, 207–212.

Mishra, V.S., McElwain, T.F., Dame, J.B., Stephens, E.B., 1992. Isolation, sequence and differential expression of the p58 gene family of *Babesia bigemina*. Mol. Biochem. Parasitol. 53, 149–158.

Mosqueda, J., McElwain, T.F., Stiller, D., Palmer, G.H., 2002. *Babesia bovis* merozoite surface antigen 1 and rhoptry-associated protein 1 are expressed in sporozoites, and specific antibodies inhibit sporozoite attachment to erythrocytes. Infect. Immun. 70, 1599–1603.

Niu, Q.L., Luo, J.X., Guan, G.Q., Liu, Z.J., Ma, M.L., Liu, A.H., Gao, J.L., Ren, Q.Y., Li, Y.Q., Qiu, J.X., Yin, H., 2009. Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. Exp. Parasitol. 121, 64–68.

Oosthuizen, M.C., Allsopp, B.A., Troskie, M., Collins, N.E., Penzhorn, B.L., 2009. Identification of novel *Babesia* and *Theileria* species in South African giraffe (*Giraffa camelopardalis*, Linnaeus, 1758) and roan antelope (*Hippotragus equinus*, Desmarest 1804). Vet. Parasitol. 163, 39–46.

Paparini, A., Ryan, U.M., Warren, K., McInnes, L.M., de Tores, P., Irwin, P.J., 2012. Identification of novel *Babesia* and *Theileria* genotypes in the

Author's personal copy

endangered marsupials, the woylie (*Bettongia penicillata ogilbyi*) and boodie (*Bettongia lesueur*). Exp. Parasitol. 131, 25–30.

Rodriguez, S.D., Palmer, G.H., McElwain, T.F., McGuire, T.C., Ruef, B.J., Chitko-McKown, C.G., Brown, W.C., 1996. CD4 T helper lymphocyte responses against *Babesia bigemina* rhoptry-associated protein-1. Infect. Immun. 64, 2079–2087.

Skuce, P.J., Mallon, T.R., Taylor, S.M., 1996. Molecular cloning of a putative rhoptry associated protein homologue from *Babesia divergens*. Mol. Biochem. Parasitol. 77, 99–102.

Suarez, C.E., Palmer, G.H., Jasmer, D.P., Hines, S.A., Perryman, L.E., McElwain, T.F., 1991a. Characterization of the gene encoding a 60-kilodalton *Babesia bovis* merozoite protein with conserved and surface exposed epitopes. Mol. Biochem. Parasitol. 46, 45–52.

Suarez, C.E., McElwain, T.F., Stephens, E.B., Mishra, V.S., Palmer, G.H., 1991b. Sequence conservation among merozoite apical complex proteins of *Babesia bovis*, *Babesia bigemina* and other apicomplexa. Mol. Biochem. Parasitol. 49, 329–332.

Suarez, C.E., Palmer, G.H., Hines, S.A., McElwain, T.F., 1993. Immunogenic B-cell epitopes of *Babesia bovis* rhoptry-associated protein 1 are distinct from sequences conserved between species. Infect. Immun. 61, 3511–3517.

Suarez, C.E., Thompson, S.M., McElwain, T.F., Hines, S.A., Palmer, G.H., 1994. Conservation of oligopeptide motifs in rhoptry protein from different genera of erythroparasitic protozoa. Exp. Parasitol. 78, 246–251.

Suarez, C.E., Palmer, G.H., Hotzel, I., Hines, S.A., McElwain, T.F., 1998a. Sequence and functional analysis of the intergenic regions separating babesial rhoptry-associated protein-1 (*rap-1*) genes. Exp. Parasitol. 90, 189–194.

Suarez, C.E., Palmer, G.H., Hötzel, I., McElwain, T.F., 1998b. Structure, sequence, and transcriptional analysis of the *Babesia bovis rap-1* multigene locus. Mol. Biochem. Parasitol. 93, 215–224.

Suarez, C.E., Palmer, G.H., Florin-Christensen, M., Hines, S.A., Hötzel, I., McElwain, T.F., 2003. Organization, transcription, and expression of rhoptry associated protein genes in the *Babesia bigemina rap-1* locus. Mol. Biochem. Parasitol. 127, 101–112.

Suarez, C.E., Laughery, J.M., Bastos, R.G., Johnson, W.C., Norimine, J., Asenzo, G., Brown, W.C., Florin-Christensen, M., Goff, W.L., 2011. A novel neutralization sensitive and subdominant RAP-1-related antigen (RRA) is expressed by *Babesia bovis* merozoites. Parasitology 18, 1–10.

Terkawi, M.A., Amornthep, A., Ooka, H., Aboge, G., Jia, H., Goo, Y.K., Nelson, B., Yamagishi, J., Nishikawa, Y., Igarashi, I., Kawazu, S.I., Fujisakik, K., Xuan, X., 2009. Molecular characterizations of three distinct *Babesia gibsoni* rhoptry-associated protein-1s (RAP-1s). Parasitology 136, 1147–1160.

Uilenberg, G., 2006. *Babesia* – a historical overview. Vet. Parasitol. 138, 3–10.

Ushe, T.C., Palmer, G.H., Sotomayor, L., Figueroa, J.V., Buening, G.M., Perryman, L.E., McElwain, T.F., 1994. Antibody response to a *Babesia bigemina* rhoptry-associated protein 1 surface-exposed and neutralization-sensitive epitope in immune cattle. Infect. Immun. 62, 5698–5701.

Wright, I.G., Casu, R., Commins, M.A., Dalrymple, B.P., Gale, K.R., Goodger, B.V., Riddles, P.W., Waltisbuhl, D.J., Abetz, I., Berrie, D.A., Bowles, Y., Dimmock, C., Hayes, T., Kalnins, H., Leatch, G., McCrae, R., Montague, P.E., Nisbet, I.T., Parrodi, F., Peters, J.M., Scheiwe, P.C., Smith, W., Rode-Bramanis, K., White, M.A., 1992. The development of a recombinant *Babesia* vaccine. Vet. Parasitol. 44, 3–13.

Yin, H., Lu, W., Luo, J.X., 1997. Babesiosis in China. Trop. Anim. Health Prod. 29, 11S–15S.

Yokoyama, N., Suthisak, B., Hirata, H., Matsuo, T., Inoue, N., Sugimoto, C., Igarashi, I., 2002. Cellular localization of *Babesia bovis* merozoite rhoptry-associated protein 1 and its erythrocyte-binding activity. Infect. Immun. 70, 5822–5826.

Yokoyama, N., Okamura, M., Igarashi, I., 2006. Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. Vet. Parasitol. 138, 22–32.

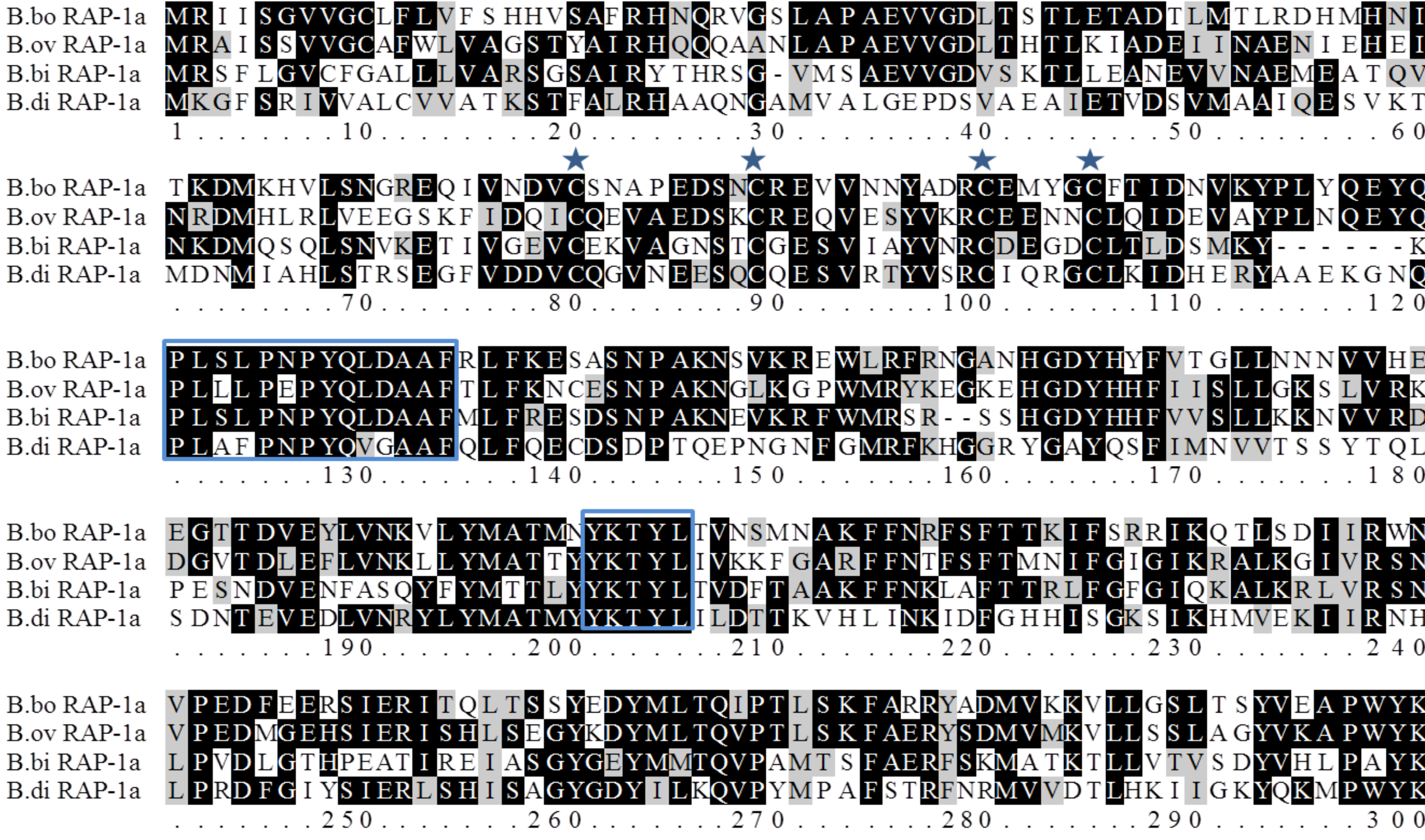
Zhou, J.L., Jia, H.L., Nishikawa, Y., Fujisaki, K., Xuan, X.N., 2007. *Babesia gibsoni* rhoptry-associated protein 1 and its potential use as a diagnostic antigen. Vet. Parasitol. 145, 16–20.

Supplementary table 1. Sequences of the primers used in this study.

Amplification primers			Sequencing primers		
Internal reference	Name	Sequences 5'-3'	Name	Internal Reference	Sequences 5'-3'
rap-1-BQ1 F2	AP1	CCH TAC CAR TTG GAY GCT GCS TTC	-	-	-
rap-1-BQ1 R3	AP2	GTA AGT YTT GTA RTW SRT GGT RGY CAT	-	-	-
61up	AP4	GAG GCA GCC ACA GTG ACT AC	SQ-61up1	2+4_ 61up	GCT TAG AGC GAC CAT CCG CGA
			SQ-61up2	4^{4th}_61up	GTT CAT GGT GTC AGT GGC CTA TGT ATG
			SQ-61up3	4^{5th}_61up	AGT GTT GAA GCG CAT TGG GCG CTT A
61down	AP3	GAC TGA CGA TGA GGT TCT GGT	SQ-61down1	6_61down	TCC GAC GTG CTA CAT TCT CAC
			SQ-61down2	6^{4th}_61down	TTA CTG CAT AAG TGA CCT ATT GAC TGG
			SQ-61down3	6^{5th}_61down	GAC TCT GTA TGT TTA CAG CAC GAG GC
67up	AP6	GAG GCA GCC ACA GTG ACT TT	SQ-67up1	5+6_67up	CAT AGT CGC TGA GCC TGG CAA G
			SQ-67up2	6^{4th}_67up	ATC CTA TAT GTG GCT GTG TTG CCT A
			SQ-67up3	6^{5th}_67up	ACA TGA GGG TGA GCT TGC CAA TGT CT
67down	AP5	GAC TGA CAA TGA GGT GCT GAA	SQ-67down1	3_ 67down	CTG TTG TTA ACG GTT CGC GCT GA
			SQ-67down2	4^{4th} _67down	CCA ATC CTC TTA TCT AAC GTT AAC G
			SQ-67down3	4^{5th}_67down	TGC AGT ACA CGT AGT ATG TAT ACG GC
86up	AP8	CCG CGC GAA TGC CAG TAG GG	SQ-86	86seq	CTA TCA ACA TTG CAG GAG AG
86down	AP7	CCC TAC TGG CAT TCG CGC GG	SQ-86down	2+5_86down	CCA TCA CTA CAG GAA GGG ATG
			SQ-86down2	IG1.3	AAG GCG GGT TAG AGA TCC TAC GGA T

Amplification primers			Sequencing primers		
Internal reference	Name	Sequences 5'-3'	Name	Internal reference	Sequences 5'-3'
bup	AP9	TCG AAG CCT AGT CGC CAA GTT	SQ-bup2	IG1.8	AGA CAG GTG CAT CGC GTA TGT GTG A
bdown	AP10	ATG ACC ATG GGA AGG ATC AAC GC	-	-	-
rap-1a up	AP11	ATG AGA AGC TTC GCG GGT G	-	-	-
rap-1a down	AP12	TCA GGG GCA GCT GCC TCC TC	-	-	-
rap1-b up	AP13	ATG GCA CCA GTG TTG AAG CGC	-	-	-
rap1-b down	AP14	CTA TCC CCT AAC TTG GCG ACT A	-	-	-
rap-1c up	AP15	ATG GTT TGC CAC AGC TTT GTC	SQ-cup	cupseq	AAG GAA GCT TAG CAT TTC GG
p128R[a]	p128R	TAT ACC ACA TCG ATG AAC C	SQ-p128R	p128seq	TTT TAG TTC ACC AGA TCA TG
IG-1.8 up	AP17	GCC GTG CCG TGT TTC CCG TTA	-	-	-
rap-1a67speup	AP18	AAG GAA CAG TGA TTG TCG TC	-	-	-
IG-0.8 down	AP19	ACG GTT GTC ATA ACT CTT GGC CGT C	-	-	-

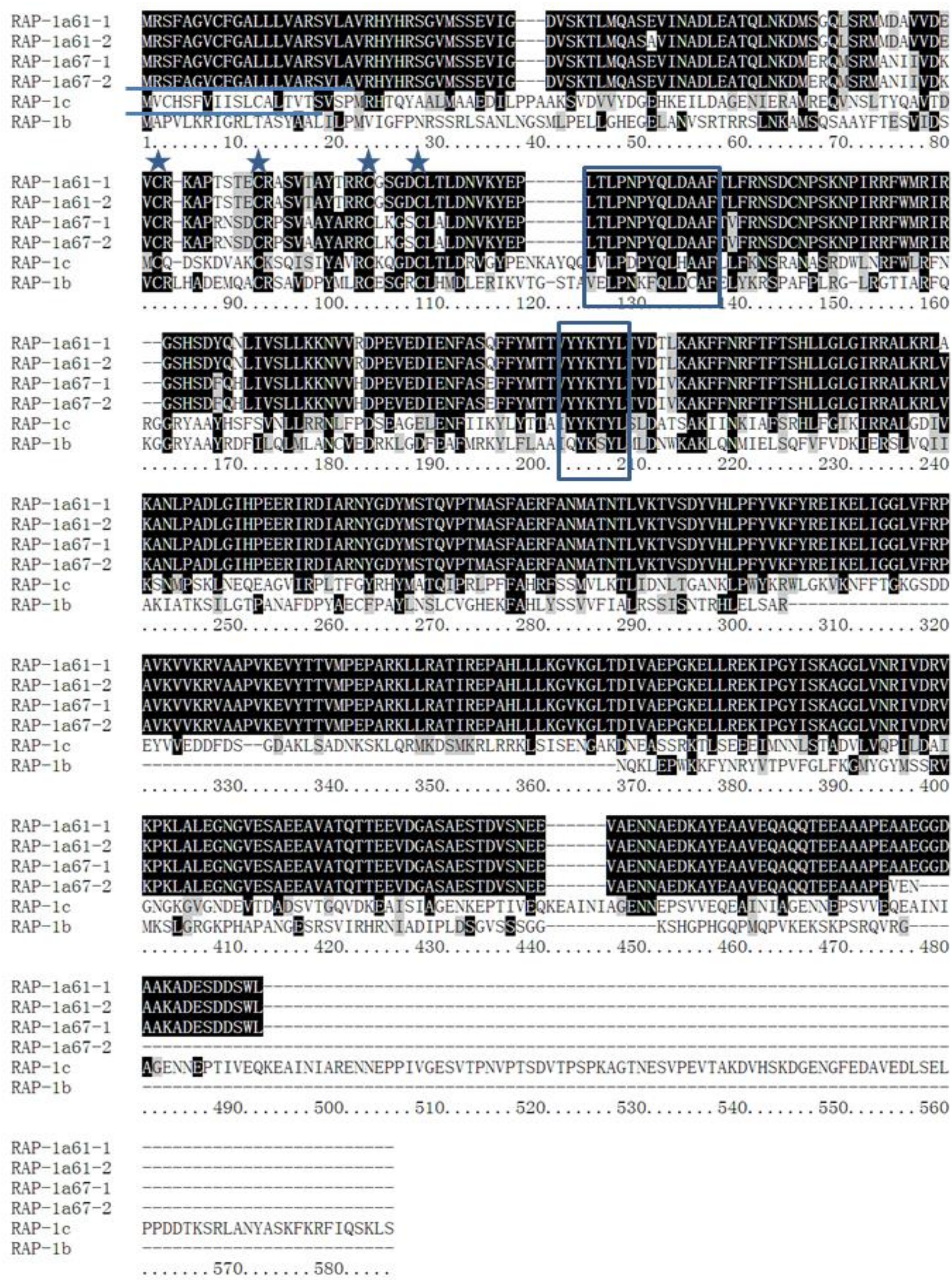
[a] p128R was cited from Suarez et al., 2003.



Supplementary figure 1. **Alignment of RAP-1a partial amino acids sequences from *Babesia* species infecting ruminants.** Conserved 4 cysteine residues are marked with asterisks and two conserved amino acid regions indicated with boxes. It showed also several shorter conserved oligopeptide motifs in the first 300 amino acids. *Babesia ovis* (GenBank Accession number: AAA27805), *B. bigemina* (GenBank Accession number: AAA65583), *B. bovis* (GenBank Accession number: AAB84267), *B. divergens* (Accession number: CAA89970).

```
rap-1a67       ATGAGAAGCTTCGCGGGTGTATGTTTTGGTGCGCTCTTGCTCGTAGCAAGAAGCGTTTTGGCTGTACGCCATTACCACCG
rap-1a67/61-2  ATGAGAAGCTTCGCGGGTGTATGTTTTGGTGCGCTCTTGCTCGTAGCAAGAAGCGTTTTGGCTGTACGCCATTACCACCG
rap-1a61       ATGAGAAGCTTCGCGGGTGTATGTTTTGGTGCGCTCTTGCTCGTAGCAAGAAGCGTTTTGGCTGTACGCCATTACCACCG
rap-1a67/61-1  ATGAGAAGCTTCGCGGGTGTATGTTTTGGTGCGCTCTTGCTCGTAGCAAGAAGCGTTTTGGCTGTACGCCATTACCACCG
               1.......10........20........30........40........50........60........70........80

rap-1a67       TTCTGGTGTGATGTCATCTGAAGTCATAGGCGATGTGTCTAAGACGTTGATGCAGGCCAGTGAGGTGATCAATGCTGATT
rap-1a67/61-2  TTCTGGTGTGATGTCATCTGAAGTCATAGGCGATGTGTCTAAGACGTTGATGCAGGCCAGTGAGGTGATCAATGCTGATT
rap-1a61       TTCTGGTGTGATGTCATCTGAAGTCATAGGCGATGTGTCTAAGACGTTGATGCAGGCCAGTGCGGTGATCAATGCTGATT
rap-1a67/61-1  TTCTGGTGTGATGTCATCTGAAGTCATAGGCGATGTGTCTAAGACGTTGATGCAGGCCAGTGAGGTGATCAATGCTGATT
               ........90.......100.......110.......120.......130.......140.......150.......160

rap-1a67       TGGAAGCTACACAGCTAAACAAGGATATGGAAAGGCAGATGTCACGCATGGCGAACATCATCGTAGACAAGGTTTGCCGC
rap-1a67/61-2  TGGAAGCTACACAGCTAAACAAGGATATGGAAAGGCAGATGTCACGCATGGCGAACATCATCGTAGACAAGGTTTGCCGC
rap-1a61       TGGAAGCTACCCAGCTTAACAAGGACATGAGTGGGCAGTTGTCGCGCATGATGGACGCCGTCGTAGACGAAGTTTGCCGC
rap-1a67/61-1  TGGAAGCTACACAGCTAAACAAGGATATGGAAAGGCAGATGTCACGCATGGCGAACATCATCGTAGACGAAGTTTGCCGC
               .......170.......180.......190.......200.......210.......220.......230.......240

rap-1a67       AAGGCTCCAAGGAACAGTGATTGTCGTCCATCCGTAGCCGCTTACGCTCGTCGTTGCCTTAAAGGTTCTTGTTTGGCATT
rap-1a67/61-2  AAGGCTCCAAGGAACAGTGATTGTCGTCCATCCGTAGCCGCTTACGCTCGTCGTTGCCTTAAAGGTTCTTGTTTGGCATT
rap-1a61       AAGGCTCCTACAAGCACTGAGTGCCGTGCATCCGTGACCGCTTACACTCGTCGTTGCGGTAGTGGCGACTGTTTGACCCT
rap-1a67/61-1  AAGGCTCCTACAAGCACTGAGTGCCGTGCATCCGTGACCGCTTACACTCGTCGTTGCGGTAGTGGCGACTGTTTGACCCT
               .......250.......260.......270.......280.......290.......300.......310.......320

rap-1a67       GGACAACGTGAAGTACGAACCATTGACGCTGCCGAATCCTTACCAATTGGATGCGGCATTCACTGTATTCAGGAACAGTG
rap-1a67/61-2  GGACAACGTGAAGTACGAACCATTGACGCTGCCGAATCCTTACCAATTGGATGCGGCATTCACTGTATTCAGGAACAGTG
rap-1a61       GGACAACGTGAAGTACGAACCATTGACGCTGCCGAATCCTTACCAGTTGGATGCGGCATTCACGCTGTTCAGGAACAGTG
rap-1a67/61-1  GGACAACGTGAAGTACGAACCATTGACGCTGCCGAATCCTTACCAGTTGGATGCGGCATTCACGCTGTTCAGGAACAGTG
               .......330.......340.......350.......360.......370.......380.......390.......400

rap-1a67       ACTGTAACCCGTCTAAGAACCCCATCAGGCGTTTCTGGATGCGTATTAGAGGCAGCCACAGTGACTTTCAGCACCTCATT
rap-1a67/61-2  ACTGTAACCCGTCTAAGAACCCCATCAGGCGTTTCTGGATGCGTATTAGAGGCAGCCACAGTGACTACCAGAACCTCATC
rap-1a61       ACTGTAACCCGTCTAAGAACCCCATCAGGCGTTTCTGGATGCGTATTAGAGGCAGCCACAGTGACTACCAGAACCTCATC
rap-1a67/61-1  ACTGTAACCCGTCTAAGAACCCCATCAGGCGTTTCTGGATGCGTATTAGAGGCAGCCACAGTGACTACCAGAACCTCATC
               .......410.......420.......430.......440.......450.......460.......470.......480

rap-1a67       GTCAGTCTTTTGAAGAAGAACGTTGTGCATGATCCCGAGGTTGAGGACATCGAAAATTTCGCATCGGAGTTCTTCTACAT
rap-1a67/61-2  GTCAGTCTTTTGAAGAAGAATGTTGTGCGTGATCCCGAGGTTGAGGACATCGAAAATTTCGCGTCGCAGTTCTTCTACAT
rap-1a61       GTCAGTCTTTTGAAGAAGAATGTTGTGCGTGATCCCGAGGTTGAGGACATCGAAAATTTCGCGTCGCAGTTCTTCTACAT
rap-1a67/61-1  GTCAGTCTTTTGAAGAAGAATGTTGTGCGTGATCCCGAGGTTGAGGACATCGAAAATTTCGCGTCGCAGTTCTTCTACAT
               .......490.......500.......510.......520.......530.......540.......550.......560

rap-1a67       GACTACGGTGTACTACAAGACCTACCTTACCGTGGATATTGTAAAGGCTAAATTCTTCAACAGGTTCACTTTCACGTCCC
rap-1a67/61-2  GACTACGGTGTACTACAAGACCTACCTTACCGTTGACACGCTGAAGGCTAAATTCTTCAACAGGTTCACTTTCACGTCCC
rap-1a61       GACTACGGTGTACTACAAGACCTACCTTACCGTTGACACGCTGAAGGCTAAATTCTTCAACAGGTTCACTTTCACGTCCC
rap-1a67/61-1  GACTACGGTGTACTACAAGACCTACCTTACCGTTGACACGCTGAAGGCTAAATTCTTCAACAGGTTCACTTTCACGTCCC
               .......570.......580.......590.......600.......610.......620.......630.......640

rap-1a67       ACCTACTTGGCTTAGGCATCAGGAGGGCTTTGAAGCGTCTGGTTAAGGCCAACCTTCCAGCTGACCTTGGTATTCACCCT
rap-1a67/61-2  ACCTACTTGGCTTAGGCATCAGGAGGGCTTTGAAGCGTCTGGTTAAGGCCAACCTTCCAGCTGACCTTGGTATTCACCCT
rap-1a61       ACCTACTTGGCTTAGGCATCAGGAGGGCTTTGAAGCGTCTGGTTAAGGCCAACCTTCCAGCTGACCTTGGTATTCACCCT
rap-1a67/61-1  ACCTACTTGGCTTAGGCATCAGGAGGGCTTTGAAGCGTCTGGTTAAGGCCAACCTTCCAGCTGACCTTGGTATTCACCCT
               .......650.......660.......670.......680.......690.......700.......710.......720

rap-1a67       GAAGAGCGCATTCGCGACATAGCGCGCAACTACGGCGACTACATGTCAACTCAGGTTCCTACGATGGCATCGTTCGCTGA
rap-1a67/61-2  GAAGAGCGCATTCGCGACATAGCGCGCAACTACGGCGACTACATGTCAACTCAGGTTCCTACGATGGCATCGTTCGCTGA
rap-1a61       GAAGAGCGCATTCGCGACATAGCGCGCAACTACGGCGACTACATGTCAACTCAGGTTCCTACGATGGCATCGTTCGCTGA
rap-1a67/61-1  GAAGAGCGCATTCGCGACATAGCGCGCAACTACGGCGACTACATGTCAACTCAGGTTCCTACGATGGCATCGTTCGCTGA
               .......730.......740.......750.......760.......770.......780.......790.......800
```

Supplementary figure 2. Partial nucleotide sequence alignment showing the *rap-1a* chimera created in the course of PCR.



Supplementary figure 3. Alignment of RAP-1a, RAP-1b and RAP-1c predicted amino acid sequences of *Babesia* sp. BQ1 (Lintan). The 4 conserved cysteine residues are marked with asterisks, the two conserved amino acid sequences used to design degenerated primers indicated with boxes, and signal peptide sequences are underlined.

GGCTAAGTCAAATAATAGGCGAATCGTATTACTGTGATGGAATCGCTAGCAA
ACTATGATA**Y**TTAATGCGTCTCGTATTTCAATCCCACGTAAATCCGCCATTT
ATCGTCGCATTAGGCGTTTGAGATAATACGTTGTAGTGACGCTAGGGCATGG
CGGATGTCCGATAGGATAGCTTAATGTTGTTCATGGTGTCAGTGGCCTATGT
ATGGGTTTTATATAGTTGCTTAAGCGGATTACCGGACGTACATATGGCACTC
TAAAACCAGAAACTCGCGGTTCCGCTCCTATGACAACACCGTGGGATGAGGA
TTCATGCATCCATCCTATATGTGGCTGTGTTGCCTACCCGACGGCCAAGAGT
TATGACAACCGTGTTTCGCAGCGAAAGCCGTTACTACATTCCGCTGCTTACA
TTCTACCGTTAATTTTATACGGTAAATACATCAGACCTCTTGGGCACTTTAC
TGATGTCCGCGAGAACTTTCGCAGCGAACGTGTAGGAATACGCGTATTAGAC
ACGGGGGTGTGCTGGAGAAAATACCCCTTGCGTGTTGCTGTGCAATACTGAC
CTTGTATACGGTTCAAAGGCGTTACAATGTGATTGCTTTGCGCTTTGGTGTG
TCCAGGCAACCTGCGAGATTTCGCACTTTTCTGCATCAATGTTACGGTCAAC
TGTATTGGTTGGCAGTGACTGCGACGTCTCCTTTCACGTGAAATATAACATA
ATGTACCGTATCGTGGTAACTGAGAACTTGTAGTTCTCGTCCACGTAATTGC
AGCCAGTTCTGTAGAGGCTGTGCTTAACACACTAATAGGTGTTTGGGTTAAC
AACGGCATTAAAGAAGTCTTG

Supplementary figure 4. Intergenic region IG1 sequence (853 bp) between *rap-1a* and *rap-1b* genes. One point mutation C/T (Y) located 62 bp after *rap-1a61/rap-1a67-1* putative stop site is indicated with bold.

AATGAAGTAATTGTTGCTGGCAAGTTATTATGGCATGCAGCGTGGTTAAGGGTAG
AAGCGAGTGGAACGTAGAATTTCCCGCTCATTTATTAGGCTATACTGCATCGAAT
GCACCTTAATCGCAAACTATTGCCGTATACATACTACGTGTACTGCAATTCCACA
TGAGGCGCCGGCAGATTGCGGATCCCGTGTAGCTCAGGGTCCCTTGCGCATGAAC
CTTTGCGTGCTAATATATCCAGCATAATGCTTGATGTGCTCGACGCTTATCACCA
CATTTGCATTTGTGAAATCTATCCTGCGTATGCATATGGAATATTTGTTTTGCCT
CGTGCTGTAAACATACAGAGTCAGTTGATGTATACTAGCGAAAGCAGTTAGATGG
TGAGTTTTTCGTCACATTTCGCGTTGTGCATGTATTGATGTGTCTCCGATGATTG
CTGTTGGTGGTAGCGTAGCTGCGCGATGGGAGAGCGTCGTGAGATACCGATCACC
GGACGATCGGCGTTGATAATGGATGTTCAGATGCTATTATGGGTGTTGTGCACCT
GTTTTATTGGTGGTTATGCTTGTGGCGGGATTTCTGATATAGACGTGGTTCCCAT
CTTACGGATTCCGTGCAGACATCGAAGTACACGTATGCATTGATTGAGGGCGCAG
CCGCTCTTTCAATTGCTCGCGCCATCGCTCTCGAATGCTGCGAGCGTTTGAATGC
GCAGAATGGAGATTTTTCAGGTATGCAACGGATTGGGAATGGCTAGAAGAGTAGC
ATACGATGCTCTCATCAATTGTTCTTTTTCGTTAACGTTAGATAAGAGGATTGGA
GTGTATTAAGGTAAGTTAACATGGGGAATAGCTCTGCATAACCCTCGCTAGTGAG
CCGGTGGAGGTTGCAGACAACACGACCTCGCTGTGGGAACTTATGACGGTAAAGA
TACCAGTCAATAGGTCACTTATGCAGTAAGCACGAGGCTTGAATGCGGCACGCCC
ATTCCATGAATGAGCAAGAATTCACTGGAGACCCTAAACAGATTTCACGCATTCC
GAACGGAAGTGTCCACTTAGCACGCCAACAGGCCTTATAATTATTGTGAGTGACA
TATAGAGGAGCGATGCGAACACGCTGAGGCTTCGTAAGCGTACGGGCGCCACCTA
TCCATGTGCCATAGAAACCTAAGGAACGGTTGGAATACACCACGTGGAGAACCAG
GCAGCTGATTTTGCCACTCGCATTGAGCATTATTGCGGCGAAGAGACAAGAAACC
AACACCTCAACCTAGAATTAGACATGCCGAAAACCTCAGTGTGTGGCCATACCAT
TTGTTACTGATGGCGGGTATTTCAGTTGTCTTTCTTCATATTTCTAACATGGCGC
ACAAATGAAACAATTGAAAGCATGCTGCATATGCAGTAATGCATATAACCACAGC
GTAGTATGAGCCAGTCCGCGTGGGAGATGGACGTAAGCGTGAGTACATATCGCA
ACAAAAGCAGGTATCGACGCTAAACAGCGCGAACCGTTAACAACAGTTAAGGTGT
CACGTAATGTGGAACACACGGGAGGCGTGTCTATGCGTGAAGGCAGCCATAAGGA
GCGTTTCGCACTTTGTTGCAGCAATTGGTGTTTTAGACAGGTGCATCGCGTATGT
GTGAGAATGTAGCACGTCGGAGCATATCCATTGTGAATGCTTGTTTGTTGAGGCG
GAGATATATCGGCCGTGCCGTGTTTCCCGTTACATTTATACTAATAGCAGAGTCG
TATATACGATCCGCTTAGATTATCGTATAATAAATATTCTTCTTGTAGCCTTTCT
TTGCTGCTGCCCGTTGGTCTTCACG

Supplementary figure 5. Intergenic region IG2 sequence (1840 bp) between rap-1b and rap-1a genes.

CGAATTTTGTTACCCTTGCATACGTCATTTTCCCGTTAATCATGCAGATATT
CTATTATATCGAATGTAAAATGTTTTTAATTGTTGTATTACACCATATAGGC
TTATTTTATTGCCATCTTTATTGAACGTGTGTCATAATATTAATGAACTCAT
TATTCATCTCGTGGTATAGTGTTGTTTCTATGTGTAATTTTTTGTTTACTCT
TCGATTCTGGCCGAACACCGCCTATGCGGGAGTGTTTTTCTGTGAAACATTT
TAAGCGAAATCTCTCTCCCGGGATGAACCATATTCTTCATACAGTCGCGAAG
ATAAATTGATTATCACATTTTCACGCAAACGACTCTTAAATACATTATAGAA
TGTCGAAACGAACTGTTAAATATTTTCGTAACAATATGTTAGCTATGACGTA
ATAACCTACGACATCCGTAGGATCTCTAACCCGCCTTTGGTTGTAAACTGCG
ACTGCACGTGCGACCCCAAATGACTCTCAACTAGTTAACAGCACTGAATACG
TGACATCTCCTTTATAAACAATCACCAACACCCGATAACACTACTGACCATA
AGATGCTTACAATTGTGATTAAATGATAACGCATTTTCGATGCGTTAATAAC
AACCTTCACGCAGACAACTATAAGTAAAAGAAGTCTGGAGGAAGATCATGAG
CCGCAGGATTTGTAGATGAAACATTCAAAAAGTGCTCTAAGAATGAAACAGA
ACTTAGCAGAAAGTCAGAATGTTCGCAGAGAACGTTTTATAATTGTTAGACA
CCATCAGCCAATAGTGCTAATACACAGGCAGCGGGATCAACCGAAGGATGAC
AAAAATTACCACTGACCACATATTTCTGCAGAACCAATATGATTATAAAACT
TATATTACATGTAGTACATGTCGTAGAATACAGCAGCGATGCATCATTATTA
ATTTGATCGAAAGTAGTTACAGATGGAATCTAGTAATCCTTCTGCCAAAACA
AGAGGTTAACTTTCGATGATTGTACACATTATGGAATTATGAGAAGAAGAAG
GCAAATTACATCTCTCTGACTTCAAGTCGTTAGCGAATCCTGTCTACTTCAC
TTATTGCACATAGGTGCATCAACGTATCGTATGCAGGGTGCAATTGCCGCGC
CGTTTTACTTCGCGAGTGATTGATCGAACAACGCCGTATGGACGTGTAGAAT
ATAGATAATATCGGATGTTCAGTCTATTTACGTGTATAGCTGTGCTATACGC
CTTGGTGTGTGTCTATACTGTGAACTTCGACTAGAAAACCCCTGTCTAAATA
GCTAAAAATTCTTTAACATCCCTTCCTGTAGTG

Supplementary figure 6. Intergenic region IG3 sequence (1333bp) between rap-1a67-2 and rap-1c gene.

Article N° 2: Strong conservation of rhoptry-associated-protein-1 (RAP-1) locus organization and sequence among *Babesia* isolates infecting sheep from China (*Babesia motasi*-like phylogenetic group)

Article N° 2: Strong conservation of rhoptry-associated-protein-1 (RAP-1) locus organization and sequence among *Babesia* isolates infecting sheep from China (*Babesia motasi*-like phylogenetic group)

Qingli Niu[a,b]*, Charlotte Valentin[a,b], Claire Bonsergent[a,b], Laurence Malandrin[a,b]

INRA, UMR1300 Biology, Epidemiology and Risk Analysis in Animal Health, CS 40706, F-44307 Nantes, France.
LUNAM Universitẻ, Oniris, UMR1300 BioEpAR, F-44307 Nantes, France.

* Corresponding author: Oniris, UMR1300 INRA / Oniris, Biology, Epidemiology and Risk Analysis in animal health, Route de Gachet, La Chantrerie, CS 40706, F-44307 Nantes, France. Phone: +33 (0) 2 40 68 78 57 Fax: +33 (0) 2 40 68 77 51 - E-mail address:

qingli.niu@oniris-nantes.fr and niuqingli@caas.cn

Article has been submitted to Parasitology Research

Abstract

Rhoptry-associated-protein 1 (RAP-1) is considered as a potential vaccine candidate due to its involvement in red blood cell invasion by parasites in the genus *Babesia.* We examined its value as a vaccine candidate by studying RAP-1 conservation in isolates of *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu and *Babesia* sp. Hebei, responsible for ovine babesiosis in different regions of China. The *rap-1* locus in these isolates has very similar features to those described for *Babesia* sp. BQ1 (Lintan), another Chinese isolate also in the *B. motasi*-like phylogenetic group, namely the presence of three types of *rap-1* genes (*rap-1a*, *rap-1b* and *rap-1c*), multiple conserved *rap-1b* copies (5) interspaced with more or less variable *rap-1a* copies (6), and the 3' localization of one *rap-1c*. The isolates *Babesia* sp. Tianzhu, *Babesia* sp. BQ1 (Lintan) and (Ningxian) were almost identical (average nucleotide identity of 99.9%) over a putative locus of about 31 Kb, including the intergenic regions. *Babesia* sp. Hebei showed a similar locus organization but differed in the *rap-1* locus sequence, for each gene and intergenic region, with an average nucleotide identity of 78%. Our results are in agreement with 18S rDNA phylogenetic studies

performed on these isolates. However, in extremely closely related isolates the *rap-1* locus seems more conserved (99.9%) than the 18S rDNA (98.7%), whereas in still closely related isolates the identities are much lower (78%) compared with the 18S rDNA (97.7%). The particularities of the *rap-1* locus in terms of evolution, phylogeny, diagnosis and vaccine development are discussed.

Keywords: *Babesia motasi*-like ovine species; rhoptry-associated-protein 1; genetic diversity; multigene family

1. Introduction

Babesiosis, caused by the multiplication of intraerythrocytic protozoa from the genus *Babesia*, is usually characterized by fever, anemia, hemoglobinuria and even death in severe cases, and occurs in tropical, subtropical as well as temperate regions of the world. Ovine babesiosis is caused mainly by *B. ovis*, *B. motasi*, *B. crassa* (Uilenberg 2006). In China, seven *Babesia* isolates have been described as responsible for ovine babesiosis (Yin et al. 1997; Bai et al. 2002). These isolates have been divided into two groups, based on phylogenetic analysis of the 18S rDNA gene and ITS sequences (Liu et al. 2007; Niu et al. 2009). Most of the described isolates (*Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu, *Babesia* sp. Madang, *Babesia* sp. Hebei and *Babesia* sp. Liaoning) form a sister clade (*B. motasi*-like) with European *B. motasi* isolates. The other group consists of *Babesia* sp. Xinjiang and several *Babesia* recently described on wild ruminants in South Africa and *B. pecorum* isolated from red deer in Spain (Guan et al. 2009; Oosthuizen et al. 2009; Jouglin et al. in press).

Clinical symptoms of babesiosis become apparent when merozoites of *Babesia* species invade and replicate within the host erythrocytes and attain detectable parasitemia (Yokoyama et al. 2006). During the process of erythrocytic invasion, a series of molecules are secreted by the apical organelles (rhoptries, micronemes and dense granules), such as AMA-1 (Associated Membrane Antigen 1), TRAP (Thrombospondin-related adhesive protein) or RAP-1 (Rhoptry-Associated-Protein-1) (Lobo et al. 2012).

The RAP-1 proteins in some *Babesia* spp. and *Plasmodium* spp. have been well characterized (Sam-Yellowe 1996). Their significant role in the invasion process and the implications for vaccine design against *Babesia* infections have been best deduced in *B. bovis* and *B. bigemina* based on the inhibition of parasite invasion or growth in vitro with MAb (Yokoyama et al. 2002; Mosqueda et al. 2002), and the use of native and recombinant

RAP-1 protein to induce immunity and protect animals upon challenge (Wright et al. 1992; Norimine et al. 2003; Brown et al. 2006; Fish et al. 2008). RAP-1 is an immunogenic protein, encoded by multiple gene copies in all the *Babesia* species studied so far. The conserved features described for members of the *Babesia rap-1* family include four conserved cysteines, the presence of a signal peptide for export of the protein to the rhoptry organelle, a highly conserved 14 amino acids motif (PLTLPNPYQL**DAAF**) and several shorter conserved oligopeptide motifs (Y**YKTYL**TVD) in the first 300 amino acids, with sequence identities ranging from 30 to 45% among species (Suarez et al, 1991, 1994; Dalrymple et al. 1993; Skuce et al. 1996). Conservation of the RAP-1 amino acid sequences among different geographic strains is essential when the aim is to target a candidate antigen for subunit vaccine development. Although the RAP-1 sequences of *B. bovis* and *B. bigemina* display high sequence identity (45%) in the N-terminal part (first 300 amino acids) responsible for cross reactions between these two species in ELISA (Suarez et al. 1991), specific immunogenic B and T cell epitopes of RAP-1 recognized by RAP-1 specific antibodies are conserved among geographical strains of each species, but not between different species (McElwain et al. 1991; Shompole et al. 1994; Brown et al. 1996; Palmer et al. 1991; Rodriguez et al. 1996). In our recent companion study, we described the features of *Babesia* sp. BQ1 (Lintan) *rap-1* genes and locus organization, and their striking resemblance with the complex locus described for *B. bigemina* : three gene types (*rap-1a rap-1b* and *rap-1c*), sequence polymorphism among the *rap-1a* copies, interspacing of *rap-1a* and *rap-1b* genes, strict sequence identity of the *rap-1b* gene copies, and conservation of the overall locus organization despite important differences in gene and intergenic sequences (Niu et al. 2013).

In the present study, we analyzed the *rap-1* locus for three other ovine isolates from the *B. motasi*-like phylogenetic group (*Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu and *Babesia* sp. Hebei), to investigate gene polymorphism and the interest of this locus as a common target for the development of recombinant vaccine against different members of the *B. motasi*-like group in China.

2. Materials and methods

2.1 *B. motasi*-like isolates

Babesia sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu, and *Babesia* sp. Hebei, naturally infective to sheep were collected from the towns of Ningxian and Tianzhu in Gansu Province and Chengde town in Hebei Province (Yin et al. 1997). The parasites were isolated by inoculating field-collected blood from asymptomatic sheep into hemoprotozoa-free splenectomized sheep. When parasitemia was more than 5%, venous blood was collected into heparinized tubes and cryopreserved in liquid nitrogen at the Vector and Vector-borne disease (VVBD) laboratory of Lanzhou Veterinary Research Institute (LVRI), CAAS Lanzhou, China.

2.2 Genomic DNA samples preparation

Genomic DNA was extracted from infected blood samples of each strain using a genomic DNA Purification Kit (Gentra, USA), according to the manufacturer's instructions. The amount of DNA isolated, assessed photometrically, was 100 ng/µl. One hundred microliters of genomic DNA sample was spotted on to FTA paper (No. FT6194905, Whatman- GE Healthcare Life Sciences), air-dried, then sent to Oniris by LVRI, China and stored at room temperature until use.

The genomic DNA was then used according to the manufacturer's instructions. Briefly, a disk (1.2 mm) was punched out of the sample area on the Whatman FTA Card and transferred to a PCR tube using a Harris Micro-Punch, washed three times with 200 µl of FTA Purification Reagent (80686019, WhatmanTM-500ml), each wash being incubated at room temperature for 5 min. The disk was then washed twice in TE buffer (10mM Tris, 1mM EDTA·Na_2, pH 8.0) (V6232, Promega-500ml) for 5 min, after which it was ready for direct use in PCR amplification.

2.3 Amplification, cloning and sequencing of *rap-1* genes and partial locus from *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu isolates

Due to the close genetic relationships between the studied isolates, the primers used to amplify *Babesia* sp. BQ1 (Lintan) *rap1-a*, *b* and *c* genes were used in the present study with the following primer combinations: AP11-AP12, AP13-AP14 and AP15-AP20, respectively. (Supplementary Fig.1a, step 1). These primers have already been described (Table 1 in Niu et al. 2013). The PCR was performed with Ex *Taq* Hot-Start DNA polymerase (TaKaRa) and cycles were limited to 30 to reduce PCR chimera production.

The *rap-1* intergenic regions IG1 (*rap-1a* to *rap-1b*), IG2 (*rap-1b* to *rap-1a*) and IG3 (*rap-1a* to *rap-1c*) were amplified, respectively, with the specific primer combinations AP16/18-AP10, AP9-AP3/5 and AP11-AP20. Efficient amplification of the *Babesia* sp. Tianzhu IG3 was obtained with the primer combination AP9-AP7 (Supplementary Fig. 1, step 2). All PCR amplicons were cloned into pGEM-T Easy vector according to the manufacturer's instructions (Promega). Colonies with the expected inserts were selected by direct colony PCR using vector primers. Plasmids were then extracted (NucleoSpin plasmid extraction, Macherey-Nagel) and the inserts were sequenced. Long inserts were sequenced by successively designing internal primers.

2.4 Amplification, cloning and sequencing of *rap-1* genes and partial locus from the *Babesia* sp. Hebei isolate

The above strategy for the amplification of *rap-1* genes was unable to provide *rap-1* expected sequences for *Babesia* sp. Hebei (data not shown). A degenerate primer pair (AP21-AP22) was designed based on alignments between orthologs of *rap-1b* genes in *B. bigemina* and *Babesia* sp. BQ1 (Lintan), (Supplementary Fig.1b, step 1). Based on the sequence of the amplified internal region, specific primers (AP23-AP24) were designed to amplify putative intercalating *rap-1* genes as well as the intergenic regions (IG1, IG2). A PCR product about 5 Kb in size, corresponding to the 3'end of a *rap-1b* copy, an IG2, an intercalating gene of *rap-1a*, an IG1 and the 5' end of the following *rap-1b* copy was obtained with the primer combination AP23-AP24 (Supplementary Fig.1b, step 2). The obtained sequence was then used to design primers to amplify the entire *rap-1a* gene (AP25-AP26) and the entire *rap-1b* (AP27-AP28) (Supplementary Fig.1b, step 3). The same strategy was used to amplify the *Babesia* sp. Hebei *rap-1c* gene. Degenerate primers (AP29-AP30), based on alignments between the *rap-1c* genes in *B. bigemina* and *Babesia* sp. BQ1 (Lintan), were designed to amplify and sequence the *rap-1c* internal region (Supplementary Fig. 1b, step 4). The primer combination AP25-AP31 was used to amplify the intergenic region upstream from *rap-1c*. AP31 was designed according to the sequence obtained by amplifying the *rap-1c* internal region. A PCR product about 3.2 Kb in size was obtained. All PCR products were cloned in pGEM-T Easy vector (Promega) and multiple clones were sequenced. The sequences of the primers used in this study are given in Table 1.

Table 1. Sequences of the primers used in this study.

Isolates	Genes	Internal reference	Name	Sequences (5'-3')	Name	Sequences 5'-3' (for sequencing)	Reference
NX[a] ,TZ[b]		rap-1a up	AP11	ATG AGA AGC TTC GCG GGT G	-	-	Niu et al.,2013
	rap-1a	rap-1a down	AP12	TCA GGG GCA GCT GCC TCC TC	-	-	Niu et al.,2013
	rap-1b	rap1-b up	AP13	ATG GCA CCA GTG TTG AAG CGC	-	-	Niu et al.,2013
		rap1-b down	AP14	CTA TCC CCT AAC TTG GCG ACT A	-	-	Niu et al.,2013
		rap-1c up	AP15	ATG GTT TGC CAC AGC TTT GTC	86up (AP8)	CCG CGC GAA TGC CAG TAG GG	Niu et al.,2013
	rap-1c	rap-1c down	AP20	TTA CGA TAA TTT GGA TTG T	-	-	Niu et al.,2013
HB[c]	rap-1b internal	HBrap-1bF	AP21	GCC ATG TCT CAA TCT GC**R** GC	-	-	In this study
		HBrap-1bR	AP22	CG**R** GTG TTG CTG AT**K** GAC GA	-	-	In this study
	rap-1b 5',3' ends and IG1, IG2	HBrap-1bup	AP23	TTC TCA GCC TAC CTA AAT TCT	HBrap-1bup1 HBrap-1bup2	GCA CGA CGC TGG TCA CTA CAT CCT AGC ACG CCA ATG CGC TT	In this study
		HBrap-1bdown	AP24	GTA CGG CTC AAC GAC GGA CCG	HBrap-1bdown1 HBrap-1bdown2	GCC CAT GAG TTC CGA TGT ATG CAC CCT TCA CGA CGA CAT	In this study
	rap-1a	HBrap-1aup-full	AP25	ATG AGA AGC TTC GTG AGT GTA	-	-	In this study
		HBrap-1adown-full	AP26	TTA AAG CCA CGA ATC AGT AGA C	-	-	In this study
	rap-1b	HBrap-1bup-full	AP27	ATG GCA CTA GTT TCT AAG CGC C	-	-	In this study
		HBrap-1bdown-full	AP28	CTA TCC TCT AAC TTG GCG AAT	-	-	In this study
	rap-1c internal	HBrap-1cF	AP29	CCT GAC CC**M** TAC CAG CT**K** CA	-	-	In this study
		HBrap-1cR	AP30	GCC ATG TA**M** TGC CTG TAA CC	-	-	In this study

	rap-1c 5'end, IG3	HBrap-1cdown	AP31	TTC AGA ACG TTG AGA GTG AAG G	-	-	In this study
	rap-1c 3'end (?)	HBrap-1cup-full	AP32	ATG GTT GGC TAC AGC TTT GTC	-	-	In this study
		HBrap-1cup	AP33	AGA GCT TTG GCC AAC ATG GTG	-	-	In this study
		p128R	p128R	TAT ACC ACA TCG ATG AAC C	-	-	Suarez et al., 2003
NX[a], TZ[b]	IG1 (61-b,67-b)	rap-1a61speup	AP16	TAC AAG CAC TGA GTG CCG TG	-	-	In this study
		rap-1a67speup	AP18	AAG GAA CAG TGA TTG TCG TC	-	-	Niu et al.,2013
		bdown	AP10	ATG ACC ATG GGA AGG ATC AAC GC			Niu et al.,2013
	IG2 (b-61,b-67)	bup	AP9	TCG AAG CCT AGT CGC CAA GTT	IG2-seq	TCT TAC GGA TTC CGT GCA GAC ATCG	Niu et al.,2013; this study
		61down	AP3	GAC TGA CGA TGA GGT TCT GGT	-	-	Niu et al.,2013
		67down	AP5	GAC TGA CAA TGA GGT GCT GAA	-	-	Niu et al.,2013
NX[a]	IG3 (67.2-c)	rap-1aup	AP11	-		-	Niu et al., 2013
		rap-1cdown	Ap20	-	2+5 _86down IG3-seq	CCA TCA CTA CAG GAA GGG ATG AAC GTT CTC TGC GAA CAT TCT GA	Niu et al., 2013 In this study
TZ[b]	IG3 (67.2-c)	bup	AP9	TCG AAG CCT AGT CGC CAA GTT	AP11, AP12	-	Niu et al.,2013
		86down	AP7	CCC TAC TGG CAT TCG CGC GG	IG3-seq, AP7	-	Niu et al.,2013; this study

a: *Babesia* sp. BQ1 (Ningxian); b: *Babesia* sp. Tianzhu; c: *Babesia* sp. Hebei

2.5 Bioinformatics analysis

All sequences obtained in this study were subjected to blast searching on the NCBI website: http://www.ncbi.nlm.gov/blast/cgi, using the program BLASTn. A search for the presence of putative signal peptides, and the cleavage site in the N-terminal part of the protein, was carried out using the software tools PrediSi (http://www.predisi.de/home.html) and SignalP (http://www.cbs.dtu.dk/services/SignalP/). Phylogenetic analysis was performed using the software MEGA 5.2. The different genes comprising the *rap-1* locus were analysed by the "neighbor-joining" method based on genetic distances. A tree was also constructed by including RAP-1 sequences from other species of *Babesia* spp., such as *B. gibsoni*, *B. caballi*, *B. ovis*, *B. bovis*, *B. orientalis* and *B. divergens,* in order to show the relationships between the isolates in our study and relate them to other RAP-1 proteins.

3. Results

As the *rap-1* gene family is multigenic and polymorphic, amplification of the conserved regions of each amplicon (gene or intergenic region) with primers resulted in a mixed population of sequences. These were systematically cloned and multiple clones were sent for sequencing to address the polymorphism among gene copies. The cloning and sequencing results obtained at each step for the three different isolates are summarized in Table 2.

Table 2. Summary of the number of inserts sequenced for each gene or intergenic regionsequencing result.

	Genes			**Intergenic regions**				
	rap-1a	*rap-1b*	*rap-1c*	IG1 (61-b)	IG1 (67-b)	IG2 (b-61)	IG2 (b-67)	IG3 (67.2-c)
Babesia sp. BQ1 (Ningxian)	15[a]-12[b]	11-9	12-10	4-3	12-10	13-12	13-12	5-4
Babesia sp. Tianzhu	12-10	13-13	12-9	8-3	12-9	2-2	12-10	1-1
Babesia sp. Hebei	20-17; (2-2)[c]; (6-4)[d]	12-12	7-4 (5'end region)	IG1: (a1-b), 2-2		IG2: (b-a1), 2-2		IG3: (a3-c), 6-4

a: Numbers of samples sent for sequencing
b: Numbers of complete sequences obtained
c: *rap-1a* sequences from partial locus amplification, containing IG1 and IG2
d: *rap-1a* sequences from partial locus amplification, containing IG3

3. 1. Sequence variation of *rap-1a* genes in the Chinese *B. motasi*-like group

3.1.1. *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu: a copycat of *Babesia* sp. BQ1 (Lintan)

Primers based on the *Babesia* sp. BQ1 *rap-1a* gene sequence (AP11-AP12) (Niu et al., 2013) were used to obtain amplicons of the expected size (partial *rap-1a* sequences of 1367 pb). The *rap-1a* sequences from both isolates were similar to the *rap1a61* and *rap-1a67* described for *Babesia* sp. BQ1 (Lintan). To obtain the complete *rap-1a* sequence, the 3' end of the *rap-1a* gene was amplified and sequenced with the intergenic regions downstream from the genes (primers AP16-AP10 for *rap-1a61* and AP18-AP10 for *rap-1a67-1* in both isolates; AP11-AP20 for *rap-1a67-2* in the *Babesia* sp. BQ1 (Ningxian) isolate and AP9-AP7 for *rap-1a67-2* in the *Babesia* sp. Tianzhu isolate).

In summary, both *rap-1a61* and *rap-1a67* genes were also found in the *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu isolates (Table 3). Two different copies of *rap-1a67* were also discovered, *rap-1a67-1* (1425bp) and *rap-1a67-2* (1380bp), which were perfectly identical with their orthologs in the *Babesia* sp. BQ1 (Lintan) isolate (Fig. 1) (Niu et al. 2013). The *rap-1a61* copies presented a polymorphism at two nucleotide positions from the putative start site: 683 nt with C/T (non synonymous substitution A/V) and 1248 nt with C/T (synonymous mutation G/G) resulting in the aa combinations AG or VG at these two positions. However, the additional polymorphic site described in *Babesia* sp. BQ1 (Lintan) at nucleotide 143 (A or C) (Niu et al. 2013) was not found in the *rap-1a61* genes of these two isolates (only C). Two *rap-1a61* gene copy types were therefore present at the *rap-1* locus in *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu: *rap-1a61-2* (combination CTT) and a new one, *rap-1a61-3* (CCC) (Table 3). Except for these three point mutations, the *rap-1a61*gene sequences were identical in all three isolates, with 1425 bp encoding a putative 474 amino acids protein for *rap-1a61*.

Table 3. Position and nature of the point mutations that differ in different *rap-1a* gene copies as well as the ratio of each sequence type obtained from the whole gene amplification and cloning for each isolate, and their rough ratio in the *rap-1* locus.

***rap-1a* type**								
	Position: nucleotide n°			Position: amino acid n°			Name	Ratio
	143	683	1248	48	228	416		
Babesia sp. BQ1 (Lintan)[a]	A	T	C	E	V	G	*rap-1a67*	8/27
	A	C	C	E	A	G	*rap-1a61-1*	9/27
	C	T	T	A	V	G	*rap-1a61-2*	10/27
Babesia sp. BQ1 (Ningxian)	A	T	C	E	V	G	*rap-1a67*	2/12
	C	T	T	A	V	G	*rap-1a61-2*	7/12
	C	C	C	A	A	G	*rap-1a61-3*	3/12
Babesia sp. Tianzhu	A	T	C	E	V	G	*rap-1a67*	3/10
	C	T	T	A	V	G	*rap-1a61-2*	4/10
	C	C	C	A	A	G	*rap-1a61-3*	3/10

a: Data was cited from Niu et al., 2013.

3.1.2. Three different *rap-1a* copies exist in the *Babesia* sp. Hebei isolate

Amplification of the *rap- 1a* genes, using the primers AP11-AP12, was never successful in *Babesia* sp. Hebei. The central region of the *Babesia* sp. Hebei *rap-1b* gene was successfully amplified and sequenced using degenerate primers (AP21-AP22) designed by aligning the *rap-1b* gene in related species (*Babesia* sp. BQ1 (Lintan) and *B. bigemina*) (Supplementary Fig.1b, step 1). As the *rap-1a* and *rap-1b* genes have been previously described as interspaced, new primers (AP23-AP24) designed from the *rap-1b* central region, were used to amplify the regions downstream and upstream of the *rap-1b* gene which supposedly contain the *rap-1a* gene between two *rap-1b* genes (Supplementary Fig.1b, step 2). Two identical sequences of an entire *rap-1a* gene (1401bp) located between two *rap-1b* genes were obtained from two clones and blasted with *Babesia* sp. BQ1 (Lintan) and *B. bigemina rap-1a* genes. Polymorphism of the *rap-1a* sequence was addressed by amplification with a newly designed primer combination (AP25-AP26), followed by cloning and sequencing (Supplementary Fig.1b, step 3). Two *rap-1a* copy types were found: the predominant one (14 clones/17) corresponded to the first *rap-1a* sequence discovered (named *rap-1a1*, 1401 pb encoding a putative protein of 466 amino acids), and the less abundant one (3 clones/17) with a size of 1416 bp (*rap-1a2*) encoded a putative 471 amino acids protein (Fig. 1).

Amplification and sequencing of the 5' end region of *rap-1c* and the intergenic region 3 (IG3) (AP25-AP31) revealed a third *rap-1a* gene type upstream of *rap-1c* (Supplementary Fig.1b, step 5). The sequences in 4 clones were identical and were named *rap-1a3,* with 1392 bp encoding a putative 463 amino acids protein (Table 2, Fig. 1).The sequences of the three *rap-1a* gene types were perfectly identical from nt 1-1226 onwards. *rap-1a2* differed by an insertion of 15 nucleotides from position 1227 nt to 1241 nt (GGCTGCTTCTGCCGA, encoding EAASA) (Supplementary Fig. 2). In *rap-1a3,* the sequence differed (deletions and substitutions) from position 1371 to the putative stop site of *rap-1a3* (Supplementary Fig. 2).

The molecular features of the *rap-1* family members, namely the 4 cysteines and the patches of conserved motifs, including PLTLPNPYQL**DAAF** and Y**YKTYL**TVD in the first 300 amino acids, were conserved (Fig. 1).

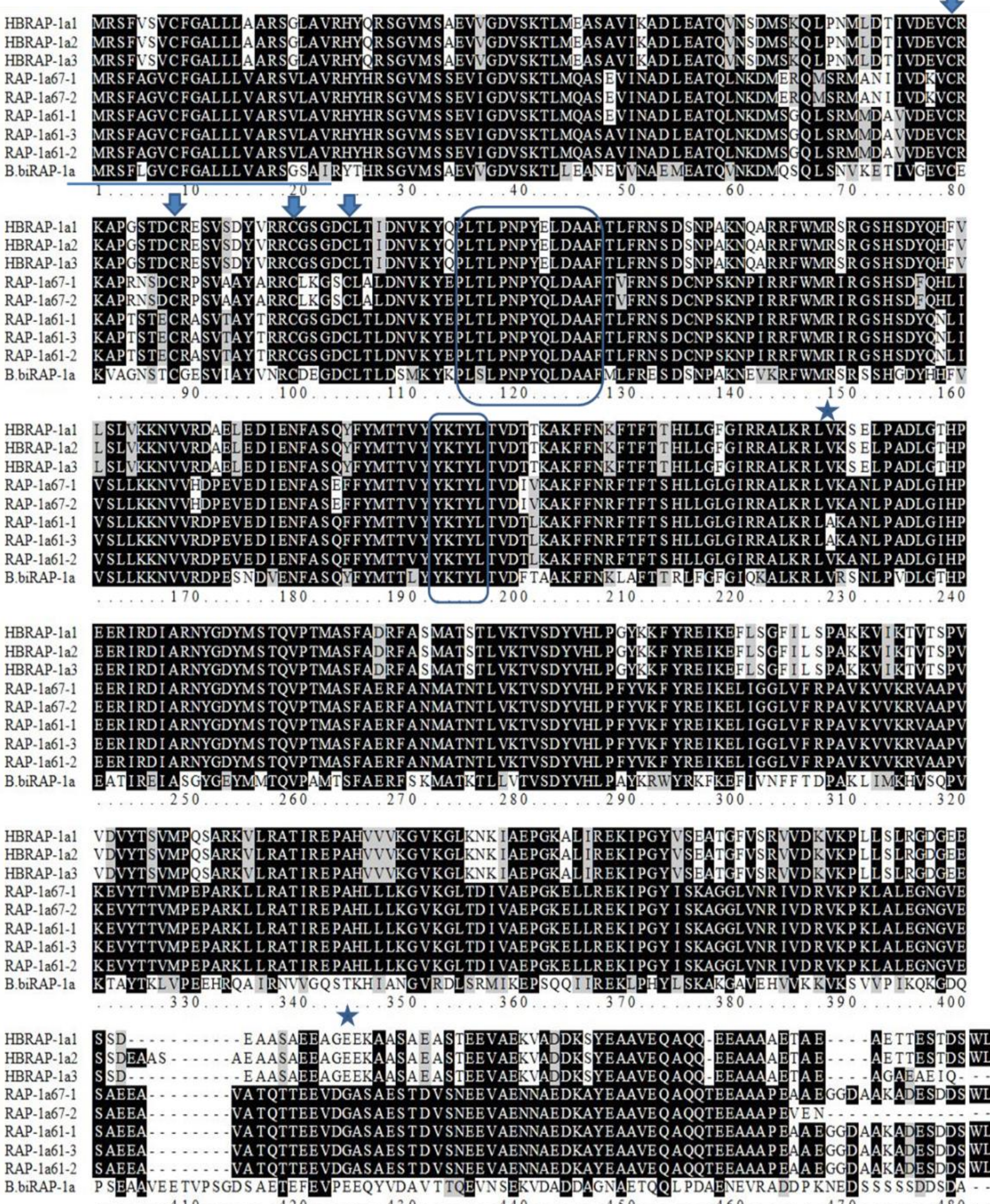


Fig.1. Alignment of predicted amino acid sequences of rap-1a copies. HBRAP-1a1: *Babesia* sp. Hebei RAP-1a1 (GenBank accession number: KJ205334), HBRAP-1a2: *Babesia* sp. Hebei RAP-1a2 (GenBank accession number: KJ205335), HBRAP-1a3: *Babesia* sp. Hebei RAP-1a3 (GenBank accession number: KJ205337). RAP-1a67-1: RAP-1a67-1 from *Babesia* sp. BQ1 (Lintan and Ningxian), *Babesia* sp. Tianzhu, (respective GenBank accession numbers: AGV15812, KJ205323, KJ205330), RAP-1a67-2: RAP-1a67-2 from *Babesia* sp. BQ1 Lintan, Ningxian and *Babesia* sp. Tianzhu, (respective GenBank accession numbers: AGV15811, KJ205326, KJ205332). RAP-1a61-1: RAP-1a61-1 from *Babesia* sp. BQ1 (Lintan) (GenBank accession numbers: AGV15809). RAP-1a61-2: RAP-1a61-2 from *Babesia* sp. BQ1 (Lintan and Ningxian), *Babesia* sp. Tianzhu (respective GenBank accession numbers: AGV15808, KJ205322, KJ205328). RAP-1a61-3: RAP-1a61-3 from *Babesia* sp. BQ1 Ningxian and *Babesia* sp. Tianzhu (respective GenBank accession numbers: KJ205324, KJ205329). B.biRAP-1a: *B. bigemina* RAP-1aα1 gene (GenBank accession number: AAA65583). The 4 conserved cysteine residues are marked with arrows, the two conserved amino acid sequences indicated with boxes, and predicted signal peptide sequences are underlined, the positions of three point mutations in different RAP-1a61 types are showed with asterisks.

3. 2. Multiple and conserved *rap-1b* copies within each *B. motasi*-like isolate

3. 2. 1. *rap-1b* copies in *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu isolates are identical to the *Babesia* sp. BQ1 (Lintan) isolate

The full *rap1- b* gene was amplified using the primer combination AP13-AP14, then cloned and sequenced. All sequenced inserts were identical to the previously described *rap-1b* gene of *Babesia* sp. BQ1 (Lintan): 1203 bp in length, with putative start and stop sites encoding a putative 400 amino acids protein (Fig. 2). The presence of multiple *rap-1b* gene copies interspaced with *rap-1a* gene copies, within the *rap-1* locus, was attested when the intergenic regions from *rap-1a* and *rap-1b* genes were amplified (see Supplementary Fig. 1).

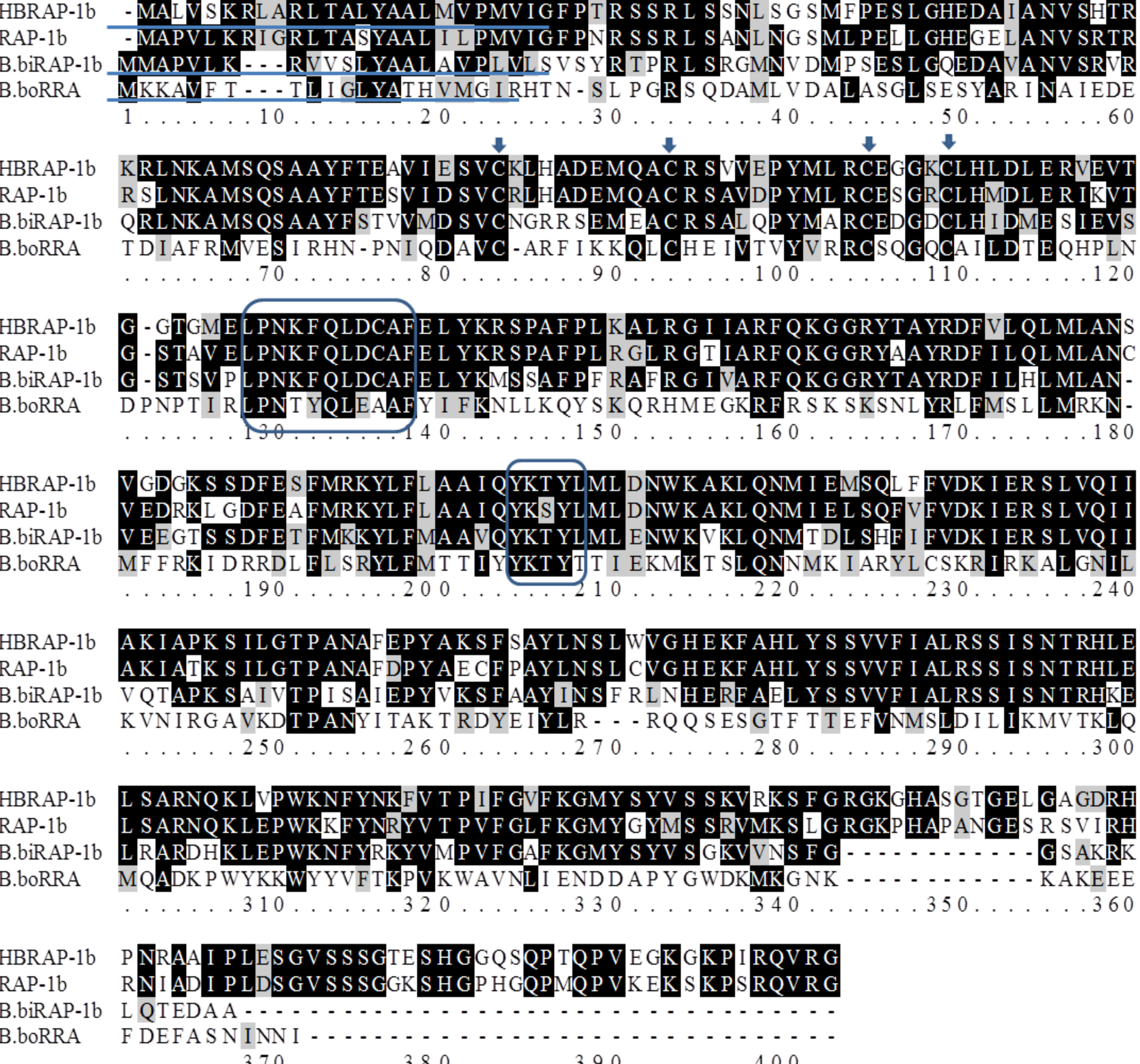


Fig.2. Alignment of RAP-1b predicted amino acid sequences. HBRAP-1b: *Babesia* sp. Hebei RAP-1b (GenBank accession number: KJ205336); RAP-1b: RAP-1b from *Babesia* sp. BQ1 (Lintan and Ningxian), *Babesia* sp. Tianzhu (respective GenBank accession numbers: AGV15813, KJ205325, KJ205331). B.biRAP-1b: *B. bigemina* RAP-1b gene (GenBank accession number: AAB72095). BboRRA: *B. bovis* RRA sequence (GenBank accession number: XM001610900). The 4 conserved cysteine residues are marked with arrows, the two conserved amino acid sequences indicated with boxes, and predicted signal peptide sequences are underlined.

3. 2. 2. Multiple *rap-1b* copies in the *Babesia* sp. Hebei isolate are conserved but differ from those of the three other *B. motasi*-like isolates

Partial *rap-1b* sequences were obtained from amplifications with primers AP21-AP22 and AP23-AP24. A primer combination specific to *Babesia* sp. Hebei *rap-1b* (AP27-AP28) was designed to amplify the full-length *rap-1b* gene, and 12 different clones were sequenced to investigate the genetic diversity of the *rap-1b* copies. The twelve insert sequences were identical and blasted with *Babesia* sp. BQ1 (Lintan) and *B. bigemina rap-1b.* The putative gene length was identical to that of *rap-1b* genes from other *B. motasi*-like isolates (1203 bp encoding a putative 400 amino acid protein), despite sequence differences (86% and 79% identity at the nucleotide and protein levels, respectively) (Fig. 2, Table 4).

3. 3. *rap-1c*: only one copy

3. 3. 1 *rap-1c* in the *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu isolates

Rap-1c was successfully amplified using the primer combination AP15-AP20. Multiple cloned *rap-1c* inserts were sequenced for each isolate which led us to conclude that the *rap-1c* gene was also extremely well-conserved and was of identical length to the *Babesia* sp. BQ1 (Lintan) *rap-1c* gene (1749 bp, encoding a putative 582 amino acid protein). The sequences were very similar, differing only at the 3' end by 11 nucleotide substitutions resulting in 5 amino acid modifications (Fig. 3; Supplementary Fig. 3). Most of these modifications were located in the 17-18 aa repeats 4 and 5 that were described in *Babesia* sp. BQ1 (Lintan) *rap-1c* (Fig. 4). Three other substitutions were located at nucleotide positions 1564, 1574 and 1614, the first two being non synonymous (Fig. 3; Supplementary Fig. 3).

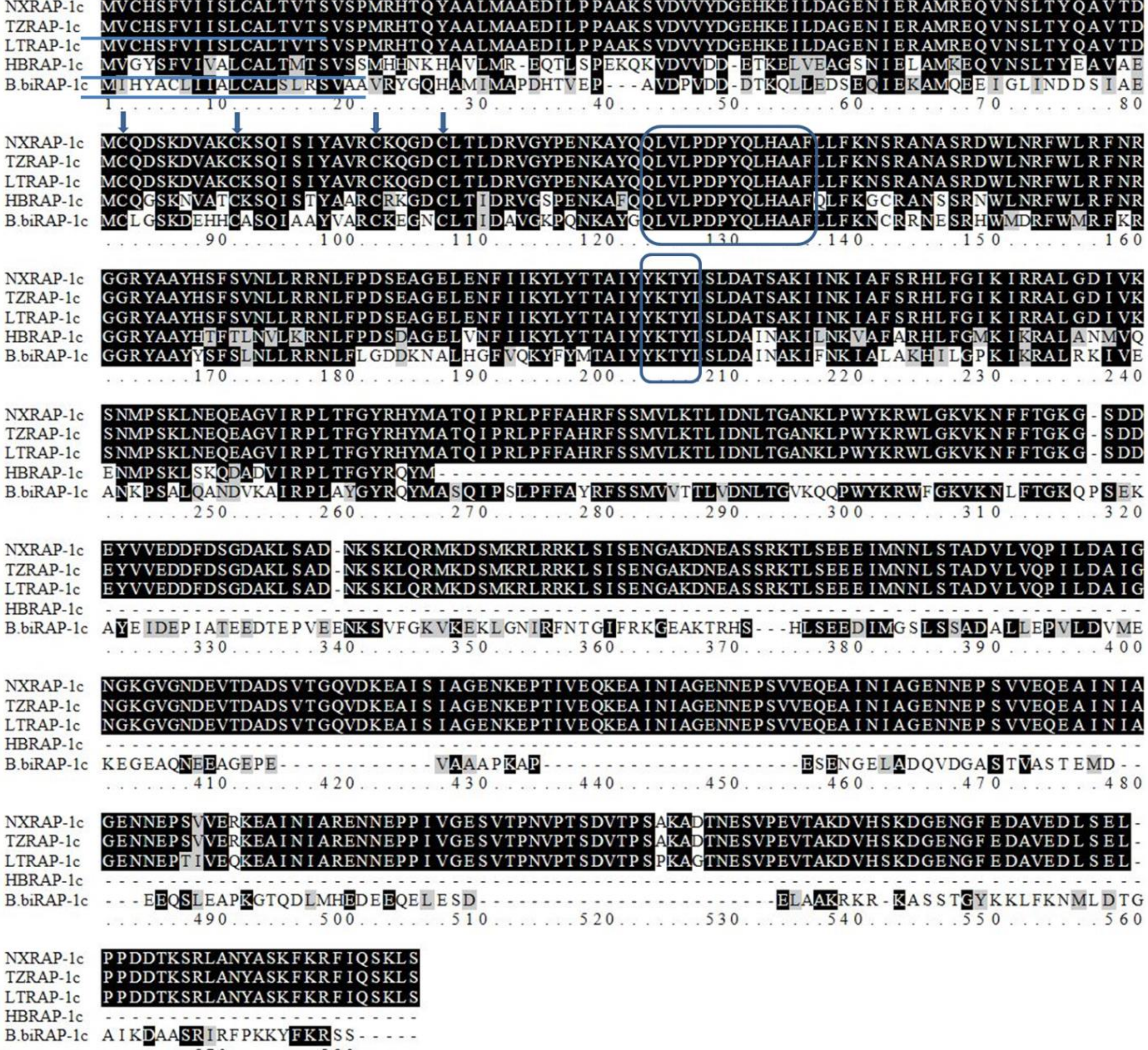
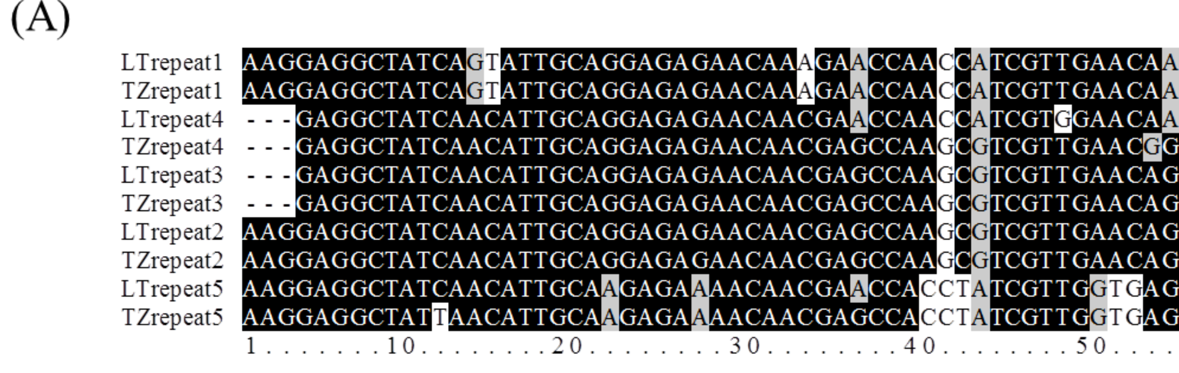


Fig.3. Alignment of RAP-1c predicted amino acid sequences.NXRAP1c: RAP-1c from *Babesia* BQ1 sp. Ningxian (GenBank accession number: KJ205327); TZRAP-1c: RAP-1c from *Babesia* sp. Tianzhu (GenBank accession number: KJ205333); LTRAP-1c: RAP-1c from *Babesia* sp.BQ1 (Lintan) (GenBank accession number: AGV15813); HBRAP-1c: partial RAP-1c (264aa) from *Babesia* sp. Hebei (GenBank accession number: KJ205338); B.biRAP-1c : *B. bigemina* RAP-1c gene (GenBank accession number: AAN84521).The 4 conserved cysteine residues are marked with arrows, the two conserved amino acid sequences indicated with boxes, and predicted signal peptide sequences are underlined.



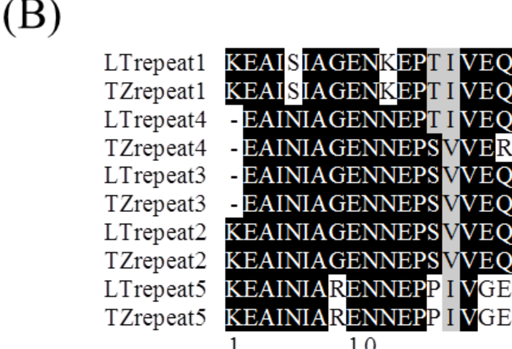


Fig.4. Alignment of nucleotide (A) and amino acid (B) repeats in the 3’ region of *rap-1c* sequence.

3. 3. 2 *Babesia* sp. Hebei *rap-1c*

The sequence of an internal region in *rap-1c* (422 bp) was obtained by amplification with the degenerate primers AP29-AP30. A specific *rap-1c* reverse primer AP31 was designed and the 5’ extremity of the gene was amplified using the primer combination AP25-AP31 (*rap-1a* to *rap-1c*). A PCR amplicon with the expected size of about 3.2 Kb (*rap-1a*: 1.4 Kb, IG3: 1.3 Kb and 5’ end of *rap-1c*: 0.5 Kb) was successfully obtained. Inserts from four clones were sequenced and gave identical results (Supplementary Fig. 1b, step 5). The strategy described for the *Babesia* sp. BQ1 (Lintan) *rap-1c* gene was applied to amplify the 3’ end of *rap-1c*, using the reverse primers p128R located downstream in the YJR070-like orf (AF026272) (Suarez et al. 1998b; Niu et al. 2013). The many attempts with the forward primers AP32 or AP33 were unsuccessful (Supplementary Fig. 1b, step 6). A partial sequence of 794 bp in the 5’ end region of *rap-1c* was obtained from AP29-AP30 and AP25-AP31 amplifications (Fig. 3).

3.4. Presence of signal peptides

A search for the presence of putative signal peptide sequences was carried out using free bioinformatics software (SignalP4-1 and PrediSi). Both software detected putative signal peptides in *rap-1a* (cleavage at aa21), and *rap-1c* (cleavage at aa18 in the *Babesia* sp. BQ1 (Ningxian) and Tianzhu isolates, and at aa21 in the *Babesia* sp. Hebei isolate). Neither software could identify a putative signal peptide for *rap-1b* in the *Babesia* sp. BQ1 (Ningxian) and Tianzhu isolates (sequence identical to that of the Lintan isolate), but a putative signal peptide was found (cleavage at aa25) in the case of *Babesia* sp. Hebei *rap-1b* with the PrediSi software but not with SignalP4-1.

3. 5. Three different intergenic regions IG1, IG2 and IG3 exist in the three isolates

The amplification of *Babesia* sp. BQ1 (Ningxian) and Tianzhu isolates with different primer combinations (AP16 /18-AP10, AP9-AP3/5, AP11-AP20, AP9-AP7), followed by cloning and insert sequencing, demonstrated the presence of three different intergenic regions of 853bp (IG1, *rap-1a* to *rap-1b)*, 1840bp (IG2, *rap-1b* to *rap-1a61* or *67-1*) and 1333bp (IG3 *rap-1a67-2* to *rap-1c)* (Supplementary Fig. 1a and Table 2). In the case of IG1 (853bp), a point mutation located 62 bp after the *rap-1a61* or *67* putative stop site distinguished IG1.1 (with a C) from IG1.2 (with a T). IG1.1 was found in *Babesia* sp. BQ1 (Ningxian) downstream from *rap-1a61-2* in 3 clones and from *rap-1a67-1* in 5 clones, and in *Babesia* sp. Tianzhu downstream from *rap-1a61-2* in two clones, from *rap-1a61-3* in one clone and from *rap-1a67-1* in 9 clones. IG1.2 was only found in *Babesia* sp. BQ1 (Ningxian), downstream from *rap-1a67-1* in 5 clones. The sequences of these intergenic regions (IG1, 2 and 3) were identical in the *rap-1* locus of three Chinese isolates from the *B. motasi*-like group, i.e. *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu and *Babesia* sp. BQ1 (Lintan) (Niu et al. 2013).

Three different intergenic regions were also discovered in the *rap-1* locus of the *Babesia* sp. Hebei isolate (primer combinations AP23-AP24, AP25-AP31): IG1 (870 bp) between *rap-1a1* and *rap-1b*, IG2 (1837 bp) between *rap-1b* and *rap-1a1*, and IG3 (1320 bp) between *rap-1a3* and *rap-1c* (Supplementary Fig. 1b and Table 2). Only a few inserts were sequenced so the sequence variability in these intergenic regions could not be investigated in depth. However, the sequences obtained for each intergenic region were identical (IG1-2 clones, IG2-2 clones and IG3-4 clones).

The A + T residue content of IG1 (55%) is almost equal to the A+T content (54%) of the downstream *rap-1b* sequence. In contrast, the A + T contents of the intergenic regions IG2 and IG3 (54% and 64%, respectively) are lower than the contents of the putative downstream coding regions (*rap-1a* with 45% and *rap-1c* with 53%), and the calculated A+T content is similar to the A+T content of the *rap-1* genes and intergenic regions in the (Lintan) strain (Niu et al. 2013).

As might be expected, the IG regions are not conserved between the *B. motasi*-like group and *B. bigemina* (20-27%), but their lengths remain similar. Within the *B. motasi*-like phylogenetic group, their sequences are moderately conserved (61-79%). For each IG, the 250-270 nt regions upstream from the *rap-1a* and *rap-1b* coding sequences are more conserved between *B. bigemina* and *B. motasi*-like isolates (63.5% for IG1 and 67.5% for IG2) than the rest of the IG (37.8% for IG1 and 36.7% for IG2). This probably corresponds to the 5' UTR region and the three previously described boxes (known as the -59 box, the -36 box and the mRNA box), supposedly involved in transcription regulation (Fig. 5) (Suarez et al. 1998b). The -59 box was well conserved in all three intergenic regions, while conservation of the -36 box was greater in IG2 and IG3 than in IG1 and conservation of the mRNA box was only apparent in IG3, and its position was variable (Fig. 5; Supplementary Fig. 4, 5 and 6).

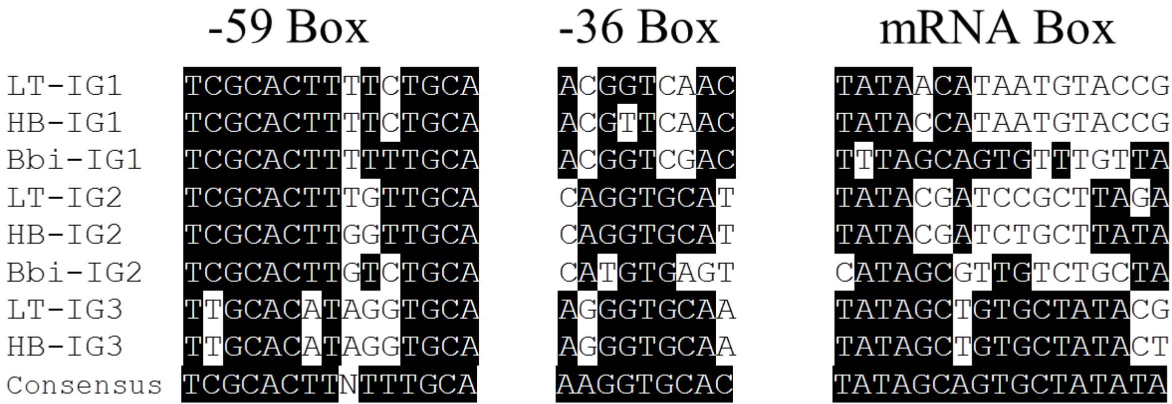


Fig. 5 Conservation of the putative rap-1 transcription regulation 5'UTR sequence in the *rap-1* intergenic regions. The position of boxes -59, -36 and mRNA are designated according to Suarez et al., 1998 and consensus sequences from Suarez et al., 1998 is indicated. LT: Lintan; HB: Hebei; Bbi: *B. bigemina*.

3. 6. *Rap-1* locus sequence comparison and phylogenetic analysis

The presence of three types of *rap-1* genes, *rap-1a*, *rap-1b* and *rap-1c*, has now been demonstrated in several *Babesia* isolates or species (Supplementary Fig. 7, 8 and 9). The gene lengths and sequences are relatively well-conserved within each of these types (identities between 55-100% for *rap-1a*, 64-100% for *rap-1b* and 53-100% for *rap-1c*), indicating that these genes are orthologs (Table 4). However, the three gene types do not share significant sequence homologies (average identities between *rap1a* and *rap-1b* 19.6%, between *rap-1b* and *rap-1c* 22.6% and between *rap-1a* and *rap-1c* 27.6% within each studied parasite), even though the typical features of this gene family (4 cysteines, conserved motifs) are conserved and they segregate into three distinct clades in the phylogenetic analysis (Fig. 6).

RAP-1b is the most conserved, exhibiting scattered sequence polymorphism over the whole molecule and long stretches of conserved sequences (up to 19 consecutive aa, Fig. 2). In RAP-1c, the 3' second half of the protein shows considerable sequence polymorphism both from residue 317 onwards and in the first 80 residues. The central part of the molecule (81-316) seems to be more conserved (Fig. 3). The localization of polymorphisms in the RAP-1a sequences is variable. The polymorphism between RAP-1a61 and RAP-1a67, within the same *rap-1* locus, is mainly limited to a 3' region between residues 64 and 107, a polymorphic region also found in *B. bigemina rap-1a* genes. Although this region is also very variable in *Babesia* sp. Hebei RAP-1a (aa 64-98), a lot of polymorphism is found in the 3' end of the protein (from residue 300), as in *B. bigemina* (Fig. 1).

All the RAP-1a sequences from *B. bigemina* and the Chinese *B. motasi*-like isolates fell into the same clade with three main sister clades, the first gathering *B. bigemina* RAP-1a, the second *Babesia* sp. Hebei RAP-1a, and the third *Babesia* sp. BQ1 (Ningxian and Lintan) RAP-1a, with RAP-1a61 separate from RAP-1a67 sequences (Fig. 6).

Table 4. Percentages of identity between the three genes and intergenic regions present in different isolates of *B. motasi*-like and *B. bigemina*.

***Babesia* sp. BQ1(Lintan)**								
	rap-1a61 (1425bp)	*rap-1a67*		*rap-1b* (1203bp)	*rap-1c* (1479bp)	IG1 (853bp)	IG2 (1840bp)	IG3 (1333bp)
		67-1 (1425bp)	67-2 (1380bp)					
Ningxian & Tianzhu	99.9[a]/99.8[b] [(c)]	100/100	100/100	100/100	99.4/99.1	100	100	100
Hebei	*rap-1a1* (1401bp) 82.1/74.9	*rap-1a2* (1416bp) 79.6/71.1	*rap-1a3* (1392bp) 78.9/71.7	(1203bp) 86.2/78.5	Partial (794bp) 83.8/72.0	(870bp) 79.2	(1837bp) 75.2	(1320bp) 61.1
B. bigemina	*aα1* (1443bp) 73.0[d]/57.0[d]	*aα1* (1443bp) 72.0[d]/57.0[d]	*aα1* (1443bp) 72.0[d]/57.0[d]	*b* (1053bp) 70.3[d]/63.7[d]	*c* (1533bp) 62.8[d]/52.9[d]	(803bp) 27.6 [d]	(1528bp) 22.3 [d]	-
***Babesia* sp. Hebei**								
	rap-1a1	*rap-1a2*	*rap-1a3*	*rap-1b*	Partial *rap-1c*	IG1	IG2	IG3
B. bigemina	68.3/55.6	61.4/54.6	67.9/55.7	70.8/64.3	65.7/56.8	20.4	22.1	-

a: Nucleotide level

b: Amino acid level

c: Percentage(%)

d: Data was cited from Niu et al., 2013.

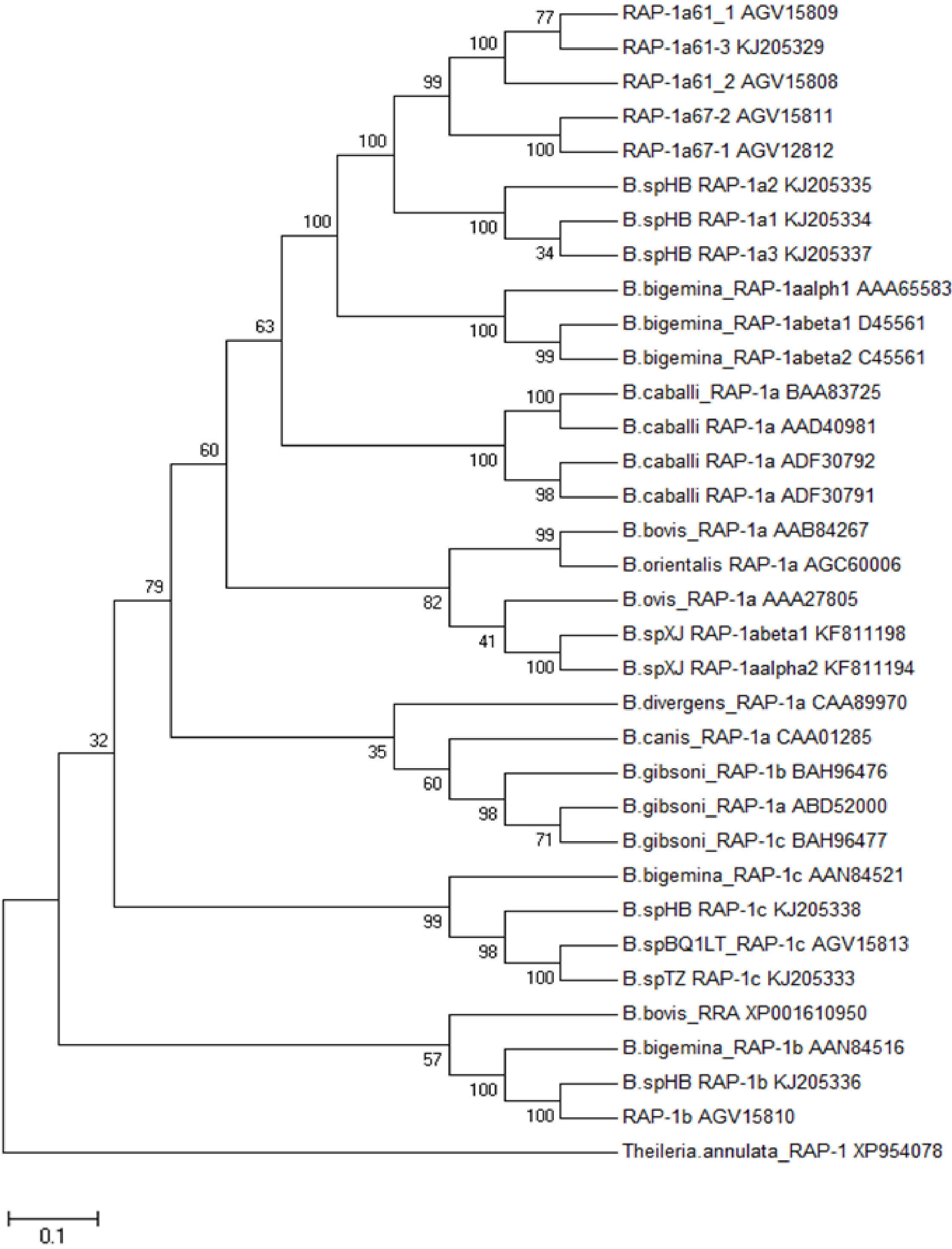


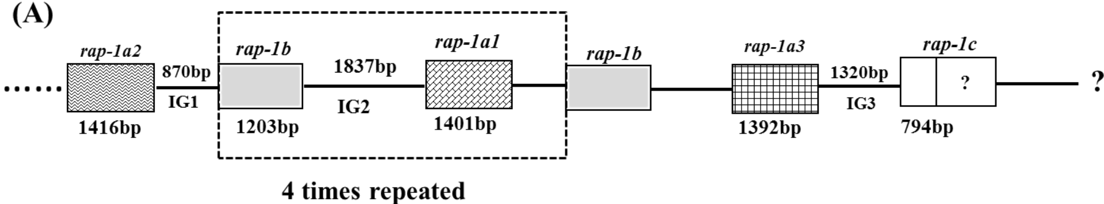
Fig. 6 Phylogenetic tree constructed with the amino acid sequences of RAP-1a61-1 (Lintan), RAP-1a61-2 (Lintan, Ningxian and Tianzhu), RAP-1a61-3 (Ningxian and Tianzhu), RAP-1a67-1, RAP-1a67-2 (Lintan, Ningxian and Tianzhu), RAP-1b (Lintan, Ningxian and Tianzhu), RAP-1a, b, c of *Babesia* sp. Hebei (HB) and of all known members of the RAP-1 family in *Babesia*. The tree was inferred using the neighbor joining method, bootstrap values of MEGA 5.2 are shown at each branch point. The *Theileria annulata* amino acid sequence was used as outlier.

3.7. Putative gene arrangement in the *Babesia* sp. BQ1 (Ningxian), Tianzhu and Hebei *rap-1* locus

It could be concluded from the amplification and sequencing data that all 4 isolates from the *B. motasi*-like group contained orthologs of the *B. bigemina rap-1a*, *rap-1b* and *rap-1c* genes.

The *rap-1* loci and genes are extremely similar in three of these isolates (*Babesia* sp. BQ1 (Lintan) and Xinjiang as well as *Babesia* sp. Tianzhu). The presence of interspaced *rap-1a* and *rap-1b* genes was also demonstrated in this study (amplifications AP16/18-AP10 and AP9-AP3/5) (Supplementary Fig. 1a), with the *rap-1a61-2* and *rap-1a 67-1* genes upstream of *rap-1b* in both isolates, and the *rap-1a61-3* gene upstream of *rap-1b* only in *Babesia* sp. Tianzhu (probably not enough clones sequenced). The localization of *rap-1a67-2* upstream of *rap-1c* was also demonstrated (primer combinations AP11-AP20 and AP9-AP7). The *rap-1c* position at the end of the locus, followed by the YJR070-like orf, was found in the *rap-1* locus of *Babesia* sp. BQ1 (Lintan), *B. bovis* and *B. bigemina* (Niu et al. 2013; Suarez et al. 1998a) and was also confirmed in the *rap-1* locus of *Babesia* sp. BQ1 (Ningxian) and Tianzhu isolates (primer combination AP8-p128R, data not shown). The *rap-1* locus organization described for *Babesia* sp. BQ1 (Lintan) (Niu et al., 2013) is therefore most probably conserved in these two other *B. motasi*-like isolates. However, the precise *rap-1a* and *rap-1b* gene numbers and the gene arrangement within the locus could not be precisely determined.

In the case of the *Babesia* sp. Hebei isolate, even though the sequences of the three gene types (*rap-1a*, *rap-1b* and *rap-1c*) differed from those of the three other *B. motasi*-like isolates, their arrangement within the locus was still conserved. We were able to demonstrate the presence of the *rap-1a1* gene between two *rap-1b* genes, localization of the *rap-1a3* gene only upstream of *rap-1c*, and the higher copy number of *rap-1a1* (14/17 clones) compared to *rap-1a2* (3/17 clones). The position of *rap-1c* at the end of the locus could not be demonstrated since numerous attempts to amplify the 3' end region of *rap-1c,* using reverse p128R with several *rap-1c* forward primers, were unsuccessful. A hypothetical locus organization that fits our data is proposed (Fig. 7).



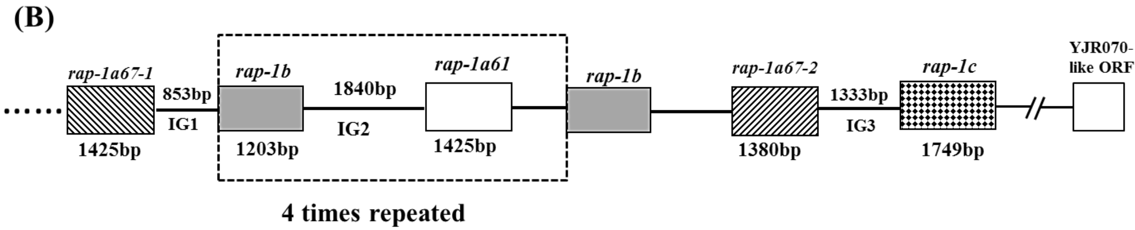


Fig. 7 Comparison of hypothetical locus organization of *Babesia* sp. Hebei (A) and *Babesia* BQ1 sp. (Lintan) (B) (Niu et al., 2013). The 4 times repeated partial locus is indicated with a dotted line box.

3. 8. Nucleotide accession numbers

All sequences have been deposited in GenBank with the following accession numbers: KJ205322 (*rap-1a61-2* and IG1.1), KJ205323 (*rap-1a67-1* and IG1.2), KJ205324 (*rap-1a61-3*), KJ205326 (*rap-1a67-2* and IG3), KJ205325 (*rap-1b* and IG2), and KJ205327 (*rap-1c*) for the *Babesia* sp. BQ1 (Ningxian) isolate; KJ205328 (*rap-1a61-2* and IG1.1), KJ205329 (*rap-1a61-3*), KJ205330 (*rap-1a67-1*), KJ205332 (*rap-1a67-2* and IG3), KJ205331 (*rap-1b* and IG2), and KJ205333 (*rap-1c*) for the *Babesia* sp. Tianzhu isolate; KJ205334 (*rap-1a1* and IG1), KJ205335 (*rap-1a2*), KJ205336 (*rap-1b* and IG2), KJ205337 (*rap-1a3* and IG3), and KJ205338 (partial *rap-1c*) for the *Babesia* sp. Hebei isolate.

4. Discussion

Several virulent isolates of *Babesia* have recently been described in sheep and/or goats in China, and phylogenetic analysis of the 18S rDNA gene has shown that some of them belong to the *B. motasi* phylogenetic group (*B. motasi*-like) (Guan et al. 2009). Vaccine development, based on the parasite’s asexual growth cycle at the blood stage, might be a good control strategy for babesiosis (Gohil et al. 2013), and rhoptry associated protein-1 (RAP-1) is considered as a potential vaccine candidate due to its role in red blood

cell invasion. We therefore investigated the sequence polymorphism of *rap-1* in three isolates from the *B. motasi*-like group: *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu and *Babesia* sp. Hebei. We had shown in a previous study that the *rap-1* locus in *Babesia* sp. BQ1 (Lintan), another Chinese isolate from the *B. motasi*-like group, was complex and similar to that of *B. bigemina.* Within a locus extending over about 31 Kb, 6 *rap-1a* genes were found interspaced with 5 *rap-1b* genes, with a 3' terminal gene of *rap-1c* type (Niu et al. 2013). Interestingly, the *rap-1* locus in the three isolates studied here share the same features. These features (especially the presence of *rap-1b* and *rap-1c* gene types) seem to be restricted to a phylogenetic clade, based on 18S rDNA, which groups *B. bigemina* and *B. motasi* (Lack et al. 2012), as well as *B. crassa*, *B. ovata* and *B. major* (Schnittger et al. 2012). Only the *rap-1a* gene type has been found in the other *Babesia* species analyzed. However, the *rap-1b* and *rap-1c* gene types share few sequence similarities with the *rap-1a* gene type so they may have been missed. Nevertheless, none of these gene types has been found in the *B. bovis* genome sequence, and only one other distantly related gene, called RRA (RAP-1 related antigen), has been described recently and shares the *rap-1* family features (Suarez et al. 2011). The new *rap-1b* and *rap-1c* sequences now available should allow primer design in the conserved sequences and facilitate the search for orthologous genes in different *Babesia* species.

The presence of duplicated and conserved *rap-1a* and *rap-1b* genes, as well as their intergenic regions, in each of the *B. motasi*-like isolates, could indicate strong selective pressure on these genes. Mechanisms such as gene conversion probably homogenize the different copies of *rap-1a* and *rap-1b* genes, as well as the intergenic regions. Such events can result in an apparent slowing down of the mutation rate (Innan 2009).

Even if the transcription of *rap-1b* and *rap-1c* genes has been demonstrated in the case of *B. bigemina*, neither their translation, nor their localization in the rhoptries has yet been proven (Suarez et al. 2003). The presence of moderately conserved 5' UTR regions upstream of the *rap-1a*, *rap-1b* and *rap-1c* genes suggests that all these genes are also transcribed in isolates from the *B. motasi*-like group, as they are in *B. bigemina*. In the case of *rap-1b*, a putative signal peptide has been found only in *Babesia* sp. Hebei and *B. bigemina*, so both the translation and export of RAP-1b to the rhoptry still needs to be confirmed.

All the *B. motasi*-like isolates studied here and in a previous work (Niu et al. 2013) are phylogenetically extremely close, their 18S rDNA genes sharing between 98 and 99% similarity. The identities of the isolates which had provided the genomic DNA sent on an

FTA card from China were confirmed by re-sequencing the 18S rDNA gene. The sequencing results were consistent with previously published data (Liu et al. 2007). Phylogenetic studies based on this gene placed *Babesia* sp. Hebei on a separate branch from *Babesia* sp. Tianzhu, *Babesia* sp. BQ1 (Lintan and Ningxian) (Lack et al. 2012; Gou et al. 2013). This phylogenetic separation was confirmed by sequencing the *rap-1* locus, which is almost identical (about 99.9%) over about 31 Kb in *Babesia* sp. Tianzhu, *Babesia* sp. BQ1 (Ningxian and Lintan) (Table 4). In contrast, the same locus in *Babesia* sp. Hebei clearly differs from the three other isolates, displaying on average 78% similarity at the nucleotide level (including IG regions) over the whole locus. However, recent phylogenetic analyses of these isolates with other markers have produced different phylogenetic trees, placing *Babesia* sp. BQ1 (Lintan) close to *Babesia* sp. Tianzhu and separating *Babesia* sp.BQ1 (Ningxian) and *Babesia* sp. Hebei from this group (ITS in Niu et al. 2009; 28S rDNA in Gou et al. 2013). Other studies of Chinese *Babesia* and *Theileria* isolates phylogeny, based on RPS8 (40S ribosomal protein 8, 561 bp analyzed, Tian et al. 2013a), cytochrome b gene (a mitochondrial marker, 550 bp analyzed, Tian et al. 2013b) and the cytochrome C oxidase subunit III (a mitochondrial marker, 552 bp analyzed, Tian et al. 2013c), did not include *Babesia* sp. Hebei in the analyses. These studies showed *Babesia* sp. BQ1 (Lintan) being grouped with *Babesia* sp. Tianzu, while *Babesia* sp. (Ningxian) was more or less closely related depending on the marker used. Thus the phylogenetic relationships existing between these Chinese *B. motasi*-like isolates still need to be clarified.

The *rap-1* locus as a whole (*rap-1a*, *rap-1b*, *rap-1c* and the corresponding intergenic regions) is an informative marker but its interest is probably limited to studies of the phylogenetic group which includes *B. motasi*, *B. bigemina*, *B. major*, *B. ovata* and *B. crassa* (Schnittger et al. 2012). The inclusion of coding regions (*rap-1b*) as well as the more rapidly evolving intergenic regions IG1 and IG2 would reveal ancient and recent evolutionary events. For a broader phylogenetic analysis within the *Babesia* genus, the *rap-1a* gene is also an informative marker, being present in all species analyzed so far. The fact that the *rap-1* genes, as well as the intergenic regions, are more conserved than the 18S rDNA gene in extremely closely related isolates (100% for *rap-1a, rap-1b* and IG regions in *Babesia* sp. Tianzhu, *Babesia* sp. BQ1 (Lintan and Ningxian) compared to 98.7% for 18S rDNA), but are less conserved in other still very closely related isolates (78% for the whole *rap* genome compared to 97.7% for the 18S rDNA) is intriguing. It can probably be related to the evolution of this locus, composed of duplicated genes, where the apparent

mutation rate has probably been slowed down by genetic conversion and concerted evolution (Innan 2009). In the course of evolution and for unknown reasons (modification of the selective pressure, sub-functionalization with a selective advantage resulting in fixation of the new genes within the population…), these processes become less efficient and mutations accumulate until the process of homogenization prevails again. In any case, the result is a clear grouping and separation of isolates, which could be of interest in phylogenetic studies.

The presence of multiple conserved copies of *rap-1b* could be useful for the molecular diagnosis of ovine babesiosis. The presence of 5 copies of this gene would indeed improve its detection and the absence of polymorphism within the locus would ensure that all copies were amplified. Due to the presence of conserved and polymorphic regions in this gene among *B. motasi*-like isolates, a general ovine babesiosis test (excluding *Babesia* sp. Xinjiang) could be set up as well as more specific tests, either by designing specific PCR primers or a PCR-RFLP to discriminate isolates after a *rap-1b* non specific amplification.

The rather high conservation of RAP-1a within a locus as well as among different isolates suggests that the development of a common vaccine, based on the conserved protein region among geographically distinct isolates, is achievable. However, even if the translation of *rap-1a* has been demonstrated in the case of *B. bigemina* and can be presumed in the case of isolates from the *B. motasi*-like group, this still remains to be proved, together with the role of these proteins in red blood cell invasion and the potential of the conserved gene region to elicit protective immunity. We are currently devoting our studies to the functional and immunological relevance of the *B. motasi*-like *rap-1* genes.

Acknowledgements

This study was supported financially by a grant from the European Union cooperation project PIROVAC (KBBE-3-245145) and the « Région Pays de Loire » (France).

References

Bai Q, Liu G, Liu D, Ren J, Li X (2002) Isolation and preliminary characterization of a large *Babesia* sp. from sheep and goats in the eastern part of Gansu Province, China. Parasitol Res 88:S16–S21.

Brown WC, McElwain TF, Ruef BJ, Suarez CE, Shkap V, Chitko-McKown CG, Tuo W, Rice-Ficht AC, Palmer GH (1996) *Babesia bovis* rhoptry-associated protein 1 is immunodominant for T helper cells of immune cattle and contains T-cell epitopes conserved among geographically distant *B. bovis* strains. Infect Immun 64:3341-3350.

Brown WC, Norimine J, Goff WL, Suarez CE, McElwain TF (2006) Prospects for recombinant vaccines against *Babesia bovis* and related parasites. Parasit Immunol 28:315–327.

Dalrymple BP, Casu RE, Peters JM, Dimmock CM, Gale KR, et al. (1993) Characterisation of a family of multi-copy genes encoding rhoptry protein homologues in *Babesia bovis*, *Babesia ovis* and *Babesia canis*. Mol Biochem Parasitol 57:181-192.

Fish L, Leibovich B, Krigel Y, McElwain T, Shkap V (2008) Vaccination of cattle against *B. bovis* infection with live attenuated parasites and non-viable immunogens. Vaccine 26:29-33.

Gohil S, Herrmann S, Günther S, Cooke BM (2013) Bovine babesiosis in the 21st century: Advances in biology and functional genomics. Int J Parasitol 43:125-132.

Gou H, Guan G, Ma M, Liu A, Liu Z et al (2013) Phylogenetic analysis based on 28S rRNA of *Babesia* spp. in ruminants in China. Exp Appl Acarol 59:463-472.

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.

Innan H (2009) Population genetic models of duplicated genes. Genetica 137:19-37.

Jouglin M, Fernandez-de-Mera IG., De la Cotte N, Ruiz-Fons F, Gortazar C, et al (in press) Isolation and characterization of *Babesia pecorum* sp. nov. from farmed red deer (*Cervus elaphus*). Vet Res

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.

Liu AH, Yin H, Guan GQ, Schnittger L, Liu ZJ, et al (2007) At least two genetically distinct large *Babesia* species infective to sheep and goats in China. Vet Parasitol 147:246-251.

Lobo CA, Rodriguez M, Cursino-Santos JR (2012) *Babesia* and red cell invasion. Curr Opin Hematol 19:170-175.

McElwain TF, Perryman LE, Musoke AJ, McGuire TC (1991) Molecular characterization and immunogenicity of neutralization-sensitive *Babesia bigemina* merozoite surface proteins. Mol Biochem Parasitol 47:213-222.

Mosqueda J, McElwain TF, Stiller D, Palmer GH (2002) *Babesia bovis* Merozoite Surface Antigen 1 and Rhoptry-Associated Protein 1 are expressed in sporozoites, and specific antibodies inhibit sporozoite attachment to erythrocytes. Infect Immun 70:1599-1603.

Niu Q, Luo J, Guan G, Liu Z, Ma M, Liu A, Gao J, Ren Q, Li Y, Qiu J, Yin H (2009) Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. Exp Parasitol 121:64-68.

Niu Q, Bonsergent C, Guan G, Yin H, Malandrin L (2013) Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group). Vet Parasitol 198:24-38.

Norimine J, Mosqueda J, Suarez C, Palmer GH, McElwain TF, Mbassa G et al (2003) Stimulation of T-helper cell gamma interferon and immunoglobulin G responses specific for *Babesia bovis* rhoptry-associated protein 1 (RAP-1) or a RAP-1 protein

lacking the carboxy-terminal repeat region is insufficient to provide protective immunity against virulent *B. bovis* challenge. Infect Immun 71:5021-5032.

Oosthuizen MC, Allsopp BA, Troskie M, Collins NE, Penzhorn BL (2009) Identification of novel *Babesia* and *Theileria* species in South African giraffe (*Giraffa camelopardalis*, Linnaeus, 1758) and roan antelope (*Hippotragus equinus*, Desmarest 1804). Vet Parasitol 163:39-46.

Palmer GH, McElwain TF, Perryman LE, Davis WC, Reduker DR, Jasmer DP, Shkap V, Pipano E, Goff WL, McGuire TC (1991) Strain variation of *Babesia bovis* merozoite surface-exposed epitopes. Infect Immun 59:3340-3342.

Rodriguez SD, Palmer GH, McElwain TF, McGuire TC, Ruef BJ, Chitko-McKown CG, Brown WC (1996) CD4+ T-helper lymphocyte responses against *Babesia bigemina* rhoptry-associated protein 1. Infect Immun 64:2079-2087.

Sam-Yellowe TY (1996) Rhoptry organelles of the apicomplexa: their role in host cell invasion and intracellular survival. Parasitol Today 12:308-316.

Schnittger L, Rodriguez AE, Florin-Christensen M, Morrison DA (2012) *Babesia*: a world emerging. Infect Genet Evol 12:1788-809.

Shompole S, McElwain TF, Jasmer DP, Hines SA, Katende J, Musoke AJ, Rurangirwa FR, McGuire TC (1994) Identification of *Babesia bigemina* infected erythrocyte surface antigens containing epitopes conserved among strains. Parasite Immunol 16:119-127.

Skuce PJ, Mallon TR, Taylor SM (1996) Molecular cloning of a putative rhoptry associated protein homologue from *Babesia divergens*. Mol Biochem Parasitol 77:99-102.

Suarez CE, McElwain TF, Stephens EB, Mishra VS, Palmer GH (1991) Sequence conservation among merozoite apical complex proteins of *Babesia bovis*, *Babesia bigemina* and other apicomplexa. Mol Biochem Parasitol 49:329-332.

Suarez CE, Thompson SM, McElwain TF, Hines SA, Palmer GH (1994) Conservation of oligopeptide motifs in rhoptry protein from different genera of erythroparasitic protozoa. Exp Parasitol 78:246-251.

Suarez CE, Palmer GH, Hötzel I, McElwain TF (1998a) Structure, sequence, and transcriptional analysis of the *Babesia bovis rap-1* multigene locus. Mol Biochem Parasitol 93:215-224.

Suarez CE, Palmer GH, Hötzel I, Hines SA, McElwain TF (1998b) Sequence and functional analysis of the intergenic regions separating babesial rhoptry-associated protein-1 (*rap-1*) genes. Exp Parasitol 90:189-94.

Suarez CE, Palmer GH, Florin-Christensen M, Hines SA, Hötzel I, McElwain TF (2003) Organization, transcription, and expression of rhoptry associated protein genes in the *Babesia bigemina rap-1* locus. Mol Biochem Parasitol 127:101-12.

Suarez CE, Laughery JM, Bastos RG, Johnson WC, Norimine J, et al (2011) A novel neutralization sensitive and subdominant RAP-1-related antigen (RRA) is expressed by *Babesia bovis* merozoites. Parasitology 18:1-10.

Tian ZC, Liu GY, Yin H, Luo JX, Guan GQ, Luo J, Xie JR, Shen H, Tian MY, Zheng JF, Yuan XS, Wang FF (2013a) RPS8 - a new informative DNA marker for phylogeny of *Babesia* and *Theileria* parasites in China. PLoS One 8:e79860.

Tian Z, Luo J, Zheng J, Xie J, Shen H, Yin H, Luo J, Tian M, Yuan X, Wang F, Liu G (2013b) Phylogenetic analysis of *Babesia* species in China based on Cytochrome b (COB) gene. Infect Genet Evol 13:36-40.

Tian Z, Liu G, Yin H, Luo J, Guan G, Xie J, Luo J, Zheng J, Tian M, Yuan X, Wang F, Chen R, Wang H (2013c) 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.

Uilenberg G (2006) *Babesia*-a historical overview. Vet Parasitol 138: 3-10.

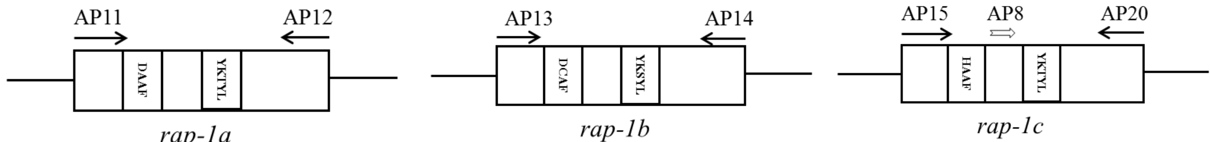
Wright IG, Casu R, Commins MA, Dalrymple BP, Gale KR, Goodger BV et al (1992) Thedevelopment of recombinant *Babesia* vaccine. Vet Parasitol 44:3-13.

Yin H, Lu W, Luo JX (1997) Babesiosis in China. Trop Anim Health Prod 29:11S-15S.

Yokoyama N, Suthisak B, Hirata H, Matsuo T, Inoue N, Sugimoto C , Igarashi I (2002) Cellular localization of *Babesia bovis* merozoite rhoptry-associated protein 1 and its erythrocyte-binding activity. Infect Immun 70:5822-5826.

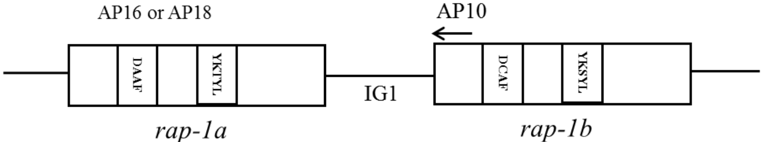
Yokoyama N, Okamura M, Igarashi I (2006) Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. Vet Parasitol 138:22-32

a). Step 1: *rap-1* genes amplification using common *rap-1a*, *b* and *c* primers with Babesia sp.BQ1 (Lintan)

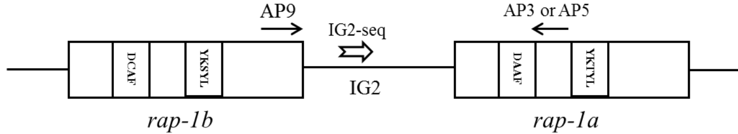


Step 2: Amplification of the intergenic regions using rap-1 specific primers

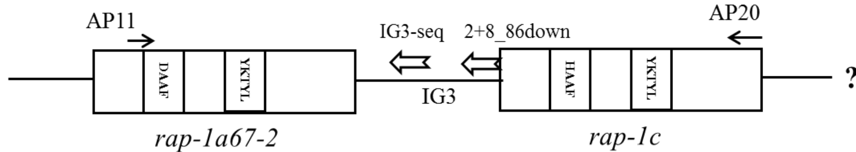
IG1: Ningxian and Tianzhu



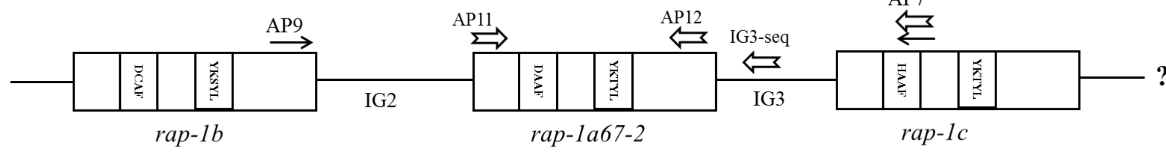
IG2: Ningxian and Tianzhu



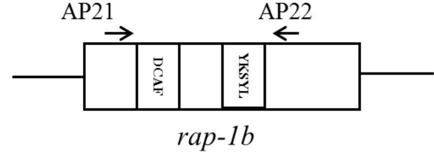
IG3: Ningxian



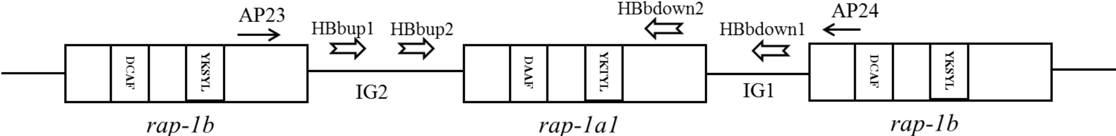
IG3: Tianzhu



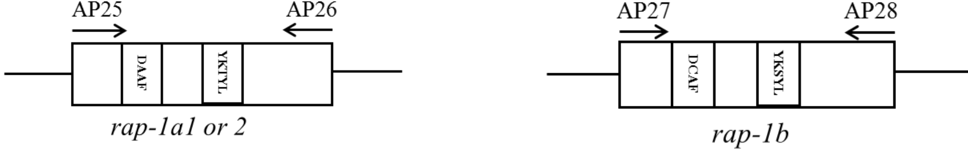
b). Step1: rap-1b internal region amplification using degenerated primers



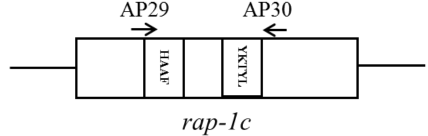
Step2: Amplification of the interspacing rap-1 genes and intergenic regions (IG1, IG 2) using rap-1b specific primers



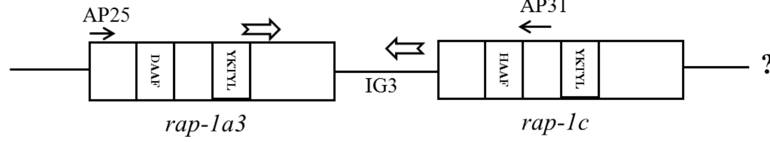
Step3: Amplification of full length rap-1a and b gene using rap-1a and b specific primers



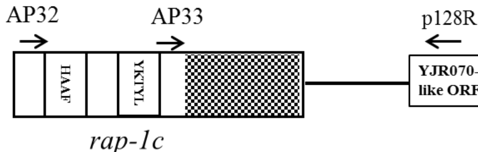
Step4: Amplification of rap-1c internal region using degenerated primers



Step5: Amplification of rap-1c beginning part and intergenic region 3 between rap-1a and c using rap-1a and c specific primers



Step6: Attempts to amplify 3' end of rap-1c



Supplementary Fig.1. Strategy for rap-1 genes and locus amplification and primers location deduced from rap-1 Babesia sp. BQ1 (Lintan). a: Strategy for rap-1 Babesia sp. BQ1 (Ningxian) and Babesia sp. Tianzhu; b: Strategy for rap-1 Babesia sp. Hebei.Primers for amplification: ⟶ Primers for sequencing.: ⇒ unknown sequence.: ▩

```
rap-1a1         TCGTCCGACGAAGCTGCTTCTGCCGA---------------AGAAGCCGGTGAGGAGAAG 1245
rap-1a2         TCGTCCGACGAAGCTGCTTCTGCCGAGGCTGCTTCTGCCGAAGAAGCCGGTGAGGAGAAG 1260
rap-1a3         TCGTCCGACGAAGCTGCTTCTGCCGA---------------AGAAGCCGGTGAGGAGAAG 1245
                **************************               *******************

rap-1a1         GCCGCCTCTGCCGAGGCCTCGACTGAAGAGGTTGCTGAGAAGGTCGCGGACGACAAATCG 1305
rap-1a2         GCCGCCTCTGCCGAGGCCTCGACTGAAGAGGTTGCTGAGAAGGTCGCGGACGACAAATCG 1320
rap-1a3         GCCGCCTCTGCCGAGGCCTCGACTGAAGAGGTTGCTGAGAAGGTCGCGGACGACAAATCG 1305
                ************************************************************

rap-1a1         TACGAGGCGGCTGTTGAGCAGGCCCAGCAGGAGGAAGCTGCCGCCGCTGAGACCGCAGAG 1365
rap-1a2         TACGAGGCGGCTGTTGAGCAGGCCCAGCAGGAGGAAGCTGCCGCCGCTGAGACCGCAGAG 1380
rap-1a3         TACGAGGCGGCTGTTGAGCAGGCCCAGCAGGAGGAAGCTGCCGCCGCTGAGACCGCAGAG 1365
                ************************************************************

rap-1a1         GCCGAAACCACTGAGTCTACTGATTCGTGGCTTTAA 1401
rap-1a2         GCCGAAACCACTGAGTCTACTGATTCGTGGCTTTAA 1416
rap-1a3         GCCGG---CGCTGAGGCCGA-AATTCAA-----TAA 1392
                ****    * ***** *     ****       ***
```

Supplementary Fig. 2 Alignment of 3' end region of three *rap-1a* gene nucleotide sequences from *Babesia* sp. Hebei, the difference of sequence indicated with bold.

```
TZrap-1c        AAAGGAGTGGGAAATGACGAAGTAACTGATGCAGACAGCGTCACAGGACAGGTAGATAAG 1260
NXrap-1c        AAAGGAGTGGGAAATGACGAAGTAACTGATGCAGACAGCGTCACAGGACAGGTAGATAAG 1260
LTrap-1c        AAAGGAGTGGGAAATGACGAAGTAACTGATGCAGACAGCGTCACAGGACAGGTAGATAAG 1260
                ************************************************************

TZrap-1c        GAGGCTATCAGTATTGCAGGAGAGAACAAAGAACCAACCATCGTTGAACAAAAGGAGGCT 1320
NXrap-1c        GAGGCTATCAGTATTGCAGGAGAGAACAAAGAACCAACCATCGTTGAACAAAAGGAGGCT 1320
LTrap-1c        GAGGCTATCAGTATTGCAGGAGAGAACAAAGAACCAACCATCGTTGAACAAAAGGAGGCT 1320
                ************************************************************

TZrap-1c        ATCAACATTGCAGGAGAGAACAACGAGCCAAGCGTCGTTGAACAGGAGGCTATCAACATT 1380
NXrap-1c        ATCAACATTGCAGGAGAGAACAACGAGCCAAGCGTCGTTGAACAGGAGGCTATCAACATT 1380
LTrap-1c        ATCAACATTGCAGGAGAGAACAACGAGCCAAGCGTCGTTGAACAGGAGGCTATCAACATT 1380
                ************************************************************

TZrap-1c        GCAGGAGAGAACAACGAGCCAAGCGTCGTTGAACAGGAGGCTATCAACATTGCAGGAGAG 1440
NXrap-1c        GCAGGAGAGAACAACGAGCCAAGCGTCGTTGAACAGGAGGCTATCAACATTGCAGGAGAG 1440
LTrap-1c        GCAGGAGAGAACAACGAGCCAAGCGTCGTTGAACAGGAGGCTATCAACATTGCAGGAGAG 1440
                ************************************************************

TZrap-1c        AACAACGAGCCAAGCGTCGTTGAACGGAAGGAGGCTATTAACATTGCAAGAGAAAACAAC 1500
NXrap-1c        AACAACGAGCCAAGCGTCGTTGAACGGAAGGAGGCTATTAACATTGCAAGAGAAAACAAC 1500
LTrap-1c        AACAACGAACCAACCATCGTGGAACAAAAGGAGGCTATCAACATTGCAAGAGAAAACAAC 1500
                ******** **** * **** ****  *********** *********************

TZrap-1c        GAGCCACCTATCGTTGGTGAGAGTGTTACACCGAATGTTCCTACGTCAGACGTTACTCCG 1560
NXrap-1c        GAGCCACCTATCGTTGGTGAGAGTGTTACACCGAATGTTCCTACGTCAGACGTTACTCCG 1560
LTrap-1c        GAACCACCTATCGTTGGTGAGAGTGTTACACCGAATGTTCCTACGTCAGACGTTACTCCG 1560
                ** *********************************************************

TZrap-1c        TCAGCTAAAGCTGATACTAATGAATCAGTACCCGAGGTAACGGCCAAGGATGTGCACTCT 1620
NXrap-1c        TCAGCTAAAGCTGATACTAATGAATCAGTACCCGAGGTAACGGCCAAGGATGTGCACTCT 1620
LTrap-1c        TCACCTAAAGCTGGTACTAATGAATCAGTACCCGAGGTAACGGCCAAGGATGTTCACTCT 1620
                *** ********* *************************************** ******

TZrap-1c        AAAGATGGGGAGAATGGGTTCGAAGATGCGGTCGAAGACCTGTCGGAGCTCCCTCCTGAT 1680
NXrap-1c        AAAGATGGGGAGAATGGGTTCGAAGATGCGGTCGAAGACCTGTCGGAGCTCCCTCCTGAT 1680
LTrap-1c        AAAGATGGGGAGAATGGGTTCGAAGATGCGGTCGAAGACCTGTCGGAGCTCCCTCCTGAT 1680
                ************************************************************

TZrap-1c        GACACTAAATCACGCTTGGCTAATTACGCATCGAAGTTCAAGCGTTTTATACAATCCAAA 1740
NXrap-1c        GACACTAAATCACGCTTGGCTAATTACGCATCGAAGTTCAAGCGTTTTATACAATCCAAA 1740
LTrap-1c        GACACTAAATCACGCTTGGCTAATTACGCATCGAAGTTCAAGCGTTTTATACAATCCAAA 1740
                ************************************************************

TZrap-1c        TTATCGTAA 1749
NXrap-1c        TTATCGTAA 1749
LTrap-1c        TTATCGTAA 1749
                *********
```

Supplementary Fig. 3. Alignment of 3' end region of *rap-1c* in Lintan, Ningxian and Tianzhu strains. The region of repeats were indicated with arrows (from position 1258 to 1521), 11 substitutions with grey background, six substitutions affected amino acids modification with underlines.

```
        1                                                                                                     100
  HBIG1 GACAAAGTCA AACAGTAAGC AAATGGTATT ACTGTGACGG AATCGCTAAC AA----TGAT ATTTAATG-- TCTCGTATTT TAGGTATACG TAATCCCGCT
LTIG1.2 GGCTAAGTCA AATAATAGGC GAATCGTATT ACTGTGATGG AATCGCTAGC AAACTATGAT ATTTAATGCG TCTCGTATTT CAATCCCACG TAAATCCGCC
LTIG1.1 GGCTAAGTCA AATAATAGGC GAATCGTATT ACTGTGATGG AATCGCTAGC AAACTATGAT ACTTAATGCG TCTCGTATTT CAATCCCACG TAAATCCGCC
B.biIG1                                                                  CTAATCTC CATCGTTTTT TACTTTTATG TTGGTCAGGT

        101                                                                                                   200
  HBIG1 AGTTTTCCTT GCATTGGGCG TTGACGATCA CACGTGTTAG TGACGCTAAG GCGGGGCGGA TGTCGTTCGA TGGGA-TAAC TTGATTTTCA TCGTGTTGTC
LTIG1.2 ATTTATCGTC GCATTAGGCG TTTGAGATAA TACGTTGTAG TGACGCTAGG GCATGGCGGA TGTC---CGA TAGGA-TAGC TTAATGTTGT TCATGGTGTC
LTIG1.1 ATTTATCGTC GCATTAGGCG TTTGAGATAA TACGTTGTAG TGACGCTAGG GCATGGCGGA TGTC---CGA TAGGA-TAGC TTAATGTTGT TCATGGTGTC
B.biIG1 GTTCATATTT GCCA-AGGCA C---CGTGCT GAGATCGCTG CGTCATTTTT GCGTGTAATA TGTCGTATAA TATGCGTCCC ATGGGTGCGC TGCTATTCGC

        201                                                                                                   300
  HBIG1 TGCGGTCGTA TCTGGGGTTG TCTATACTTG CTTAT--GCG GGACGTCGGT GATAAATGAC AGCACTGTAG AACTCAAAAC TCGCAGATCC GCTCACGTAA
LTIG1.2 AGTGGCCTAT GTATGGGTTT TATATAGTTG CTTAA--GCG GATTACCGGA CGTACAT-AT GGCACTCTAA AACCAGAAAC TCGCGGTTCC GCTCCTATGA
LTIG1.1 AGTGGCCTAT GTATGGGTTT TATATAGTTG CTTAA--GCG GATTACCGGA CGTACAT-AT GGCACTCTAA AACCAGAAAC TCGCGGTTCC GCTCCTATGA
B.biIG1 TACATGCGCT GCCAACATGG CCGTAAATAG CGTGTCTGCA GGCAACGCTG CAATAATAAA AAATGTCGTG ACTGAGTTAC GCACGCGCCA TAACGTTATA

        301                                                                                                   400
  HBIG1 CAA-CGGTGT TGCATGAGGG TTATGGGAGA CGGATTCG-- TGAATCCGTC CTATATGTGG GTGTGTTGTA TACCCGACAA GCAATAATTA CAATAACTGC
LTIG1.2 CAA-CACCGT GGGATGAGG- ---------- ---ATTCA-- TGCATCCATC CTATATGTGG CTGTGTTGCC TACCCGACGG CCAAGAGTTA TGACAACCGT
LTIG1.1 CAA-CACCGT GGGATGAGG- ---------- ---ATTCA-- TGCATCCATC CTATATGTGG CTGTGTTGCC TACCCGACGG CCAAGAGTTA TGACAACCGT
B.biIG1 CAATCAGTAT TAGAAGAGT- ---------A CAGCGTCGAC TGCAAATTTT TCGTAGGTGC GTGTGATTCC TACCCGACAG GCAGCATTCA AAGACAACGA

        401                                                                                                   500
  HBIG1 GTTTCGTAGC TAAAGCCGTT ACTCGATTCG GCTGATTACA TTCTACCTTC AAATTAATAT GTTTCATACA TCGGAACTCA TGGGCGCTTT ACTGATGGCC
LTIG1.2 GTTTCGCAGC GAAAGCCGTT ACTACATTCC GCTGCTTACA TTCTACCGTT AATTTTATAC GGTAAATACA TCAGACCTCT TGGGCACTTT ACTGATGTCC
LTIG1.1 GTTTCGCAGC GAAAGCCGTT ACTACATTCC GCTGCTTACA TTCTACCGTT AATTTTATAC GGTAAATACA TCAGACCTCT TGGGCACTTT ACTGATGTCC
B.biIG1 CTTAAGATAC TAT-GCCATC CACACATTC- -------TCA TGAGAGTGAT AAATGTTCAA AAGTAGTTCA CTGTTAATCA AATAAGAGTC AATGGTGATT
```

```
        501                                                                                                   600
  HBIG1 GCGAGACTCC GCGCAGCAAA CAAACGTCGG AATACGAATG TTAGACACGG AGGTGTAGTG GAGAAATTA- CCCCTTGCGC GTTGGTGTGC ACTAGCGGGC
LTIG1.2 GCGAGAACTT TCGCAGCGAA CGT--GTAGG AATACGCGTA TTAGACACGG GGGTGTGCTG GAGAAAATA- CCCCTTGCGT GTTGCTGTGC AATACTGACC
LTIG1.1 GCGAGAACTT TCGCAGCGAA CGT--GTAGG AATACGCGTA TTAGACACGG GGGTGTGCTG GAGAAAATA- CCCCTTGCGT GTTGCTGTGC AATACTGACC
B.biIG1 GCATTGACCT TTGCTGTAAG GATGTGTATG AACGCCGCCG CTAGACACGG CAGCGCAATT GAGAAAAGAG ACATTTGAGT ATTGCTACAC AATGTTGCAA

        601                                                                                                   700
  HBIG1 TTCAATAGGT TTTAAAAGTG TTACAC-ATG TGATTGCTTT GTGTTATTGT GTGTCCAGGC AACGTGCGGG TTTTCGCACT TTTCTGCATC AC--TTTCTC
LTIG1.2 TTGTATACGG TTCAAAGGCG TTACA--ATG TGATTGCTTT GCGCTTTGGT GTGTCCAGGC AACCTGCGAG ATTTCGCACT TTTCTGCATC AA--TGTT--
LTIG1.1 TTGTATACGG TTCAAAGGCG TTACA--ATG TGATTGCTTT GCGCTTTGGT GTGTCCAGGC AACCTGCGAG ATTTCGCACT TTTCTGCATC AA--TGTT--
B.biIG1 TTATAAGCGT TGGGTGTATG ATGAATGATG CGATTCTTTT GCGTTTGGAT TTGTCTAGGC CACGTGCGAG GTTTCGCACT TTTTTGCATC AACGTATCTT

        701                                                                                                   800
  HBIG1 ACGTTCAACC GTATTGGTTG GCAATGATTG CGACGTCTCC TTCCATGTGA AATATACCAT AATG-----T ACCGTATCGT GGGAATCAAG AACTTGTAGT
LTIG1.2 ACGGTCAACT GTATTGGTTG GCAGTGACTG CGACGTCTCC TTTCACGTGA AATATAACAT AATG-----T ACCGTATCGT GGTAACTGAG AACTTGTAGT
LTIG1.1 ACGGTCAACT GTATTGGTTG GCAGTGACTG CGACGTCTCC TTTCACGTGA AATATAACAT AATG-----T ACCGTATCGT GGTAACTGAG AACTTGTAGT
B.biIG1 ACGGTCGACT ATATCGGTCG GCAACGGTTC CTACGCCTTC TTCTTTGTAT AATATAGCTC ATTGCGGTGT ACTTTATTGT ATGAGGTACG ATTTAGCAGT

        801                                                                                   891
  HBIG1 TCTTGTTCTC GTTATTGCAG CCAGTTCTGT AGAGGCTGTG CTTAACACAC TAATAGGTGT TTGGGTTAAC AACGGCATCA AAGAAGTCGT G
LTIG1.2 TCTCGTCCAC GTAATTGCAG CCAGTTCTGT AGAGGCTGTG CTTAACACAC TAATAGGTGT TTGGGTTAAC AACGGCATTA AAGAAGTCTT G
LTIG1.1 TCTCGTCCAC GTAATTGCAG CCAGTTCTGT AGAGGCTGTG CTTAACACAC TAATAGGTGT TTGGGTTAAC AACGGCATTA AAGAAGTCTT G
B.biIG1 GTTTGTTTCC GTCCTTGCAT CCAAATTTGT GCAGACTACA CTTAACACAC TGATAGGTCT TTAAGTTATA CACCAGATTC CGAAGGTC
```

Supplementary Fig. 4. Alignment of nucleotide sequences of intergenic region 1 (IG1). LTIG1.1 and LTIG1.2: *Babesia* sp. BQ1 (Lintan) *rap-1* intergenic region 1.1 and 1.2 (GenBank accession number: KC953701 and KC953700); HBIG1: *Babesia* sp. Hebei *rap-1* intergenic region 1 (GenBank accession number: KJ 205334); B.biIG1: *B. bigemina rap-1* intergenic region 1(GenBank accession number: AF021246). The positions of three conserved boxes at 5' UTR sequence were indicated with different colors, -59box (yellow), -36box (gray), mRNA box (green).

```
           1                                                                                                      100
   LTIG2   AATGAAGTAA TTGTTGCTGG CAAGTTATTA TGGCATGCAG CGTGGTTAAG GGTAGAAGCG AGTGGAACGT AGAATTTCCC GCTCATTTAT TAGGCTATAC
   HBIG2   AATGAAGTCA TTGTTGCTGG CGAGTTATGA TCGCATAAAG CGTGGTTAGG GCTATAACCG GGTGGATCGT GGTATTTTC- GCTAAGTTGT GAAGCTCTA-
 B.biIG2

           101                                                                                                    200
   LTIG2   TGCATCGAAT GCACCTTAAT CGCAAACTAT TGCCGTATAC ATACTACGTG TACTGCAATT CCACATGAGG CGCCGGCAGA TTGCGGATCC CGTGTAGCTC
   HBIG2   TGCAACGAAT GCACCTTGAT CGCAAATTAT TACCGTATAA ATAC---GTG TACTGCAATT TCGCATGAGG CGCAGGCCGA TTTCTGATCG CATGGAGCTG
 B.biIG2                                                                    GAGAGCGG ACCCTGCCCT TAGAATACGG GATGATTATT

           201                                                                                                    300
   LTIG2   AGGGTCCCTT GCGCATGAAC CTTTGCGTGC TAATATATCC AGCATAATGC TTGATGTGCT CGACGCTTAT CACCACATTT GCATTTGTGA AATCTATCCT
   HBIG2   AGGCTCCATC GATCATGGAC TTTTGCGTGC TAATATGTTC GGCATAATGC TAGATGTGCA CGACGCTGGT CACTACATTT GCACTTGTGA AATGTATTTT
 B.biIG2   CGGAAGGTAA AATCA--AAT ATTAGCCAG- -AATGGGTTC CCTAACCCGG TCACTGAGAA CAATGCAGGA CATTCCGTTA CATTTAGAGA TATA------

           301                                                                                                    400
   LTIG2   GCGTATGCAT ATGGAATATT TGTTTTGCCT C---GTGCTG TAAACATAC- -AGAGTCAGT TGATGTATAC TAGCGAAAGC AGT--TAGA- -TGGTGAGTT
   HBIG2   GC-TATGTAC ATCAACTATT TCTTTTGCAT CTTCGTGCTG TGAAATTCAG TAGAGTCGTT TGATATATAC TAACGATATC AGTAATAGAC ATGGTGAGCT
 B.biIG2   ---------- ATGAAGTAAC GGTTTCGGGA C---ATGCCC TTATCGCATG AAGCGTCGTT AGGTGGATA- -AAAGACTGC C-TCGTATCC TTCTCGATCT

           401                                                                                                    500
   LTIG2   --TTTCGTCA CATTTCGCGT TGTGCATGTA TTGATGTGTC TCCGATGATT GCTGTTGGTG GTAGCGTAGC TGCGCGATGG GAGAGCGTCG TGAGATACCG
   HBIG2   CATTTCTGC- CATTTGTCGT TGTGCATGTG TTGTTTGCCG TGCGATGATT CGTCTCGCTG TTAGCTGCAT TTGGCGATAT GAGAGGGTTG GGCGGTATCG
 B.biIG2   C--TTCTTCA TCTATA-CGA TAAACGTTTC CT-TAGGCTG TCCCATTTT- ---GTCACCT ATTGTATAAT AT---GAATG GGGGAATTCT GAAAGTGATG

           501                                                                                                    600
   LTIG2   ATCACCGGAC GATCGGCGTT GATAATGGAT GTTCAGATGC TATTATGGGT GTTGTGCACC -TGTTTTATT GGTGGTTATG CTTGTGGCGG GATTTCTGAT
   HBIG2   ATCA------ -----GCGTT GATAGGCGGT GATGAGATGG TATTATTGGT TTCGTGTATC GTGTTTTATT ACCGCTTCGC TTTGTAGCAG GATTCCTGAC
 B.biIG2   CACAG----- ----GGCGAT TGTG------ GATCATATCC AGGCTTAGGT AGCTTGCGCG TAGATATTCG GGCGGCGTGA TTTTCGAAAC GATTCC-GAG

           601                                                                                                    700
   LTIG2   ATAGACGTGG TTCCCATCTT ACGGATTCCG TGCAGACATC GAAGTACACG TATGCATTGA TTGAGGGCGC AGCCGCTCTT TCAATTGCTC GCGCCATCGC
   HBIG2   GTACACCGAA TTCCAATCAT GGGAATTCCG CGCAGACGTC GAGGCGCGCA TATGCATAGA TTGTTGGCGC AGTTGGTGTT TCTACAGCCT GTGTCATCTT
 B.biIG2   TTGTAAGCGA CTATAGTCA- ---------- --CTAACTTC G-------CG TTGG--TAAA TCATGAAATC CACCTGTATA TGTACAACAT TTTGTTAAGC
```

```
         701                                                                                                   800
   LTIG2 TCTCGAATGC TGCGAGCGTT TGAATGCGCA GAATGGAGAT TT--TTCAGG TATGCAACGG ATTGGGAATG GCTA-GAAGA GTAGCATACG ATGCTCTCAT
   HBIG2 TCTGGAATGG TGCGGGCGTT TGATAGCTCA GAAGGGAGCT GTGTTCCAAG AATCCGACGG ATTGGGAACG ACTACAAAAA ACAGTTCACT ATGGTCGCGC
 B.biIG2 TCTGCGGTGT TGAGCGCGGT -GCGAGCACG GTTGTGTCTT GG----CCTG AATGTTGTGG TT----GTTG GTCATAGATA TTTGGGCGTA TTGAACGCG-

         801                                                                                                   900
   LTIG2 CAATTGTTCT TTTTCGTTAA CGTTAGATAA GAGGATTGGA GTGTATTAAG GTAAGTTAAC ATGGGGAATA GCT-CTGCA- TAACCCTC-G CTAGTGAGCC
   HBIG2 CACTTGATTA GTTTGATTAA CTACAGATAT GGACGCTTGA GTGTACTACA GTACCGTAAG ATTAGTTATT GCTGCAGCGC TGGCCCTCCG CTAATAAGTC
 B.biIG2 ----TGATGT TTAGCGGCGT CGATCGAAAA TCTCGCAGCA GTGCTCC--- ---------- ---------- ----CAGCAT TCATACAAAA CTAG--GGGC

         901                                                                                                  1000
   LTIG2 GGTGGAGGTT GCAGACAACA CGACCTCGCT GTGGGAACTT ATGACGGTAA AGATACCAGT CAATAGGTCA CTTATGCAGT AAGCACGAGG CTTGAATGCG
   HBIG2 GGTGGAGGCA -CAGGCAAGA TTCGAAGGCT GTGAGAACAT ACGGGTGTGA CGATATCGTT AAATTAGTCA ACTATGCAGT AAGAACGCGG CGTGAAGGCA
 B.biIG2 GGTGTCG--- ---------- -TACATAGC- --GACAACTG CAGC--GTAA AGGCGTCA-T CATTAAGCGT GCTGAGTGGT CAGTATA--- CATGGCCGTA

         1001                                                                                                 1100
   LTIG2 GCACGCCCAT TCCATGA-AT GAGCAAGAAT TCACTGGAGA CCCTAAACAG ATTTCACGCA TTCCGAACGG AAGTGTCCA- -CTTAGCACG CCAACAGGCC
   HBIG2 GTACACCCAT CTCATGGGAT GAGCTAGCGT TTATCGGCGA GCTCAATCAG CTTTAACGCA GTCAAAACGG ATGTGTTCG- -CCTAGCACG CCAATGCGCT
 B.biIG2 GTCTACGCAT GACGCGG-AG GAACAGAACT AC-----CTA ATTCACGCAA CTT---CGAC TTAAGAACGC ATGAAATAAA ACACATCAAG CCAAC--GCA

         1101                                                                                                 1200
   LTIG2 TTATAATTAT TGTGAGTGAC ATATAGAGGA GCGATGCGAA CACGCTGAGG CTTCGTAAGC GTACGGGCGC CACCTATCCA TGTGCCATAG AAACCTAAGG
   HBIG2 TAATAGATAC TGTGAATTAC ATACAGAGCA GTGATGCGAA TA-GCCTAGG CTTCGTAACC GTACGTGCGC CACGTATCCA TCCGCCATAG AAACGTAAGA
 B.biIG2 TTAGCGGTAA CG--CATTAC CTCAAG---- ---------- ------GTGG CATAAATTAC --AAGTG-GA CACGTTTACA TAAACCAGTG TCAACATTGA

         1201                                                                                                 1300
   LTIG2 AACGGTTGGA ATACACCACG TGGAGAACCA GGCAGCTGAT TTTGCCACTC GCATTGAGCA TTATTGCGGC GAAGAGACAA GAAACCAACA CCTCAACCTA
   HBIG2 CACGATTAAC GCAGGC-ACG CTGCTAATTA ACAATCTGAT TTTGCGACTC --ATGGAGCA TGATCAGGGC GTTCACCTGA GGAATCAAAA CGCCAACCTA
 B.biIG2 TAAGA---AC GTTCAAAACG C----AAAAA CGCGCCCAAT AATGT-ACAC ------AAAA TGCATACGCC GTTCTGACGA GGAATCATCA ACTCAACCTG

         1301                                                                                                 1400
   LTIG2 GAATTAGACA TGCC-GAAAA CCTCAGTGTG TGGCCATACC ATTTGTTACT GATGGCGGGT ATTTCAGT-T GTCTTTCTTC ATATTTCTAA CATGGCGCAC
   HBIG2 GAAGTAGACA TGCTTGATGA CCTCGGTGTG TGGCCATAC- ATTTGTTCCT GATGGCGGGT ATGTCAGC-T GTCTTTCCTC TTATTTTGAA TATGGCGGAC
 B.biIG2 GAATTAGACA CGGTAAAGAA GTTCAGTGTG TGGC-ACAC- ATTCGCCCTC GAGGGGAATG ATTTCACCCA GTTGTACGCA TTGCTTATGA TATGGTACAC
```

```
        1401                                                                                                     1500
LTIG2   AAATGAAACA ATTGAAAGCA TGCTGCATAT GCAGTAATGC ATATAACCAC AGCGTAGTAT GAGCCAGTCC GCGTGGGAGA TGGGACGTAA GCGTGAGTAC
HBIG2   AAATAAAAGT CTTGAAAGCA TACTGCATAT GCAGTAATGC GTGTACTAAG AATGTAGTAT GAATCATGCC GCGTGG-CGA CGGTATGTAA CTGTGAGTGC
B.biIG2 GTACGGCAGT GT-GAAAGAT TTAATAATGT ATCCTTATAT ATATCAGGAT GCCGTAGAAG GTGCGTGTCA CTACGCGAAA GAGCCGGTGA GCATGAATGC

        1501                                                                                                     1600
LTIG2   ATATCGCAAC AAAAGCAGGT ATCGACGCTA AACAGCGCGA ACCGTTAACA ACAGTTAAGG TGTCACGTAA TGTGGAACAC ACGGGAGGCG TGTCTATGCG
HBIG2   CTATGGCAAC AAAAGCAGGT ATCGACGTTA AACAGCGCGA ACTGTTTATT AGTAGTATGC TGTCACGTAA TGTGCGACAC ACGGGAGGCG TGTCTATGCG
B.biIG2 GTAGCGCAAT ATGGGCGGAT GTGAACGCCA AGGGTTATGA ATAGCGTGTA AACAGTAAAA TATTATGAGG CATCCGTCAC ATAGAGCGTG TGTCTAGAGC

        1601                                                                                                     1700
LTIG2   TGAAGGCAGC CATAAGGAGC GTTTCGCACT TTGTTGCAGC AATTGGTGTT TTAGACAGGT GCATCGCGTA TGTGTGAGAA TGTAGCACGT CGGAGCATAT
HBIG2   GTTATGCAGC CATAACTAGC GTTTCGCACT TGGTTGCAGT AATGGCTGTT TTAGACAGGT GCATCGCGTA TGTGTGAGAA TGTAGCACGT CGGAGCACAT
B.biIG2 AGCCGTCGTC TACACTCACT GCATCGCACT TGTCTGCAGT AACGCTTGTG TTAGACATGT GAGTCGCGTA TGTGTGAGAA TATGGCACAT TGGCGCATAA
        1701                                                                                                     1800
LTIG2   CCATTGTGAA TGCTTGTTTG TTGAGGCGGA GATATATCGG CCGTGCCGTG TTTCCCGTTA CATTTATACT AATAGCAGAG TCGTATATAC GATCC-GCTT
HBIG2   TCACTGTGTC TACTTGATTG TGGAAGCGGA TATATATCGG CTGTGCCGTG TTTCCCGTTA CATTTATATT AATAGCAGAG TCGTCTATAC GATCT-GCTT
B.biIG2 GCACTCCCAA TAAGTGATTG TGAACGCGGA ATTAGGTCGG CCGTGCCGT- TTTTCCGTTA GAATAATATT TCAAGCAGAT TCGTCTAATC GTTCCTGCTG

        1801                                                                                 1885
LTIG2   AGATTATCGT ATAATAAATA ---TTCTTCT TGTAGCCTTT --CTTTGCTG CTGC--CCGT TGGTCTTCAC G
HBIG2   ATATTATCGT ATAATAAATA ---ATCTTCT CGTAGCCTTT --GCTTACTG CTGTTTCCGT TGCAGTTCAC G
B.biIG2 TCACTATCGT AATACACATA GCGTTGGTCT GCTAACGTTT TGGTGAGCAA CATTCCCATT TGCGTTACAA ATGAAGGATC TTCCT
```

Supplementary Fig. 5 Alignment of nucleotide sequences of intergenic region 2 (IG2). LTIG2: *Babesia* sp. BQ1 (Lintan) *rap-1* intergenic region 2 (GenBank accession number: KC953702); HBIG2: *Babesia* sp. Hebei *rap-1* intergenic region 2 (GenBank accession number: KJ205336); B.biIG2: *B. bigemina rap-1* intergenic region 2 (GenBank accession number: AF021246). The positions of three conserved boxes at 5' UTR sequence were indicated with different colors, -59box (yellow), -36box (gray), mRNA box (green).

```
       1                                                                                                  100
LTIG3  CGAATTTTGT TACCCTTGCA TACGTCATTT TCCCGTTAAT CATGCA--GA TATTCTATTA TATCGAATGT AAAATGTTTT TAATTGTTGT ATTACACCAT
HBIG3       ATTCGC CATCAGTGTA TCGTATAATA TATTTTTATT AATGCGTTGA TTTCTTGAAA TATCGCA-GA CAATCTTTTT AATTTAAGTT CTTTTATCAT

       101                                                                                                200
LTIG3  ATAGGCTTAT TTTATTGCCA TCTTTATTGA ACGTGTGTCA TAATATTAAT GAACTCATTA TTCATCTCGT GGTATAGTGT TGTTTCTATG TGTAATTTTT
HBIG3  GTACAATACT ACTA-----A GCTTTTTTGA T------TCG --ATGTTTGC GAACTCATTT TCAGTTTC-- ---TTTCAGA TGTTTGTGTC ATTTTTTTGT

       201                                                                                                300
LTIG3  TGTTTACTCT TCGATTCTGG CCGAACACCG CCTATG-CGG GAGTGTTTTT CTGT----GA AACATTT--T AAGCGAAATC TCTCTCCCGG GATGAACCAT
HBIG3  TGTTAAGCCT CCCTATCAG- CTGCTAACCC GCTATGTCGT GATTCTCTTT TTGTTGCAGG AAAATTTCCT CAATACGGAC ATTTTAACGA GTTGTACCGT

       301                                                                                                400
LTIG3  ATTCTTCATA CAGT----CG CGAAGATAAA --TTGATTAT CACATTTT-C ACGCAAACGA CTCTTAAATA CATTATAGAA TGTCGAAACG AACTGTTAAA
HBIG3  ATTCTTTATG AGATGGAGTG CCACACTAAA CGAAGATGGT CTCATTTTGT TTGCATGCGA AT--TTAATG CAGCATTGAA AGGTTAAAAT GAGTATTGAA

       401                                                                                                500
LTIG3  TATTTTCGTA ACAATATGTT AGCTATGACG TAATAACCTA CGACATCCGT AGGATCTCTA ACCCGCCTTT -GGTTGTAAA C-TGCGACTG CACGTGCGAC
HBIG3  TTTTGTA--A ACGATAGGTT ATCTGTGATA CAATATTC-A TGCCATTGGT ATCTCCTTTG ACCCAAGTTT CAGTTGCGAA GGTGCTAATC CAGGTTAGGA

       501                                                                                                600
LTIG3  CCCAAATGAC TCTCAACTAG TTAACAGCAC TGAATACGTG ACATCTCCTT TATAAACAAT CACCAACACC CGATAACACT ACTGACCATA AGATGCTTAC
HBIG3  AA-AATTGGC CGTCTAGTAT TTAACAACGT TTAATAAAT- -CATTTCCCT TTAACACCAA AACCACTACT ACATTCCACC ACTGATGATT AAATGTTTAC

       601                                                                                                700
LTIG3  AATTGTGATT AAATGATAAC GCATTTTCGA TGCGTTAATA ACAACCTTCA CGCAGACAAC TATAAGTAAA AGAAGTCTGG AGGAAGATCA TGAGCCGCAG
HBIG3  TACTGTGATT AAATGAGAAA GCATTTCTGC AGCGTTACGT GCAATATGCA GGCCTACAAC TATTAGTAAA CGAATACTAG AGAATCATCT TGCTTCGCAG
```

```
       701                                                                                                  800
LTIG3  GATTTGTAGA TGAAACATTC AAAAGTGCT  CTAAGAATGA AACA-GAACT TAGCAGAAAG TCAGAATGTT CGCAGAGAAC GTTTTATAAT TGTTAGACAC
HBIG3  AATTTGTAAG TGAAGCATTA GAAATGTGCT CATATAACTA AACACAAACT TATAAGAAGG ACATCATTTA TTCACAGATC ATTTTATTAT CGTTAGACAC

       801                                                                                                  900
LTIG3  CATCAGCCAA TAGTGCTAAT ACACAGGCAG C-GGGATCAA CCGAAGGATG ACAAAAATTA CCAC-TGACC ACATATTTCT GCAGAACCAA TATGATTATA
HBIG3  TATCCACCAC CAATGCTAAC ACACACGCGG AAGGGATCAG CGGAAGGTCC ACAAAATATA CCAAGTAAAC ACACATTTCT GCAAATATGA TATGCTTATG

       901                                                                                                  1000
LTIG3  AAACTTATAT TACATGTAGT ACATGTCGTA GAATACAGCA GCGATGCATC ATTATTAATT TGATCGAAAG TAGTTACAGA TGGAATCTAG TAATCCTTCT
HBIG3  AATGGTATAA TACAAGTTGT ATATGTCGTA GAATATAGCA CTGATACATT ATGATTACTT TGACACAAGG CACTTACACA TGGAGCATGG CAAACCTTCT

       1001                                                                                                 1100
LTIG3  GCCAAAACAA GAGGTTAACT TTCGATGATT GTACACATTA TGGAATTATG AGAAGAAGAA GGCAAATTAC ATCTCTCTGA CTTCAAGTCG TTAGCGAATC
HBIG3  ATTGAAACTA GAGAATGCAA TGGGATTACA GTACACATTA CGTAATCATG AGAAGGAGAA AGCGAATCT- ATCGCTGTAT TTTTGAATCG TTCGCGAATC

       1101                                                                                                 1200
LTIG3  CTGTCTACTT CACTTATTGC ACATAGGTGC ATCAACGTAT CGTATGCAGG GTGCAATTGC CGCGCCGTTT TACTTCGCGA GTGATTGATC GAACAACGCC
HBIG3  TTGTCTAGAT GAGTTATTGC ACATAGGTGC ATCAACCTAT CGAATACAGG GTGCAATATC CTCGCCGCTT TACTTTGCAA CTGATTAATC GAATTAAGCC

       1201                                                                                                 1300
LTIG3  GTATGGACGT GTAGAATATA GATAATATCG GATGTTCAGT CTATTTACGT GTATAGCTGT GCTATACGCC TTGGTGTGTG TCTATACTGT GAACTTCGAC
HBIG3  GCATTGACTT GTAGAACATA GATTAATTTT GAAGTTTAGT GTAGTTACGT GTATAGCTGT GCTATACTCG TTGGTGTGTG TCTACACTGT ACACATCCAT

       1301                                             1354
LTIG3  TAGAAAACCC CTGTCTAAAT AGCTAAAAAT TCTTTAACAT CCCTTCCTGT AGTG
HBIG3  TGGGGAAACC CTGTCTAGAT ATTTAAAAGT TGTTTCACA- CCCTTCGTGT AAAG
```

Supplementary Fig. 6. Alignment of nucleotide sequences of *rap-1* intergenic region 3 (IG3). LTIG3: *Babesia* sp. BQ1 (Lintan) *rap-1* intergenic region 3 (GenBank accession number: KC953703); HBIG3: *Babesia* sp. Hebei *rap-1* intergenic region 3 (KJ205337). The positions of three conserved boxes at 5' UTR sequence were indicated with different colors, -59box (yellow), -36box (gray), mRNA box (green).

```
                   1                                                                                                      100
B.birap-1a         ATGAGGAGCT TCTTGGGTGT GTGTTTTGGA GCTCTCTTGC TCGTAGCAAG GAGCGGTTCT GCTATTCGCT ATACTCACCG TTCGGGTGTT ATGTCAGCAG
LT,NX,TZrap-1a67   ATGAGAAGCT TCGCGGGTGT ATGTTTTGGT GCGCTCTTGC TCGTAGCAAG AAGCGTTTTG GCTGTACGCC ATTACCACCG TTCTGGTGTG ATGTCATCTG
LT,NX,TZrap-1a61   ATGAGAAGCT TCGCGGGTGT ATGTTTTGGT GCGCTCTTGC TCGTAGCAAG AAGCGTTTTG GCTGTACGCC ATTACCACCG TTCTGGTGTG ATGTCATCTG
  HBrap-1a         ATGAGAAGCT TCGTGAGTGT ATGTTTTGGT GCGCTCTTGC TCGCAGCAAG AAGCGGTTTG GCTGTGCGCC ATTACCAGCG TTCTGGTGTG ATGTCAGCTG

                   101                                                                                                    200
B.birap-1a         AGGTGGTTGG AGATGTGTCC AAGACCTTGC TGGAAGCCAA TGAGGTTGTC AATGCTGAAA TGGAAGCAAC TCAGGTCAAC AAAGATATGC AAAGTCAATT
LT,NX,TZrap-1a67   AAGTCATAGG CGATGTGTCT AAGACGTTGA TGCAGGCCAG TGAGGTGATC AATGCTGATT TGGAAGCTAC ACAGCTAAAC AAGGATATGG AAAGGCAGAT
LT,NX,TZrap-1a61   AAGTCATAGG CGATGTGTCT AAGACGTTGA TGCAGGCCAG TGCGGTGATC AATGCTGATT TGGAAGCTAC CCAGCTTAAC AAGGACATGA GTGGGCAGTT
  HBrap-1a         AAGTCGTAGG CGATGTGTCT AAGACTTTGA TGGAGGCCAG TGCGGTGATC AAGGCTGATT TGGAGGCTAC CCAGGTTAAC AGTGACATGA GCAAGCAGTT

                   201                                                                                                    300
B.birap-1a         GTCTAATGTT AAGGAGACCA TTGTTGGTGA GGTCTGCGAG AAAGTTGCTG GAAACTCTAC CTGCGGTGAG AGCGTAATTG CCTATGTTAA CCGTTGTGAT
LT,NX,TZrap-1a67   GTCACGCATG GCGAACATCA TCGTAGACAA GGTTTGCCGC AAGGCTCCAA GGAACAGTGA TTGTCGTCCA TCCGTAGCCG CTTACGCTCG TCGTTGCCTT
LT,NX,TZrap-1a61   GTCGCGCATG ATGGACGCCG TCGTAGACGA AGTTTGCCGC AAGGCTCCTA CAAGCACTGA GTGCCGTGCA TCCGTGACCG CTTACACTCG TCGTTGCGGT
  HBrap-1a         GCCTAACATG TTGGACACCA TCGTAGACGA GGTTTGCCGC AAGGCTCCTG GAAGCACTGA TTGCCGTGAA TCCGTGAGCG ATTACGTTCG TCGTTGCGGC

                   301                                                                                                    400
B.birap-1a         GAGGGCGATT GTCTGACGCT TGACAGCATG AAGTACAAGC CGTTGAGTCT GCCAAATCCT TACCAGTTGG ACGCTGCCTT CATGCTTTTC AGGGAAAGTG
LT,NX,TZrap-1a67   AAAGGTTCTT GTTTGGCATT GGACAACGTG AAGTACGAAC CATTGACGCT GCCGAATCCT TACCAATTGG ATGCGGCATT CACTGTATTC AGGAACAGTG
LT,NX,TZrap-1a61   AGTGGCGACT GTTTGACCCT GGACAACGTG AAGTACGAAC CATTGACGCT GCCGAATCCT TACCAGTTGG ATGCGGCATT CACGCTGTTC AGGAACAGTG
  HBrap-1a         AGCGGCGACT GCTTGACCAT TGACAACGTG AAGTACCAGC CATTGACGCT GCCTAACCCT TACGAGTTGG ATGCTGCCTT CACGTTGTTC AGGAACAGTG

                   401                                                                                                    500
B.birap-1a         ATTCTAACCC TGCGAAGAAT GAGGTGAAGC GCTTCTGGAT GCGTTCGAGG AGCAGCCACG GCGACTACCA TCACTTTGTT GTTAGCTTGT TGAAGAAGAA
LT,NX,TZrap-1a67   ACTGTAACCC GTCTAAGAAC CCCATCAGGC GTTTCTGGAT GCGTATTAGA GGCAGCCACA GTGACTTTCA GCACCTCATT GTCAGTCTTT TGAAGAAGAA
LT,NX,TZrap-1a61   ACTGTAACCC GTCTAAGAAC CCCATCAGGC GTTTCTGGAT GCGTATTAGA GGCAGCCACA GTGACTACCA GAACCTCATC GTCAGTCTTT TGAAGAAGAA
  HBrap-1a         ACTCTAACCC CGCTAAGAAC CAGGCCAGGC GTTTCTGGAT GCGTTCCAGG GGCAGCCACA GTGACTACCA GCACTTTGTT CTTAGTCTGG TGAAGAAGAA

                   501                                                                                                    600
B.birap-1a         TGTTGTACGC GACCCTGAAT CCAATGATGT TGAGAACTTC GCATCGCAGT ACTTCTACAT GACTACGTTG TACTACAAGA CTTACCTGAC CGTTGACTTT
LT,NX,TZrap-1a67   CGTTGTGCAT GATCCCGAGG TTGAGGACAT CGAAAATTTC GCATCGGAGT TCTTCTACAT GACTACGGTG TACTACAAGA CCTACCTTAC CGTGGATATT
LT,NX,TZrap-1a61   TGTTGTGCGT GATCCCGAGG TTGAGGACAT CGAAAATTTC GCGTCGCAGT TCTTCTACAT GACTACGGTG TACTACAAGA CCTACCTTAC CGTTGACACG
  HBrap-1a         TGTTGTGCGT GATGCCGAGC TTGAGGATAT TGAGAATTTC GCATCGCAGT ACTTCTACAT GACTACAGTT TACTACAAGA CCTACCTGAC CGTGGACACT
```

```
                 601                                                                                              700
B.birap-1a       ACGGCGGCTA AGTTCTTCAA CAAGCTTGCT TTCACAACTC GCCTGTTCGG TTTCGGTATC CAGAAGGCGT TGAAGCGTTT GGTTAGGAGC AACCTTCCCG
LT,NX,TZrap-1a67 GTAAAGGCTA AATTCTTCAA CAGGTTCACT TTCACGTCCC ACCTACTTGG CTTAGGCATC AGGAGGGCTT TGAAGCGTCT GGTTAAGGCC AACCTTCCAG
LT,NX,TZrap-1a61 CTGAAGGCTA AATTCTTCAA CAGGTTCACT TTCACGTCCC ACCTACTTGG CTTAGGCATC AGGAGGGCTT TGAAGCGTCT GGTTAAGGCC AACCTTCCAG
  HBrap-1a       ACCAAGGCTA AATTCTTCAA CAAGTTCACC TTCACGACCC ACTTGCTTGG CTTCGGCATC AGGAGGGCTT TGAAGCGTCT TGTTAAGTCC GAACTTCCCG

                 701                                                                                              800
B.birap-1a       TTGACCTTGG AACCCACCCT GAGGCCACCA TCCGCGAAAT AGCTAGCGGC TACGGCGAGT ACATGATGAC CCAGGTGCCT GCGATGACCT CGTTCGCTGA
LT,NX,TZrap-1a67 CTGACCTTGG TATTCACCCT GAAGAGCGCA TTCGCGACAT AGCGCGCAAC TACGGCGACT ACATGTCAAC TCAGGTTCCT ACGATGGCAT CGTTCGCTGA
LT,NX,TZrap-1a61 CTGACCTTGG TATTCACCCT GAAGAGCGCA TTCGCGACAT AGCGCGCAAC TACGGCGACT ACATGTCAAC TCAGGTTCCT ACGATGGCAT CGTTCGCTGA
  HBrap-1a       CTGACCTTGG TACCCACCCT GAGGAGCGCA TCCGTGACAT CGCTCGCAAC TACGGCGACT ACATGTCGAC TCAGGTTCCT ACGATGGCCT CCTTCGCTGA

                 801                                                                                              900
B.birap-1a       GCGTTTCTCC AAGATGGCTA CTAAGACTCT GTTGGTTACC GTCAGCGACT ACGTCCATTT GCCCGCGTAC AAGAGGTGGT ACAGGAAGTT CAAGGAATTC
LT,NX,TZrap-1a67 GCGTTTCGCC AACATGGCTA CTAACACTCT GGTGAAGACC GTGAGCGACT ACGTTCATTT GCCATTCTAC GTGAAGTTCT ACAGGGAGAT TAAGGAGCTC
LT,NX,TZrap-1a61 GCGTTTCGCC AACATGGCTA CTAACACTCT GGTGAAGACC GTGAGCGACT ACGTTCATTT GCCATTCTAC GTGAAGTTCT ACAGGGAGAT TAAGGAGCTC
  HBrap-1a       CCGTTTCGCC AGCATGGCTA CTAGCACTCT TGTGAAGACT GTCAGTGACT ACGTTCACTT GCCTGGGTAC AAGAAGTTCT ACAGGGAGAT TAAGGAGTTC

                 901                                                                                             1000
B.birap-1a       ATTGTGAACT TCTTTACTGA CCCTGCCAAG TTGATTATGA AGCACGTCTC TCAGCCTGTA AAGACTGCCT ACACAAAGCT GGTCCCCGAA GAGCACAGGC
LT,NX,TZrap-1a67 ATCGGTGGCC TCGTGTTCCG TCCGGCCGTG AAGGTGGTCA AGCGCGTCGC AGCCCCCGTG AAGGAAGTGT ACACCACGGT GATGCCCGAG CCTGCCAGGA
LT,NX,TZrap-1a61 ATCGGTGGCC TCGTGTTCCG TCCGGCCGTG AAGGTGGTCA AGCGCGTCGC AGCCCCCGTG AAGGAAGTGT ACACCACGGT GATGCCCGAG CCTGCCAGGA
  HBrap-1a       TTGAGTGGCT TCATCTTGAG CCCTGCCAAG AAGGTGATCA AGACCGTCAC CTCACCCGTG GTGGATGTCT ACACTTCTGT GATGCCCCAG TCTGCAAGGA

                 1001                                                                                            1100
B.birap-1a       AGGCTATCAG GAATGTCGTC GGTCAAAGCA CCAAGCATAT TGCCAACGGT GTACGTGATT TGTCAAGGAT GATTAAGGAG CCTAGCCAAC AAATAATTCG
LT,NX,TZrap-1a67 AACTGCTTAG AGCGACCATC CGCGAGCCCG CCCATCTCCT ATTGAAAGGT GTGAAGGGTT TGACAGACAT AGTCGCTGAG CCTGGCAAGG AACTGCTGCG
LT,NX,TZrap-1a61 AACTGCTTAG AGCGACCATC CGCGAGCCCG CCCATCTCCT ATTGAAAGGT GTGAAGGGTT TGACAGACAT AGTCGCTGAG CCTGGCAAGG AACTGCTGCG
  HBrap-1a       AGGTCCTTAG GGCTACCATC CGTGAGCCCG CCCATGTCGT CGTGAAGGGT GTAAAGGGTT TGAAGAACAA GATCGCGGAG CCCGGTAAAG CCCTGATCCG

                 1101                                                                                            1200
B.birap-1a       TGAGAAGCTG CCTCACTACC TTTCTAAGGC AAAGGGAGCC GTTGAGCACG TTGTTAAGAA GGTTAAATCC GTTGTGCCGA TAAAGCAAAA GGGCGACCAA
LT,NX,TZrap-1a67 TGAGAAGATT CCCGGCTACA TTTCTAAGGC CGGCGGTTTG GTGAACCGCA TTGTGGACAG AGTGAAGCCC AAGCTTGCAC TGGAGGGTAA CGGTGTGGAG
LT,NX,TZrap-1a61 TGAGAAGATT CCCGGCTACA TTTCTAAGGC CGGCGGTTTG GTGAACCGCA TTGTGGACAG AGTGAAGCCC AAGCTTGCAC TGGAGGGTAA CGGTGTGGAG
  HBrap-1a       TGAGAAGATC CCCGGTTACG TGTCTGAGGC CACGGGTTTT GTGAGCCGCG TTGTGGACAA GGTCAAGCCC CTGTTGAGCC TGCGCGGTGA CGGTGAGGAA
```

```
                 1201                                                                                                 1300
B.birap-1a       CCATCCGAAG CAGCTGTAGA GGAAACCGTT CCGTCTGGCG ATTCCGCGGA AACTGAATTT GAGGTCCCTG AAGAACAATA CGTCGATGCT GTTACTACTC
LT,NX,TZrap-1a67 TCGGCTGAAG AAGCGGT--- GGCTACCCAA ACCACTGAGG AG---GTGGA CGGCGCGTCT GCCGAATCTA CTGACGTTTC CAACGAAGAA GTTGCT----
LT,NX,TZrap-1a61 TCGGCTGAAG AAGCGGT--- GGCTACCCAA ACCACTGAGG AG---GTGGA CGGTGCGTCT GCCGAATCTA CTGACGTTTC CAACGAAGAA GTTGCT----
  HBrap-1a       TCGTCCGACG AAGCTGC--- TTCTGCCGAA GAAGCCGGTG AG---GAGAA GGCCGCCTCT GCCGAGGCCT C-GAC----- ---TGAAGAG GTTGCT----

                 1301                                                                                                 1400
B.birap-1a       AGGAGGTTAA CAGCGAGAAG GTTGATGCCG ACGATGCGGG TAATGCCGAA ACCCAGCAGC TTCCAGATGC AGAAAATGAA GTGCGCGCTG ATGACCCCAA
LT,NX,TZrap-1a67 --GAG---AA CAACGCCGAA GACAAAGCGT ACGAGGCGGC TGTTGAGCAG GCCCAACAGA CTGAGGAGGC AGCTGCCCCT GAGGCCGCAG AAG---GCGG
LT,NX,TZrap-1a61 --GAG---AA CAACGCCGAA GACAAAGCGT ACGAGGCGGC TGTTGAGCAG GCCCAACAGA CTGAGGAGGC AGCTGCCCCT GAGGCCGCAG AAG---GCGG
  HBrap-1a       --GAG---AA GGTCGCGGAC GACAAATCGT ACGAGGCGGC TGTTGAGCAG GCCCAGCAG- --GAGGAAGC TGCCGCCGCT GAGACCGCAG AGG---CCG-

                 1401                                        1444
B.birap-1a       AAATGAAGAT TCTTCAAGTT CTTCAGATGA TTCAGATGCG TAA
LT,NX,TZrap-1a67 TGATGCCGCG AAAGCCGATG AATCCGATGA TTC-GTGGCT CTAA
LT,NX,TZrap-1a61 TGATGCCGCG AAAGCCGATG AATCCGATGA TTC-GTGGCT CTAA
  HBrap-1a       ---------- -AAACCACTG AGTCTACTGA TTC-GTGGCT TTAA
```

Supplementary Fig. 7 Alignment of *rap-1a* gene sequences of *B. motasi*-like and *B. bigemina*. B. birap-1a: *B. bigemina rap-1aα1* gene (GenBank accession number: M60878), LT,NX,TZrap-1a67: *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu *rap-1a67-1* gene (represented sequence from LT isolate, GenBank accession number: KF039723); LT,NX,TZrap-1a61: *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu *rap-1a61-2* gene (represented sequence from LT isolate, GenBank accession number: KC953700); HBrap-1a: *Babesia* sp. Hebei *rap-1a1* gene (GenBank accession number: KJ205334).

```
                1                                                                                                           100
LT,NX,TZrap-1b     ATGGCAC CAGTGTTGAA GCGCATTGGG CGCTTAACAG CGTCATATGC GGCGTTGATC CTTCCCATGG TCATCGGTTT CCCAAACCGT TCTTCGCGCC
  HBrap-1b         ATGGCAC TAGTTTCTAA GCGCCTCGCG CGCTTAACAG CATTGTATGC TGCCCTGATG GTTCCCATGG TCATCGGTTT CCCAACCCGT TCTTCGCGTC
B.birap-1b      ATGATGGCAC CTGTTTTGAA GCGCGTCGTG T--------- CATTGTATGC TGCGCTGGCT GTGCCCCTTG TCCTCAGTGT ATCTTATCGT ACCCCGCGCT

                101                                                                                                         200
LT,NX,TZrap-1b  TGTCTGCTAA CTTGAATGGC AGCATGCTCC CCGAACTGTT GGGACATGAG GGTGAGCTTG CCAATGTCTC TCGTACTCGT CGGAGTTTAA ACAAGGCCAT
  HBrap-1b      TGTCTTCTAA CTTGAGTGGT AGCATGTTCC CCGAATCGTT GGGCCATGAG GATGCGATTG CCAATGTCTC CCATACTCGT AAGAGGCTAA ACAAAGCCAT
B.birap-1b      TGTCTAGGGG CATGAATGTG GACATGCCGT CCGAATCACT GGGACAGGAA GACGCTGTTG CCAACGTGTC GCGTGTCCGT CAAAGGTTGA ATAAAGCCAT

                201                                                                                                         300
LT,NX,TZrap-1b  GTCTCAATCT GCAGCTTATT TTACTGAATC CGTCATCGAT TCGGTCTGCA GGTTACACGC CGACGAAATG CAGGCTTGCC GGTCGGCTGT TGACCCCTAC
  HBrap-1b      GTCTCAATCT GCGGCTTATT TTACTGAAGC TGTCATCGAA TCGGTTTGCA AGTTACATGC CGACGAGATG CAGGCTTGCC GGTCCGTCGT TGAGCCGTAC
B.birap-1b      GTCTCAATCT GCGGCTTATT TCAGTACAGT AGTTATGGAT TCTGTCTGCA ACGGACGCAG GAGCGAAATG GAAGCTTGTC GTTCCGCGTT GCAACCCTAC

                301                                                                                                         400
LT,NX,TZrap-1b  ATGCTTCGTT GTGAGAGTGG CAGGTGTTTA CATATGGATC TGGAGAGGAT TAAAGTAACA GGTAGCACGG CCGTGGAACT GCCTAACAAA TTCCAACTGG
  HBrap-1b      ATGCTTCGTT GTGAGGGTGG CAAGTGTTTG CATCTTGATC TGGAGAGGGT TGAAGTAACG GGTGGGACGG GCATGGAACT GCCTAATAAA TTTCAACTAG
B.birap-1b      ATGGCTCGTT GCGAGGATGG CGACTGTTTG CATATCGACA TGGAGAGTAT TGAGGTATCA GGAAGCACTA GTGTGCCGCT ACCAAATAAA TTTCAACTGG

                401                                                                                                         500
LT,NX,TZrap-1b  ATTGCGCCTT CGAGTTGTAC AAGAGGAGCC CCGCATTTCC ATTACGGGGT TTGCGTGGGA CCATAGCCCG ATTCCAGAAG GGGGGGCGTT ACGCTGCCTA
  HBrap-1b      ATTGCGCCTT CGAGTTGTAC AAGAGGAGCC CTGCATTTCC TTTAAAGGCT TTGCGTGGGA TAATAGCCCG ATTCCAGAAG GGAGGGCGTT ACACTGCGTA
B.birap-1b      ATTGTGCCTT CGAGTTGTAC AAAATGAGCT CTGCATTTCC ATTCAGGGCT TTCCGCGGTA TCGTAGCTCG ATTCCAGAAA GGAGGACGTT ACACTGCATA

                501                                                                                                         600
LT,NX,TZrap-1b  TCGCGACTTC ATTCTCCAAC TTATGCTAGC TAATTGCGTG GAAGATAGAA AGTTAGGCGA TTTTGAGGCG TTCATGAGGA AATACCTTTT TCTGGCCGCT
  HBrap-1b      TCGCGACTTC GTTCTTCAAC TTATGTTGGC CAACTCCGTG GGAGATGGAA AATCAAGCGA TTTTGAGTCG TTCATGAGGA AATACCTTTT TTTGGCAGCT
B.birap-1b      CCGCGACTTT ATATTGCATC TTATGTTGGC AAAT---GTG GAAGAAGGGA CGTCAAGCGA TTTTGAGACG TTCATGAAGA AATATCTTTT CATGGCGGCT

                601                                                                                                         700
LT,NX,TZrap-1b  ATCCAATACA AATCATATTT GATGCTGGAC AACTGGAAGG CGAAGCTTCA GAATATGATA GAATTATCGC AATTTGTTTT TGTAGATAAA ATTGAACGGT
  HBrap-1b      ATCCAATACA AAACATATTT GATGCTAGAT AACTGGAAGG CGAAGCTTCA GAATATGATC GAAATGTCGC AATTATTTTT TGTAGATAAA ATTGAACGGT
B.birap-1b      GTCCAATACA AGACTTATTT GATGTTGGAA AACTGGAAAG TGAAACTCCA GAATATGACT GACTTATCGC ATTTTATATT TGTGGACAAA ATTGAACGAT
```

```
                701                                                                                                    800
LT,NX,TZrap-1b  CCTTGGTTCA AATAATTGCA AAAATTGCAA CGAAGTCCAT TTTAGGGACG CCAGCGAATG CGTTTGATCC TTACGCAGAA TGTTTCCCTG CTTACCTCAA
  HBrap-1b      CCTTGGTTCA AATAATTGCG AAAATTGCAC CCAAGTCCAT TTTAGGAACT CCAGCGAACG CTTTTGAGCC TTACGCAAAA AGTTTCTCAG CCTACCTAAA
B.birap-1b      CTCTGGTTCA AATAATCGTG CAGACTGCAC CCAAGTCCGC TATAGTAACG CCTATAAGTG CGATTGAACC ATACGTGAAG AGCTTCGCTG CTTACATAAA

                801                                                                                                    900
LT,NX,TZrap-1b  TTCCTTATGT GTAGGGCATG AGAAATTTGC GCATTTATAT TCGTCGGTTG TCTTCATTGC TCTGAGATCG TCCATCAGCA ACACCCGCCA CTTGGAGTTA
  HBrap-1b      TTCTTTATGG GTCGGCCATG AGAAATTTGC GCATTTGTAT TCGTCGGTTG TGTTCATTGC TCTTCGATCG TCCATCAGTA ACACTCGCCA CTTGGAGTTA
B.birap-1b      TTCGTTCCGA CTCAATCATG AGCGATTTGC GGAATTATAT TCTTCAGTAG TCTTCATTGC GCTTAGGTCG TCAATCAGCA ACACTCGCCA CAAGGAATTA

                901                                                                                                   1000
LT,NX,TZrap-1b  AGCGCAAGGA ATCAAAAGCT CGAACCTTGG AAGAAGTTTT ACAATAGATA TGTTACGCCA GTTTTTGGCC TATTCAAGGG TATGTATGGC TACATGTCCA
  HBrap-1b      AGCGCAAGGA ATCAAAAGCT CGTACCTTGG AAGAACTTTT ACAATAAATT TGTTACGCCA ATTTTTGGCG TATTCAAGGG TATGTATAGC TACGTGTCCA
B.birap-1b      AGAGCAAGGG ATCACAAGCT GGAGCCCTGG AAAAACTTTT ACCGGAAATA CGTGATGCCG GTATTTGGCG CATTTAAGGG CATGTACAGC TACGTGTCAG

                1001                                                                                                  1100
LT,NX,TZrap-1b  GCAGGGTTAT GAAATCGCTT GGGAGGGGCA AACCTCATGC ACCAGCTAAT GGAGAGTCGA GATCAGTCAT CAGACATCGA AATATAGCTG ATATACCGTT
  HBrap-1b      GCAAGGTTAG GAAATCGTTT GGGAGGGGCA AAGGTCATGC ATCAGGTACT GGAGAGTTGG GAGCGGGCGA CAGACATCCA AATAGAGCTG CTATACCGTT
B.birap-1b      GTAAGGTCGT GAACTCGTTT GGCGGG---- ---------- -TCAGCTAAG AGAAAGTTGC AAACAGAGGA C-GCCGCCTA G

                1101                                                                                                  1200
LT,NX,TZrap-1b  AGACTCAGGT GTGTCGTCTT CTGGTGGTAA ATCCCACGGA CCGCATGGTC AGCCTATGCA ACCCGTCAAA GAAAAATCGA AGCCTAGTCG CCAAGTTAGG
  HBrap-1b      AGAATCTGGA GTTTCATCTT CTGGAACTGA ATCGCACGGA GGGCAGAGTC AGCCTACGCA ACCCGTCGAA GGAAAAGGGA AGCCTATTCG CCAAGTTAGA
B.birap-1b

                1201
LT,NX,TZrap-1b  GGATAG
  HBrap-1b      GGATAG
B.birap-1b
```

Supplementary Fig. 8 Alignment of *rap-1b* gene sequences of *B. motasi*-like and *B. bigemina*. LT,NX,TZrap-1b: *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu *rap-1b* gene (represented sequence from LT isolate, GenBank accession number: KC953702); HBrap-1b: *Babesia* sp. Hebei *rap-1b* gene (GenBank accession number: KJ205336). B. birap-1b: *B. bigemina rap-1b* gene (GenBank accession number: AF021247).

```
             1                                                                                                       100
  LTrap-1c   ATGGTTTGCC ACAGCTTTGT CATTATCTCA TTATGCGCGC TGACCGTGAC CTCTGTGTCA CCCATGCGCC ACACGCAGTA TGCTGCTCTC ATGGCGGCAG
NX,TZrap-1c  ATGGTTTGCC ACAGCTTTGT CATTATCTCA TTATGCGCGC TGACCGTGAC CTCTGTGTCA CCCATGCGCC ACACGCAGTA TGCTGCTCTC ATGGCGGCAG
B.birap-1c   ATGATTCACT ACGCTTGCCT CATTATCGCA CTTTGCGCCT TGTCGCTGCG CTCTGTAGCC GCAGTGCGCT ACGGCCAGCA TGCCATGATC ATGGCCCCTG
  HBrap-1c   ATGGTTGGCT ACAGCTTTGT CATTGTCGCA TTATGCGCAC TGACAATGAC CTCTGTGTCA TCTATGCACC ACAATAAGCA TGCTGTTCTC ATGAG---AG

            101                                                                                                      200
  LTrap-1c   AGGACATTCT ACCGCCCGCA GCGAAGAGCG TGGATGTGGT GTATGATGGC GAACATAAGG AAATACTCGA CGCAGGTGAA AACATTGAAA GGGCTATGAG
NX,TZrap-1c  AGGACATTCT ACCGCCCGCA GCGAAGAGCG TGGATGTGGT GTATGATGGC GAACATAAGG AAATACTCGA CGCAGGTGAA AACATTGAAA GGGCTATGAG
B.birap-1c   ACCACACTGT CGAGCCTGCG GTTGATCCCG T--------- ---CGACGAC GACACCAAGC AGCTTCTCGA AGACAGCGAA CAGATTGAGA AGGCCATGCA
  HBrap-1c   AACAGACACT ATCGCCTGAA AAGCAGAAAG TAGATGTGGT ---TGATGAC GAGACGAAGG AACTTGTTGA AGCAGGTAGT AACATTGAAT TGGCAATGAA

             201                                                                                                     300
  LTrap-1c   GGAACAGGTC AATAGCTTAA CATACCAAGC TGTCACCGAT ATGTGCCAAG ACAGCAAGGA CGTTGCGAAA TGCAAATCGC AAATTTCTAT ATATGCGGTA
NX,TZrap-1c  GGAACAGGTC AATAGCTTAA CATACCAAGC TGTCACCGAT ATGTGCCAAG ACAGCAAGGA CGTTGCGAAA TGCAAATCGC AAATTTCTAT ATATGCGGTA
B.birap-1c   GGAGGAAATC GGGCTGATCA ACGATGACTC CATTGCCGAA ATGTGCCTCG GCAGCAAGGA CGAGCACCAT TGCGCATCGC AAATCGCTGC CTACGTTGCG
  HBrap-1c   GGAACAGGTC AATAGCCTAA CATACGAAGC TGTCGCTGAA ATGTGCCAAG GCAGCAAAAA CGTGGCGACA TGCAAATCGC AAATATCTAC ATATGCTGCA

             301                                                                                                     400
  LTrap-1c   CGTTGCAAGC AAGGTGATTG CCTCACTCTC GATCGTGTTG GCTACCCGGA AAACAAGGCA TACCAGCAAC TCGTTTTACC TGACCCATAC CAGCTGCATG
NX,TZrap-1c  CGTTGCAAGC AAGGTGATTG CCTCACTCTC GATCGTGTTG GCTACCCGGA AAACAAGGCA TACCAGCAAC TCGTTTTACC TGACCCATAC CAGCTGCATG
B.birap-1c   CGTTGCAAGG AGGGCAACTG CCTCACCATC GATGCCGTGG GAAAACCGCA AAACAAGGCC TACGGACAGC TCGTCCTGCC TGACCCCTAC CAGCTTCACG
  HBrap-1c   CGTTGCAGAA AAGGTGATTG CCTCACCATC GATCGTGTTG GCTCCCCGGA AAACAAGGCT TTCCAGCAAC TCGTTTTGCC TGACCCATAT CAGCTGCATG

             401                                                                                                     500
  LTrap-1c   CGGCATTCCT GCTGTTCAAA AACAGCCG-C GCGAATGCCA GTAGGGACTG GCTGAACAGA TTCTGGTTGA GATTCAACAG AGGCGGTCGT TACGCTGCGT
NX,TZrap-1c  CGGCATTCCT GCTGTTCAAA AACAGCCG-C GCGAATGCCA GTAGGGACTG GCTGAACAGA TTCTGGTTGA GATTCAACAG AGGCGGTCGT TACGCTGCGT
B.birap-1c   CCGCGTTCCT GTTGTTCAAG AACTGCCGGC GC-AACGAGA GCAGGCACTG GATGGACAGG TTCTGGATGC GTTTCAAGAG GGGAGGGCGT TACGCTGCTT
  HBrap-1c   CTGCCTTCCA ACTGTTCAAG GGCTGTCG-C GCGAATTCCA GTAGGAACTG GCTGAACAGA TTTTGGTTGA GATTCAACCG AGGTGGTCGT TACGCTGCCT

             501                                                                                                     600
  LTrap-1c   ACCATAGCTT CAGCGTCAAC CTTCTAAGGC GCAATCTATT CCCGGACTCT GAGGCAGGCG AGCTGGAGAA CTTCATCATC AAGTACCTTT ACACGACTGC
NX,TZrap-1c  ACCATAGCTT CAGCGTCAAC CTTCTAAGGC GCAATCTATT CCCGGACTCT GAGGCAGGCG AGCTGGAGAA CTTCATCATC AAGTACCTTT ACACGACTGC
B.birap-1c   ACTACAGCTT CAGCCTCAAC CTGTTGAGGC GCAACCTGTT CCTCGGCGAC GACAAAAATG CCCTGCATGG CTTCGTGCAG AAGTACTTCT ACATGACCGC
  HBrap-1c   ACCATACCTT CACTCTCAAC GTTCTGAAGC GCAATTTATT CCCGGATTCT GATGCTGGCG AGCTGGTGAA CTTCATCATC AAGTACCTTT ACACGACTGC

             601                                                                                                     700
  LTrap-1c   CATATACTAC AAAACTTACC TGTCTCTGGA TGCAACTAGC GCAAAAATTA TCAACAAAAT CGCTTTTTCA AGACACCTCT TCGGCATCAA GATTAGAAGA
NX,TZrap-1c  CATATACTAC AAAACTTACC TGTCTCTGGA TGCAACTAGC GCAAAAATTA TCAACAAAAT CGCTTTTTCA AGACACCTCT TCGGCATCAA GATTAGAAGA
B.birap-1c   CATATACTAC AAGACATACT TGTCGCTGGA TGCCATCAAC GCCAAAATAT TTAACAAAAT CGCCTTGGCG AAGCACATTC TGGGACCTAA GATCAAAAGG
  HBrap-1c   CATATACTAC AAGACTTACT TGTCGTTGGA TGCAATAAAC GCTAAAATCC TTAACAAAGT AGCTTTTGCC AGACACCTCT TCGGCATGAA GATTAAAAGA
```

```
             701                                                                                                      800
  LTrap-1c   GCTTTGGGCG ATATTGTGAA ATCAAACATG CCAAGCAAAC TCAACGAGCA GGAAGCCGGT GTGATCCGTC CTTTGACATT TGGTTACAGG CATTACATGG
NX,TZrap-1c  GCTTTGGGCG ATATTGTGAA ATCAAACATG CCAAGCAAAC TCAACGAGCA GGAAGCCGGT GTGATCCGTC CTTTGACATT TGGTTACAGG CATTACATGG
B.birap-1c   GCGTTGAGGA AGATCGTCGA GGCCAACAAG CCGAGCGCGC TCCAAGCGAA CGACGTCAAA GCCATACGCC CCCTGGCCTA CGGTTACAGG CAGTACATGG
  HBrap-1c   GCTTTGGCCA ACATGGTGCA AGAAAACATG CCAAGCAAAC TCAGCAAGCA GGATGCTGAT GTGATCCGTC CTTTGACATT CGGTTACAGG CAGTACATGG

             801                                                                                                      900
  LTrap-1c   CTACCCAAAT TCCGCGTCTT CCTTTCTTTG CCCATCGATT CTCTTCCATG GTTCTCAAGA CACTTATCGA CAACCTCACT GGAGCTAATA AGCTGCCGTG
NX,TZrap-1c  CTACCCAAAT TCCGCGTCTT CCTTTCTTTG CCCATCGATT CTCTTCCATG GTTCTCAAGA CACTTATCGA CAACCTCACT GGAGCTAATA AGCTGCCGTG
B.birap-1c   CCAGCCAAAT CCCATCGCTT CCGTTCTTCG CCTACCGTTT CTCCTCGATG GTCGTCACCA CTCTCGTGGA CAACCTCACC GGCGTTAAGC AGCAGCCGTG
  HBrap-1c   C

             901                                                                                                     1000
  LTrap-1c   GTATAAACGG TGGCTTGGTA AAGTGAAAAA CTTCTTTACC GGTAAAGGGT CCGACGAC-- -GAATATGTT GTTGAGGACG ACTTTGATAG CGGTGACGCA
NX,TZrap-1c  GTATAAACGG TGGCTTGGTA AAGTGAAAAA CTTCTTTACC GGTAAAGGGT CCGACGAC-- -GAATATGTT GTTGAGGACG ACTTTGATAG CGGTGACGCA
B.birap-1c   GTACAAGCGT TGGTTTGGAA AGGTGAAGAA CCTCTTCACC GGTAAGCAAC CCAGCGAAAA GGCCTACGAA ATTGACGAAC CCATCGCCAC CGAGGAGGAT
  HBrap-1c

             1001                                                                                                    1100
  LTrap-1c   AAGCTTAGCG CTGACAATAA GTCGAAGCTT CAGAGAATGA AAGATTCCAT GAAGAGACTA AGAAGGAAGC TTAGCATTTC GGAAAACGGT GCTAAGGACA
NX,TZrap-1c  AAGCTTAGCG CTGACAATAA GTCGAAGCTT CAGAGAATGA AAGATTCCAT GAAGAGACTA AGAAGGAAGC TTAGCATTTC GGAAAACGGT GCTAAGGACA
B.birap-1c   A--CCGAGC- CTGTCGAGGA AAATAAG--T CAGTTTTTGG CAAGGTCAAG GAAAAG-CTT GGCAACATCC GCTTCAACAC GGGCATTTTC CGAAAGGG--
  HBrap-1c

             1101                                                                                                    1200
  LTrap-1c   ATGAAGCTTC ATCAAGAAAA ACGCTC---T CAGAAGAAGA AATTATGAAT AACCTATCCA CAGCAGATGT GCTCGTTCAA CCGATTCTGG ATGCCATAGG
NX,TZrap-1c  ATGAAGCTTC ATCAAGAAAA ACGCTC---T CAGAAGAAGA AATTATGAAT AACCTATCCA CAGCAGATGT GCTCGTTCAA CCGATTCTGG ATGCCATAGG
B.birap-1c   -TGAGGCAAA GACTCGTCAC TCGCACCTTT CCGAAGAGGA CATTATGGGC AGCTTGTCGT CGGCTGATGC GCTGCTTGAG CCGGTGCTCG ATGTCATGGA
  HBrap-1c

             1201                                                                                                    1300
  LTrap-1c   CAATGGTAAA GGAGTGGGAA ATGACGAAGT AACTGATGCA GACAGCGTCA CAGGACAGGT AGATAAGGAG GCTATCAGTA TTGCAGGAGA GAACAAAGAA
NX,TZrap-1c  CAATGGTAAA GGAGTGGGAA ATGACGAAGT AACTGATGCA GACAGCGTCA CAGGACAGGT AGATAAGGAG GCTATCAGTA TTGCAGGAGA GAACAAAGAA
B.birap-1c   GAAAG---AG GGCGAAGCTC AGAACGAGGA GGCAGGTGAA CCCGAAGTGG CCG---CGGC TCCCAAGGCG CCTGAGAGTG AAAATGGCGA ACTCGCAGAC
  HBrap-1c

             1301                                                                                                    1400
  LTrap-1c   CCAACCATCG TTGAACAAAA GGAGGCTATC AACATTGCAG GAGAGAACAA CGAGCCAAGC GTCGTTGAAC A-GGAGGCTA TCAACATTGC AGGAGAGAAC
NX,TZrap-1c  CCAACCATCG TTGAACAAAA GGAGGCTATC AACATTGCAG GAGAGAACAA CGAGCCAAGC GTCGTTGAAC A-GGAGGCTA TCAACATTGC AGGAGAGAAC
B.birap-1c   CAAGTAGATG GTGCTTCAAC TGTGGC-CTC AACG-----G AAATGGACGA AGAGCAGTCT CTCGAAGCTC CTAAAGGCAC GCAGGATTTG ATGCACGAA-
  HBrap-1c
```

```
             1401                                                                                                  1500
  LTrap-1c   AACGAGCCAA GCGTCGTTGA ACAGGAGGCT ATCAACATTG CAGGAGAGAA CAACGAACCA ACCATCGTGG AACAAAAGGA GGCTATCAAC ATTGCAAGAG
NX,TZrap-1c  AACGAGCCAA GCGTCGTTGA ACAGGAGGCT ATCAACATTG CAGGAGAGAA CAACGAGCCA AGCGTCGTTG AACGGAAGGA GGCTATTAAC ATTGCAAGAG
B.birap-1c   GATGAG--GA GCAG----GA ACTCGAATCT GACGAATTGG CTGCCAAGAG GAA-GCGTAA AGCTTCAA-G CAC---AGGC TACAAAAAGC TCTTCAAGAA
  HBrap-1c

             1501                                                                                                  1600
  LTrap-1c   AAA-----AC AACGAACCAC CTATCGTTGG TGAGAGTGTT ACACCGAATG TTCCTACGTC AGACGTTACT CCGTCACCTA AAGCTGGTAC TAATGAATCA
NX,TZrap-1c  AAA-----AC AACGAGCCAC CTATCGTTGG TGAGAGTGTT ACACCGAATG TTCCTACGTC AGACGTTACT CCGTCAGCTA AAGCTGATAC TAATGAATCA
B.birap-1c   CATGTTGGAC ACCGGTGCAA TTAAGGATGC CGCTAGC--C GCATCAGGT- TCCCGAAGAA GTACTTCAAA CGATCGTCGT AA
  HBrap-1c

             1601                                                                                                  1700
  LTrap-1c   GTACCCGAGG TAACGGCCAA GGATGTTCAC TCTAAAGATG GGGAGAATGG GTTCGAAGAT GCGGTCGAAG ACCTGTCGGA GCTCCCTCCT GATGACACTA
NX,TZrap-1c  GTACCCGAGG TAACGGCCAA GGATGTGCAC TCTAAAGATG GGGAGAATGG GTTCGAAGAT GCGGTCGAAG ACCTGTCGGA GCTCCCTCCT GATGACACTA
B.birap-1c
  HBrap-1c

             1701                                                           1762
  LTrap-1c   AATCACGCTT GGCTAATTAC GCATCGAAGT TCAAGCGTTT TATACAATCC AAATTATCGT AA
NX,TZrap-1c  AATCACGCTT GGCTAATTAC GCATCGAAGT TCAAGCGTTT TATACAATCC AAATTATCGT AA
B.birap-1c
  HBrap-1c
```

Supplementary Fig. 9 Alignment of *rap-1c* gene sequences of *B. motasi*-like and *B. bigemina*. LTrap-1c: *Babesia* sp. BQ1 (Lintan) *rap-1b* gene (GenBank accession number: KF039724); NX,TZrap-1c: *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu *rap-1c* gene (represented sequence from NX isolate, GenBank accession number: KJ205327); HBrap-1c: *Babesia* sp. Hebei partial *rap-1c* gene (GenBank accession number: KJ205338). B. birap-1c: *B. bigemina rap-1c* gene (GenBank accession number: AF026272).

Article N° 3: Expression of *rap-1* genes in *Babesia* sp. BQ1 (Lintan) (*B. motasi*-like phylogenetic group), a pathogen of sheep in China

Article N° 3: **Expression of *rap-1* genes in *Babesia* sp. BQ1 (Lintan) (*B. motasi*-like phylogenetic group), a pathogen of sheep in China**

Qingli Niu[a,b], Claire Bonsergent[a,b], Guiquan Guan[c], Laurence Malandrin[a,b], Emmanuelle Moreau[a,b*]

a.INRA, UMR1300 Biology, Epidemiology and Risk Analysis in Animal Health, CS 40706, F-44307 Nantes, France.
b.LUNAM Université, Oniris, UMR BioEpAR, F-44307 Nantes, France.
c.State Key Laboratory of Veterinary Etiological Biology, LVRI, Lanzhou, China.

* Corresponding author: Oniris, UMR INRA / Oniris 1300, Biology, Epidemiology and Risk Analysis in animal health, Route de Gachet, La Chantrerie, BP 40706, F-44307 Nantes, France. Phone: +33 (0) 2 40 68 78 57 Fax: +33 (0) 2 40 68 77 51.
E-mail address: Emmanuelle.moreau@oniris-nantes.fr

Additional data to be acquired before publication. English has not been corrected

Abstract

Babesia sp. BQ1 (Lintan) is one of the parasites isolated from infected sheep in China that belongs to the *B. motasi*-like phylogenetic group. Within this group, the rhoptry-associated-protein 1 (*rap-1*) locus has a complex organization of 12 genes from three main types: 6 *rap-1a* variants intercalated with 5 identical copies of *rap-1b* and a 3' ending *rap-1c* unique gene. In the present study, transcription analysis performed by standard RT-PCR demonstrated that the three different *rap-1* gene types and the four *rap-1a* variants were transcribed by the parasite cultivated in vitro. Peptides, specific for each *rap-1* type gene, were selected in putative linear B-epitopes and used to raise polyclonal rabbit antisera. Using these sera, the expression pattern of RAP-1 proteins was found to be the same in parasites cultivated in vitro or collected from acute infection; only RAP-1a61 was detectable in merozoite extracts. However, ELISA performed with these peptides and sera from infected sheep demonstrated the presence of antibodies against each of the peptides, suggesting that the corresponding proteins were produced in vivo in sufficient amounts to stimulate a host immune response. The kinetics of response was different according to the protein, with an early and transient response to RAP-1a, and a late but lasting response to RAP-1b and RAP-1c. The role of such an apparent sequential protein expression as an immune evasion strategy developed by these parasites is being discussed.

Keywords : *Babesia*, *rap-1*, protein expression, immune evasion

Introduction

Ovine babesiosis is an important tick-borne disease of sheep and goats, which is caused by the multiplication of intraerythrocytic protozoa of the genus *Babesia*, mainly by *Babesia ovis*, *B. motasi* and *B. crassa* (Uilenberg, 2006). Several other parasites responsible of ovine babesiosis have been recently reported in different regions of China (Guan *et al*., 2001, 2002; Bai *et al*., 2002; Liu *et al*., 2007). Most of them, *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu, *Babesia* sp. Hebei, *Babesia* sp. Madang and *Babesia* sp. Liaoning, belong to the same phylogenetic group defined on the 18S rDNA gene sequences. Due to their close phylogenetic position and their similar morphological features with the european *B. motasi* (Liu *et al*., 2007; Niu *et al*., 2009), this group is referred to as *B. motasi*-like. The ranking of these parasites as one or many species, different or not from *B. motasi* is still under debate due to several conflicting phylogenetic reconstruction within this group depending on the phylogenetic marker used (Niu *et al*., 2009; Gou *et al*., 2013; Tian *et al*., 2013a, b and c; Niu *et al*., 2014a in press). On the other hand, *Babesia* sp. Xinjiang together with parasites described from wild ungulates in South Africa (Oosthuizen *et al*., 2009) and *B. pecorum* from red deer in Spain (Jouglin *et al*., 2014 in press) forms a sister clade to the *B. ovis*, *B. bovis*, *B. orientalis* and *B. occultans* clade (Lack *et al*., 2012).

The clinical symptoms of ovine babesiosis are usually characterized by fever, depression, hemolytic anaemia, icterus, hemoglobinuria and even death in severe cases of sheep parasitized by *Babesia* in the field (Bai *et al*., 2002; Guan *et al*., 2012c). The virulence of *Babesia* sp. BQ1 (Lintan) and (Ningxian) is moderate and severe respectively, higher than that of *Babesia* sp. Xinjiang which often led to inapparent infection (Liu *et al*., 2007; Guan *et al*., 2010a). The cases of ovine babesiosis were reported in many provinces, mainly in the northwest of China, and recent sero-epidemiological surveys confirmed the presence of parasites from the two afore-mentionned clades in 22 provinces of China (Guan *et al*., 2012a; Wang *et al*., 2013).

At present, two conventional chemotherapeutic drugs (imidocarb and diminazene aceturate) and several new anti-*Babesia* drugs (e.g. Triclosan; Nerolidol; Artesunate; Epoxomicin; Gossypol and Atovaquone) are used to control and treat babesiosis (Mosqueda *et al*., 2012). But the drawbacks of drug treatment, such as drug-resistance and drug residues in animals' products, have prompted the search for more effective ways

against babesiosis (de Waal and Combrink, 2006, Mosqueda *et al.*, 2012). Therefore, the efforts are now mainly being directed towards the development of recombinant vaccines against a number of *Babesia* species, targeting antigens to either block the parasite transmission between ticks and vertebral hosts, or block the parasite multiplication within the vertebrate host. However, a combination of both is thought practical and ideal control strategies of babesiosis (Gohil *et al.*, 2013).

During the asexual multiplication of *Babesia* in the vertebrate host, the presence of free parasites in the blood stream, the merozoites, represents a brief period during the parasite life cycle (Sun *et al.*, 2011), during which the parasite can be reached by the host immune system, unprotected by the erythrocyte membrane. Therefore, development of vaccines targets either the merozoite surface proteins or the proteins secreted from the apical complex organelles (micronemes, rhoptries, and dense granules) to achieve the erythrocyte invasion. Some of these proteins, including Rhoptry-Associated-Protein-1 (RAP-1), Apical Membrane Antigen-1 (AMA-1), and the Spherical Body Proteins (SBP-1, 2, 3), their functions and their implications for vaccine design against *Babesia* infections have been described mainly in *B. bigemina* and *B. bovis* (Yokoyama *et al.*, 2006). RAP-1 is a protein of about 40-60kDa localized in the rhoptries. The presence of *rap-1* genes has been demonstrated in all *Babesia* species examined so far (*B. bovis, B. bigemina, B. divergens, B. canis, B. caballi, B. ovis, B. gibsoni, B. motasi*-like parasites, and *Babesia* sp. Xinjiang) (Suarez *et al.*, 1991; Dalrymple *et al.*, 1993; Skuce *et al.*, 1996; Zhou *et al.*, 2007; Bhoora *et al.* 2010; Niu *et al.*, 2014b, in press). RAP-1 is an immunogenic protein detected in all the asexual growth stages (Yokoyama *et al.*, 2002) and expressed also in sporozoite stage (Mosqueda *et al.*, 2002). The role of RAP-1 in the erythrocyte invasion process has been studied and RAP-1 specific antibodies inhibit *Babesia* growth in vitro, parasite adhesion to erythrocyte or invasion (Figueroa and Buening, 1991; Mosqueda *et al.*, 2002; Yokoyama *et al.*, 2002). As immunization with RAP-1 can also induce protection upon challenge, RAP-1 is considered as a potential candidate for the development of recombinant vaccines against babesiosis (Brown *et al.*, 1998, 1999; Brown and Palmer, 1999; Wright *et al.*, 1992; Machado *et al.*, 1999).

In *Babesia*, *rap* is described as a multigene family with tandemly arranged genes and conserved protein features (Suarez *et al.*, 1998). In the sheep parasite *Babesia* sp. BQ1 (Lintan), the *rap-1* locus is complex, as in *B. bigemina* (Suarez *et al.*, 2003), with 12 genes of three different types (*rap-1a*, *rap-1b* and *rap-1c*), 5 copies of identical *rap-1b* interspacing 6 different *rap-1a* copies (Niu *et al.*, 2013). To determine the potential of

these different genes as vaccine candidate, we studied the expression of these different genes in vitro as well as in vivo, and the presence of antibodies directed against these proteins during the course of infection.

Materials and methods

Origin and in vitro cultivation of *Babesia* sp. BQ1 (Lintan)

Babesia sp. BQ1 (Lintan) was originally isolated from a sheep infested with adult *Haemaphysalis qinghaiensis* ticks from Lintan, Gansu Province in China (Guan *et al*., 2002). A clonal line of *Babesia* sp. BQ1 (Lintan) from sheep blood samples was isolated using in vitro continuous cultures as described previously (Malandrin *et al*., 2009; Guan *et al*., 2010b) and stabilates are preserved in liquid nitrogen at the Lanzhou Veterinary Research Institute (LVRI), China and at Oniris, France.

The infected blood samples of *Babesia* sp. BQ1 (Lintan) cryopreserved in liquid nitrogen were thawed in a water bath at 37°C, washed with RPMI1640 (Lonza, Belgium) plus FBS 20%. The cultures were performed essentially as described previously (Malandrin *et al*., 2004; Guan *et al*, 2010b), in 24-well culture plates and fed every 2-3 days by removing 1.5 ml of the supernatant medium overlying the erythrocyte layer without disturbing the RBC layer and replacing it with an equal volume of fresh medium (RPMI1640 + FBS 20%). A subculture was then prepared and expanded when parasitemia reached 2% by transferring the entire contents of the culture well of infected RBC into 25 cm^2 flasks in 10 ml of fresh medium.

Infection of sheep with *Babesia* sp. BQ1 (Lintan)

Sheep used in this experiment were infected with *Babesia* sp. BQ1 (Lintan) as described by Guan *et al* (2010c). All sheep were free of *Babesia* spp. before infection and kept in a sheep-fold with tick-free food and water available ad libitum. Briefly, 2 Tan mutton sheep (N° 3216 and 2007) were experimentally infected with the original strain isolated in China and one “Vendean” sheep (3446) with 10^9 red blood cells infected with a monoclonal line of *Babesia* sp. BQ1 by in vitro culture. Sera were collected before infection (DPI 0) and every week then every month after infection (Guan *et al*., 2010c).

Total RNA extraction, cDNA synthesis and RT-PCR analysis

When the cultures reached a parasitemia of about 10%, the iRBC were harvested for total RNA extraction using the standard TRIzol reagent protocol (Life Technologies) according to the manufacturer's instructions. Total RNA was then treated with DNAse I (Invitrogen). For the reverse transcription, cDNA was then synthesized with oligo $(dT)_{20}$ primer using Superscript III First Strand Synthesis System for RT-PCR kit (Invitrogen) following the manufacturer's recommendation. Briefly, to prepare RNA/ primer mixture in a 0.2 ml tube: 7µl of treated RNA, 1µl of 50µM oligo (dT) 20, 1µl of 10 mM dNTPs and with 1µl DEPC water were mixed to the final volume of 10µl. The mixture was incubated at 65˚C for 5 min. The cDNA Synthesis mixture was then prepared with 2µl of 10×RT buffer, 4 µl of 25 mM MgCl2, 2 µl of 0.1M DTT, 1 µl of RNAse OUT and 1µl of Superseript III in a 0.2 ml tube, and added 10 ul of cDNA Synthesis mixture to each RNA/primer mixture to final volume of 20 µl. The reactions were performed in a thermocycler under the following incubation conditions: 25˚C for 10 min, 50˚C for 50 min, then 1µl of RNAseH/tube was added and incubated 37 for 20min. Two microlitres of the cDNA were used in PCRs and the remainder was stored at -20°C.

PCR amplification of intergenic region (IG2) was be performed to check the absence of DNA contamination in the RNA extracts. The pair of primer used (Table 1) could amplify a 1.7 kb fragment of the intergenic region (IG2) described by Niu et al., 2013 on genomic DNA, but no amplification on cDNA should occur. Each *rap-1a*, *rap-1b* and *rap-1c* genes were then amplified by PCR using cDNA (Table 1). Three different *rap-1a* genes (*rap-1a61*, *rap-1a67-1* and *rap-1a67-2*) were amplified using specific primers. Genomic DNA was used as PCR and amplicon size control.

The plasmid DNA inserts (full-length *rap-1a* and *rap-1b*) or PCR product (*rap-1c*) were sequenced by the "GATC" company using the vector primers.

Table 1 Primer combinations, amplification conditions and product that cloned and sequenced of three *rap-1* gene types for RT-PCR in this study.

Genes	Primer name	Primer sequence 5’ to 3’	DNA polymerase	*Tm*	Cycles	Size of product	Sequencing
IG2 and partial 5’ of rap-1a61	n° 10_11bup2	TCT TAC GGA TTC CGT GCA GAC ATC G	Phusion High Fidelity	60	35	1723 bp	-
	61down	GACTGACGATGAGGTTCTGGT					
rap-1a	*rap-1a*downend	TTA GAG CCA CGA ATC ATC GG	Phusion High Fidelity	60	30	1425 bp	5 clones
	*rap-1a*up	ATG AGA AGC TTC GCG GGT G					
rap-1a67-2	*rap-1a*up		Ex *Taq* Hot Start	58	30	1380 bp	-
	*rap-1a67-2*down	TTA GTT TTC AAC CTC AGG G					
rap-1a61	*rap-1a61*speup	TAC AAG CAC TGA GTG CCG TG	Phusion High Fidelity	61	30	1178 bp	-
	*rap-1a*downend	TTA GAG CCA CGA ATC ATC GG					
rap-1a67-1	*rap-1a*downend		Phusion High Fidelity	59	30	1178 bp	-
	*rap-1a67-1*speup	AAG GAA CAG TGA TTG TCG TC					
rap-1b	*rap-1b*up	ATG GCA CCA GTG TTG AAG CGC	Phusion High Fidelity	66	35	1203 bp	4 clones
	*rap-1b*down	CTA TCC CCT AAC TTG GCG ACT A					
rap-1c	*rap-1c*up	ATG GTT TGC CAC AGC TTT GTC	Ex *Taq* Hot Start	54	40	1749 bp	PCR product
	*rap-1c*down	TTA CGA TAA TTT GGA TTG T					

Synthesis of specific peptides and production of polyclonal antibodies

To produce antibodies against each RAP-1 (RAP-1a61, RAP-1a67, common RAP-1a61/67, RAP-1b and RAP-1c), B-cell epitopes were *in silico* predicted using the program Bepipred Linear Epitope Prediction (http://tools.immuneepitope.org/tools/bcell/iedb_input) (Figure 1A). At positions corresponding to putative B-cell epitopes, five specific peptides for each RAP-1 protein were then designed based on the alignment of amino-acid sequence of putative RAP-1a, b and c ORF (http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::clustalw-multialign using the software ClustalW 2.0.12) (Figure 1B). The synthetic peptides used for immunizations were: RAP-1a61: SGQLSRMMDAV-cys; RAP-1a67-ERQMSRMANII-cys; RAP-1a: LTLPNPYQL-cys; RAP-1b: cys-VTGSTAVELP and RAP-1c: cys-ILPPAAKSVDVVYD (Figure 1B). Peptides and antibodies against each peptide were produced by the company GeneCust Europe (Luxembourg). The synthesized peptides were coupled to the carrier protein (KLH) via either an N-terminal (RAP-1a) cysteine or C-terminal (RAP-1b and c) cysteine added to the peptides sequence (GeneCust). Two rabbits have been immunized with each peptide.

Immunoglobulins G were then purified from each immunized rabbit sera and from a negative rabbit serum (R9133, Sigma) by affinity using HiTrap Protein A HP columns (17-0402-01, GE Healthcare) according to the manufacturer's protocol. IgG were then dialyzed against PBS during 48h and proteins concentrations were determined with the bicinchronic acid assay (BC Assay Protein Quantitation Small Kit, Interchim).

The sera or the purified immunoglobulins were used in the western blot analysis against either merozoite antigens and in ELISA.

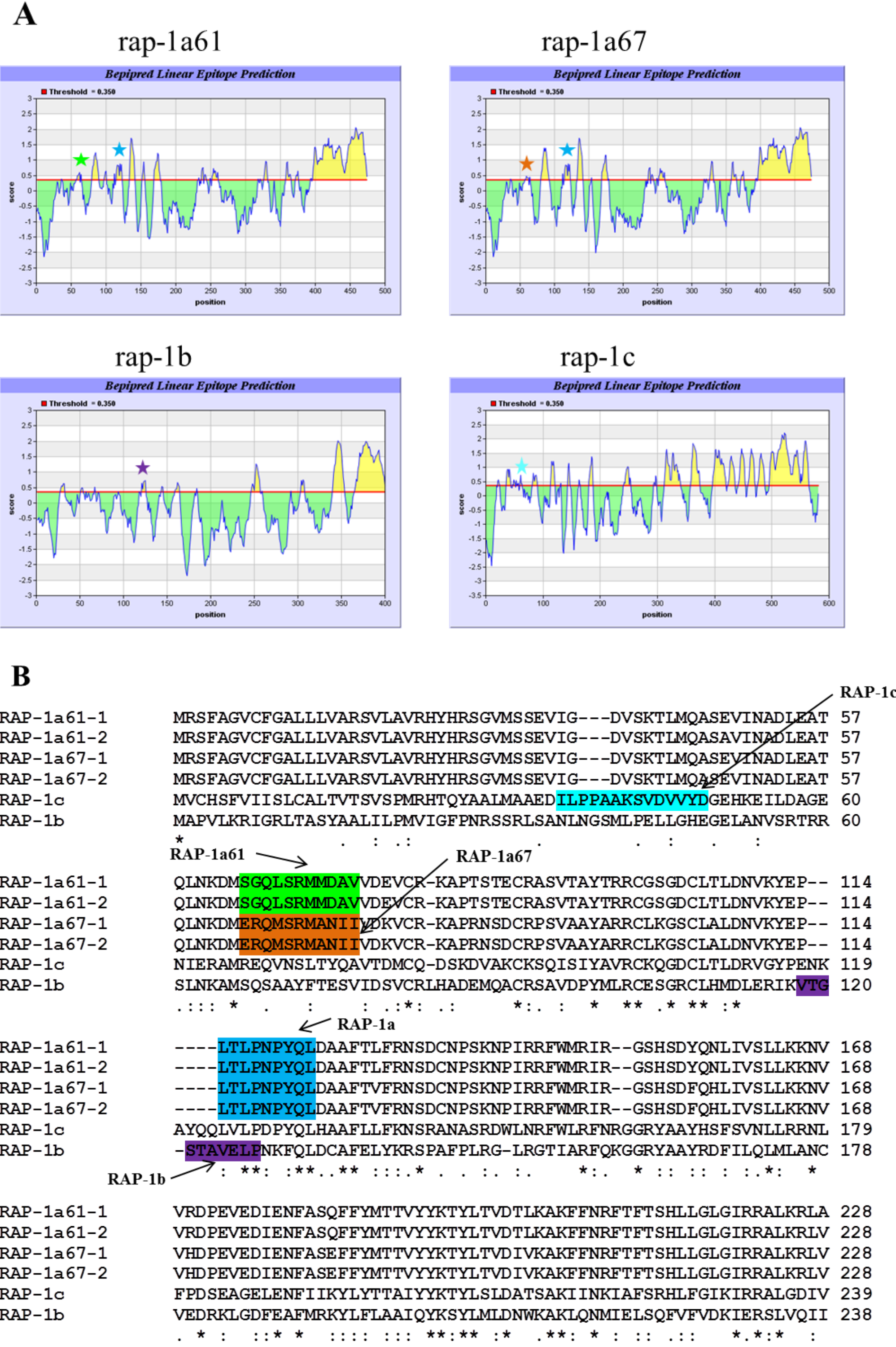


Figure 1. B-cell epitope prediction on the different RAP-1 proteins (**A**): B-cell epitope prediction of four RAP-1 proteins by using method "Bepipred Linear Epitope Prediction". Predicted B epitopes were indicated with yellow. The location of specific peptide for each protein used in this study was indicated with asterisk, a blue asterisk: common peptide of RAP-1a; red asterisk: RAP-1a61; pink asterisk: RAP-1a67; green asterisk: RAP-1b; dark yellow asterisk: RAP-1c. **(B):** Partial alignment of RAP-1a, b and c for peptide design, showing the sequence (different colors, corresponding to A) of the specific peptide used in this study.

PCR, cloning and recombinant protein expression

The full or almost full length *rap-1a* (*rap-1a61* and *67) rap-1b* and *rap-1c* genes were amplified using primers modified to include *XhoI* and *XbaI* (New England BioLabs) enzyme restriction sites (underlined) in their 5'end (Table 2).

The quality of the amplification products was controlled by electrophoresis in 1% agarose gels and visualized under UV light. The crude PCR products were then cloned into the pGEM-T Easy vector according to the manufacturer's protocol (Promega). Plasmid DNA with expected size for each *rap-1 gene* type was extracted (NucleoSpin plasmid extraction, Macherey-Nagel) and the inserts were sequenced using vector primers. Representative plasmids were selected and digested with *Xho I* and *Xba I* restriction enzymes, cloned into digested pBAD- *Myc*-His B expression vector (Invitrogen) according to the manufacturer's protocol resulting in the generation of pBAD-*Myc*-His B-RAP-1 as histidine-tagged fusion proteins, which contain the 1369 bp of *rap-1a* and the full-length *rap-1b* and *c* fragments, respectively. Two positive clones with expected size for each *rap-1* type were cultured in LB medium (LB broth, L3522-1KG, Sigma). Plasmids were then extracted and sequenced. The culture was cryopreserved at -80°C for further expression analysis of recombinant proteins.

For the expression of recombinant proteins, pBAD/Myc-His/lacZ (120kDa) was used as positive control, the conserved bacteria with recombinant plasmid of *rap-1* genes was grown in liquid LB at 37°C, 225 rpm, incubated overnight. Five hundred microliters of cultured bacteria were added to 50 ml of fresh LB medium supplemented with carbenicillin (50 μg/ml) in an erlenmeyer flask for 2 h at 37°C, 225rpm. When the optical density (OD) at 590 nm reached 0.5~0.8 (T0), the bacterial culture was split into three 50 ml tubes, with 10 ml per tube. The protein expression was induced by the addition of L-Arabinose at 20% (Acros organics) with final concentrations of either 0.2% or 1%, with a non-induced control. Samples of the 3 cultures were collected every hour for 4 hours after induction. At each time point, OD was measured at 590 nm, 1 ml of culture was transferred into a separate 1.5 ml tube with addition of 10 μl Halt™ protease inhibitors cocktail (Thermo Scientific), and centrifuged at 5,000 g for 5 min at 4°C. The pelleted bacteria were frozen at -20°C until use. Cell pellets were then suspended in 1X SDS-PAGE sample loading buffer (62.5 mM Tris-HCl [pH 6.8], 2% SDS, 5% β-mercaptoethanol, 10% glycerol, 0.02% bromophenol blue) and samples were boiled for 5 min and centrifuged briefly prior to SDS-PAGE. Recombinant protein expression was then analyzed by western blotting.

Table 2 Primer combinations, amplification conditions and product that cloned and sequenced of three *rap-1* gene types for recombinant protein in this study.

Genes	Primer name	Primer sequence 5' to 3'	DNA polymerase	*Tm*	Cycles	Size of product	Expressed vector	Sequencing
rap-1a (*61/ 67*)	rec-*rap-1a*-fw	GTCCTCGAGGATGAGAAGCTTCGCGGGTG	Ex *Taq* Hot Start	64	30	1369 bp	pBAD-*Myc*-His B	2 clones *rap-1a61* type
	rec-*rap-1a*-rev	CCGTCTAGAACCTCAGGGGCAGCTGCCTC						2 clones *rap-1a67* type
rap-1b	rec-*rap-1b*-fw	GTCCTCGAGTATGGCACCAGTGTTGAAGC		60	35	1200 bp		2 clones
	rec-*rap-1b*-rev	GCGTCTAGAAGTCCCCTAACTTGGCGACT-3						
rap-1c	rec-*rap-1c*-fw	GCGCTCGAGCATGGTTTGCCACAGCTTTG		53	35	1746 bp		2 clones
	rec-*rap-1c*-rev	GCGTCTAGACGCGATAATTTGGATTGTATAAAAC						

The sequence with underline: "CTCGAG" *XhoI* restriction enzyme; "TCTAGA" *XbaI* restriction enzyme

Western blotting analysis

Thirty microliters of bacterial extracts (recombinant protein analysis) or 180 µg of *Babesia* sp. BQ1 (Lintan) Merozoite Antigens (BQMA) prepared from in vitro culture as described by Guan *et al*., 2010c were separated by SDS-PAGE using 10% or 12% polyacrylamide gel.

The proteins were transferred on nitrocellulose (NC) membranes of 0.45 µm pore size (BioRad) for 2 h at 144 mA (0.8mA/cm^2 of gel). The membrane was then incubated in a blocking solution (5 % skimmed milk powder in Tris-buffered saline (pH 7.6) with 0.1% Tween-20 (TBST)), shaking overnight at +4°C.

The NC membrane was incubated for 2 h with either 1-Mouse monoclonal Anti-polyHistidine antibody (H1029-2ML, Sigma; dilution 1:5,000) (recombinant protein expression control experiments), 2- sera from immunized and negative control rabbit sera (dilution 1/100) or 3- with sera from *Babesia* sp. BQ1 (Lintan) infected sheep (dilution 1/100). After 3 washes with TBST, membranes were incubated 1 hour respectively with 1-Goat anti mouse IgG Alkaline Phosphatase conjugate (Millipore, dilution 1:5,000), 2- anti-rabbit IgG-Alkaline Phosphatase conjugate (sigma, dilution 1/5000) or 3- monoclonal anti Goat/Sheep IgG-Alkaline phosphatase conjugate (Sigma, dilution 1/1000). The membranes were washed three times with TBST and placed into a 5-bromo-4-chloro-3-indolyl phosphate/Nitro blue tetrazolium (BCIP/NBT) Liquid Substrate System (B1911-100ML, Sigma) for 20 to 30 min. The approximate molecular weights of the revealed proteins were evaluated by comparing their migrations with the migration of the standard molecular weight marker (26619-250µl, PageRuler™, Plus Prestained Protein Ladder, Thermo Scientific).

ELISA procedure

According to the optimum coating concentration of peptides defined by the manufacturer of previous experiments, 100 µl of 4 µg/ml of *rap-1* peptides or 2,5 µg/ml of BQMA in coating buffer (0.1M carbonate/bicarbonate, pH 9.6) were distributed and adsorbed in microplates of 96 well flat-bottom at 37 °C for 1 hour and then at 4°C overnight. After three washings with PBST (0.1% Tween-20), the plates were blocked with 150 µl of a solution at 2% gelatin in PBST, at 37 °C for 30 min. After washing, IgG from rabbit sera (1 µg/ml or 2 mg/ml) or *Babesia* sp. BQ1 (Lintan) infected sheep sera (dilution 1/10 or 1/100) were distributed in duplicate and the plates were incubated for 1 hour at 37°C. After three washings, 100 µl of monoclonal anti-rabbit IgG (γ-chain-specific) clone

RG-96- Peroxydase conjugate (Sigma, dilution 1/10000), or 100 µl of monoclonal anti-Goat/sheep IgG-peroxidase, (Sigma, dilution 1/1000) were added to each well and incubated for 1 hour at 37°C. After three washings, 100 µl of 1-step^{TM} Ultra TMB-ELISA (34028-250ml, Thermo Scientific) were added to each well and incubated for 20 to 30 min at room temperature. The reaction was then stopped by adding 100 µl of 2 M H_2SO_4 and the plates were read at 450 nm with an ELISA automat (MRX microplate reader, USA).

Statistical analysis

Comparison of antibody production before and after *Babesia* sp. BQ1 infection has been performed using non-parametric Wilcoxon test.

Results

All *rap-1* types of the *Babesia* sp. BQ1 (Lintan) were transcribed

Transcription of genes can be regulated by sequences present in the 5' region of the genes. Conserved motifs that could be involved in *rap-1* transcription regulation were therefore searched in the three different intergenic regions. The three boxes described by Suarez et al. in 1998, called the -59 box, the -36 box and the mRNA box, were found with a high degree of sequence similarities for the -59 box and the -36 box (Figure 2). Except for the IG3 region, the localization and sequence of the mRNA box were not as conserved. The transcription of *rap-1* genes was analyzed using iRBC produced by in vitro cultures of *Babesia* sp. BQ1 (Lintan) and a RT-PCR specific for each *rap-1* gene type (Figure 3). The absence of amplification of the intergenic region IG2 from the prepared cDNA confirmed the absence of DNA contaminations in the RNA extract (data not shown). Amplicons were obtained for each primer combination, indicating that *rap-1a61*, *rap-1a67-1, rap-1a67-2*, *rap-1b* and *rap-1c* were transcribed by this parasite in vitro (Figure 3). The amplification patterns from genomic DNA (Figure 3, lane 2 or 5) and cDNA extract (Figure 3, lane 1 or 4) were identical using common (*rap-1a61/67-1*) and specific primer pairs (Figure 3A, *rap-1a61*, *67*, Figure 3B, *rap-1b* and Figure 3C, *rap-1c*). The sequencing of inserts from 5 clones for *rap-1a*, 4 clones for *rap-1b* and the PCR product for *rap-1c* was performed and confirmed the absence of introns in *rap-1* sequences. Sequences corresponding to the *rap-1a61-1* type (1 clone) and the *rap-1a61-2* type (4 clones) were found, but no *rap-1a67-1* type of sequence was discovered. From these results, we can conclude that all 4 *rap-1a*

gene types are transcribed in vitro.

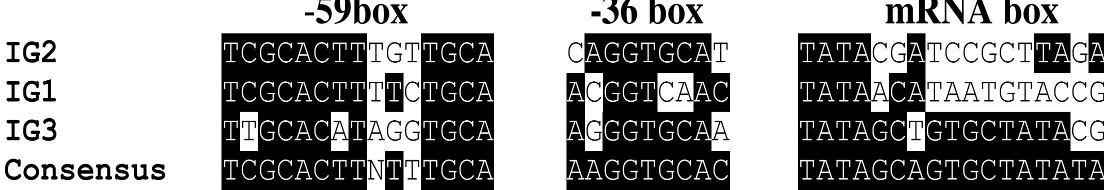


Figure 2. Conservation of the putative *rap-1a* transcription regulation 5'UTR sequence in the *rap-1a* (IG1), *rap-1b* (IG2) and *rap-1c* (IG3) intergenic regions. The consensus sequences position of boxes -59, -36 and mRNA are designated according to Suarez et al., 1998. The sequences could be searched on NCBI website, GenBank accession number: KC953700 or KC953701 for IG1; KC953702 for IG2 and KC953703 for IG3.

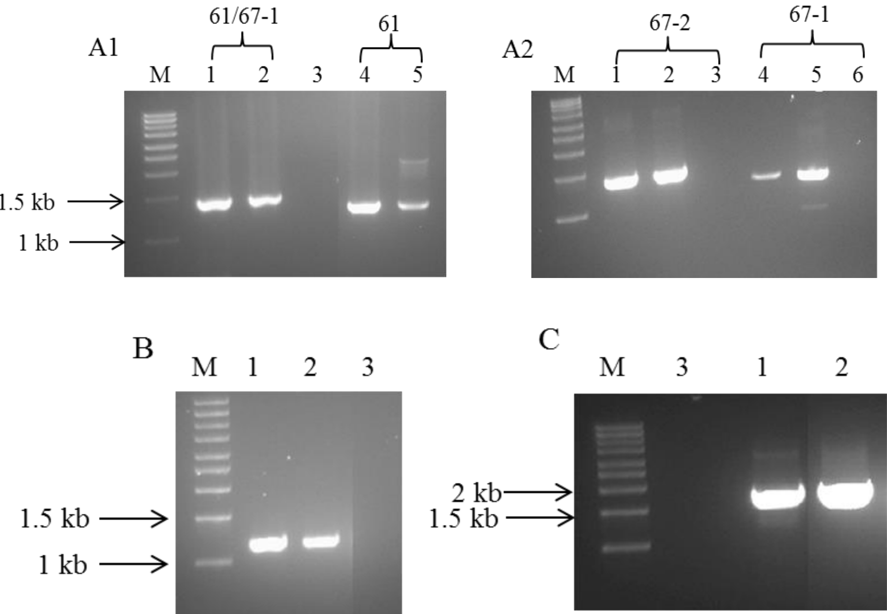


Figure 3. Transcriptional analysis of the *Babesia* sp. BQ1 Lintan *rap-1* genes.
(A): RT-PCR with *rap-1a* primers. (A1): full-length amplification of *rap-1a61* and *rap-1a67-1* using common primer pair (*rap1-a*-up/*rap-1a*-downend), lane 1: cDNA; lane 2: gDNA; and specific amplification for *rap-1a61* type with primer combination (*rap-1a61*spe-up1/*rap-1a*-downend), lane 4: cDNA; lane 5: gDNA. Lane 3: blank control. (A2): *rap-1a67-2* full-length amplification using primer paire (*rap-1a*-up/*rap-1a67-2*-downend), lane 1: cDNA; lane 2: gDNA; lane 3: blank control; and specific amplification for *rap-1a67-1* type with primer combination (*rap-1a67*spe-up1/*rap-1a*-downend), lane 4: cDNA; lane 5: gDNA; lane 6: blank control. (B): RT-PCR with *rap-1b* primer for full-length amplification of gene. (C): RT-PCR with *rap-1c* primer for full-length amplification of gene. Lane 1: cDNA; lane 2: gDNA; lane 3: blank control. M: XL molecular marker.

Bioinformatic putative B-cell epitopes mapping on RAP-1 proteins

Bioinformatic prediction of linear B-cell epitopes of three different RAP-1 proteins was performed using the software on website http:// tools.immuneepitope.org/tools/bcell/iedb_input. The putative B-cell epitopes were located mainly in the C-terminal part of the different RAP-1 putative proteins (Figure 1A), especially in the case of RAP-1c in which repeated similar sequences (Niu *et al*., 2013) were found to each bear B-epitope. Fewer B-cell epitopes have been predicted in the N-terminal region of these proteins, especially in the case of RAP-1b. It was difficult to find an sequence specific of each RAP-1a type, that was also confidently predicted as a B-epitope, and only one location was possible to design a long enough specific peptide. Therefore, another RAP-1a peptide was selected (LTLPNPYQL, aa 115-123 in blue in Figure 1B), that had a greater chance to represent a true B-epitope, but was identical for the two RAP-1a types (RAP-1a61 and RAP-1a67).

Rabbit antibodies recognized specifically the different RAP-1 peptides

To validate the specificity of rabbit IgG, the absence of cross-reactions between antibodies and the different peptides was tested by ELISA; IgG from rabbit sera recognized the peptide with which rabbit have been immunized (Figure 4) and no cross-reactions between peptides have been observed.

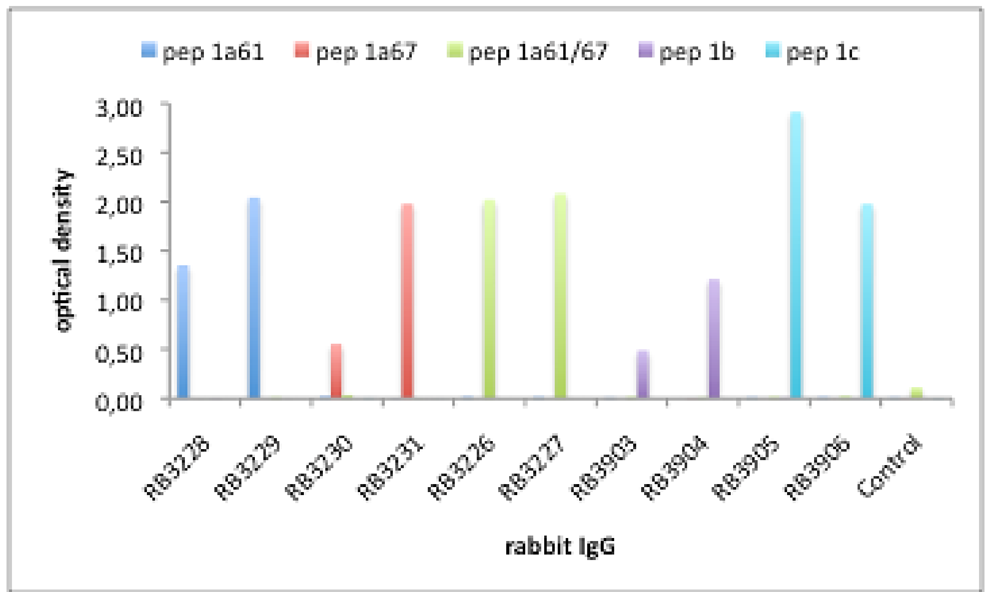


Figure 4. Absence of cross reactivities between IgG from Rabbit sera immunized with peptide or negative control and the differents peptides. Elisa plate were coated with each peptide (2µg/ml) and incubated with each IgG sample (1 µg/ml). RB3228 and RB3229 are rabbits immunized with pep 1a61, RB3230 and RB3231 with pep 1a67, RB3226 and RB3227 with pep 1a61/67, RB3903 and RB3904 with pep 1b, RB3905 and RB3906 with pep 1c and control is not immunized.

Analysis of RAP-1 protein expression in vitro and in vivo

The reactivity of each immunized rabbit serum with BQMA (merozoite antigens produced in vitro from a clone of *Babesia* sp. BQ1 Lintan) has been analyzed by western blot (Figure 5A). A strong reaction with the expected size of RAP-1 (about 52 kDa) was observed with one serum specific of pep 1a61 (RB3228, lane 1) and the 2 sera specific of pep 1a61/67 (RB3226 and RB3227, lanes 9 and 11 respectively). A band of the same size has been also observed with sera from *Babesia* sp. BQ1 (Lintan) infected sheep (N°3216 and 2007, lanes 24 and 26 respectively in Figure 5A**,** and 3446, data not shown). In contrast, no reaction has been observed with sera from rabbit immunized with pep 1a67, pep 1b and pep 1c. A same result has been observed using in vivo infected red blood cells as antigens (iRBC collected from a sheep experimentally infected with the original strain of *Babesia* sp. BQ1 (Lintan) 4 days after infection, with a parasitemia of 10%) (Data not shown).

Analysis of antibody production against the different RAP-1 proteins in the course of infection

The kinetics of antibody production against each RAP-1 peptides have been studied by ELISA using each peptide as antigens and sera from the 3 infected sheep collected before (DPI0) and after infection (at 14, between 30 and 42 and between 77 and 90 days post infection) (Figure 5B). A statistically significant increase of antibody produced against each RAP-1 was observed after infection, but with a different kinetics. The production of antibodies against pep 1a61 and pep 1a67 increased early during the infection (DPI14) and then decreased to return to the DPI0 level at DPI77-90 for pep 1a61. In contrast, a gradual increase of antibody produced against pep 1c was observed all along the infection. Finally, for pep 1b, antibodies were produced from DPI14 and then slightly decreased to remain stable.

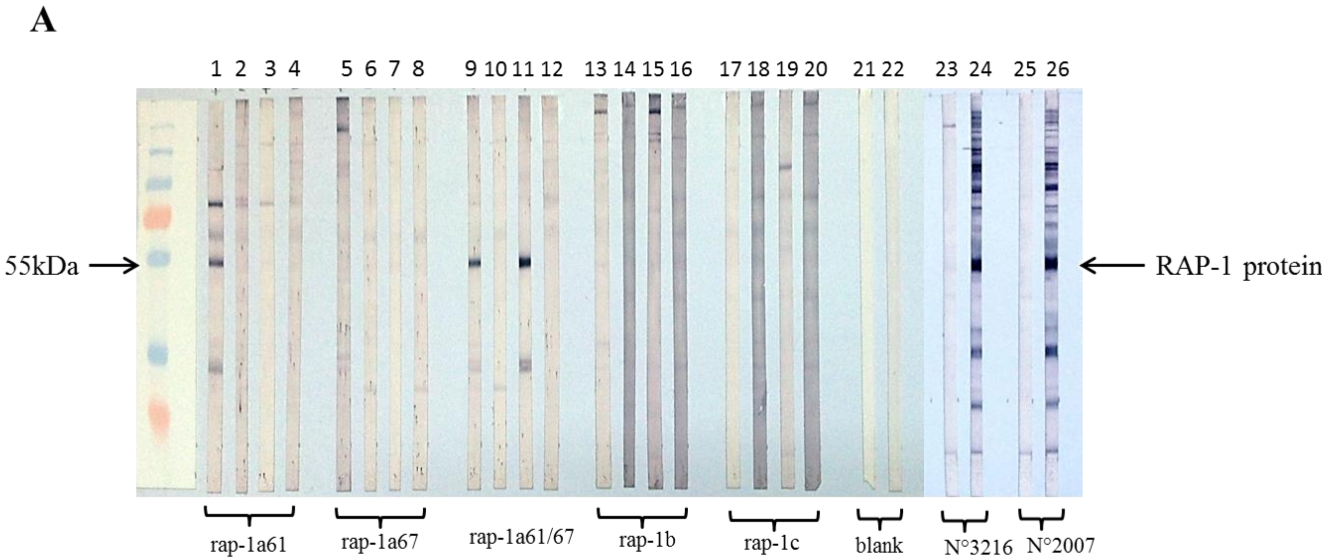


Figure 5A. Analysis of the translation of each RAP-1 protein. Western blot analysis of the reactivity between sera from rabbit immunized with RAP-1 peptides or from one *Babesia* sp. BQ1 infected sheep (N°3216, 2007) and *Babesia* sp BQ1 merozoite antigens. Lanes 1,3,5,7,9,11,13,15,17,19,24 and 26: positive sera; Lanes 2,4,6,8,10,12,14,16,18,20,23,25: negative sera; Lanes 21 and 22: blank for rabbit and sheep. Lanes 1 and 2 (RB3228); 3 and 4 (RB3229); 5 and 6 (RB3230); 7 and 8 (RB3231); 9 and 10 (RB3226); 11 and 12 (RB3227); 13 and 14 (RB3903); 15 and 16 (RB3904); 17 and 18 (RB3905); 19 and 20 (RB3906); 23 and 24 (N°3216); 25 and 26 (N°2007): the sera from the same animal.

B

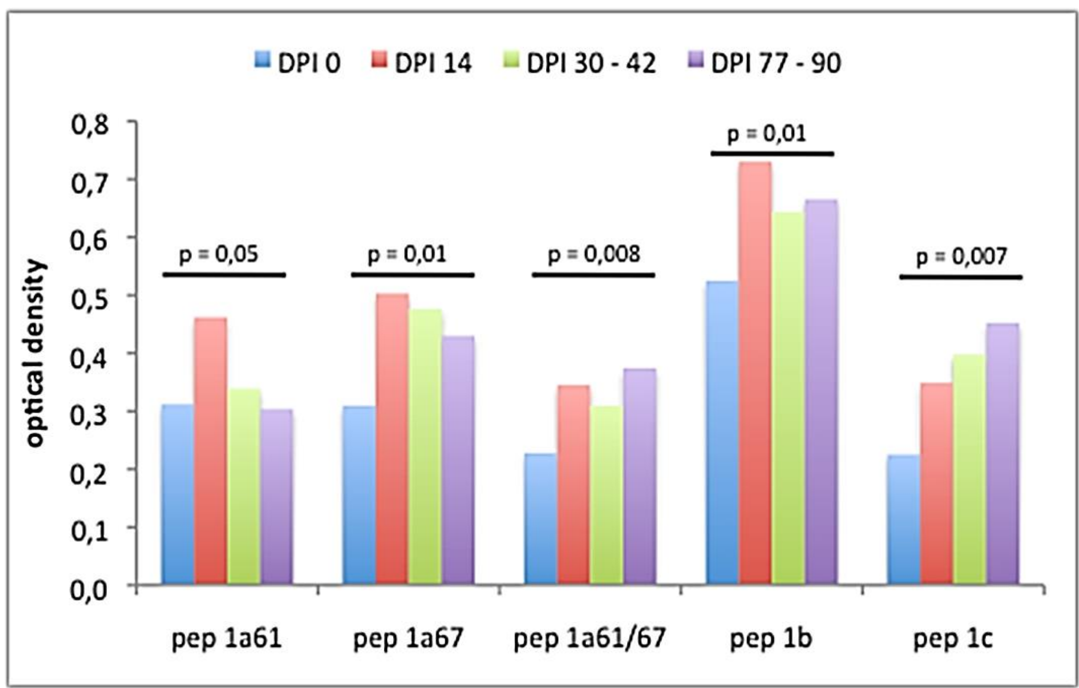


Figure 5B. Analysis of the translation of each RAP-1 protein. Kinetic of antibody production against each RAP-1 peptide during a sheep infection with *Babesia* sp. BQ1. ELISA plate has been coated by each peptide (4 µg/ml) and incubated with sera collected before infection (DPI 0) and after infection (14, between 30 and 42 and between 77 and 90 days after infection: DPI14, DPI30-42, DPI77-90 respectively) from sheep 2007, 3216 and 3446. Results have been expressed as mean of optical density (OD) obtained for each sheep. OD before and after infection has been compared with the Wilcoxon test.

Production of the recombinant proteins

Each recombinant RAP-1 proteins have been produced in the *E. coli* system. A single open reading frame (ORF) of *rap-1a (61* and *67)*, *rap-1b* and *rap-1c* has been inserted in a plasmid. To verify the quality of the insertion, each fragment was amplified using specific primers to each other (Figure 6A), showing that the amplicons correspond with the ORF of each protein (1369 bp for *rap-1a*, 1200 bp for *rap-1b*, 1746 bp for *rap-1c*). The sequences of the insert have been also verified (data not shown).

After induction of each recombinant protein expression by L-arabinose, the lysates of *E. coli* has been analyzed by western blot using an anti-His antibody (Figure 6B). Recombinant RAP-1a61-6xHis, RAP-1a67-6xHis and RAP-1b-6xHis have been produced with the expected molecular weight (52kDa of the 2 RAP-1a, 45kDa of RAP-1b). For

RAP-1c, no expression has been revealed even if the positive control *lac* Z has been expressed with success (data not shown).

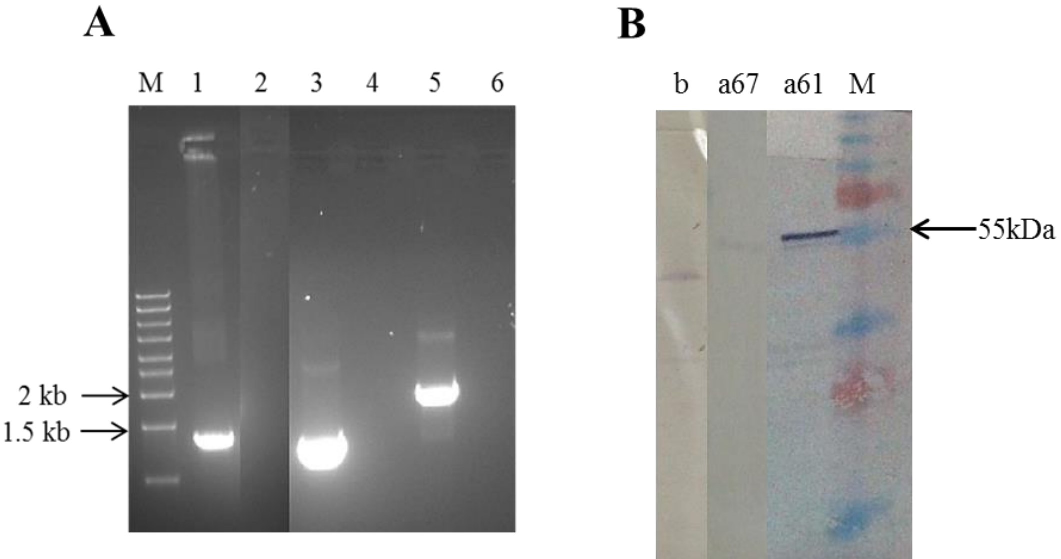


Figure 6. A: PCR analysis for recombinant protein production. PCR amplification fragment (1369 bp) of the *rap-1a* gene (lane 1); 1200 bp of the *rap-1b* gene (lane 3); 1746 bp of the *rap-1c* gene (lane 5); Lane 2, 4, 6: blank control for *rap-1a*, *b* and *c*, respectively; M: XL molecular marker. **B**: Western blot analysis of recombinant protein expression of *Babesia* sp. BQ1 (Lintan) RAP-1a, b using anti-His antibodies.

Discussion

When *rap-1a* locus organization was studied (*B. bovis*, *B. bigemina*, *B. ovis*, *B. canis*, *B. motasi*-like group), the *rap-1* family was shown to be composed of multiple gene copies arranged tandemly, separated by intergenic (IG) regions that are thought containing a promoter region to regulate *rap-1* gene transcription. Three conserved areas (designated -59, -36, and mRNA boxes) located in the IG regions have been described at the 5' UTR region of the *rap-1* genes in *B. bovis*, *B. bigemina*, *B. ovis* and *B. canis* (Suarez *et al.*, 1998). The role of this region in the *rap-1* gene expression has also been demonstrated in *B. bovis* (Suarez *et al.*, 2004). The three conserved sequence boxes were also found in the three *rap-1* intergenic regions (IG1, 2 and 3) of *Babesia* sp. BQ1 (Lintan) by sequences analysis and comparison (Figure 2). Two of these 3 boxes (-59 and -36 boxes) have a high degree of sequence similarities with other *Babesia* species, while the mRNA box varies more in sequence and localization in this 5' UTR region. The feature conservation of putative regulatory boxes among the *Babesia* could be therefore an explanation that all

rap-1 gene types (*rap-1a, b* and *c*) are transcribed by the in vitro cultivated parasite of *Babesia* sp. BQ1 (Lintan) (Figure 3). About the *rap-1a* transcription, we demonstrated either by RT-PCR with specific primers for *rap-1a67-1* and *rap-1a67-2*, or by sequencing the RT-PCR cloned amplicon for *rap-1a61-1* and *rap-1a61-2*, that all four *rap-1a* gene types were transcribed by the in vitro cultivated parasite. The absence of a *rap-1a67-1* sequence in the five clones sequenced may be associated with its fewer proportion in *rap-1* locus (one *rap-1a67-1* type located at the beginning of the locus against four *rap-1a61* types (Niu *et al*., 2013). In the case of *B. bigemina*, whose *rap-1* locus gene composition and organization are very similar, the transcription of *rap-1a*, *rap-1b* and *rap-1c* genes was also demonstrated, and transcription of at least three of the four *rap-1a* variants also occur (Mishra *et al*., 1992; Suarez *et al*., 2003).

The translation investigation of the *rap-1* genes in *B. bigemina* by western blot using lysates of merozoites obtained from in vitro culture or in vivo infected red blood cells as antigen and monoclonal anti-RAP-1a and mouse anti-RAP-1b or anti-RAP-1c peptides revealed that only RAP-1a is translated. If at least three of the *rap-1a* variants were shown to be transcribed, no detectable translation products were produced for the *rap-1aβ2* variant (Suarez *et al*., 2003). Similar results have been obtained in the present study with *Babesia* sp. BQ1 (Lintan): only one rabbit serum immunized with pep 1a61 and the 2 rabbit sera immunized with pep 1a61/67 revealed an antigen with the expected molecular weight of RAP-1a (near to 55 kDa) in *Babesia* sp. BQ1 (Lintan) merozoite lysate (Figure 6A), as well as in infected red blood cells from an infected sheep. These results suggest that only *rap-1a61* has been translated and that *rap-1a67*, *rap-1b* and *rap-1c* have been transcribed but not translated or not enough to produce detectable products. However, we have demonstrated too that sera from 3 *Babesia* sp. BQ1 (Lintan) - infected sheep contained antibodies against the 5 peptides (pep 1a61, pep 1a67, pep 1a61/67, pep 1b and pep 1c) (Figure 6B), suggesting that these sheep have been immunized by all RAP-1 protein variants during infection and so, that the 4 *rap-1* genes have been translated. This discrepancy between these results could be explained in two ways, that both need to be analyzed. First, the sera may not be able to recognize the peptides in the whole proteins. As seen with the anti-pep 1a61 produced sera, only one of the two sera recognized RAP-1a on the western blot (Figure 5A), even if they both reacted with the peptide used for immunization in ELISA (Figure 4). It is therefore possible that three of the sera directed specifically against RAP-1a61 (RB3229) and RAP-1a67 (RB3230 and RB3231) did not recognize the RAP proteins on the western blot. The prediction of these peptides to

confidently represent putative B epitopes was also low (Figure 1A). The recognition of all four RAP-1 recombinant proteins by the same sera need to be controlled by western blot to eliminate this possibility. Second, the translation of the different RAP protein may evolve in the course of infection. As seen in figure 5B, antibody production against RAP-1a61 and against RAP-1a67 has increased between DPI0 and DPI14 and then decreased, while it plateaued or increased all along the infection period studied for RAP-1b and RAP-1c respectively (Figure 6B). So, RAP-1a may be produced mainly in the beginning of *Babesia* sp. BQ1 (Lintan) infection while RAP-1b and RAP-1c may be produced later during the same infection. The sequential expression of RAP-1 proteins during infection could be due to a selective pressure exerted by the host immune system. The absence of RAP-1b and RAP-1c in infected red blood cells from an infected sheep (Figure 6A) could be explained by the use of parasite collected during the first days after infection. And in vitro, in the absence of host immune pressure, only RAP-1a could be translated. In their study on *rap-1* expression in *B. bigemina*, Suarez et al. also used merozoite lysate from a calf with an acute infection, so in the first stage of infection (Suarez *et al.*, 2003), while the sera we used were collected during the subsequent persistent low parasitemic infection. However, the sequential expression of the RAP-1 proteins in the course of infection must be confirmed by using other specific but more immunogenic peptides for each RAP-1 type in ELISA, and by analysing the reaction of sera collected at different times during infection against the whole four recombinant proteins.

B epitope mapping on each RAP-1 protein using amino acid sequences and method "Bepipred Linear Epitope Prediction" indicated the presence of putative B-cell epitopes located mainly in the C-terminal region of the RAP-1 proteins and few B cell epitopes in the N terminal region (Figure 1A). Using monoclonal antibodies and immune sera from cattle, Suarez *et al* (1993) have also shown that the N-terminal sequence of *B. bigemina* and *B. bovis* RAP-1 contained poorly immunogenic B-cell epitopes. Moreover, the sequence conservation of *rap-1a* genes in theN-terminal part of the protein among different *Babesia* species suggests this region could bear a significant function in erythrocyte invasion (Hötzel *et al.*, 1996, 1997). These results suggest that antibodies could be directed mainly against the C terminal region of RAP-1 proteins which might not be involved in the protein function. In contrary, no or low antibodies might be produced against the functional N terminal part of the proteins. The sequential expression of RAP-1 proteins due to a selective pressure exerted by the host immune system and the localization of B cell epitopes in the C-terminal region without important function could represent a mechanism

developed by *Babesi*a sp. BQ1 (Lintan) and in general by *Babesia* sp. to escape to the immune response of the host and to persist in it.

B cell epitope mapping and manually sequence analysis have revealed that in RAP-1c, many B-cell epitopes were located at almost identical repeat motifs in the C-terminal region. These repeats were not detected in the case of RAP-1a and RAP-1b (our work and Niu et al., 2014a, in press). Proteins with repetitive amino acids domains have functions often associated with either immune evasion or cell invasion process in intracellular parasites including *Babesia*, *Theileria and Plasmodium* (Mendes *et al*., 2013). Indeed, repeats of a motifs involved in the interaction with the cell receptor could facilitate the cell invasion (Mendes *et al*., 2013). However, the presence of B-cell epitopes localized in the repetitive motifs in the C-terminal region of RAP-1c, the low overall immunogenicity of RAP-1b and the possible sequential production of each RAP-1 proteins during *Babesia* sp. BQ1 (Lintan) infection suggests that these proteins could have different role in host-parasite interactions, potentially at different stages of the infection process. RAP-1c with its repeat motifs bearing B epitopes could distract the host immune system during the persistent phase of infection helping the parasite survival and evasion from the immune system of the host, preserving the functional B cell epitopes in N terminal region.

RAP-1a has been found in all examined *Babesia* species, including three *Babesia* parasites of sheep: *Babesi*a sp BQ1 (Lintan) and three other Chinese *B. motasi*-like isolates (Niu *et al*., 2014a submitted), *Babesia* sp. Xinjiang (Niu *et al*., 2014b submitted) and *B. ovis* (Dalrymple *et al*., 1993). A conserved domain corresponding to a B cell epitope, with an amino acid sequence LTLPNPYQL (aa 115-123**,** Figure 1B), has been predicted in the N-terminal region of the 2 different *Babesia* sp BQ1 (Lintan) RAP-1a, and used to produce RAP-1a specific antibodies. The produced antibodies have been shown to recognize RAP-1a protein in western blot (Figure 5A) and antibodies against this peptide are produced steadily in the course of infection (Figure 5B). This domain is also perfectly conserved in other sheep *Babesia* (*B. canis*, *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu) or with slight variations in other sheep parasites: LVLPNPYQL in *Babesia* sp. Xinjiang, LTLPNPYEL in *Babesia* sp. Hebei and LSLPNPYQL in *B. ovis.* (Variable amino acids are underlined). The high conservation of this peptidic motif in the sheep *Babesia* species could be the signature of its important function in the host-parasite interaction, especially in invasion process. This should be confirmed by studying the ability of rabbit IgG directed against the peptide 1a61/67 to inhibit the in vitro growth of *Babesia* sp. BQ1 (Lintan).

In conclusion, because of the early expression of *rap-1a* during *Babesia* sp. BQ1

(Lintan) infection and the presence of a highly conserved B cell epitope between the 3 major sheep *Babesia* studied, this peptide could be an interesting vaccine candidate.

However, several results or hypothesis from our study must be confirmed or studied further:

1. Recognition of the recombinant proteins rRAP-1a, rRAP-1b and rRAP-1c by the peptide specific rabbit antibodies

2. The sequential production of RAP-1a, RAP-1b and RAP-1c during infection by western blot or ELISA using different specific peptides

3. The role of the different RAP-1 proteins in the red blood cell invasion process by testing the ability of IgG against the different RAP-1 to inhibit the in vitro *Babesia* sp. BQ1 (Lintan) growth. The ability of the antibodies directed against the conserved RAP-1a motif to inhibit the growth of other *Babesia* spcies could also be evaluated.

It should be also very interesting to compare the functions of RAP-1a, RAP-1b and RAP-1c in the parasite – host interaction, in particular by investigating the function of repetitive B cell epitopes described in the C terminal part of RAP-1c.

Acknowledgements

This study was supported financially by a grant from the European Union cooperation project PIROVAC (KBBE-3-245145), the «Région Pays de la Loire» (France), and INRA.

References

Bai Q, Liu G, Liu D, Ren J, Li X. 2002. Isolation and preliminary characterization of a large *Babesia* sp. from sheep and goats in the eastern part of Gansu Province, China. *Parasitol Res*. 88, S16–S21.

Bhoora R, Quan M, Zweygarth E, Guthrie AJ, Prinsloo SA, Collins NE. 2010. Sequence heterogeneity in the gene encoding the rhoptry-associated protein-1 (RAP-1) of *Babesia caballi* isolates from South Africa. *Vet Parasitol*. 169: 279-88.

Brown WC, McElwain TF, Hötzel I, Suarez CE, Palmer GH.1998. Helper T-cell epitopes encoded by the *Babesia bigemina rap*-1 gene family in the constant and variant domains are conserved among parasite strains. *Infect Immun*. 66: 1561-1569.

Brown WC, McElwain TF, Palmer GH, Chantler SE, Estes DM. 1999. Bovine CD4 (+) T-lymphocyte clones specific for rhoptry-associated protein1 of *Babesia bigemina* stimulate enhanced immunoglobulin G1 (IgG1) and IgG2 synthesis. *Infect Immun*. 67: 155-164.

Brown WC, Palmer GH. 1999. Designing blood-stage vaccines against *Babesia bovis* and *B. bigemina*. *Parasitol. Today* 15: 275-281.

de Waal DT, Combrink MP. 2006. Live vaccines against bovine babesiosis. *Vet Parasitol*. 138: 88-96.

Dalrymple BP, Casu RE, Peters JM, Dimmock CM, Gale KR, Boese R, Wright IG. 1993. Characterization of a family of multicopy genes encoding rhoptry protein homologues in *Babesia bovis*, *Babesia ovis* and *Babesia canis*. *Mol Biochem Parasitol*. 57: 181-192.

Figueroa JV, Buening GM.1991. In vitro inhibition of multiplication of *Babesia bigemina* by using monoclonal antibodies. *J Clin Microbiol*. 29: 997-1003.

Gohil S, Herrmann S, Günther S, Cooke BM. 2013. Bovine babesiosis in the 21st century: advances in biology and functional genomics. *Int J Parasitol*. 43: 125-132.

Gou H, Guan G, Ma M, Liu A, Liu Z, Ren Q, Li Y, Yang J, Chen Z, Yin H, Luo J. 2013. Phylogenetic analysis based on 28S rRNA of *Babesia* spp. in ruminants in China. *Exp Appl Acarol*. 59: 463-472.

Guan GQ, Yin H, Luo JX. 2001. The primary studies on morphology and pathogenicity on an unidentified large ovine *Babesia*. *Chin. J. Vet. Sci. Technol*. 11: 35-36.

Guan GQ, Yin H, Luo JX, Lu WS, Zhang QC, Gao YL, Lu BY. 2002. Transmission of *Babesia* sp. to sheep with field-collected *Haemaphysalis qinghaiensis*. *Parasitol Res*. 88: S22-24.

Guan G, Moreau E, Liu J, Hao X, Ma M, Luo J, Chauvin A, Yin H. 2010a. *Babesia* sp. BQ1 (Lintan): molecular evidence of experimental transmission to sheep by *Haemaphysalis qinghaiensis* and *Haemaphysalis longicornis*. *Parasitol Int*. 59: 265-267.

Guan G, Moreau E, Brisseau N, Luo J, Yin H, Chauvin A. 2010b. Determination of erythrocyte susceptibility of Chinese sheep (Tan mutton breed) and French sheep (Vendéen breed) to *Babesia* sp. BQ1 (Lintan) by in vitro culture. *Vet Parasitol*. 170: 37-43.

Guan GQ, Chauvin A, Rogniaux H, Luo JX, Yin H, Moreau E. 2010c. Merozoite proteins from *Babesia* sp. BQ1 (Lintan) as potential antigens for serodiagnosis by ELISA. *Parasitology.* 137: 927-938.

Guan G, Ma M, Liu A, Ren Q, Wang J, Yang J, Li A, Liu Z, Du P, Li Y, Liu Q, Zhu H, Yin H, Luo J. 2012. A recently identified ovine *Babesia* in China: serology and sero-epidemiology. *Parasitol Int.* 61: 532-537.

Hötzel I, Brown WC, McElwain TF, Rodríguez SD, Palmer GH. 1996. Dimorphic sequences of *rap-1* genes encode B and CD4+ T helper lymphocyte epitopes in the *Babesia bigemina* rhoptry associated protein-1. *Mol Biochem Parasitol*. 81: 89-99.

Hötzel I, Suarez CE, McElwain TF, Palmer GH. 1997. Genetic variation in the dimorphic regions of RAP-1 genes and *rap-1* loci of *Babesia bigemina*. *Mol Biochem Parasitol*. 90: 479-89.

Jouglin M, Fernandez-de-Mera IG., De la Cotte N, Ruiz-Fons F, Gortazar C, et al 2014. Isolation and characterization of *Babesia pecorum* sp. nov. from farmed red deer (*Cervus elaphus*). *Vet Res* (in press)

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: 353-63.

Liu AH, Yin H, Guan GQ, Schnittger L, Liu ZJ, Ma ML, Dang ZS, Liu JL, Ren QY, Bai Q, Ahmed JS, Luo JX. 2007. At least two genetically distinct large *Babesia* species infective to sheep and goats in China. *Vet Parasitol*. 147: 246-251.

Machado RZ, McElwain TF, Pancracio HP, Freschi CR, Palmer GH. 1999. *Babesia bigemina* : immunization with purified rhoptries induces protection against acute parasitemia. *Exp Parasitol*. 93: 105-108.

Malandrin L, L'Hostis M, Chauvin A. 2004. Isolation of *Babesia divergens* from carrier cattle blood using in vitro culture. *Vet Res*. 35: 1-9.

Malandrin L, Jouglin M, Moreau E, Chauvin A. 2009. Individual heterogeneity in erythrocyte susceptibility to *Babesia divergens* is a critical factor for the outcome of experimental spleen-intact sheep infections. *Vet Res*. 40: 125.

Mendes TA, Lobo FP, Rodrigues TS, Rodrigues-Luiz GF, daRocha WD, Fujiwara RT, Teixeira SM, Bartholomeu DC. 2013. Repeat-enriched proteins are related to host cell invasion and immune evasion in parasitic protozoa. *Mol Biol Evol*. 30: 951-63.

Mishra VS, McElwain TF, Dame JB, Stephens EB. 1992. Isolation, sequence and differential expression of the p58 gene family of *Babesia bigemina*. *Mol Biochem Parasitol*. 53: 149-158.

Mosqueda J, McElwain TF, Stiller D, Palmer GH. 2002. *Babesia bovis* merozoite surface antigen 1and rhoptry-associated protein 1 are expressed in sporozoites, and specific antibodies inhibit sporozoite attachment to erythrocytes. *Infect Immun*. 70: 1599-1603.

Mosqueda J, Olvera-Ramirez A, Aguilar-Tipacamu G, Canto GJ. 2012. Current advances in detection and treatment of babesiosis. *Curr Med Chem*. 19: 1504-1518.

Niu QL, Luo JX, Guan GQ, Liu ZJ, Ma ML, Liu AH, Gao JL, Ren QY, Li YQ, Qiu JX, Yin H. 2009. Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. *Exp Parasitol*. 121: 64-68.

Niu Q, Bonsergent C, Guan G, Yin H, Malandrin L. 2013. Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group). *Vet Parasitol*. 198: 24-38.

Niu Q, Valentin C, Bonsergent C, Malandrin L .2014a. Strong conservation of rhoptry associated-protein-1 (RAP-1) locus organization and sequence among *Babesia* isolates infecting sheep from China (*Babesia motasi*-like phylogenetic group). *Parasitol Res*.(in press)

Niu Q., Marchand J, Yang C, Bonsergent C, Guan G, Yin H, Malandrin L. 2014b. Rhoptry-Associated-Protein (*rap-1*) genes in the sheep pathogen *Babesia* sp. Xinjiang: multiple copies differing by 3’ end repeated sequences. *PloS One*. (in press).

Oosthuizen MC, Allsopp BA, Troskie M, Collins NE, Penzhorn BL. 2009. Identification of novel *Babesia* and *Theileria* species in South African giraffe (*Giraffa camelopardalis*, Linnaeus, 1758) and roan antelope (*Hippotragus equinus*, Desmarest 1804). *Vet Parasitol*. 163: 39-46.

Skuce PJ, Mallon TR, Taylor SM. 1996. Molecular cloning of a putative rhoptry associated protein homologue from *Babesia divergens*. *Mol Biochem Parasitol*. 77: 99-102.

Suarez CE, Palmer GH, Jasmer DP, Hines SA, Perryman LE, McElwain TF. 1991. Characterization of the gene encoding a 60-kilodalton *Babesia bovis* merozoite protein with conserved and surface exposed epitopes. *Mol Biochem Parasitol*. 46: 45-52.

Suarez CE, Palmer GH, Hines SA, McElwain TF. 1993. Immunogenic B-cell epitopes of *Babesia bovis* rhoptry-associated protein 1 are distinct from sequences conserved between species. *Infect Immun*. 61: 3511-3517.

Suarez CE, Palmer GH, Hötzel I, Hines SA, McElwain TF. 1998. Sequence and functional analysis of the intergenic regions separating babesial rhoptry-associated protein-1 (*rap-1*) genes. *Exp Parasitol*. 90: 189-194.

Suarez CE, Palmer GH, Florin-Christensen M, Hines SA, Hötzel I, McElwain TF. 2003. Organization, transcription, and expression of rhoptry associated protein genes in the *Babesia bigemina rap-1* locus. *Mol Biochem Parasitol*. 127: 101-12.

Suarez CE, Palmer GH, LeRoith T, Florin-Christensen M, Crabb B, McElwain TF. 2004. Intergenic regions in the rhoptry associated protein-1 (*rap-1*) locus promote exogenous gene expression in *Babesia bovis*. *Int J Parasitol*. 34: 1177-1184.

Sun Y, Moreau E, Chauvin A, Malandrin L. 2011. The invasion process of bovine erythrocyte by *Babesia divergens*: knowledge from an in vitro assay. *Vet Res*.11: 42:62.

Tian ZC, Liu GY, Yin H, Luo JX, Guan GQ, Luo J, Xie JR, Shen H, Tian MY, Zheng JF, Yuan XS, Wang FF. 2013a. RPS8--a new informative DNA marker for phylogeny of *Babesia* and *Theileria* parasites in China. PLoS One. 8:e79860.

Tian Z, Luo J, Zheng J, Xie J, Shen H, Yin H, Luo J, Tian M, Yuan X, Wang F, Liu G. 2013b. Phylogenetic analysis of *Babesia* species in China based on Cytochrome b (COB) gene. *Infect Genet Evol*. 13: 36-40.

Tian Z, Liu G, Yin H, Luo J, Guan G, Xie J, Luo J, Zheng J, Tian M, Yuan X, Wang F, Chen R, Wang H. 2013c. 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.

Uilenberg G. 2006. *Babesia*-a historical overview. *Vet Parasitol* 138: 3-10.

Wang JM, Ma ML, Liu AH, Ren QY, Li AY, Liu ZJ, Li YQ, Yin H, Luo JX, Guan GQ. 2013. A sero-epidemiological survey of Chinese *Babesia motasi* for small ruminants in China. *Parasitol Res*.112:2387-2391.

Wright IG, Casu R, Commins MA, Dalrymple BP, Gale KR, Goodger BV, Riddles PW, Waltisbuhl DJ, Abetz I, Berrie DA, Bowles Y, Dimmock C, Hayes T, Kalnins H, Leatch, G, McCrae T, Montague PE, Nisbet IT, Parrodi F, Peters JM, Scheiwe PC, Smith W, Rode-Bramanis K, White MA.1992.The development of a recombinant *Babesia* vaccine. *Vet Parasitol*. 44:3-13.

Yokoyama N, Suthisak B, Hirata H, Matsuo T, Inoue N, Sugimoto C, Igarashi I. 2002. Cellular localization of *Babesia bovis* merozoite rhoptry-associated protein 1 and its erythrocyte-binding activity. *Infect Immun*. 70: 5822-5826.

Yokoyama N, Okamura M, Igarashi I. 2006. Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. *Vet Parasitol*. 138: 22-32.

Zhou JL, Jia HL, Nishikawa Y, Fujisaki K, Xuan XX. 2007. *Babesia gibsoni* rhoptry-associated protein 1 and its potential use as a diagnostic antigen. *Vet Parasitol*. 145: 16-20.

Article N° 4: Rhoptry-Associated-Protein (*rap-1*) genes in the sheep pathogen *Babesia* sp. Xinjiang: multiple copies differing by 3' end repeated sequences

Article N° 4: Rhoptry-Associated-Protein (*rap-1*) genes in the sheep pathogen *Babesia* sp. Xinjiang: multiple copies differing by 3' end repeated sequences

Qingli Niu[a,b], Jordan Marchand[a,b], Congshan Yang[c], Claire Bonsergent[a,b], Guiquan Guan[c], Hong Yin[c], Laurence Malandrin[a,b]*

a. INRA, UMR1300 Biology, Epidemiology and Risk Analysis in Animal Health, CS 40706, F-44307 Nantes, France.
b. LUNAM Université, Oniris, UMR1300 BioEpAR, F-44307 Nantes, France.
c. State Key Laboratory of Veterinary Etiological Biology, LVRI, Lanzhou, China.

*** Corresponding author: Oniris, UMR INRA/Oniris 1300, Biology, Epidemiology and Risk Analysis in animal health, Route de Gachet, La Chantrerie, CS 40706, F-44307 Nantes, France. Phone: +33 02 40 68 78 57 Fax: +33 02 40 68 77 51. E-mail addresses:** laurence.malandrin@oniris-nantes.fr and philippe.maestri@gmail.com

Article has been submitted to PLOS ONE

Abstract

Sheep babesiosis occurs mainly in tropical and subtropical areas. In China, several *Babesia* isolates infecting sheep and goats have been described which belong to two phylogenetic groups, *B. motasi*-like and *Babesia* sp. Xinjiang. In order to develop recombinant subunit vaccines to target *Babesia* species infecting small ruminants, we focused on RAP-1 (Rhoptry-Associated-Protein 1) as this protein is known to be involved in the red blood cell invasion process. Seven different *rap-1* genes were amplified in *Babesia* sp. Xinjiang, using degenerate primers designed from conserved motifs. In all seven genes, the 5' regions exhibited identical sequences over 936 nt, and the 3' regions differed at 28 positions over 147 nt, defining two types of genes designated *α* and *β*. The remaining 3' part varied from 72 to 360 nt in length, depending on the gene. This region consists of a succession of two to ten 36 nt repeats, which explains the size differences. Even if the nucleotide sequences varied, 6 repeats encoded the same stretch of amino acids. B-epitope mapping of the protein sequences revealed that each of these repeats contained a putative B-epitope, and that most of the predicted B-epitopes were located at the C-terminal end of the protein. Transcription of at least three α and two β genes was demonstrated by standard RT-PCR. The features of this multigene multirepeats protein family in *Babesia* sp. Xinjiang are discussed in terms of evolution, function, diagnosis and vaccine development.

Keywords: *Babesia*, sheep, rhoptry-associated-protein 1, multigene family, amino acid repeats

Introduction

Babesia is a tick-transmitted obligate intraerythrocytic protozoan parasite, belonging to the phylum Apicomplexa, which infects a broad variety of vertebrate hosts [1]. Babesiosis, the hemolytic disease caused by *Babesia,* is responsible for significant mortality and morbidity rates and enormous economic losses to the international trade of animals and the livestock industry in tropical and subtropical regions worldwide [2, 3]. Several *Babesia* species are known to infect small ruminants, e.g. *Babesia ovis*, *B. motasi* and *B. crassa* [4]. Recently, many more *Babesia* isolates from small ruminants have been described in China, most of them closely related to *B. motasi* (*Babesia* sp. BQ1 (Lintan and Ningxian)), *Babesia* sp. Tianzhu, *Babesia* sp. Madang, *Babesia* sp. Liaoning and *Babesia* sp. Hebei) [5]. One of them, however, *Babesia* sp. Xinjiang, belongs to a separate and new phylogenetic clade, together with other recently described *Babesia* from wild ruminants in South Africa and *B. pecorum* isolated from red deer in Spain [5-8]. This *Babesia* was isolated from a splenectomized sheep that had been experimentally infested with adult *Rhipicephalus sanguineus* and *Hyalomma anatolicum anatolicum* collected from farmed sheep and goats in the Xinjiang autonomous region [9]. This isolate exhibited different morphological characteristics to the above-mentioned *B. motasi*-like isolates: i.e. typical piriform but more slender pairs than the other Chinese isolates with an average size of 2.42 (± 0.35) µm × 1.06 (± 0.22) µm [10]. The widespread distribution of *Babesia* sp. Xinjiang-related parasites in China (over 50 prefectures in 22 provinces) was demonstrated in a sero-epidemiological survey [11].

Clinical symptoms of babesiosis become apparent when the parasite invades and replicates within the host erythrocytes and attains detectable parasitemia. These symptoms are characterized by fever and mild hemolytic anemia, whereas severe cases exhibit depression and hemoglobinuria with subsequent organ and systemic failure resulting in death [12]. Diagnostic and control strategies for babesiosis, introduced in recent years, are still based on treatment with toxic chemotherapy agents and/or vaccination with live attenuated strains of the parasite [13, 14]. However, drawbacks associated with drug residues, rapidly developing drug resistance, poor induced protection against challenge with virulent parasites, and reversion of live attenuated strains to virulence, have prompted a search for more effective ways of dealing with babesiosis [15]. An effort is now being made to develop multivalent recombinant subunit vaccines that target parasite multiplication within the host and, more precisely, the steps involved in erythrocyte

invasion [16]. Potential vaccine candidates include Rhoptry Associated Protein-1 (RAP-1), Apical Membrane Antigen-1 (AMA-1), and Spherical Body Proteins (SBP-1, 2, 3), all of which are secreted or exocytosed by the apical complex organelles (rhoptries, micronemes and spherical bodies or dense granules) to achieve erythrocyte invasion. These candidates have been identified and characterized mainly in *B. bigemina* and *B. bovis* [17, 18]. One such protein, the rhoptry protein RAP-1, is an immunogenic protein whose role in red blood cell invasion was inferred from a series of erythrocyte adhesion or invasion trials and a test of *Babesia* growth inhibition in vitro using RAP-1 antibodies, although the precise function of RAP-1 is still unknown [3, 19-24].

Rap-1 genes have been identified in all *Babesia* species examined to date [25-33]. The organization of these *rap-1* genes is characterized by the presence of a tandem arrangement of multiple gene copies. A simple arrangement with only two almost identical copies of *rap-1a* is found in *B. bovis* [27]. In contrast, the *rap-1* locus in *B. bigemina* and in the sheep parasite *Babesia* sp. BQ1 (Lintan) contains three types of *rap-1* genes, (*rap-1a*, *rap-1b* and *rap-1c)*, with multiple copies of polymorphic *rap-1a* and conserved *rap-1b* arranged in tandem [29, 33]. All babesial RAP-1a proteins exhibit well-defined molecular features such as strict conservation of 4 cysteine residues at the N-terminus, a 14 amino acids motif and several shorter conserved motifs, and the presence of a signal peptide [34-37].

The aim of our study was to amplify and sequence the *rap-1* genes present in the sheep parasite *Babesia* sp. Xinjiang. As multiple genes were detected, their transcription was also analyzed by standard reverse transcription PCR. Implications in terms of gene evolution, diagnostics and vaccine development are discussed.

Materials and methods

Parasite

A monoclonal line of *Babesia* sp. in Xinjiang was derived from in vitro culture by limiting dilution, as described previously [38].

Genomic DNA extraction

Parasite genomic DNA was extracted from in vitro cloned cultivated parasites using a genomic DNA Purification Kit (Gentra, USA), according to the manufacturer's instructions. The amount of isolated DNA, estimated photometrically, was 100 ng/μl.

Cloning and sequencing of the central region of *Babesia* sp. Xinjiang *rap-1* genes

Rap-1 partial sequences of *Babesia* sp. Xinjiang were initially amplified by PCR using the degenerate primers *rap-1*F2/R3 (Table 1). These primers were designed from conserved motifs (PLSLPNPYQLDAAF and YKTYL) of *rap-1* genes from ruminant-infecting *Babesia* species [33, 37]. Amplification of a central *rap-1* region of about 250 bp was expected. The amplification conditions, using Phusion High-Fidelity DNA polymerase (Thermo Scientific), were: initial denaturation for 2 min at 98°C, followed by 40 cycles of 10s at 98°C, 1 min 30s at 50°C, 1 min 30s at 72°C, and final extension for 10 min at 72°C. Once the presence of an amplicon of the expected size had been detected on agarose gel, the crude PCR product was cloned into Zero-Blunt vector (Zero Blunt PCR Cloning Kit, Invitrogen), then transformed into TOP10 *E. coli* cells (according to the manufacturer's instructions).

Three different types of *rap-1* genes (*rap-1a*, *rap-1b* and *rap-1c*) can be found in the *rap-1* locus of some species, namely *B. bigemina* and *Babesia* sp. BQ1 (Lintan) [29, 33]. As the above degenerate primers could not amplify the *rap-1b* type of gene [33], we designed new degenerate *rap-1b*-BXJ-F/R primers using two known *rap-1b* sequences to amplify the central region of the *rap-1b* gene (Table 1). PCR amplification was performed as described above, and the PCR products were cloned and sequenced.

Cloning and sequencing of *rap-1* 3' and 5' ends and intergenic regions

As *rap-1* genes are known to be tandemly repeated in a head to tail arrangement, specific primers (*rap-1a*-BXJ-inverse F/R) were designed from the unique *rap-1* internal sequence obtained previously, and used outwards from the internal region to amplify the ends and intergenic regions of the genes (Table 1). Multiple PCR products were expected, with sizes of about 3 kb (no intercalating gene), 5 kb (one intercalating gene) or more (depending on the number of *rap-1* genes in the locus), based on the average published lengths of *rap-1* genes and their intergenic regions. Two PCR amplicons were obtained by using Ex *Taq* Hot-Start DNA polymerase (TaKaRa): a major amplicon of about 2.7 kb and a minor one of about 5.4 kb. The crude PCR product was cloned into pGEM-T Easy vector. Then vector primers and successively designed internal primers (Table 1-F1-seq, F2-seq and R-seq) were used to sequence the entire insert.

Table 1. List and sequences of primers used in this study.

Application	Primer name	Forward primer 5'-3'	Reverse primer 5'-3'
Central region	*rap-1*F2/R3	CCHTACCARTTGGAYGCTGCSTTC	GTAAGTYTTGTARTWSRTGGTRGYCAT
	rap-1b-BXJ	ACTGGATTGYGCCYTCGAG	AATAWGWYTTGTATTGGAYAGC
5', 3' ends IG regions	*rap-1a*-BXJ-inverse	GTACCTCGTAAGCAAGCTCC	TCCACGTCAGTAGCATCAGG
Internal sequencing	F-seq/R-seq	F1 GAGGTATCTCAAGAATAAGGC F2 CAGCTACGTTGATGCCCCGT	GCGAAGTGAGGACAGTGTCG
Full-*rap1a*	*rap1a*-full	ATGAAGGCTTTCATCAGCGT	GCCTTATTCTTGAGATACCTC
Variable region *rap-1aα*	*rap1a*-F-875/*rap1aα*-R ***rap1a*-F-915**/*rap1aα*-R	GCTACGTTGATGCCCCGTGG **GTTCAAGGAGTACCTTGCC**	GCAGGAATCTCCTGCCACC
Variable region *rap-1aβ*	*rap1a*-F-875/*rap1aβ*-R ***rap1a*-F-915**/*rap1aβ*-R		CCTCTTTAGCTCTGGGCAATAC
IG after α type IG after β type	*rap1aα*-F/*rap1a*-R *rap1aβ*-F/*rap1a*-R	GGTGGCAGGAGATTCCTGC GTATTGCCCAGAGCTAAAGAGG	CGAATACGCTGATGAAAGCC
Tandem arrangement of gene copies	*rap1aα*-F/ *rap1aβ*-R (α-β) *rap1aβ*-F/ *rap1aβ*-R (β-β)		CCTCTTTAGCTCTGGGCAATAC
	rap1aα-F/*rap1aα*-R (α-α) *rap1aβ*-F/ *rap1aα*-R (β-α)		GCAGGAATCTCCTGCCACC

Amplification and sequencing of whole *rap-1* genes

Putative start and stop sites of *rap-1* genes were determined from the sequences obtained for the 3' and 5' ends and the intergenic regions. Primers located at the ends of each putative gene were designed (*rap1a*-full-F/R) and used to amplify entire *rap-1* genes with the Phusion High-Fidelity DNA polymerase (Thermo Scientific). After cloning in Zero-Blunt vector and selecting colonies with inserts of the right size, sequencing was performed with the vector primers (M13F/M13R).

A region of variable size, depending on the insert, and a polymorphism at the 3' end of the *rap-1* genes were discovered, so primers were designed to specifically amplify the variable region and estimate the number of different *rap-1* gene copies. Two sets of primers framing the variable region were designed and used in four different combinations (Table 1, variable region). PCR was carried out with Phusion High-Fidelity DNA polymerase, and the number of cycles was limited to 30 to reduce PCR chimera production [39, 40].

Sequencing the intergenic regions

Since multiple *rap-1* genes were discovered, the intergenic regions were investigated for sequence variability. Two PCRs were performed using the primer combinations *rap1aα*-F/*rap1a*-R and *rap1aβ*-F/*rap1a*-R (Table 1). Cloning and sequencing were performed as already described.

Tandem arrangement of *rap-1* gene copies in the genome

Because numerous different copies of *rap-1* genes were discovered, their presence was analyzed as genes arranged in tandem at the same locus, rather than at separate loci. Specific primers of two different gene types were used in four different combinations to amplify the different tandemly-arranged copies (Table 1, locus part).

Transcription analysis of *rap-1a* genes

Total RNA was extracted from an in vitro culture of *Babesia* sp. Xinjiang by standard TRIzol reagent protocol (Life Technologies), and treated with DNAse I (Invitrogen). cDNA was then synthesized using the Superscript III First Strand Synthesis System for RT-PCR kit (Invitrogen). A reverse transcription reaction was performed following the

manufacturer's recommendations for first-strand synthesis using an Oligo (dT) and random hexamers primers. We named cDNA from Oligo (dT) sample A and cDNA from random hexamers sample B. Absence of DNA contamination in the RNA extracts A and B was controlled using the primer combinations *rap1aα*-F/rap1a-R and *rap1aβ*-F/rap1a-R. These two primer pairs could amplify a 1.3 kb fragment of intergenic region from genomic DNA, but nothing from cDNA.

The PCR was then performed on cDNA samples A and B using Ex *Taq*, and the genomic DNA was used as template to check the presence of the expected multiple amplicons. Two primer combinations (*rap1a*-F-875/*rap1aα*-R and *rap1a*-F-875/*rap1aβ*-R) were used to amplify the variable size regions of the different copies, and thus the transcription analysis of different *rap-1* copies, in a single reaction. The primer combination *rap1a*-F-875/*rap1aα*-R was used to amplify the *rap-1aα* type and the *rap1a*-F-875/*rap1aβ*-R primer combination was used to amplify the *rap-1aβ* type. These primers are located on either side of the variable region. A PCR reaction with 30 cycles was performed to limit chimera production during the reaction.

Bioinformatics analysis

Two different software for signal peptide prediction were used to search for the presence of putative signal peptides (SignalP, http://www.cbs.dtu.dk/services/SignalP/; PrediSi, http://www.predisi.de/home.html). Multiple sequences alignment was analyzed on the website http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::clustalw-multialign using the software ClustalW 2.0.12: Clustalw: Multiple alignment. Phylogenetic analysis was performed with MEGA 5.2 software. B-cell epitope prediction was performed on the website http:// tools.immuneepitope.org/tools/bcell/iedb_input.

Results

Babesia sp. Xinjiang genome contains a *rap-1a* gene

Amplification of the *rap-1* gene central region using the degenerate primers RAP-1F2/R3 (Table 1) produced three different amplicons (200, 250 and 700 bp). An amplicon of about 250 bp had been expected for the central *rap-1* gene region. The crude PCR product was cloned and 94 colonies with the expected insert size were selected by colony PCR. Inserts from 16 clones had an identical sequence of 229 bp, which blasted with *B.*

bovis, *B. ovis, B. caballi* and *B. bigemina rap-1a* sequences on the NCBI website (http://www.ncbi.nlm.gov/blast/cgi). This preliminary result suggested the presence of a single *rap-1a* gene. The *rap-1b* degenerate primers RAP-1b-BXJF/R produced an amplicon of unexpected size (700bp). This amplicon was nevertheless cloned and its sequence was confirmed to be unrelated to any *rap-1* gene.

The multiple *rap-1a* gene copies at the *Babesia* sp. Xinjiang *rap-1* locus are arranged in tandem

PCR amplification of the extremities of the *rap-1* gene with the primer combination *rap-1a*-BXJ-inverse F/R (Table 1) produced a 2.7 kb amplicon that could correspond in size to the 3' end of a *rap-1a* copy, the *rap-1a* intergenic region and the 5' end of the following *rap-1a* copy. Direct colony PCR with vector primers (T7/Sp6) was performed on 20 colonies after cloning and transformation, and two clones with the expected 2.7 kb fragments were selected and sequenced. For both 2.7 kb sequences, the 3' and 5' sequences outwards from the central region blasted with *rap-1a* genes on the NCBI website, indicating the presence of at least two tandemly-arranged *rap-1a* copies. Since the two inserts had sequences differing in size by 36 bp, the presence of at least a third *rap-1a* copy was hypothesized. Otherwise, the sequences in both inserts were identical over about 2100 bp upstream of the central region (including the intergenic region) and over 376 bp downstream. A region of variable size and sequence could already be identified in the 3' end of these putative genes.

A putative start site was found in the 5' region (559 bp from the primer *rap-1a*-BXJ-inverse R) and a putative stop site in the 3' region (847 bp or 883 bp from the primer *rap-1a*-BXJ-inverse F), thus delimiting putative genes of 1407 bp or 1443 bp and intergenic regions of 1362 bp.

The *Babesia* sp. Xinjiang *rap-1a* gene is present in multiple and polymorphic copies

The polymorphism of the multiple *rap-1a* gene copies was addressed by designing primers to amplify the full *rap-1a* gene from the conserved 5' and 3' ends of the putative genes. The 1.4 kb amplicon was cloned and the inserts from 34 selected clones were sequenced. Alignments were obtained manually, as the different intermediate regions were difficult to align automatically due to the large size differences and the presence of almost

identical sequences within these regions.

Three distinct regions in the different gene copies could be distinguished: a perfectly conserved 5' region of 936 bp, an intermediate region of variable length and sequence, and a relatively conserved 3' region of 147 bp with two groups of sequences differing by 28 point mutations (Figure 1). The two *rap-1a* gene copy types were named *rap-1aα* and *rap-1aβ,* according to the two sets in the 3' variable region.

Table 2. Sequence of repeats on nucleotide and amino acid level.

Repeat number	Nucleotidic sequence	Peptidic sequence	*rap-1a* copy type
1	ATCGCCATTCCAACAAAAGACTTCTTTGAGAACAAG	IAIPTKDFFENK	α and β
2	ATCGGTGCTCCCACCAAAGATTTCTTTGAGAACAAG	IGAPTKDFFENK	α and β
3	ATCGGTGCTCCCACCAAAGATTTCTTCGAGAACAAG	IGAPTKDFFENK	α and β
4	ATCGGTGCTCCCACCAAAGACTTCTTCGAGAACAAG	IGAPTKDFFENK	β
5	TTAGGACCTCGTGCCAAAGATTTCTTTGAAAACAAG	LGPRAKDFFENK	β
6	ATTAGCGCTCCTACCAAAGTTTTCTTCGAGAACAAG	ISAPTKVFFENK	β
7	GTAGGACCTCGTACCAAAGATTTCTTTGAAAACAAG	VGPRTKDFFENK	α and β
8	GTTGGTGCTCCCACAAAAGATTTCTTTGAAAACAAG	VGAPTKDFFENK	β
9	TTGGGCCCTCAGGCCAAAAATTTTTTCGAGAGTAAA	LGPQAKNFFESK	β last
10	ATTGGCGCTCCTACCAAGGACTTCTTTGAAAACAAG	IGAPTKDFFENK	α
11	ATTGGCGCTCCTACCAAGGACTTCTTTGAGAACAAG	IGAPTKDFFENK	α
12	ATTGGTGCTCCCACCAAGGACTTCTTTGAAAACAAG	IGAPTKDFFENK	α
13	ATTGGTGCTCCCACCAAGGAATTTTTCCAGAACAAG	IGAPTKEFFQNK	α last

Table 3. Arrangement of the repeats in the variable region of *rap-1a* gene copies in *Babesia* sp. Xinjiang. The repeats sequence types are indicated with numbers corresponding to table 2, and identical colors used for different repeat numbers indicate identical amino acid sequences.

Repeat type and arrangement										Number of repeats	*rap-1a* copy name	Gene length
1	2	3	3	7	10	11	12	11	13	10	*rap1-aα1*	1443 bp
1	2	3	3	3	7	10	11	13		9	*rap1-aα2*	1407 bp
1	2	3	3	7	10	11	13			8	*rap1-aα3*	1371 bp
1	3	3	3	7	10	11	13			8	*rap1-aα4*	1371 bp
1	13									2	*rap1-aα5*	1155 bp
1	2	3	4	5	6	7	8	9		9	*rap1-aβ1*	1407 bp
1	4	5	6	7	8	9				7	*rap1-aβ2*	1335 bp

```
rap-1aα2  ATGAAGGCTT TCATCAGCGT ATTCGTTAGC TCTTTCTTGC TTGTAGCAAG GAGTGCTTAC GCTGTTCGTC ACTATCAGAA TGTTACCTCT CTGTCTCCTT
rap-1aβ1  ATGAAGGCTT TCATCAGCGT ATTCGTTAGC TCTTTCTTGC TTGTAGCAAG GAGTGCTTAC GCTGTTCGTC ACTATCAGAA TGTTACCTCT CTGTCTCCTT
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  CGGAGGTCAT CGGTAATGTG AGTGAGAGCT TGGCAGCTGG CGACACTGTC CTCACTTCGC AAATGACCGC CCGCGAAATC AGCCACGACA TGAGGAATTA
rap-1aβ1  CGGAGGTCAT CGGTAATGTG AGTGAGAGCT TGGCAGCTGG CGACACTGTC CTCACTTCGC AAATGACCGC CCGCGAAATC AGCCACGACA TGAGGAATTA
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  CTTGATGAGC GTACGCGAAG GCTTTGTTGA TGAAATATGC ATGCAAGTAC CTCAAGATGC CAAGTGCGCC GAAGCTGTGA GCGCCTACGT TAGCCGCTGT
rap-1aβ1  CTTGATGAGC GTACGCGAAG GCTTTGTTGA TGAAATATGC ATGCAAGTAC CTCAAGATGC CAAGTGCGCC GAAGCTGTGA GCGCCTACGT TAGCCGCTGT
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  ATGTCCTACG GCTGCCTGAA GGTTGACAGC GTAAGGTACC CTATGGATGC TGAATACCAG CCACTCGTTC TCCCCAACCC CTACCAGTTG GATGCCGCCT
rap-1aβ1  ATGTCCTACG GCTGCCTGAA GGTTGACAGC GTAAGGTACC CTATGGATGC TGAATACCAG CCACTCGTTC TCCCCAACCC CTACCAGTTG GATGCCGCCT
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  TCACCCTATT CAGGAACAGC GAGTCCAACC CTTCTAAGAA TGTTATTAAG CGCCCATGGA TGCGTTTCCG CAATGGAGGC AGCCACGGTG ACTACCACAA
rap-1aβ1  TCACCCTATT CAGGAACAGC GAGTCCAACC CTTCTAAGAA TGTTATTAAG CGCCCATGGA TGCGTTTCCG CAATGGAGGC AGCCACGGTG ACTACCACAA
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  CTTCTTGGTT AGCCTAATTA GTGTGAGCAT CATTCGGGAA CCTGATGCTA CTGACGTGGA GTACCTCGTA AGCAAGCTCC TCTTCATGGC TACTCTCTAC
rap-1aβ1  CTTCTTGGTT AGCCTAATTA GTGTGAGCAT CATTCGGGAA CCTGATGCTA CTGACGTGGA GTACCTCGTA AGCAAGCTCC TCTTCATGGC TACTCTCTAC
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  TACAAGACTT ACCTCATTGT GGATGCAACC AAGGCAAAGA TCTTCAACAA ATTGTCCTTC GCCACCCACA TTTTCGGTTC TGGCATCAGG AAGGTACTGA
rap-1aβ1  TACAAGACTT ACCTCATTGT GGATGCAACC AAGGCAAAGA TCTTCAACAA ATTGTCCTTC GCCACCCACA TTTTCGGTTC TGGCATCAGG AAGGTACTGA
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  AGACCATCGT AAGGTCGAAC GTTCCTGAGG ATGTCGGAGA GTACAGCATC GAGCGTATTC GTCATCTCAC CGGCAGCTAC AGCGACTACA TGTTGACCCA
rap-1aβ1  AGACCATCGT AAGGTCGAAC GTTCCTGAGG ATGTCGGAGA GTACAGCATC GAGCGTATTC GTCATCTCAC CGGCAGCTAC AGCGACTACA TGTTGACCCA
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********

rap-1aα2  CGTTCCCCAG CTCTCATCGT TTGCTCACCG CTACTCACGT ATGGTTGTAA AGGTAGTCCT TTCTGCCCTT ACCAGCTACG TTGATGCCCC GTGGTACAAG
rap-1aβ1  CGTTCCCCAG CTCTCATCGT TTGCTCACCG CTACTCACGT ATGGTTGTAA AGGTAGTCCT TTCTGCCCTT ACCAGCTACG TTGATGCCCC GTGGTACAAG
          ********** ********** ********** ********** ********** ********** ********** ********** ********** **********
```

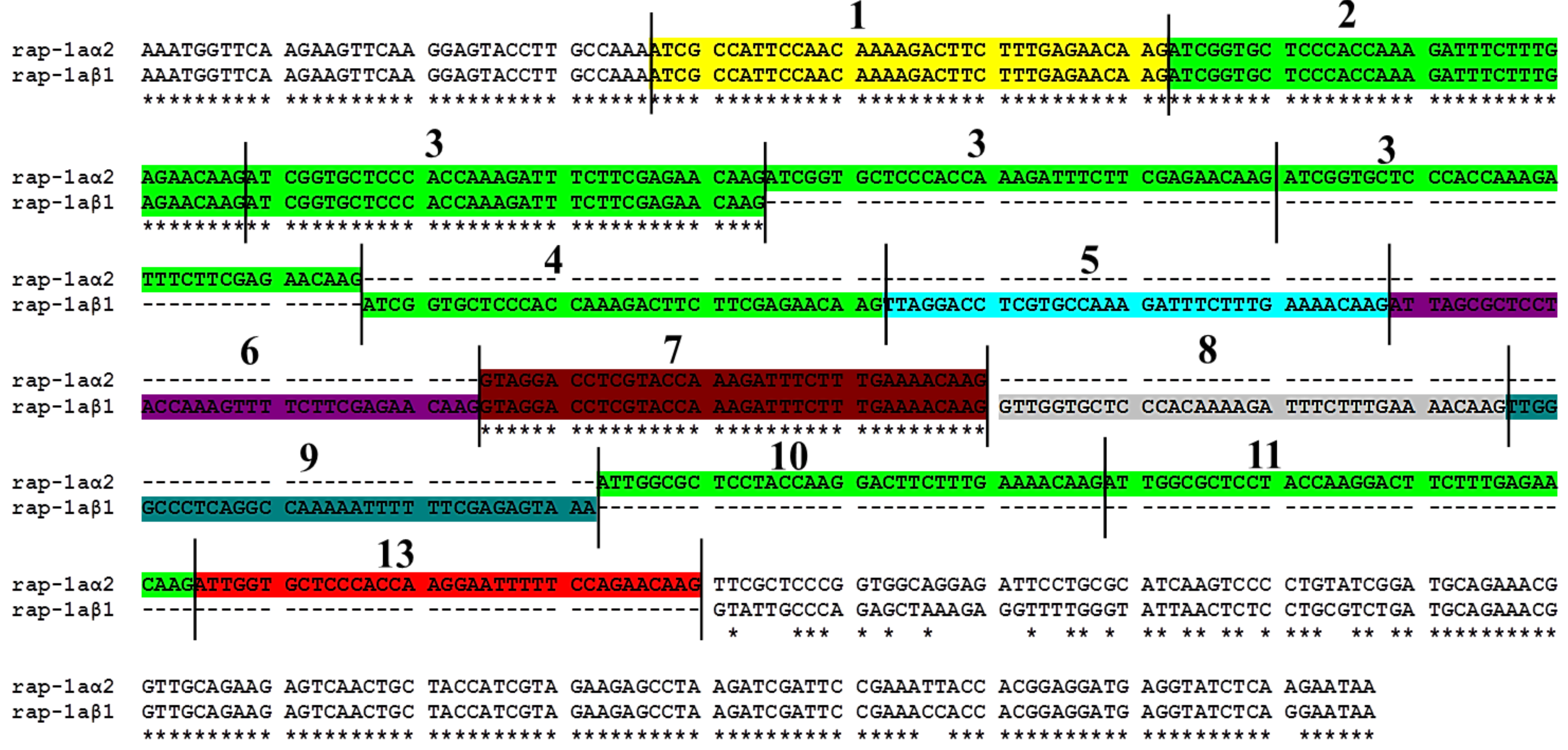


Figure 1.Nucleotide sequence alignments of the *α* and *β* types of *rap-1a* genes. The repeat numbers and colors correspond to those used in table 3.

The different *rap-1a* gene copies differ in the number and sequence of 36 nt repeats (variable region)

The number and polymorphism of *rap-1a* gene copies in *Babesia* sp. Xinjiang were further examined by designing specific *rap-1aα* and *rap-1aβ* primers to frame the variable region. PCR cycles were limited to 30 to reduce the potential risk of generating chimera, associated with the use of multi-template DNA. Two primer combinations were used for each gene copy type (*rap-1aα* and *rap-1aβ*) which generated amplicons differing by 40 bp. Multiple amplicons were obtained then separated and visualized on 2% agarose gel (Figure 2). For *rap-1aα*, three amplicons with approximate sizes of 500 bp and comparable intensities were visualized as well as a smaller one of about 150 bp. For *rap-1β*, two main amplicons approximately 400 bp in size were visualized, the smaller one being of greater intensity. Each gene copy type (*α* and *β*), showed the same pattern with the two primer combinations, and the expected size differences. Reproducibility of the PCR amplification was checked by performing three different PCRs (two are shown in Figure 2). This variable region was further analyzed for size and sequence by cloning and sequencing the crude PCR products for each of the four primer combinations.

Ninety seven sequences were analyzed, 53 for the *rap-1aα* gene type and 44 for the *rap-1aβ* gene type. Manual alignments revealed that the variable region consisted of a variable number of 36 nt repeats. The variable region in the different *rap-1aα* copies consisted of 2 repeats (72 bp), 8 repeats (288 bp), 9 repeats (324 bp) or 10 repeats (360 bp). The *rap-1aβ* copies showed arrangements of 7 (252 bp) or 9 repeats (324 bp). The total length of the putative *rap-1a* gene copies therefore varied between 1155 and 1443 bp.

Analysis of the repeat sequences revealed closely related or identical sequences in the different gene copies, which explained the difficulties encountered with automatic alignment. Thirteen different sequences of 36 nt were found. Four of these sequences were only found in *rap-1aα*, 5 only in *rap-1aβ*, and 4 in both gene copy types (Table 2). The sequences between the repeats varied from 1 nt to 15 nt. Seven different *rap-1a* gene copies could be distinguished from the sequencing data i.e. 5 copies of *rap-1aα* (*α1* to *α5*) and 2 copies of *rap-1aβ* (*β*1 and *β*2). The arrangement of the repeats in the 7 different sequences is shown in Table 3. In general, each repeat occurred only once in a putative gene, except repeat 3 which could be repeated up to three times consecutively.

Thirteen of the 97 sequences analyzed were found to be unique and differed either in the sequence of one repeat or arrangements of the different repeats. These sequences were discarded as representing potential polymerase replication errors or chimera production during PCR, respectively.

The repeats appeared to be more highly conserved when translated into amino acids. Six of the 13 nucleotidic repeats encoded the same peptide IGAPTKDFFENK, despite up to 7 substitutions, all of them in the third codon position (Table 2). Eight different peptide repeats could therefore be distinguished, with 4 residues conserved in all repeated sequences, and 3 positions with only two possible residues. The most highly conserved residues were located at the second 3' half of the repeats. Depending on the arrangement of the different nucleotide repeats, the putative *rap-1aα* proteins consisted of a succession of 5 to 7 strictly identical peptides separated by peptide repeat 7, which had a closely related sequence (3 different residues in the 5' end of the repeat). The variable region in the different gene α copies consisted of only 4 different peptide repeats. Consequently, RAP-1aα3 and RAP-1aα4 were identical at the amino acid level. The peptide sequences of the repeats in *rap-1aβ* were more diverse, with 7 different peptide repeats, 4 of them unique to the *rap-1aβ* type. The *α* and *β* types could therefore be distinguished by differences in the 147 bp 3'-end region, and by *β*- specific peptide repeats in the variable region.

All sequences obtained in this study were submitted to GenBank under the following accession numbers: KF811193 (*rap-1aα1*), KF811194 (*rap-1aα2* and IG), KF811195 (*rap-1aα3*), KF811196 (*rap-1aα4*), KF811197 (*rap-1aα5*), KF811198 (*rap-1aβ1*) and KF811199 (*rap-1aβ2*).

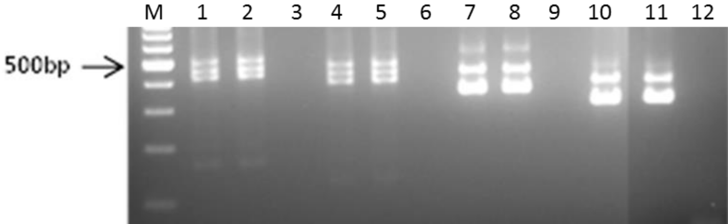


Figure 2. Amplification of the variable length region of *rap-1aα* and *rap-1aβ* gene types. Lanes 1-6: r*ap-1aα;* lanes 7-12: *rap-1aβ*. Lanes 1 and 2: two separate amplifications using primers *rap1a*-F-875/*rap1aα*-R, lane 3: blank control; lanes 4 and 5: duplicate amplifications using primers *rap1a*-F-915/*rap1aα*-R, lane 6: blank control; lanes 7 and 8: duplicate amplifications using primers *rap1a*-F-875/*rap1aβ*-R, lane 9: blank control; lanes 10 and 11: duplicate amplifications using primers *rap1a*-F-915/*rap1aβ*-R, lane 12: blank control. M: 100 bp

The two *rap-1a* gene types are located at the same locus

To determine whether *rap-1aα* and *rap-1aβ* were located at the same locus, the *rap-1a* type specific primers were used in different combinations to amplify the tandem arrangements *α* to *α* (*rap1aα-F/rap1aα-R*), *α* to *β* (*rap1aα-F/rap1aβ-R), β* to *β* (*rap1aβ-F/rap1aβ-R*) and *β* to *α* (*rap1aβ-F/ rap1aα-R*). Amplicons were obtained with the expected size of 2.6 kb except for the *β to β* arrangement, despite numerous attempts (data not shown). Sequencing of the 2.6 kb amplicons, obtained for each primer combination after cloning, revealed the tandem arrangements: *α-α1*, *α-β1*, and *β-α2.* We could conclude that the two *rap-1a* gene types were located at the same locus and arranged in tandem.

Conserved intergenic regions exist in the *Babesia* sp. Xinjiang *rap-1* locus

Each intergenic region was then analyzed for sequence variability (11 clones for the intergenic region located downstream from *rap-1aα*, and 11 clones for the intergenic region downstream from *rap-1aβ*) using the primer combinations *rap1aα*-F/*rap1a*-R and *rap1aβ*-F/*rap1a*-R, respectively. Intergenic regions of 1362 bp were obtained in all cases (Supplementary figure 1). Three point mutations were detected: one located 311 bp after the putative stop site, another at 858 bp and the third at 1006 bp. Four combinations of the substitutions were detected downstream from *rap-1aα*: TAA, TAC, CAC and CAA, whereas only two combinations (TAA and TCA) were found downstream from *rap-1aβ*. TAA combinations had already been found between *α* and *α1*, CAC between *α* and *β1*, and TCA between *β* and *α2* (sequencing data from section 3.5).

Phylogenetic analysis

A phylogenetic tree, based on all babesial RAP-1 and recently published RAP-1 Related Antigen (RRA) sequences deposited in GenBank, was constructed by neighbor-joining method using the software MEGA5.2. RAP-1a sequences from the RAP-1aα2 and RAP-1aβ1 gene types of *Babesia*. sp Xinjiang were selected for this tree. The RAP-1 sequence from *Theileria annulata* was included as an outgroup (figure 3). Results indicated that the RAP-1a sequences from *Babesia* sp. Xinjiang formed a clade with the RAP-1a sequences from *B. ovis*, *B. bovis* and *B. orientalis*, and clustered with all the babesial RAP-1a type sequences (for an explanation of *B. gibsoni rap-1* gene nomenclature, see [33]).

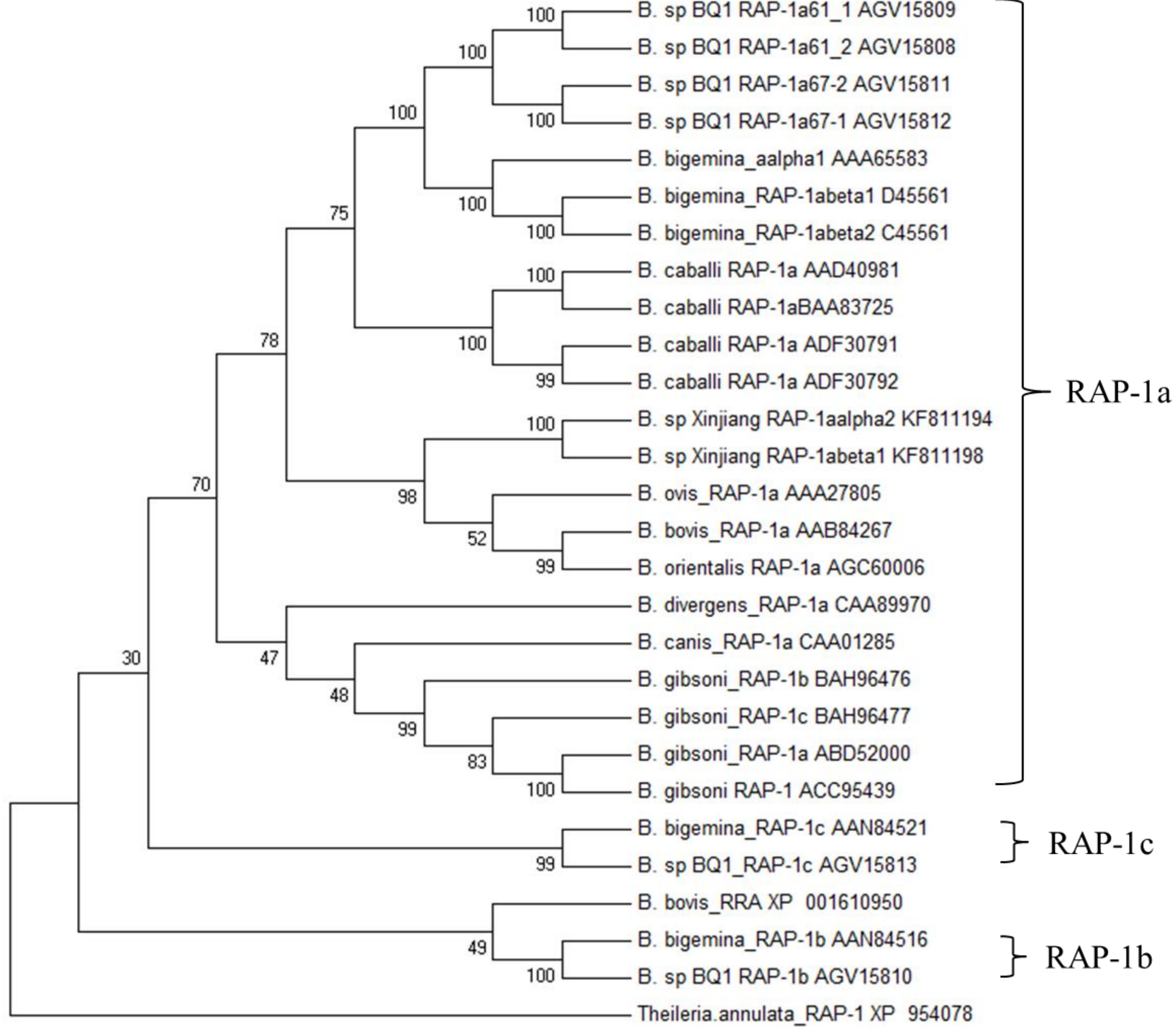


Figure 3. Phylogenetic tree based on the amino acid sequences of *Babesia* sp. Xinjiang RAP-1aα2, RAP-1aβ1 (nucleotide sequence accession number) and of all known members of the RAP-1 family in *Babesia*. The tree was inferred using the neighbor joining method, bootstrap values of MEGA 5 are shown at each branch point. Numbers above the branch demonstrate bootstrap support from 1000 replications. The *Theileria annulata* RAP-1 sequence was used as outlier.

At least 5 different copies of *rap-1a* are transcribed in vitro

To determine whether both *rap-1aα* and *rap-1aβ* gene copies were transcribed in vitro, combinations of the specific *α* and *β* primers were used to frame the variable region (*rap1a*-F-875/*rap1aα*-R and *rap1a*-F875/*rap1aβ*-R). Primers amplifying the intergenic region starting from the two neighboring genes were included in the PCR to control DNA contamination in the RNA extracts and indicated the absence of DNA from both RNA preparations after PCR amplification (data not shown). Identical RT-PCR results were obtained for two cDNA sample preparations (A and B) (Figure 4, lanes 2-3 and 6-7). The amplification patterns for the genomic DNA (Figure 4, lanes 1 and 5) and RNA extracts (Figure 4, lanes 2-3 and 6-7) were also identical with three amplicons corresponding to *rap-1α* copies and two amplicons corresponding to *rap-1aβ1* and *rap-1aβ2*. It was deduced that at least three *rap-1aα* gene types, plus *rap-1aβ1*and *rap-1aβ2,* were transcribed in vitro.

Conserved motifs that might be involved in the regulation of *rap-1a* transcription were researched in the intergenic region and aligned with the corresponding regions in different *Babesia* species (Figure 5). The three boxes, described by Suarez et al. in 1998 and known as the -59 box, the -36 box and the mRNA box, were found with a high degree of sequence similarities.

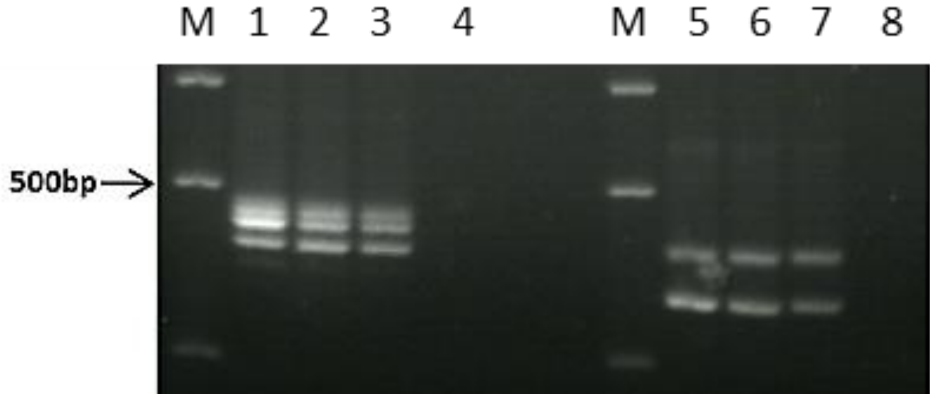


Figure 4. Transcription analysis of *Babesia* sp. Xinjiang *rap-1a* genes. RT-PCR with *rap-1aα* primers *rap1a-F-875/rap1aβ-R* for repeats amplification, lane 1-4: gDNA, cDNA (A), cDNA (B), blank; RT-PCR with *rap-1aβ* primers *rap1a-F-875/rap1aα- R* for repeats amplification, lane 5-8: gDNA, cDNA (A), cDNA (B), blank. M: 5000 bp molecular marker.

	-59 box	-36 box	mRNA box
1	TCGCACTTAGCTGCA	GAGGTGCAG	TATAGCAGTGCTATATA
2	TCGCACTTNTTTGCA	AAGGTGCAC	TATAGCAGTGCTATATA
3	TCGCACTCGCTTGCA	GAGGTGCAG	TATAGCAGTGCTATATA

Figure 5. Conservation of the putative *rap-1a* transcription regulation 5'UTR sequence in the *rap-1a* intergenic regions. The position of boxes -59, -36 and mRNA are designated according to [27]. 1: *Babesia* sp. Xinjiang; 2: Consensus from [27]; 3: *B. ovis*. The *Babesia* sp. Xinjiang underlined sequences correspond to sequence identical to the consensus.

Bioinformatic putative B epitopes mapping

The *rap-1a* gene sequences were further analyzed by bioinformatic prediction of linear B epitopes (website http:// tools.immuneepitope.org/tools/bcell/iedb_input). The localization, lengths, and scores of the different predicted B epitopes are given in Figure 6 for two RAP-1a proteins, namely RAP-1aα2 and RAP-1aβ1. The major epitope was located at the 3' end of the protein, with a very long stretch of 35-37 residues. Interestingly, most of the peptides corresponding to the repeated blocks were predicted to be putative B epitopes, whatever the peptide sequences (IGAPTKDF, VGPRTKDF, LGPQ for example).

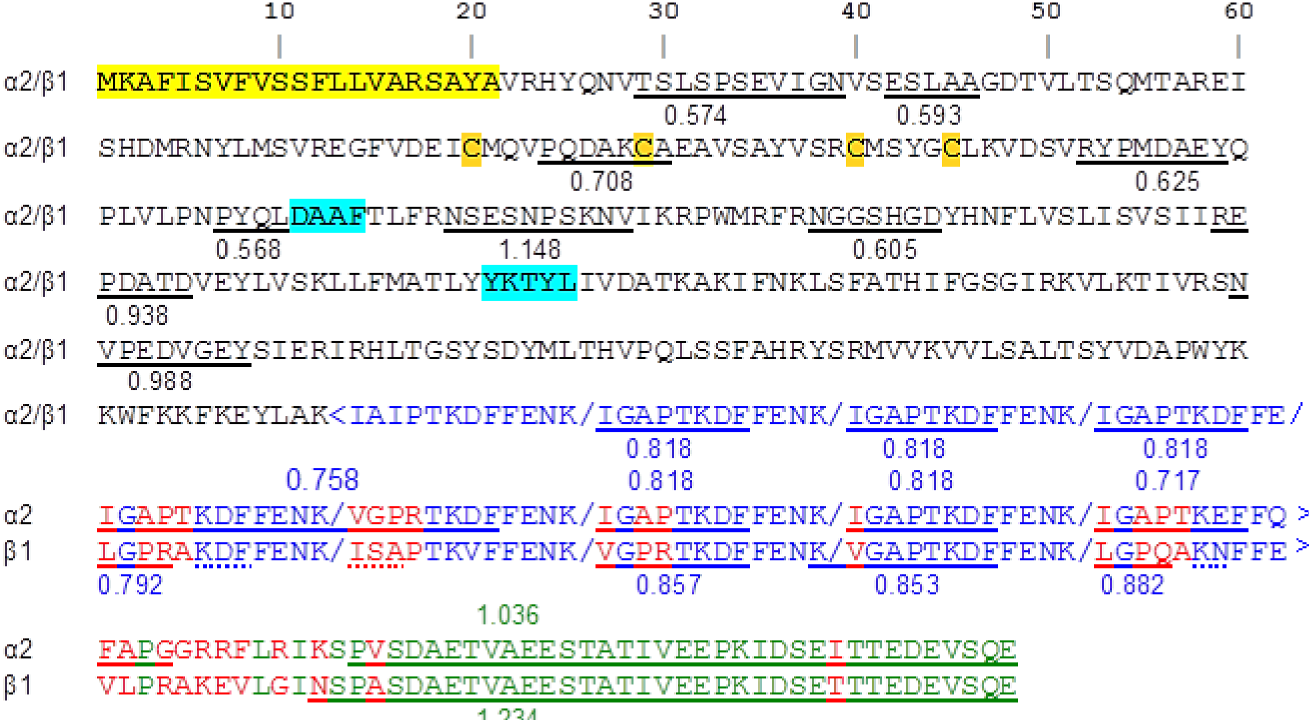


Figure 6. Amino acid alignment of RAP-1aα2 and RAP-1aβ1 and B epitope prediction by using method "Bepipred Linear Epitope Prediction". The 5' conserved region is indicated in black, the variable region with repeats in blue and the 3' end in green. When the two sequences are identical, only the consensus is indicated as α2/β1. Amino acids that differ between the two sequences are indicated in red. The variable region is delimited with < > and each repeat separated by /. The signal peptide sequence is highlighted in yellow, the 4 conserved cysteines in orange and the conserved motifs in light blue. Predicted B epitopes are underlined and the mean score for each epitope is indicated either bellow or over the sequence, when α2 and β1 sequences are dissociated. A score threshold of 0.35 was used and an epitope is defined as at least 4 successive amino acids with a score value over 0.35. Shorter putative B epitopes are underlined with a dotted line as they are in a repeat or close to a longer epitope.

Discussion

In this study, we sequenced *rap-1* genes from the sheep hemoparasite *Babesia* sp. Xinjiang. Seven copies, all of the *rap-1a* gene type, were sequenced and separated into 2 distinct types, named *rap-1aα* (5 copies) and *rap-1aβ* (2 copies) according to the 3' end sequence. These copies, despite having perfectly conserved 5' sequences, varied in sequence and size due to the 3'-end presence of a variable number of more or less degenerated repeats of 36 bp. The head to tail arrangement of the *α* and *β* gene types, and the conserved intergenic regions between them were demonstrated. Transcription of at least 5 of the 7 copies from *α* and *β* types was also demonstrated. As highlighted by B epitope mapping, the 3' end of the protein including the repeats region exhibited high antigenicity.

Rap-1 genes belong to a multigene family with fewer members than the well-known *B. bovis* families, such as *ves* (about 150 members) and *smorf* (44 members) [41, 42]. As yet, multiple members have only been demonstrated in a few species (2 in *B. bovis* [27], 4-5 in *B. ovis* [30], 11 in *B. bigemina* [29] and 12 in *Babesia* sp. BQ1 [33]), and strongly suspected for others (*B. divergens* and *B. caballi*) [37, 43]. Although members of multigene families are usually located on the same locus, some can be scattered on different chromosomes, such as the members of the *smorf* family [42]. Even if the tandem head to tail arrangements of the two types have been demonstrated in *Babesia* sp. Xinjiang, scattered localization of the tandem copies, as in the *smorf* members of *B. bovis*, cannot be excluded.

Whereas the *B. bovis* genome contains two identical copies of the *rap-1a* gene, the genomes of the two sheep parasites *B. ovis* and *Babesia* sp. Xinjiang have been shown to carry at least 4 and 7 copies respectively. Duplicated genes may be fixed in a population if parasite fitness is increased by having more than one gene copy [44]. As the regulatory regions localized in the intergenic regions are also duplicated in *Babesia* sp. Xinjiang, the gene dosage is probably increased and seems to be advantageous for parasite survival. Moreover, as demonstrated in *B. bovis*, the head to tail arrangement with two intergenic regions significantly increases protein expression [45]. Thus, the multiple tandemly arranged copies of *rap-1a* are probably maintained to cope with the need for large amounts of these proteins at specific stages of intraerythrocytic development. In the case of *Babesia* sp. Xinjiang, we show that at least 5 *rap-1a* copies were transcribed from both *α* and *β* gene types. Our purpose in this study was not to quantify the transcripts from different genes, but only to show that they were expressed. We did not study *rap-1* expression in

Babesia sp. Xinjiang, since the design of peptides to produce the required protein-specific antibodies is precluded by the close relatedness of the different gene copies. In the *B. bovis ves* family, only one gene copy is expressed at a time as a mechanism of immune evasion [41], whereas in the *smorf* family, several smorf proteins are being expressed at any one time [42].

The presence of two different *rap-1* gene types, *α* and *β,* differing mainly in the 3' repeated part of the terminal region, was demonstrated in our study. The strong sequence conservation of *α* type gene copies is probably the result of concerted evolution through a intergenic gene conversion mechanism. Gene conversion is favored in multigene families and results in sequence homogenization between the different gene copies [44]. At one point in evolution, the *β* gene type seems to have been excluded from the gene conversion process and has accumulated point mutations that could have led to the acquisition of a slightly different and novel function compared to the *α* gene type. A similar trend has been observed in the *rap-1* gene locus of *B.ovis*. Four *rap-1a* gene copies have almost identical sequences, differing in the number of 36 bp repeats. In the fifth gene copy, the entire sequence is different, with only 72% identity in the 5' more conserved region, the absence of clear repeats, and no significant identity in the 3' sequence [30]. The accumulation of point mutations in this fifth copy is probably due to an absence of conversion of this gene in *B. ovis*, whereas in *Babesia* sp. Xinjiang, gene conversion probably continues to homogenize the two identified *β* gene copies, and thereby reduce the effective mutation rate.

Multigene families play a crucial role in the biology of parasites, facilitating their persistence and mediating functions such as invasion of the host cell and immune escape. The 7 *rap-1a* gene copies differ mainly in the 3' region, due to the presence of a different number (2 to 10) of more or less degenerate 36 bp repeats (59% to 100% sequence similarity). The presence of repeats in coding sequences is not rare, with about 14% of all proteins containing one or more repeats [46], and a higher proportion of proteins with repeats occurring in intra-cellular Protozoa than in extracellular or free-living Protozoa [47]. In *Babesia* sp. Xinjiang *rap-1* gene copies, these repeats can be classified as high-complexity repeats, due to their length and complex amino acids composition [46]. These repeats do not exhibit the usual characteristics, namely the under-representation of large and/or hydrophobic amino acids (with the exception of leucine), and over-representation of small and/or polar amino acids and the amino acids involved in N- or O-glycosylation (Asn, Ser, Thr) [46-48]. The repeats are more conserved in the *α* type *rap-1a* copies than in

the β type, and the most prevalent repeat sequence (IGAPTKDFFENK) is encoded by 6 nucleotidic sequences that differ by 7 different codons at the third position. Four amino acids (the two basic Lys and the two non polar Phe) appear to be highly conserved and even if the amino acid of the repeat changes, it tends to be replaced by another with the same polarity. Analysis of repeats in the proteome of Protozoa revealed a lower ratio of degenerate/perfect repeats in protein families involved in host cell adhesion, which suggests a role of perfect repeats in this function. The same study showed an opposite tendency for the proteins involved in immune evasion and antigenic variation [47], the diversification of sequences being imposed by the immune system. Thus, the high conservation of the repeats in the *Babesia* sp. Xinjiang *rap-1a* gene, as well as in the gene copies, suggests that such repeats may be linked to functional constraints such as protein-protein interaction or adhesion. The IGAPTKDFFENK repeat is not only conserved in *Babesia* sp. Xinjiang *rap-1a* genes, but also in the *rap-1a* sequences of another sheep parasite, *B. ovis* [30], which belongs to a sister phylogenetic group. Whereas this amino acid motif is repeated and conserved in *Babesia* sp. Xinjiang and *B. ovis*, the remaining portion of their *rap-1a* genes differs in both the 5' usually more conserved region (66% (nt) and 56.7% aa) as well as in the 3' terminal region (49% (nt) and 20% aa). This level of conservation suggests that such repeats have a similar function in both pathogens. However, B-epitope mapping suggests that most of these repeats correspond to putative B-epitopes, as well as the long 3' stretch of mostly polar amino acids (67% of the amino acids content in this region). Tandem repeat proteins have been shown to be immunodominant in *Trypanosoma cruzi*, as well as in *Leishmania infantum* and *Plasmodium* [49-51], and the number of repeats in *T. cruzi* FRA protein has been found to increase antibodies affinity but not antigenicity [52]. The role of repetitive regions of tandem repeat proteins as a decoy to divert the immune response from functional regions has long been suggested. In the case of RAP-1a proteins, the function might be linked to the more conserved 5' region of the protein, and the 3' end including the repetitive regions would act as a smoke screen. However, with such a role, a greater genetic diversity of the repeats would be expected, and not their conservation in different species. The role of RAP-1 proteins in the general process of red blood cell invasion by *Babesia* parasites has been demonstrated. Their erythrocyte-binding activity and their role in parasite attachment to the erythrocytes have been demonstrated [23, 24], but their precise role in invasion remains to be elucidated [18]. Comparative sequence analysis of *rap-1a* sequences from different *Babesia* species may

help in targeting the conserved parts of the protein potentially involved in protein-protein interactions.

As mentioned earlier, protein repeats are usually extremely immunogenic and can therefore be useful in serological diagnostics [51, 52]. In the case of *Babesia* sp. Xinjiang *rap-1* genes, the IGAPTKDFFENK motif could be a good target to develop an ELISA based test since it bears a putative B-epitope, it is repeated at least 27 times in the different gene copies, and in some gene copies, the repeats are contiguous, a configuration that has been shown to increase antigenicity [52]. However, the repeated region in the *rap-1a* gene of another sheep parasite, *B. ovis*, contains an identical motif [30]. Such a diagnostic tool could be used in China, because of the different geographical distributions of these two sheep parasites [53], but it would not be possible to differentiate between the two parasites. Other putative B-epitopes localized in the 5' conserved region or in the more specific 3' end terminal part of the protein could be targeted, as 7 copies of the target would, in any case, improve the level of serological detection. However, the recognition of such peptides by the sera of naturally infected animals first needs to be analysed, to determine if they are immunogenic and represent interesting potential diagnostic targets. *Rap-1* has been used as a marker for both the serological and molecular diagnosis of cattle babesiosis [54, 55], *B. gibsoni* infection in dogs [32] and *B. caballi* infection in horses [56]. In the serological detection of cattle babesiosis, the C-terminal region of the protein proved to be more specific than the entire protein [57, 58].

The role of RAP-1 in erythrocyte invasion establishes this protein as a potential vaccine candidate. Even if RAP-1 is recognized by sera from *B. bovis* infected cattle that are immune to challenge, and even if it is recognized by memory CD4+-T helper cells with a predominant response against aa 114 to 316, it is still insufficient to provide protective immunity [19, 59]. However, this absence of protection could be due to structural differences between recombinant and native proteins. A study of *B. divergens* vaccine development also showed that the presence of hydrophobic moieties in recombinant proteins is a critical factor to induce a protective immune response [60]. The development of an effective recombinant vaccine will require not only a better understanding of the mechanism of immunity, but also the use of a combination of proteins, RAP-1 probably being one of them.

Acknowledgements

This study was supported financially by a grant from the European Union cooperation project PIROVAC (KBBE-3-245145), the «Région Pays de la Loire» (France), and INRA.

References

1. Telford 3rd SR, Gorenflot A, Brasseur P, Spielman A (1993) Babesial infection in human and wildlife. In: Kreier, J.P. (Ed.), Parasitic Protozoa vol. 5. Academic Press, San Diego, CA, pp. 1-47.

2. De waal DT (2000) Global importance of piroplasmosis. J Protozool Res.10: 106-127.

3. Brown WC, Palmer GH (1999) Designing blood-stage vaccines against *Babesia bovis* and *B. bigemina*. Parasitol Today 15: 275-281.

4. Uilenberg G (2006) *Babesia*-a historical overview. Vet Parasitol 138: 3-10.

5. Liu AH, Yin H, Guan GQ, Schnittger L, Liu ZJ, et al. (2007) At least two genetically distinct large *Babesia* species infective to sheep and goats in China. Vet Parasitol 147:246-251.

6. Niu Q, Luo J, Guan G, Liu Z, Ma M, et al. (2009) Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. Exp Parasitol 121: 64-68.

7. Oosthuizen MC, Allsopp BA, Troskie M, Collins NE, Penzhorn BL (2009) Identification of novel *Babesia* and *Theileria* species in South African giraffe (*Giraffa camelopardalis*, Linnaeus, 1758) and roan antelope (*Hippotragus equinus*, Desmarest 1804). Vet Parasitol 163: 39-46.

8. Jouglin M, Fernandez-de-Mera IG., De la Cotte N, Ruiz-Fons F, Gortazar C, et al (in press) Isolation and characterization of *Babesia pecorum* sp. nov. from farmed red deer (*Cervus elaphus*). Vet Res

9. Guan GQ, Yin H, Luo JX, Lu WS, Zhang QC, et al. (2001) Isolation of a large ovine *Babesia* sp. in Xinjiang, China (in Chinese). Chinese Journal of Veterinary Science of Technology 31: 35-36.

10. 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.

11. Guan G, Ma M, Liu A, Ren Q, Wang J, et al. (2012) A recently identified ovine *Babesia* in China: serology and sero-epidemiology. Parasitol Int 61: 532-537.

12. Wright IG, Goodger BV (1988) Pathogenesis of Babesiosis. pp, 100-118. In. M. Ristic. (ed.), Babesiosis of domestic animals and man. CRC. Press. Boca Raton. Fla.

13. Wright LG (1990) Immunodiagnosis of and immunoprophylaxis against the haemoprotozoa *Babesia* spp. and *Anaplasma* spp. in domestic animals. Rev Sci Tech OIE 9: 345-356.

14. Kuttler KL (1990) Pharmacotherapeutics of drugs used in treatment of anaplasmosis and babesiosis. J Am Vet Assoc 176: 1103-1108.

15. De Waal DT, Combrink MP (2006) Live vaccines against bovine babesiosis. Vet Parasitol 138: 88-96.

16. Gohil S, Herrmann S, Günther S, Cooke BM (2013) Bovine babesiosis in the 21st century: Advances in biology and functional genomics. Int J Parasitol 43:125-132.

17. Yokoyama N, Okamura M, Igarashi I (2006) Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. Vet Parasitol 138: 22-32.

18. Lobo CA, Rodriguez M, Cursino-Santos JR (2012) *Babesia* and red cell invasion. Curr Opin Hematol 19: 170-175.

19. Norimine J, Mosqueda J, Suarez C, Palmer GH, McElwain TF, et al. (2003) Stimulation of T-helper cell gamma interferon and immunoglobulin G responses specific for *Babesia bovis* rhoptry-associated protein 1 (RAP-1) or a RAP-1 protein lacking the carboxy-terminal repeat region is insufficient to provide protective immunity against virulent *B. bovis* challenge. Infect Immun 71: 5021-5032.

20. Brown WC, McElwain TF, Hötzel I, Suarez CE, Palmer GH (1998) Helper T-cell epitopes encoded by the *Babesia bigemina rap-1* gene family in the constant and variant domains are conserved among parasite strains. Infect Immun 66: 1561-1569.

21. Brown WC, McElwain TF, Palmer GH, Chantler SE, Estes DM (1999) Bovine CD4(+) T-lymphocyte clones specific for rhoptry-associated protein 1 of *Babesia bigemina* stimulate enhanced immunoglobulin G1 (IgG1) and IgG2 synthesis. Infect Immun 67: 155-164.

22. Figueroa JV, Buening GM (1991) In vitro inhibition of multiplication of *Babesia bigemina* by using monoclonal antibodies. J Clin Microbiol 29: 997-1003.

23. Mosqueda J, McElwain TF, Stiller D, Palmer GH (2002) *Babesia bovis* merozoite surface antigen 1 and rhoptry-associated protein 1 are expressed in sporozoites, and specific antibodies inhibit sporozoite attachment to erythrocytes. Infect Immun 70: 1599-1603.

24. Yokoyama N, Suthisak B, Hirata H, Matsuo T, Inoue N, et al. (2002) Cellular localization of *Babesia bovis* merozoite rhoptry-associated protein 1 and its erythrocyte-binding activity. Infect Immun 70 5822-5826.

25. Goff WL, Davis WC, Palmer GH, McElwain TF, Johnson WC, et al. (1988) Identification of *Babesia bovis* merozoite surface antigens by using immune bovine sera and monoclonal antibodies. Infect Immun 56: 2363-2368.

26. Wright IG, Casu R, Commins MA, Dalrymple BP, Gale KR, et al. (1992) The development of a recombinant *Babesia* vaccine. Vet Parasitol 44: 3-13.

27. Suarez CE, Palmer GH, Hötzel I, McElwain TF (1998) Structure, sequence, and transcriptional analysis of the *Babesia bovis rap-1* multigene locus. Mol Biochem Parasitol. 93: 215-224.

28. Kappmeyer LS, Perryman LE, Hines SA, Baszler TV, Katz JB, et al. (1999) Detection of equine antibodies to *Babesia caballi* by recombinant *B. caballi* rhoptry-associated protein 1 in a competitive-inhibition enzyme-linked immunosorbent assay. J Clin Microbiol 37: 2285-2290.

29. Suarez CE, Palmer GH, Florin-Christensen M, Hines SA, Hötzel I, et al. (2003) Organization, transcription, and expression of rhoptry associated protein genes in the *Babesia bigemina rap-1* locus. Mol Biochem Parasitol 127: 101-112.

30. Dalrymple BP, Casu RE, Peters JM, Dimmock CM, Gale KR, et al. (1993) Characterisation of a family of multi-copy genes encoding rhoptry protein homologues in *Babesia bovis*, *Babesia ovis* and *Babesia canis*. Mol Biochem Parasitol 57: 181-192.

31. Skuce PJ, Mallon TR, Taylor SM (1996) Molecular cloning of a putative rhoptry associated protein homologue from *Babesia divergens*. Mol Biochem Parasitol 77: 99-102.

32. Zhou JL, Jia HL, Nishikawa Y, Fujisaki K, Xuan XN (2007) *Babesia gibsoni* rhoptry-associated protein 1 and its potential use as a diagnostic antigen. Vet Parasitol 145: 16-20.

33. Niu Q, Bonsergent C, Guan GQ, Yin H, Malandrin L (2013) Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group). Vet Parasitol 198: 24-38.

34. Suarez CE, Palmer GH, Jasmer DP, Hines SA, Perryman LE, et al. (1991a) Characterization of the gene encoding a 60-kilodalton *Babesia bovis* merozoite protein with conserved and surface exposed epitopes. Mol Biochem Parasitol 46: 45-52.

35. Suarez CE, McElwain TF, Stephens EB, Mishra VS, Palmer GH (1991b) Sequence conservation among merozoite apical complex proteins of *Babesia bovis, Babesia bigemina* and other apicomplexa. Mol Biochem Parasitol 49:329-332.

36. Suarez CE, Thompson SM, McElwain TF, Hines SA, Palmer GH (1994) Conservation of oligopeptide motifs in rhoptry protein from different genera of erythroparasitic protozoa. Exp Parasitol 78: 246-251.

37. Dalrymple BP, Peters JM, Böse R, Wright IG (1996) A polymerase chain reaction method for the identification of genes encoding members of the Bv60/p58 family of rhoptry protein homologues in the genus *Babesia*. Exp Parasitol 84: 96-100.

38. Malandrin L, Jouglin M, Moreau E, Chauvin A (2009) Individual heterogeneity in erythrocyte susceptibility to *Babesia divergens* is a critical factor for the outcome of experimental spleen-intact sheep infections.Vet Res 40: 25.

39. Meyerhans A, Vartanian JP, Wain-Hobson S (1990) DNA recombination during PCR. Nucleic Acids Res 18: 1687-1691.

40. Kanagawa T (2003) Bias and artifacts in multitemplate polymerase chain reactions (PCR). Biosci Bioeng 96: 317-323.

41. 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.

42. Ferreri LM, Brayton KA, Sondgeroth KS, Lau AO, Suarez CE, et al (2012) Expression and strain variation of the novel "small open reading frame" (*smorf*) multigene family in *Babesia bovis*. Int J Parasitol 42:131-138.

43. Bhoora R, Quan M, Zweygarth E, Guthrie AJ, Prinsloo SA, et al (2010) Sequence heterogeneity in the gene encoding the rhoptry-associated protein-1 (RAP-1) of *Babesia caballi* isolates from South Africa. Vet Parasitol 169:279-288.

44. Innan H (2009) Population genetic models of duplicated genes. Genetica 137:19-37.

45. Suarez CE, Palmer GH, LeRoith T, Florin-Christensen M, Crabb B, et al. (2004) Intergenic regions in the rhoptry associated protein-1 (*rap-1*) locus promote exogenous gene expression in *Babesia bovis*. Int J Parasitol 34:1177-11184.

46. Marcotte EM, Pellegrini M, Yeates TO, Eisenberg D (1999) A census of protein repeats. J Mol Biol. 293:151-60.

47. Mendes TA, Lobo FP, Rodrigues TS, Rodrigues-Luiz GF, daRocha WD, et al. (2013) Repeat-enriched proteins are related to host cell invasion and immune evasion in parasitic protozoa. Mol Biol Evol 30:951-63.

48. Fankhauser N, Nguyen-Ha TM, Adler J, Mäser P (2007) Surface antigens and potential virulence factors from parasites detected by comparative genomics of perfect amino acid repeats. Proteome Sci 5:20.

49. Goto Y, Carter D, Reed SG (2008) Immunological dominance of *Trypanosoma cruzi* tandem repeat proteins. Infect Immun 76:3967-3974.

50. Goto Y, Coler RN, Guderian J, Mohamath R, Reed SG (2006) Cloning, characterization, and serodiagnostic evaluation of *Leishmania infantum* tandem repeat proteins. Infect Immun 74:3939-3945.

51. Lima-Junior JC, Jiang J, Rodrigues-da-Silva RN, Banic DM, Tran TM, et al. (2011) B cell epitope mapping and characterization of naturally acquired antibodies to the *Plasmodium vivax* merozoite surface protein-3α (PvMSP-3α) in malaria exposed individuals from Brazilian Amazon.Vaccine 29:1801-1811.

52. Valiente-Gabioud AA, Veaute C, Perrig M, Galan-Romano FS, Sferco SJ, et al. (2011) Effect of repetitiveness on the immunogenicity and antigenicity of *Trypanosoma cruzi* FRA protein. Exp Parasitol 127:672-679.

53. Yin H, Lu WS, Luo JX, Zhang QC, Lu WX, Dou HF (1997) Isolation and morphological observation of *Babesia motasi* and *Babesia ovis* in China. Chin J Vet Sci Technol 10:7-9. [In Chinese].

54. Mtshali MS, Mtshali PS (2013) Molecular diagnosis and phylogenetic analysis of *Babesia bigemina* and *Babesia bovis* hemoparasites from cattle in South Africa. BMC Vet Res 9:154.

55. Ybañez AP, Sivakumar T, Ybañez RH, Vincoy MR, Tingson JA, et al. (2013) Molecular survey of bovine vector-borne pathogens in Cebu, Philippines. Vet Parasitol 196:13-20.

56. Awinda PO, Mealey RH, Williams LB, Conrad PA, Packham AE, et al. (2013) Serum antibodies from a subset of horses positive for *Babesia caballi* by competitive enzyme-linked immunosorbent assay demonstrate a protein recognition pattern that is not consistent with infection. Clin Vaccine Immunol 20:1752-1757.

57. Boonchit S, Xuan X, Yokoyama N, Goff WL, Waghela SD, et al. (2004) Improved enzyme- linked immunosorbent assay using C-terminal truncated recombinant antigens of *Babesia bovis* rhoptry-associated protein-1 for detection of specific antibodies. J Clin Microbiol 42:1601-1604.

58. Boonchit S, Alhassan A, Chan B, Xuan X, Yokoyama N, et al. (2006) Expression of C-terminal truncated and full-length *Babesia bigemina* rhoptry-associated protein 1 and their potential use in enzyme-linked immunosorbent assay. Vet Parasitol 137:28-35.

59. Norimine J, Suarez CE, McElwain TF, Florin-Christensen M, Brown WC (2002) Immunodominant epitopes in *Babesia bovis* rhoptry-associated protein 1 that elicit memory CD4(+)-T-lymphocyte responses in *B. bovis*-immune individuals are located in the amino-terminal domain. Infect Immun 70:2039-2048.

60. Delbecq S, Hadj-Kaddour K, Randazzo S, Kleuskens J, Schetters T, et al. (2006) Hydrophobic moeties in recombinant proteins are crucial to generate efficient saponin-based vaccine against Apicomplexan *Babesia divergens*. Vaccine 24:613-621.

```
GGCAGCTTCAAAGCGTACATCGAAGACAACACCATTTCGTTAGGAAAAGTGCTCATACTCGGCTCTACGATGTAACGCTTGTG
TTTCTTCCTTTTTGATAGCTGCCAAAGTTCCATGAAACAATAACATGATTGTTTAGTAATCGGATACCCTGAGTTTAATATAA
GAAATATTGCCACCAGTTCGCGAAGATGGCTATACGAATGCAGGGTTGTCGGAAACGAGCACTATCGAAGGAAAATGGATGAT
TGCGTTTAATAATATTTAAGAACCTAAAGTTTAAATTCTTGTAAACAACACATATTTATGGYCTGAAATAATGCATATGAAGC
ATTCCTATAAGATTTGAAAAGTCAGAATGATTAAACCGTGAATACTTAAGGTTTAGTTAATTTATTAATGTTAATATTAATAG
TCATTTTATTTGTGAAGTGCTTAACGTCGCTCACTTATCGTATTATAGATATTTTGTGATATGCATTATATGTTATATTGTCG
GTTTGAACTATATGGTACAAATTAATATTTGTTCCCATGAGCGAATAGTTTATTATGTGATATATGTATGGTTTGAGTGGCGC
CTCACGTGGTCATCGCTGTTTATTATTATCGACTATGTCATATAGAATTCGGATGCCGCCATTGATACGAATTTATGGATATA
AGGTCGCACAGTTCATTAGACCAATTCCTTTGGACACAGGCAACCGAGGGGCGTAGGGTCATTTCCCTATCGAAATTTATAGA
ATATATGGCAATGTATTAAATGTAATTATATACCATTGAAATTTTGGATGTGCATTGACAACCGGTGAGTGTGTTGATATTTT
TAGACAATGTGCGAGTGCGGCACGATAMTTGAAAATACCCCTGTGACGATGTAGTCCACTGCCACACACATTTATTACGCAGT
GTTGAAGAATAATAGACAGAGGTTTATATTTTGCGTTGATGAGTAATATGTGTAGTCAGCCAAAGCCAAAAGGAATCTCGGAG
TCGATGTGTMTATTAGTTAGTGTAAATATAATTCTAGTTAGCATATGAATGTAATTTCTTGATTTTTCAAGCAATATTTTGGA
AAAGCAAGGTTGATAAAGGCTGTCTATTTTGATAAATCCACTTTTTCTAGCTGAATCGCACTTAGCTGCACTTCCTGTAATCC
TACGAGGTGCAGGTTTAGCGTGTGAAAGGGCGTCTTTCAGCTGCACACTAAACATTCAACAAACATACTGAGTCGGTAGACTC
CAATATCAATACCTGTTTCGTTAATTATAGCAGTGCTATATACGCGGTAGTGTGTTGATATCCTAGAGAAGTCATATTATTGT
CTAATAACTACACATTCATTGCCTGGTGTTAAGA
```

Supplementary figure 1. Intergenic region IG sequence (1362 bp) downstream from *rap-1aα* gene. Three point mutations C/T (Y) located 311 bp, A/C (M) located 858 bp and A/C (M) located 1006 bp after *rap-1aα* putative stop site is indicated with gray background.

GENERAL DISCUSSION

Piroplasmosis (Babesiosis and Theileriosis) is caused by tick-borne haemoprotozoan parasites of the genera *Theileria* and *Babesia*, which many species of them are highly pathogenic for ruminants (Telford et al., 1993). In China, piroplasmosis resulted in heavy economic losses in the endemic regions. China is a large country in the world with one-fifth of the world population and with the various climate zones (hot and wet in the south, cold and dry in the north). The climatic feature attracts large number of tick species and cases of tick-borne diseases are increased in China (Yin and Luo, 2007). The ruminant grazing areas are mainly distributed in the north of China including Gansu, Ningxia, Inner Mongolia, Qinghai, Sichuan and Shaanxi, and the products of ruminants are the major source of income in these regions for the local farmers and herdsmen. According to statistics, China's mutton output was 3.76 million tons in 2008. However, this output value decreased by 1.8 % over the previous year (Bureau of Veterinary, Ministry of Agriculture, People' s Republic of China 2013), in particular due to these diseases. The piroplasmosis of ruminant in China is considered to be one of the most frequent of the tick-borne diseases and to have important economic impact (Li et al., 2011). Surveys have indicated that piroplasmosis is prevalent in many provinces and more serious in lambs. The infection rate and mortality of piroplasmosis of sheep and goats reached average 28-95% and 43-85% respectively in some regions (Yin et al., 2002b; Luo and Yin, 1997; Guo et al., 2005, 2007). Small ruminants are often infected simultaneously by two or three *Babesia* and *Theileria* species in the most prevalent regions when the parasites share the same transmission vector (Sun et al., 2008; Niu et al., 2012).

The control and treatment for this disease encounter very troublesome problems due to variously complex reasons, such as: drawbacks of chemotherapy (drug-resistance, drug residues) and the lack of complete comprehension of the biology of *Babesia*, in particular the interaction between *Babesia* and its invertebrate and vertebrate hosts. Vaccine development is still at present the best option to contribute to the control of babesiosis of small ruminant in China as well as in the world. A potential vaccine antigen against babesiosis must have different properties: block the infection and the multiplication of the parasites, desirably act on merozoites but also sporozoites, induce a protective immune response, show no or weak polymorphism among geographically distant parsites. Among the few described proteins of *Babesia* sp., RAP-1 is thought to have an essential role in the process of erythrocyte invasion by Apicomplexan parasites, is expressed by merozoites and sporozoites and so, represented a potential vaccine candidate (Brown et al., 1998; Brown and Palmer, 1999; Yokoyama et al., 2006). However, polymorphism of RAP-1 described

for bovine *Babesia* could be a brake for the development of RAP-1 based vaccine (Mishra et al., 1991, 1992; Hötzel et al., 1996).

My PhD thesis is a preliminary study of the RAP-1 potential as subunit vaccine candidate against the two major prevalent sheep *Babesia* species, *B. motasi*-like (Lintan isolate) and *Babesia* sp. Xinjiang in China. Since we had no sequence data for these species, it focused on 3 points:

1- gene characterization (sequence and copy number) and locus organization,

2- analysis of *rap-1* genes polymorphism, with the few isolates available,

3- analysis of *rap-1* expression, ideally at the levels of transcription and protein expression.

The general discussion of my thesis will focus on three aspects: the taxonomic position and phylogenetic relationships of Chinese *Babesia* isolates, the function of RAP-1 and the potential of RAP-1 as vaccine candidate.

1. Taxonomy of *Babesia* spp. responsible of small ruminant in China

In China, seven *Babesia* isolates responsible of small ruminants were desribed. They were divided into two distinct groups according to the morphology, pathogenicity, transmission vectors and molecular biology data based on the 18s rRNA and ITS (Liu et al., 2007; Niu et al., 2009a). Six of them, *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu, *Babesia* sp. Madang, *Babesia* sp. Hebei, and *Babesia* sp. Liaoning belong to the *B. motasi*-like phylogenetic group that forms a sister clade with the European isolates of *B. motasi*. *Babesi*a sp. Xinjiang forms a separate clade close to other *Babesia* such as *B. ovis*, *B. bovis*, and *B. orientalis.* This clade also gathers *Babesia* parasites recently described at the molecular level in wild ruminants in South Africa as well as the newly described species *B. pecorum* isolated from deer in Spain (Oosthuizen et al., 2009; Jouglin et al., in press) (figure 1).

The *rap-1* genes belong to a multigenic family and the multiple gene copies when characterized are arranged in head to tail tandem. However the number of *rap-1* gene copies, the *rap-1* gene types as well as the *rap-1* locus organization were very different among *Babesia* species. Only two almost identical *rap-1a* genes composed a simple

arrangement in *B. bovis* (Suarez et al., 1998a) while 11 *rap-1* gene copies with 3 types (5 *rap-1a*, 5 *rap-1b* and 1 *rap-1c*) composed an extremely complex locus arrangement in *B. bigemina* (Suarez et al., 2003).

In our study, we have characterized the *rap-1* genes, intergenic regions and locus organization in four isolates of the *B. motasi*-like group (Lintan, Ningxian, Tianzhu and Hebei isolates) and *Babesia* sp. Xinjiang. We demonstrated that *rap-1* gene types and locus organization are really different between the 2 groups (table 6). Three *rap-1* gene types (*rap-1a*, *rap-1b* and *rap-1c*), were found in the *B. motasi*-like group, while only *rap-1a* gene type, was found in *Babesia* sp. Xinjiang. Even if the *rap-1a* gene type was found in both *Babesia* species, the similarity of *rap-1a* gene sequences was very low with only 32% identity between members of the two different phylogenetic groups.

Seven *rap-1a* gene copies composed the *rap-1* locus of *Babesia* sp. Xinjiang and were separated by conserved intergenic regions. The sequences of the 7 *rap-1a* copies differed only by the 3' end repeats and the sequence after the repeats to the putative stop site, the sequences showed high homologies (57% at 5'-end conserved region) with *B. ovis*, the highest identities (100% in table 7) of a repeat (IGAPTKDFFENK) were fount with *B. ovis*. Futhermore, the presence of repeats at the 3' end of the gene seems to be a common feature in several other *Babesia* species such as *B. orientalis*, *B. gibsoni*, *B. canis*, and *B. caballi* (table 7).

In *Babesia* sp. BQ1 (Lintan) and in the 3 others isolates studied of the *B. motasi*-like group, 12 gene copies with three *rap-1* gene types (6 *rap-1a, 5 rap-1b* and *1 rap-1c*) were found, arranged tandemly and separated by three different intergenic regions. This locus organization and the sequences of *rap-1a* as well as *rap-1b* and *rap-1c* are closely related to those of *B. bigemina*. Comparison between the 18S rDNA gene (figure 1 and 13) and the *rap-1* gene (types and locus organization) is consistent and indicates the relatedness between the 4 isolates of the *B. motasi*-like group, and between *B. motasi*-like and *B. bigemina*. It confirms also that *Babesia* sp. Xinjiang makes with *B. orientalis* as a sister clade very different than the precedent one (figure 13).

However, within the *B. motasi*-like group, the four isolates can be divided into two subgroups, according to the phylogenetic analysis of RAP-1, one includes *Babesia* sp. BQ1 Lintan, *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu and the other one *Babesia* sp. Hebei. The *rap-1* gene types and locus organization as well as arrangement in the locus shared the almost identical *rap-1* sequence and the same locus features among *Babesia* sp. BQ1 Lintan, *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu. The sequences of *rap-1*

genes in *Babesia* sp. Hebei were more different from the sequences of *rap-1* genes in the three others isolates and was thus placed on a branch separately. However, the sequences of RAP-1 (100%) among Lintan, Ningxian and Tianzhu isolates are even more conserved than the sequences of their 18S rDNA gene (98.7%), and the sequence of *rap-1* in Hebei isolate compared with other three isolates is much less conserved than the 18S rDNA gene (78 % and 97.7 % respectively).

Other reports on phylogenetic analyses of these isolates using other markers indicated *Babesia* sp. BQ1 (Lintan) branched in the same clade with *Babesia* sp. Tianzhu, while *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Hebei were separated and located in the same clade (28S rRNA in Gou et al., 2013 (figure 14); ITS in Niu et al., 2009 (figure 15). Recent study of Chinese *Babesia* and *Theileria* isolates phylogeny, based on molecular marker: 40S ribosomal protein 8 (RPS8, Tian et al., 2013a), cytochrome b gene (COB, Tian et al., 2013b) and the cytochrome C oxidase subunit III (COXIII, Tian et al. 2013c), did not include *Babesia* sp. Hebei in the analyses, the similar result with 28S and ITS analysis for other three isolates was demonstrated that *Babesia* sp. BQ1(Lintan) and *Babesia* sp. Tianzhu were located the same branch, while *Babesia* sp. BQ1 (Ningxian) is sperateded. Thus, the different evolution relatedness among these five isolates of *B. motasi*-like group is probably related with the differently assessed molecular marker.

Babesia sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu isolates came from the same region in China, the Gansu Province, but still separated by about 500-700 km, while *Babesia* sp. Hebei comes from a more distant region, Hebei province , at about 1500 km from the Gansu Province. Could geographic distance explain this difference among these isolates? Or do they represent two different species? Further studies are needed to verify the questions by studying their biological characteristics, such as transmitted vector, virulence.

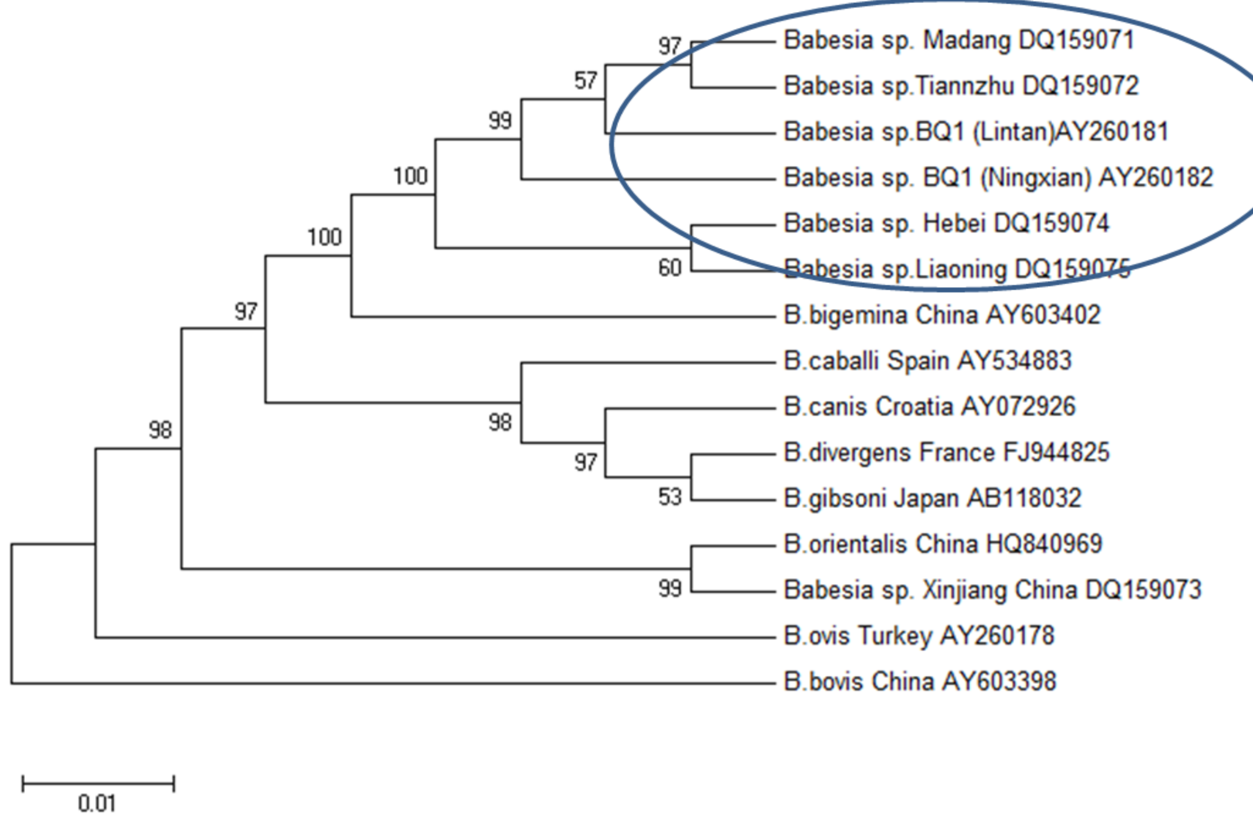


Figure 13. Phylogenetic analysis based on the 18S rRNA gene, including 18s rRNA sequences of ovine babesia representatives and other babesia species having *rap-1* sequences. The tree was inferred using the neighbor joining method of MEGA 5.2, bootstrap values are shown at each branch point. Numbers above the branch demonstrate bootstrap support from 1000 replications. The model of nucleotide substitution used corresponds to the method of "maximum composite likelihood", The scale used to represent the branch lengths is expressed in the same unit as that used to calculate the distances between sequences. All sites of the alignment containing insertions-deletions, missing data were eliminated from the analysis (option "complete deletion "). *B. motasi*-like group was indicated with oval.

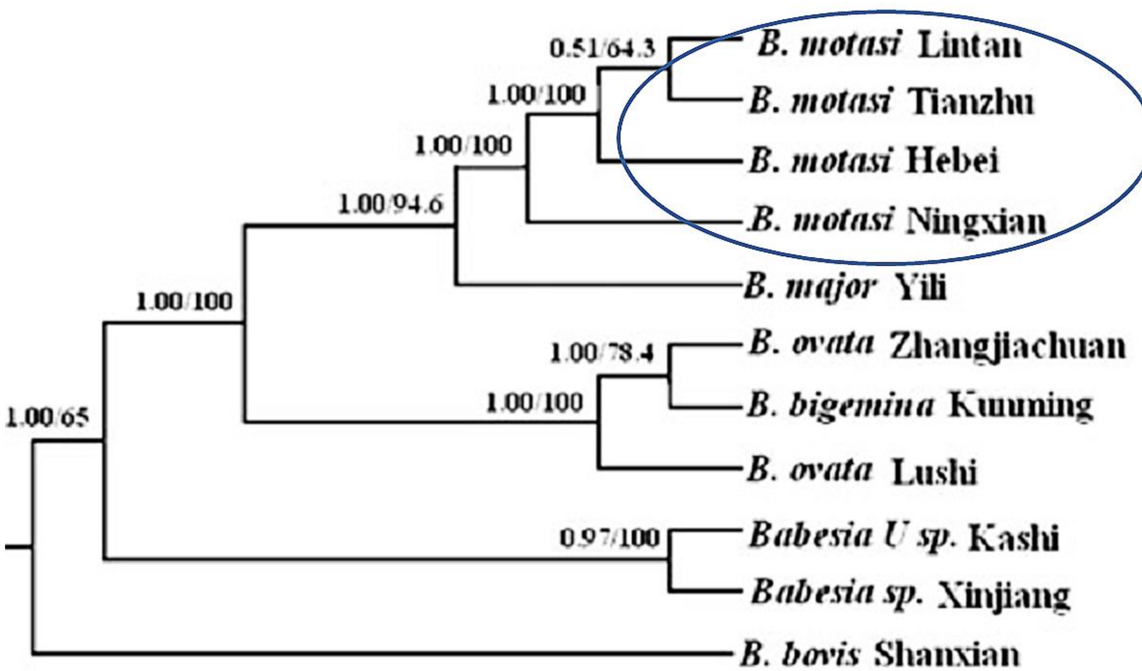


Figure 14. Phylogenetic trees of several *Babesia* isolates based on their 28S rRNA genes. This tree was calculated using the BI and MP methods (the tree cited from Gou *et al*., 2013).

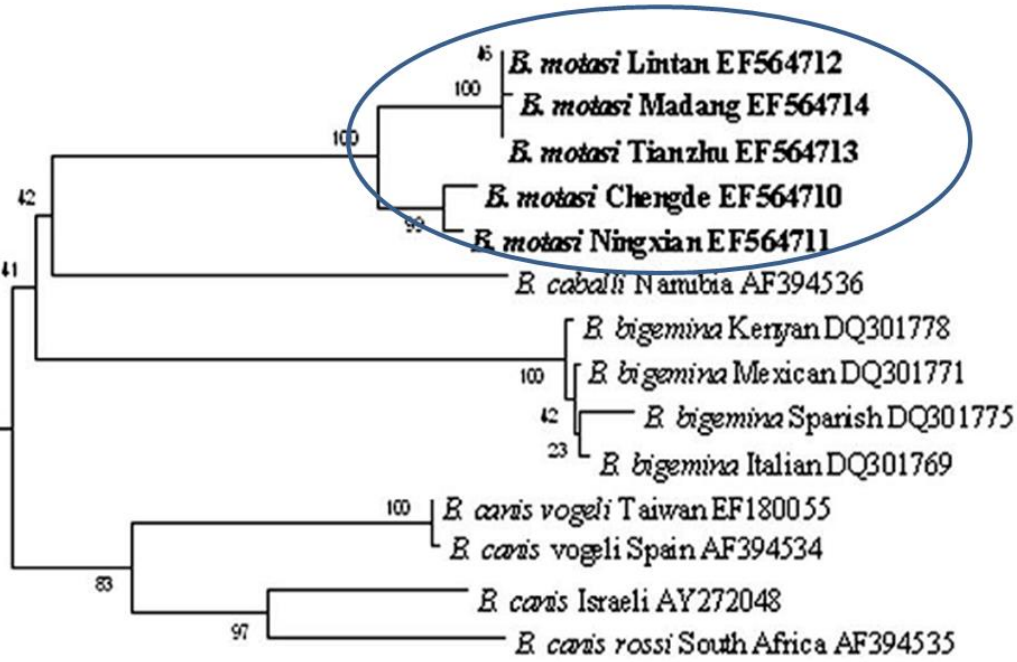


Figure 15. Phylogenetic tree depicting the relationship among several *B. motasi*-like Chinese isolates and other *Babesia* species based on ITS region sequences, the "*B. motasi* Chengde" is Hebei isolate (the tree cited from Niu *et al*., 2009a)

Table 6. The summary of *rap-1* genes, gene numbers, intergenic regions and locus in two different sheep *Babesia* species in China.

<table>
<tr><th rowspan="3">Babesia species</th><th rowspan="3">Isolates</th><th colspan="5">rap-1a</th><th rowspan="3">rap-1b</th><th rowspan="3">rap-1c</th><th rowspan="3">IG1</th><th rowspan="3">IG2</th><th rowspan="3">IG3</th><th rowspan="3">Gene numbers</th><th rowspan="3">Size of locus</th></tr>
<tr><th colspan="3">rap-1a61</th><th colspan="2">rap-1a67</th></tr>
<tr><th>a61-1</th><th>a61-2</th><th>a61-3</th><th>a67-1</th><th>a67-2</th></tr>
<tr><td rowspan="5">B. motasi-like</td><td>LT</td><td>2</td><td>2</td><td>-</td><td>1</td><td>1</td><td>5</td><td>1</td><td>5</td><td>5</td><td>1</td><td>12</td><td>≈31 kb</td></tr>
<tr><td>NX</td><td>-</td><td>2</td><td>2</td><td>1</td><td>1</td><td>5</td><td>1</td><td>5</td><td>5</td><td>1</td><td>12</td><td>≈31 kb</td></tr>
<tr><td>TZ</td><td>-</td><td>2</td><td>2</td><td>1</td><td>1</td><td>5</td><td>1</td><td>5</td><td>5</td><td>1</td><td>12</td><td>≈31 kb</td></tr>
<tr><td rowspan="2">HB</td><td colspan="3">rap-1a1</td><td>rap-1a2</td><td>rap-1a3</td><td rowspan="2">5</td><td rowspan="2">1</td><td rowspan="2">5</td><td rowspan="2">5</td><td rowspan="2">1</td><td rowspan="2">12</td><td rowspan="2">≈30 kb</td></tr>
<tr><td colspan="3">4</td><td>1</td><td>1</td></tr>
<tr><td rowspan="2">Babesia sp. Xinjiang</td><td rowspan="2">XJ</td><td colspan="3">rap-1aα</td><td colspan="2">rap-1aβ</td><td rowspan="2">-</td><td rowspan="2">-</td><td colspan="3" rowspan="2">6</td><td rowspan="2">7</td><td rowspan="2">≈18 kb</td></tr>
<tr><td colspan="3">5</td><td colspan="2">2</td></tr>
</table>

LT: Lintan
NX: Ningxian
TZ: Tianzhu
HB: Hebei
XJ: Xinjiang
IG: Intergenic region

2. Function of RAP-1

Rhoptries contain numerous RAP-1 proteins that homologs exist in the different genera of Apicomplexan parasites (*Plasmodium* sp., *Babesia* sp. …), suggesting an important function of these proteins in the parasite biology and in particular in the invasion process of host cells (Dlugonska, 2008). The erythrocytes invasion process of merozoite was thought similar between *Babesia* and *Plasmodium*. A function of RAP-1 proteins during RBC invasion was reflected by the conservation of peptide motifs between RAP-1 and others functional proteins. For example, 17 conserved peptide motifs have been found between the *Babesia* RAP-1 and *Plasmodium* AMA-1 (Pf83 and Pk66). These conserved peptide domains were thought mediating common mechanisms required for erythrocyte parasitism (Suarez et al., 1994b, Palmer and McElwain, 1995).

In the genus of *Babesia*, all members of RAP-1 family shared some well-defined common features, whatever the species of *Babesia* and its vertebrate host (ruminant, equid or dog), including sheep *Babesia* as we have shown in our study: the 4 cysteine residues at the N-terminus are strictly conserved and 14 amino acids as well as several shorter oligopeptide motifs are also conserved in the first 300 amino acid, in the N terminal region. They share overall sequence identity ranging from 30 to 45% (Suarez et al., 1991a, b, 1994; Dalrymple et al., 1996). Furthermore, RAP-1 proteins are expressed not only by merozoites but also by sporozoites and sexual stages (Mosqueda et al., 2002a; Vichido et al., 2008). The remarkable conservation of N terminal sequences between different RAP-1 within and among different *Babesia* species implied also a functionally important role of RAP-1 and in particular the N terminal region, in erythrocyte invasion process and parasite survival (Suarez et al., 1994, Hötzel et al., 1996; Norimine et al., 2003).

In contrast, variable amino acid sequences and repetitive motifs at C-terminal are observed in several RAP-1 of many *Babesia* species, including RAP-1a of *B. ovis* and *Babesia* sp. Xinjiang as well as RAP-1c but not RAP-1a and RAP1b of *B. motasi*-like isolates studied (table 7). In a recent paper, Mendes et al (2013) suggest that the repetitive regions of a protein could have important function in the cellular invasion for intracellular parasite. It could be the same for *Babesia* RAP-1a or RAP-1c in which C-terminal repetitive motifs could be also involved in the RBC penetration.

Analysis of transcription has revealed that all *rap-1* gene types in *B. motasi*-like and at least five of the seven *rap-1a* genes described in *Babesi*a sp. Xinjiang are transcribed in in vitro cultivated merozoites. This could be explained by the presence of three conserved

regulatory boxes in intergenic regions in both sheep Chinese *Babesia* (figure 16). Indeed, the transcription of *rap-1* gene was regulated by three conserved putative regulatory nucleotide motifs located in IG regions of *B. bovis* and these motifs were conserved among all the homologous IG regions of *Babesia* species (Suarez et al., 1998b).

	-59 box	**-36 box**	**mRNA box**
LT-IG1	TCGCACTTTTCTGCA	ACGGTCAAC	TATAACATAATGTACCG
HB-IG1	TCGCACTTTTCTGCA	ACGTTCAAC	TATACCATAATGTACCG
LT-IG2	TCGCACTTTGTTGCA	CAGGTGCAT	TATACGATCCGCTTAGA
HB-IG2	TCGCACTTGGTTGCA	CAGGTGCAT	TATACGATCTGCTTATA
LT-IG3	TTGCACATAGGTGCA	AGGGTGCAA	TATAGCTGTGCTATACG
HB-IG3	TTGCACATAGGTGCA	AGGGTGCAA	TATAGCTGTGCTATACT
XJ-IG	TCGCACTTAGCTGCA	GAGGTGCAG	TATAGCAGTGCTATATA
B.o-IG	TCGCACTCGCTTGCA	GAGGTGCAG	TATAGCAGTGCTATATA
Consensus	TCGCACTTNTTTGCA	AAGGTGCAC	TATAGCAGTGCTATATA

Figure 16. Conservation of the putative *rap-1* transcription regulation 5'UTR sequence in the *rap-1* intergenic regions (IG1: *rap-1a* to *b*; IG2: *b* to *a*; IG3: *a* to *c*). The position of boxes -59, -36 and mRNA are designated according to Suarez et al., 1998b. LT: *Babesia* sp. BQ1 (Lintan); HB: *Babesia* sp. Hebei; XJ: *Babesia* sp. Xinjiang; B. o: *B. ovis*; Consensus from Suarez et al., 1998b.

The expression of different RAP-1 proteins was analyzed using sera from three *Babesia* sp. BQ1 (Lintan)-infected sheep by ELISA, indicating that all *rap-1* genes are not only transcribed but also probably translated. Furthermore, RAP-1a, b and c were probably sequentially produced during infection (RAP-1c is expressed later than RAP-1a and RAP-1b). This result is really different than *B. bigemina* in which all *rap-1* genes are transcribed but only RAP-1a is produced (Suarez et al. 2003). However, in the study performed with *B. bigemina*, the in vivo expression study was performed at an early stage of infection (acute phase), while we demonstrate by ELISA a late expression in the course of infection. Using in vivo produced merozoites during the acute phase, our results by western blot perfectly corroborate those with *B. bigemina* : no RAP-1b and RAP-1c produced. In *Babesia* sp. BQ1 (Lintan) and also *B. bigemina*, the 3 RAP-1 probably have different functions, as suggested in *B. gibsoni* (Terkawi et al., 2009). Indeed, in this parasite, the 3 BgRAP-1 described were differently localized once the parasite has entered the RBC (2 of them were restricted within the merozoite whereas one was found secreted into the iRBC and in the supernatant of *B. gibsoni* culture) and only C terminal region of 2 BgRAP-1 was antigenic,

suggesting divergent roles for the 3 BgRAP-1. As RAP-1c is the only protein with repeats in C-terminal region and is mainly produced later during the infection (mainly 2-3 months after infection), RAP-1c could facilitate the *Babesia* cycle when the parasitemia is very low (detectable only by culture and not in blood smear), by increasing the interaction between parasite and host cells.

Table 7. Repeats in *rap-1* genes from different *Babesia* species. Identical amino acids repeats were found in different species, namely the IGAPTKDFFENK found in *Babesia* sp. Xinjiang and *B. ovis*. Interestingly, the motif PTKDFF (in bold in table 7) was extremely conserved and could be found in repeats from 44 different species, some of them being distant phylogenetically. B-cell epitopes were indicated with underlines.

***Babesia* sp. Xinjiang**		***B. canis***	***B. orientalis***
(313-420) IAI**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** VGPRTKDFFENK IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGAPTKEFFQNK	(313-408) IAI**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** LGPRAKDFFENK ISAPTKVFFENK VGPRTKDFFENK VGA**PTKDFFENK** LGPQAKNFFESK	CAA01285 (314-385) IHEPTKEFFVNK IHEPTKEFFVNK IHEPTKEFFVNK IHEPTKEFFVNK LHEPTKEFFVNK LHEPTKEFFSNM	AGC60006 (314-440) IGAPTKEFFVDT HEK-TKDFFENK VAKPTKEFFVNN HEK-TKDFFENK VAKPTKEFFVNN HEK-TKDFFENK VGKPTKEFFVNN HEK-TKEFLDNN VGKPTKEFFVDT HEK-TKDFFENK VAKPTKNFFSKI
***B. ovis* (60.4)** AAA27804 (336-420) MHERTKDFFENK IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** LPEHTKDFFENK IGAPIKESFENL	AAA27811(2 or 3) (336-468) MHERTKDFFENK IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** IGA**PTKDFFENK** LPERTKDFFENK LPEHTKDFFENK LPEHTKDFFENK IGAPIKESFENL	***B. bovis*** ACM44006 (314-497) VTQPTKKFIEDTNEVTKNYLKAN VAEPTKKFMQDTHEKTKGYLKEN VAEPTKTFFKEAPQVTKHFFDDN TGQPTKEFFREAPQATKHFLDEN IGQPTKEFFREAPQVTKHFLDEN IAQPTKEFFRDVPQVTKKVITEN IAQPTKEFLKEVPHTTMKVLNEN IAQPAKEIIHEFGTGAKNFISAA	
B. caballi ADF30792 (326-385) VYT**PTKDFFQNK** VGR**PTKDFFQNK** VGR**PTKDFFQNK** VGR**PTKDFFQNK** VGRPTKDFFRKA ADF30793 (326-385) VYT**PTKDFFQNK** VAE**PTKDFFQNK** VGR**PTKDFFQNK** VGR**PTKDFFQNK** VGRPTKDFFQKA	ADF30791 (326-421) VYTPTKNFFQNK VAPKTKDIFQNE IAPKTKHFIKNE LAPKTKDFLQNK VARPTKDFFENK VRHPTKDFLENK VGHPTTEFFKNK VGQPTKDFLKND AF092736 (311-485) MVNRVFIPTKKFFNKEIREPSKALKEKVSTDTKDL FENKIGQGTVDFFNKEIRDPSKALKEKVSNDAKDL FENKIGQGTVDFINNEIRDPSKALIRKVSTGAEDL FENKIGQGTVDFINNEIRDPGKALIRKVYTEADDL FENKIGQGTVDFINKEIRDPSKALIRKVSTEADNL	Isolate 418 (326-421) VYTPTKTFLQNK VGRPTRDFLQNK VGRPTKDFXQNK VGRPTKDFLQNK VGRPTRDFLQNK VGRPTKDFLQNK VGHPTKEFFENK VGQPTKDFLKNG	Isolate 105 (326-398) VYT**PTKDFFQNK** VAE**PTKDFFQNK** VGR**PTKDFFQNK** VGR**PTKDFFQNK** VAKPTKTFFKEE VAPKDKNLLEND
B. gibsoni ABD52000 (343-419) FRKAGEATKAT FQKFGEDTKAG FQKFGTATADK FQKMNEATKSG FQRFGEATKAG FQKFGEVTKAG FNKLGNVGKDG	ACC95439 (343-419) FRKAGEATKAT FQKFGEDTKAG FQRFGTDTADK FQKINEATKAK FQRFGEATKAG FQKFGEATKAG FNKLGNVGNDG	AB480715 (315-422) VHKPAKSFVFDK MYGPAKEYYWTQ VHHPFKKYYVER VRDPVKNYYWDQ IHEPAKKLYKNR IHEPVTKFYVNQ IHEPAKKFYKEG IHKPVKRLYMKG FHEPVAKLFTET	***Babesia* sp. BQ1** RAP-1c AGV15813 (420-507) KEAISIAGENKEPTIVEQ KEAINIAGENNEPSVVEQ -EAINIAGENNEPSVVEQ -EAINIAGENNEPTIVEQ KEAINIARENNEPPIVGE

Table 8. Percentage identity of RAP-1a amino acid sequences of *B. motasi*-like, *Babesia* sp. Xinjiang and *B. ovis* deduced after CLUSTAL W alignment.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
100	99.6	99.8	94.3	93.5	**32.7**	**33.1**	**34.0**	**34.0**	**39.6**	**33.5**	**35.4**	74.2	73.5	73.7	**31.4**	1	LTRAP-1a61-1 KC953701
	100	99.8	94.3	93.5	**32.9**	**33.3**	**34.2**	**34.2**	**39.8**	**33.8**	**35.6**	74.7	73.9	74.1	**31.4**	2	LTRAP-1a61-2 KC953700
		100	94.1	93.2	**32.7**	**33.1**	**34.0**	**34.0**	**39.6**	**33.5**	**35.4**	74.5	73.7	73.9	**31.2**	3	NXRAP-1a61-3 KJ205324
			100	99.3	**31.4**	**31.8**	**32.7**	**32.7**	**38.0**	**32.3**	**34.0**	71.9	71.1	71.3	**31.2**	4	LTRAP-1a67-1 KF039723
				100	**32.5**	**32.5**	**32.7**	**32.7**	**38.0**	**32.9**	**34.0**	71.5	71.5	71.5	**32.2**	5	LTRAP-1a67-2 KC953703
					100	99.4	100	100	100	94.7	94.4	**33.5**	**34.4**	**33.7**	**39.6**	6	XJRAP-1aα1 KF811193
						100	100	100	100	93.4	93.7	**33.5**	**34.8**	**33.7**	**40.8**	7	XJRAP-1aα2 KF811194
							100	100	100	95.4	94.4	**34.4**	**35.7**	**34.4**	**41.9**	8	XJRAP-1aα3 KF811195
								100	100	95.4	94.4	**34.4**	**35.7**	**34.4**	**41.9**	9	XJRAP-1aα4 KF811196
									100	96.6	96.6	**39.1**	**40.6**	**39.1**	**47.7**	10	XJRAP-1aα5 KF811197
										100	100	**34.5**	**34.6**	**34.8**	**39.7**	11	XJRAP-1aβ1 KF811198
											100	**36.7**	**36.9**	**36.7**	**41.9**	12	XJRAP-1aβ2 KF811199
												100	100	98.5	**33.3**	13	HBRAP-1a1 KJ205334
													100	98.5	**32.9**	14	HBRAP-1a2 KJ205335
														100	**33.5**	15	HBRAP-1a3 KJ205337
															100	16	B. ovis RAP-1a60.5 AAA27805
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		

LT: *Babesia* sp. BQ1 (Lintan)
NX: *Babesia* sp. BQ1 (Ningxian)
HB: *Babesia* sp. Hebei
XJ: *Babesia* sp. Xinjiang

3. RAP-1 as vaccine candidate

RAP-1 is a protein of about 60 kDa localized in the apical complex of the parasite and more specifically in the rhoptries. It has been targeted as a vaccine candidate (Brown and Palmer, 1999; Wright et al., 1992). The RAP-1 is a highly immunogenic protein and the capability of both RAP-1a of *B. bigemina* and *B. bovis* to induce partial immune protection against challenge in cattle has been demonstrated using purified native *B. bigemina* RAP-1 (McElwain et al., 1991; Rodríguez et al., 1996), *B. bovis* RAP-1 (Commins et al., 1985) and recombinant *B. bovis* RAP-1 (Wright et al., 1992; Fish et al., 2008). RAP-1 is detectable in merozoites as well as in sporozoites during the asexual intraerythrocytic cycle in *B. bovis* (Mosqueda et al., 2002a). The RAP-1 specific sera from immunized rabbit were able to neutralize the binding of sporozoites of *B. bovis* to erythrocytes effectively (Mosqueda et al., 2002a), and RAP-1 specific monoclonal antibodies blocked binding of soluble RAP-1 to merozoites and inhibited merozoite growth in vitro (Yokoyama et al., 2002). All together, these features of RAP-1 are advantageous to develop a vaccine to block early stages of infection before erythrocyte invasion.

The main brake for the development of RAP-1 based vaccine is its polymorphism. *Rap-1* genes belong to multigene family consisting of multiple gene copies tandemly arranged in all *Babesia* species (2 in *B. bovis*, 5 in *B. ovis*, 7 in *Babesia* sp. Xinjiang, 11 in *B. bigemina* and 12 in *B. motasi*-like) and several *rap-1* gene types has been identified (*rap-1a* in all *Babesia* plus *rap-1b* and *rap-1c* in *B. bigemina* and *B. motasi*-like) (Suarez et al., 1998, 2003; Dalrymple et al., 1993a; Niu et al., 2013). The *rap*-1 genes are polymorphic, for example: sequence with low identity ranging from 30 to 45% among RAP-1a genes of different *Babesia* species (Suarez et al., 1991a, b, 1994; Dalrymple et al., 1996); average identities with only 19.6% between *rap1a* and *rap-1b*, 22.6% between *rap-1b* and *rap-1c* and 27.6% between RAP-1a and RAP-1c in the *B. motasi*-like species, (similar values for *B. bigemina).*

The structure of RAP-1 proteins, especially the in N-terminal regions, as well as their sequences are rather conserved at the species level among geographically distant isolates. For example, in *B. bovis*, two identical or very similar copies of *rap-1a* are present in the locus and their polymorphism is very limited among different strains, with a maximum of 9 amino acid substitutions. In *B. bigemina*, RAP-1 variants contain blocks of sequence dimorphism in the amino (NT-1 and NT-2) and in the carboxy terminal ends (CT-1, CT-2

and CT-3). Sequences encoding NT-1 and NT-2 regions are well conserved among strains (at least 99% nucleotide identity) (Hötzel et al., 1996, 1997). The problem of polymorphism with a multigene family is not only between strains, but starts with the polymorphism within the locus among the different copies. For *Babesia* sp. Xinjiang, the sequence is identical among copies over the first 300 and the last 30 amino acids, with few differences in the sequences of the repeats. Within the *B. motasi*-like group, the polymorphism between the *rap-1a* copies is limited and absent between the *rap-1b* copies. Even geographically distant isolates from the same Province in China demonstrate an absence of polymorphism. Only for the more distant isolate *Babesia* sp. Hebei, the polymorphism is more important (73% identity between RAP-1a type, 64% between RAP-1b and 53% between RAP-1c), but as discussed before, it may belong to a different species. For the sequences of *rap-1* genes in Hebei isolate, even though the sequences in *rap-1* genes are different with the others isolates studied, the comparison of *rap-1* genes of this isolate with other *B. motasi*-like isolates highlighted many conserved motifs, most of them located in the N-terminal region. To conclude, polymorphism for RAP-1 may not be a critical brake to develop a vaccine.

B-cell epitopes in RAP-1 of *B. bovis* and *B. bigemina* have been mapped by monoclonal antibodies and immune sera from cattle, showing the B-cell epitopes mainly located in the C-terminal part of the proteins (Suarez et al., 1993; Hötzel et al., 1996, 1997). In the present study, the prediction of B-cell epitopes in all RAP-1 proteins using software also demonstrated the presence of B-cell epitopes located mainly in the polymorphic C-terminal region, including the repeats. Only few B cell-epitopes were predicted in the conserved N-terminal region. As detailed before (part 2 of the discussion), the N-terminal part of the protein is thought to bear functional RAP-1 domains and the sequence similarity of N-terminal regions among *Babesia* species supports this idea (Mishra et al., 1992). This suggests that antibodies directed against critical epitopes in this region could efficiently block erythrocyte invasion and therefore inhibit parasite multiplication and growth (Hötzel et al., 1996). However the lack of strong immunogenic B-cell epitopes in the N-terminal conserved region may represent the way the parasite persists in the face of the host humoral immune response (Suarez et al., 1993).

As this region induced a high humoral response, the variable C-terminal region among *Babesia* species could be a useful antigen for specific diagnostic in epidemiological investigation purpose. The C-terminal truncated region of RAP-1a protein was used as specific antigen to differentiate infection between *B. bovis* and *B. bigemina* in an ELISA

assay (Boonchit et al., 2002, 2004, and 2006). In contrast, N terminal region contains highly conserved motifs which could be used for vaccination. In our study, even if the RAP-1 sequences from Hebei isolate are different from other isolates, the motif of 14 aa (included “LTLPNPYEL”) was conserved. This motif was also predicted as possessing a B-cell epitope. Vaccination using this functional B epitope could induce a high antibody production which is not necessarily obtained during natural infection and which could block the invasion process of the extracellular merozoites. Futhermore, as *Babesia* is an intracellular parasite, the Th1 cellular immune response mediated by the NO produced by IFNγ-activated macrophages is essential as protective immune response (Brown et al, 2006a). Norimine et al (2002, 2003) have shown that N terminal domain and not C terminal domain of *B. bovis* RAP-1 contained immunodominant T cell epitopes antigenically conserved between different strains of *B. bovis* which stimulated Th1 cells IFNγ production.

Finally, RAP-1a and mainly its N terminal region should be targeted as vaccine candidate against sheep Chinese *Babesia*. Indeed, this protein is produced by all *Babesia*, suggesting an important role for the parasite in the RBC invasion process. It contained highly conserved motifs in different geographical strains inside *B. motasi*-like group. Furthermore, it seemed to be immunogenic and produced early during the infection (from DPI14). These features are essential to block the parasite multiplication.

CONCLUSIONS
PERSPECTIVES

The *rap-1* gene sequences, gene copy number and locus organization for *B. motasi*-like and *Babesia* sp. Xinjiang, two *Babesia* species mainly distributed in China, were described for the first time. Our study is to further clarify the relationship of Chinese ovine *Babesia* isolates for *B. motasi*-like and *Babesia* sp. Xinjiang. Based on the sequence of the *rap-1* gene, Chinese sheep *Babesia* can be divided into 2 groups, *B. motasi*-like group (*Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian), *Babesia* sp. Tianzhu, *Babesia* sp. Hebei) and *Babesia* sp. Xinjiang group, which is in agreement with the results obtained with the 18S rDNA gene, ITS and recently described 28S rDNA, gene, RPS8 (40S ribosomal protein 8), COB (cytochrome b gene) and COX3 (the cytochrome C oxidase subunit III). The gene types and locus organization of *B. motasi*-like were similar with *B. bigemina*, while *Babesia* sp. Xinjiang *rap-1a* genes showed similar features with *rap-1a* genes in *B. ovis*, including degenerated repeats in 3' end region. Multiple copies of *rap-1a* and *rap-1b* interspacing of each other by their multiple intergenic regions and a single *rap-1c* copy located at the 3' end of *rap-1* locus were determined in *B. motasi*-like, while only *rap-1a* gene types with 7 copies divided into *rap-1aα* (5 copies) and *rap-1aβ* (2 copies) occurs in *Babesia* sp. Xinjiang. The *rap-1a* genes within the *B. motasi*-like group are more conserved compared with the 4 *B. bigemina rap-1a* genes which sequences are described as dimorphic at both N- and C- terminal. The variable region between the two *rap-1a* gene types (*rap-1a61* and *rap-167*) is limited at mainly N-terminal in *B. motasi* like group, especially in *Babesia* sp. BQ1 (Lintan), *Babesia* sp. BQ1 (Ningxian) and *Babesia* sp. Tianzhu. In *Babesia* sp. Hebei, the sequences of three *rap-1a* genes were more conserved compared with other isolates *rap-1a*. The *rap-1b* genes are perfectly conserved within each isolates as well as among three isolates. In *Babesia* sp. Xinjiang, the sequences characterization of *rap-1a* gene copies was similar with *B. ovis*, which are four *rap-1a* copies almost identical with regular repeats; the fifth copy was with unobvious repeats and low identity with other four copies. The presence of clear repeats was observed in 5 *rap-1aα* of *Babesia* sp. Xinjiang, while two *rap-1aβ* copies indicated less regular repeats than *rap-1aα*.

Multigene families usually play crucial roles in the biology of parasite. The different *rap-1* gene types and their proportion in the whole locus might imply their different roles for parasites survival and different function during the invasion process as well as in host-parasite interactions.

1. The phylogenic relationship analysis based on the RAP-1 genes has been indicated the Chinese sheep *Babesia* isolates divided into two distinct groups (*B. motasi*-like and *Babesia* sp. Xinjiang), consistent with other molecular markers. However, within *B. motasi*-like group, whether the isolates of this group consisting in two or more distinct species remains to be clarified. To clarify completely the classification of these parasites, additional information on biological characteristics, especially for *Babesia* sp. Tianzhu and *Babesia* sp. Hebei are needed, such as the tick vector, the virulence and infectivity for small ruminant.

2. We have proven that all RAP-1 types are expressed using sera from infected sheep. However, the precise cellular localization of *B. motasi*-like RAP-1 should be determined by rabbit antibodies against each of RAP-1 specific peptides. Specific antibodies should be produced against the repeats in RAP-1c protein to determine the function of these repeats during infection by inhibition test to confirm the function of these repeats in the parasite-RBC interaction.

3. It would be interesting to investigate if the conserved B-cell epitope region (peptide 1a61/67: LTLPNPYQL) can induce immune response to RAP-1 among sheep *Babesia* species (*B. motasi*-like, *B. ovis* and *Babesia* sp. Xinjiang). Cooperation will be set up with other laboratory of "PIROVAC" project participant. Most notably, we will contact the Portuguese group (for *B. ovis*) and the Chinese group (*Babesia* sp. Xinjiang) and send the two antibodies produced against this peptide. Two experiments will be carried out: (1). These two antibodies will be used to recognize RAP-1 proteins in different sheep *Babesia* species using merozoite extracts from in vitro culture; (2). If the RAP-1 proteins are expressed, these antibodies will be used to test the effect on inhibition growth in different *Babesia* species.

4. For the vaccine purpose, each recombinant RAP-1 protein should be produced and used to immunize sheep to study if these antigens can induce protective immune response against challenge in sheep. Their antibodies will be used to perform in vitro growth and invasion inhibition test.

5. Since *Babesia* species are intracellular parasites, protective immunity against *Babesia* sp. is usually mediated by Th1-like responses to produce IFN-γ, can provoke the lysis of the infected red blood cells (iRBC). The cellular immune response is therefore important to eliminate infected red blood cells. In cattle *Babesia* species, both IgG1 and IgG2 antibodies, which are complement-fixing and opsonizing antibodies in cattle, are produced in response to immunization in cattle (Norimine et al., 2003). Furthermore, immunization with *B. bigemina* and *B. bovis* RAP-1 protein in RIBI adjuvant induced a predominant T-cell response that was characterized by type 1 cytokine expression (IFN-γ) (Rodríguez et al., 1996; Norimine et al., 2002). Th1 response and T-cell epitopes mapping should be explored in all RAP-1 types from sheep *Babesia* species.

REFERENCES

AbouLaila M, Sivakumar T, Yokoyama N, Igarashi I. 2010. Inhibitory effect of terpene nerolidol on the growth of *Babesia* parasites. *Parasitol Int.* 59: 278-282.

Ahmed J, Yin H, Schnittger L, Jongejan F. 2002. Ticks and tick-borne diseases in Asia with special emphasis on China. *Parasitol Res.* 88: S51-55.

Ahmed JS, Luo J, Schnittger L, Seitzer U, Jongejan F, Yin H. 2006. Phylogenetic position of small-ruminant infecting piroplasms. *Ann N Y Acad Sci*. 1081: 498-504.

Aikawa M, Miller LH, Johnson J, Rabbege J. 1978. Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J Cell Biol*. 77: 72-82.

Aktaş M, Altay K, Dumanli N. 2005. Development of a polymerase chain reaction method for diagnosis of *Babesia ovis* infection in sheep and goats. *Vet Parasitol*. 133: 277-281.

Aktaş M. 2014. A survey of ixodid tick species and molecular identification of tick-borne pathogens. *Vet Parasitol*. 200:276-83.

Alani AJ, Herbert IV. 1988a. The pathogenesis of *Babesia motasi* (Wales) infection in sheep. *Vet Parasitol*. 27: 209-220.

Alani AJ, Herbert IV. 1988b. The morphometrics of *Babesia motasi* (Wales) and its transmission by *Haemaphysalis punctate* (Canestrni and Fanzago, 1877) to sheep. *Vet Parasitol*. 30: 87-95.

Al-Khedery B, Allred DR. 2006. Antigenic variation in *Babesia bovis* occurs through segmental gene conversion of the *ves* multigene family, within a bidirectional locus of active transcription. *Mol Microbiol*. 59: 402-414.

Almazán C, Lagunes R, Villar M, Canales M, Rosario-Cruz R, Jongejan F, de la Fuente J. 2010. Identification and characterization of *Rhipicephalus* (*Boophilus*) *microplus* candidate protective antigens for the control of cattle tick infestations. *Parasitol Res*. 106: 471-479.

Almazán C, Moreno-Cantú O, Moreno-Cid JA, Galindo RC, Canales M, Villar M, de la Fuente J. 2012. Control of tick infestations in cattle vaccinated with bacterial membranes containing surface-exposed tick protective antigens. *Vaccine.* 30: 265-272.

Altay K, Dumanli N, Aktas M. 2007. Molecular identification, genetic diversity and distribution of *Theileria* and *Babesia* species infecting small ruminants. *Vet Parasitol*. 147: 161-165.

Altay K, Aktas M, Dumanli N. 2008. Detection of *Babesia ovis* by PCR in *Rhipicephalus bursa* collected from naturally infested sheep and goats. *Res Vet Sci.* 85: 116-119.

Altay K, Dumanli N, Aktas M. 2012. A study on ovine tick-borne hemoprotozoan parasites (*Theileria* and *Babesia*) in the East Black Sea Region of Turkey. *Parasitol Res*.111:149-153.

Andreotti R, Guerrero FD, Soares MA, Barros JC, Miller RJ, Léon AP. 2011. Acaricide resistance of *Rhipicephalus* (*Boophilus*) microplus in State of Mato Grosso do Sul, Brazil. *Rev Bras Parasitol Vet*. 20: 127-133.

Awinda PO, Mealey RH, Williams LB, Conrad PA, Packham AE, Reif KE, Grause JF, Pelzel-McCluskey AM, Chung C, Bastos RG, Kappmeyer LS, Howe DK, Ness SL, Knowles DP, Ueti MW. 2013. Serum antibodies from a subset of horses positive for *Babesia caballi* by competitive enzyme-linked immunosorbent assay demonstrate a protein recognition pattern that is not consistent with infection. *Clin Vaccine Immunol*. 20:1752-7.

Bai Q, Liu GY. 1990. Studies on isolation and preservation of a single species of haemotecytozoon in bovine: Isolation and cryopreservation of *Babesia ovata* in China. *Chin J Vet Med*. 16: 2-4.

Bai Q, Liu GY, Liu DK, Ren JX, Li X. 2002. Isolation and preliminary characterization of a large *Babesia* sp. from sheep and goats in the eastern part of Gansu Province, China. *Parasitol Res.* 88: S16-S21.

Bakheit MA, Seitzer U, Mbati PA, Ahmed JS. 2007. Serological diagnostic tools for the major tick-borne protozoan diseases of livestock. *Parassitologi*a. 49 Suppl 1:53-62.

Bannister LH, Mitchell GH, Butcher GA, Dennis ED. 1986. Lamellar membranes associated with rhoptries in erythrocytic merozoites of *Plasmodium knowlesi*: a clue to the mechanism of invasion. *Parasitology*. 92: 291-303.

Bannister LH, Mitchell GH. 1989. The fine structure of secretion by *Plasmodium knowlesi* merozoites during red cell invasion. *J Protozool*. 36: 362-367.

Bannister LH, Dluzewski AR. 1990. The ultrastructure of red cell invasion in malaria infections: a review. *Blood Cells*. 16: 257-292; discussion 293-297.

Bannister LH, Hopkins JM, Fowler RE, Krishna S, Mitchell GH. 2000. A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol Today*. 16: 427-433.

Bannister L, Mitchell G. 2003. The ins, outs, and roundabouts of malaria. *Trends Parasitol*. 19: 209-213.

Baravalle ME, Thompson C, de Echaide ST, Palacios C, Valentini B, Suárez CE, Christensen MF, Echaide I. 2010. The novel protein BboRhop68 is expressed by intraerythrocytic stages of *Babesia bovis*. *Parasitol Int*. 59: 571-578.

Bauza K, Malinauskas T, Pfander C, Anar B, Jones EY, Billker O, Hill AV, Reyes-Sandoval A. 2014. Efficacy of a *Plasmodium vivax* Malaria Vaccine Using ChAd63 and Modified Vaccinia Ankara Expressing Thrombospondin-Related Anonymous Protein as Assessed with Transgenic *Plasmodium berghei* Parasites. *Infect Immun*. 82:1277-1286.

Bhoora R, Quan M, Zweygarth E, Guthrie AJ, Prinsloo SA, Collins NE. 2010. Sequence heterogeneity in the gene encoding the rhoptry-associated protein-1 (RAP-1) of *Babesia caballi* isolates from South Africa. *Vet Parasitol*. 169: 279-288.

Blackman MJ, Bannister LH. 2001. Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation. Mol. *Biochem. Parasitol*. 117: 11-25.

Bock R, de Vos AJ, Kingston TG, Shiels IA, Dalgliesh RJ. 1992. Investigations of breakdowns in protection provided by living *Babesia bovis* vaccine. *Vet Parasitol.* 43: 45-56.

Boonchit S, Xuan X, Yokoyama N, Goff WL, Wagner G, Igarashi I. 2002. Evaluation of an enzyme-linked immunosorbent assay with recombinant rhoptry-associated protein 1 antigen against *Babesia bovis* for the detection of specific antibodies in cattle. *J Clin Microbiol*. 40: 3771-3775.

Boonchit S, Xuan X, Yokoyama N, Goff WL, Waghela SD, Wagner G, Igarashi I. 2004. Improved enzyme-linked immunosorbent assay using C-terminal truncated recombinant antigens of *Babesia bovis* rhoptry-associated protein-1 for detection of specific antibodies. *J Clin Microbiol*. 42: 1601-1604.

Boonchit S, Alhassan A, Chan B, Xuan X, Yokoyama N, Ooshiro M, Goff WL, Waghela SD, Wagner G, Igarashi I. 2006. Expression of C-terminal truncated and full-length *Babesia bigemina* rhoptry-associated protein 1 and their potential use in enzyme-linked immunosorbent assay. *Vet Parasitol*. 137: 28-35.

Bork S, Yokoyama N, Matsuo T, Claveria FG, Fujisaki K, Igarashi I. 2003. Growth inhibitory effect of triclosan on equine and bovine *Babesia* parasites. *Am J Trop Med Hyg*. 68: 334-340.

Boylea MJ, Wilsona DW, Richards JS, Riglar DT, Tetteh KK, Conway DJ, Ralphe SA, Bauma J, Beeson JG. 2010. Isolation of viable *Plasmodium falciparum* merozoites to define erythrocyte invasion events and advance vaccine and drug development. *Proc Natl Acad Sci U S A*. 107: 14378-14383.

Brayton KA, Lau AO, Herndon DR, Hannick L, Kappmeyer LS, Berens SJ, Bidwell SL, Brown WC, Crabtree J, Fadrosh D, Feldblum T, Forberger HA, Haas BJ, Howell JM, Khouri H, Koo H, Mann DJ, Norimine J, Paulsen IT, Radune D, Ren Q, Smith RK Jr, Suarez CE, White O, Wortman JR, Knowles DP Jr, McElwain TF, Nene VM. 2007. Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. *PLoS Pathog*. 3:1401-1413.

Brown WC, Zhao S, Woods VM, Tripp CA, Tetzlaff CL, Heussler VT, Dobbelaere DA, Rice-Ficht AC. 1993. Identification of two Th1 cell epitopes on the 77 kDa *Babesia bovis* merozoite protein (Bb-1) by use of truncated recombinant fusion proteins. *Infect Immun*. 61: 236-244.

Brown WC, Estes DM, Chantler SE, Kegerreis KA, Suarez CE. 1998. DNA and a CpG oligonucleotide derived from *Babesia bovis* are mitogenic for bovine B cells. *Infect Immun*. 66: 5423-32.

Brown WC, Palmer GH. 1999. Designing Blood-stage Vaccines against *Babesia bovis* and *B. bigemina*. *Parasitol Today*. 15: 275-281.

Brown WC, Norimine J, Knowles DP, Goff WL. 2006a. Immune control of *Babesia bovis* infection. *Vet Parasitol*. 138: 75-87.

Brown WC, Norimine J, Goff WL, Suarez CE, McElwain TF. 2006b. Prospects for recombinant vaccines against *Babesia bovis* and related parasites. *Parasite Immunol*. 28: 315-327.

Bush AO, Fernandes JC, Esch GW, Seed RJ. 2001. Parasitism: The Diversity and Ecology of Animal Parasites. Cambridge University Press, Cambridge, pp 89-91.

Callow LL, Mellors LT, McGregor W. 1979. Reduction in virulence of *Babesia bovis* due to rapid passage in splenectomized cattle. *Int J Parasitol*. 9: 333-338.

Canales M, Enriquez A, Ramos E, Cabrera D, Dandie H, Soto A, Falcon V, Rodríguez, M, de la Fuente J. 1997. Large-scale production in Pichia pastoris of the recombinant vaccine Gavac against cattle tick. *Vaccin*e. 15: 414-422.

Canales M, Almazán C, Naranjo V, Jongejan F, de la Fuente J. 2009. Vaccination with recombinant *Boophilus annulatus* Bm86 ortholog protein, Ba86, protects cattle against *B. annulatus* and *B. microplus* infestations. *BMC Biotechnol*. 9: 29.

Carcy B, Précigout E, Valentin A, Gorenflot A, Schrevel J. 1995. A 37-kilodalton glycoprotein of *Babesia divergens* is a major component of a protective fraction containing low-molecular-mass culture-derived exoantigens. *Infect Immun*. 63: 811-7.

Carcy B, Précigout E, Schetters T, Gorenflot A. 2006. Genetic basis for GPI-anchor merozoite surface antigen polymorphism of *Babesia* and resulting antigenic diversity. *Vet Parasitol*. 138: 33-49.

Chen DM. 1982. Investigations on ovine piroplasmosis. *Chin J Vet Sci Technol*. 3: 31-32.

Chesne-Seck ML, Pizarro JC, Vulliez-Le Normand B, Collins CR, Blackman MJ, Faber BW, Remarque EJ, Kocken CH, Thomas AW, Bentley GA. 2005. Structural

comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites. *Mol Biochem Parasitol*. 144: 55-67.

Chevillon C, Ducornez S, de Meeûs T, Koffi BB, Gaïa H, Delathière JM, Barré N. 2007. Accumulation of acaricide resistance mechanisms in *Rhipicephalus* (*Boophilus*) *microplus* (Acari: Ixodidae) populations from New Caledonia Island. *Vet Parasitol*. 147: 276-288.

Cicek H, Duzgun A, Emre Z, Karaer Z. 2004. Seroprevalence of *Babesia ovis* in sheep around Afyon. *Turk J Vet Anim Sci*. 28: 683-686.

Criado-Fornelio A.2007.A review of nucleic-acid-based diagnostic tests for *Babesia* and *Theileria*, with emphasis on bovine piroplasms. *Parassitologia*. 49 Suppl 1:39-44.

Commins MA, Goodger BV, Wright IG. 1985. Proteinases in the lysate of bovine erythrocytes infected with *Babesia bovis*: initial vaccination studies. *Int J Parasitol*. 15: 491-495.

Cortés A, Mellombo M, Mueller I, Benet A, Reeder JC, Anders RF. 2003. Geographical structure of diversity and differences between symptomatic and asymptomatic infections for *Plasmodium falciparum* vaccine candidate AMA1. *Infect Immun*. 71: 1416-1426.

Cowman AF, Crabb BS. 2006. Invasion of red blood cells by malaria parasites. *Cell.* 124: 755-766.

Dalrymple BP, Casu RE, Peters JM, Dimmock CM, Gale KR, Boese R, Wright IG. 1993a. Characterization of a family of multi-copy genes encoding rhoptry protein homologues in *Babesia bovis*, *B. ovis* and *B. canis*. *Mol Biochem Parasitol*. 57: 181-192.

Dalrymple BP. 1993b. Molecular variation and diversity in candidate vaccine antigens from *Babesia*. *Acta Trop*. 53: 227-238.

Dalrymple BP, Peters JM, Böse R, Wright IG. 1996. A polymerase chain reaction method for the identification of genes encoding members of the Bv60/p58 family of rhoptry protein homologues in the genus *Babesia*. *Exp Parasitol.* 84:96-100.

Dautu G, Munyaka B, Carmen G, Zhang G, Omata Y, Xuenan X, Igarashi M. 2007. *Toxoplasma gondii*: DNA vaccination with genes encoding antigens MIC2, M2AP, AMA1 and BAG1 and evaluation of their immunogenic potential. *Exp Parasitol.* 116: 273-282.

de Castro JJ. 1997. Sustainable tick and tickborne disease control in livestock improvement in developing countries. *Vet Parasitol*. 71: 77-97.

de la Fuente J, Kocan KM. 2006. Strategies for development of vaccines for control of ixodid tick species. *Parasite Immunol*. 28: 275-283.

Delbecq S, Hadj-Kaddour K, Randazzo S, Kleuskens J, Schetters T, Gorenflot A, Précigout E. 2006. Hydrophobic moeties in recombinant proteins are crucial to generate efficient saponin-based vaccine against Apicomplexan *Babesia divergens*. *Vaccine*. 24: 613-621.

Delbecq S, Auguin D, Yang YS, Löhr F, Arold S, Schetters T, Précigout E, Gorenflot A, Roumestand C. 2008.The solution structure of the adhesion protein Bd37 from *Babesia divergens* reveals structural homology with eukaryotic proteins involved in membrane trafficking. *J Mol Biol.* 375: 409-424.

de Vos S, Zeinstra L, Taoufik O, Willadsen P, Jongejan F. 2001. Evidence for the utility of the Bm86 antigen from *Boophilus microplusin* vaccination against other tick species. *Exp Appl Acarol* 25: 245-261.

de Waal DT, Combrink MP. 2006. Live vaccines against bovine babesiosis. *Vet Parasitol.* 138: 88-96.

Dlugonska H. 2008. *Toxoplasma* rhoptries: unique secretory organelles and source of promising vaccine proteins for immunoprevention of toxoplasmosis. *J Biomed Biotechnol* 2008: 632424.

Dominguez M, Echaide I, Echaide ST, Mosqueda J, Cetrá B, Suarez CE, Florin-Christensen M. 2010. In silico predicted conserved B-cell epitopes in the merozoite surface antigen-2 family of *B. bovis* are neutralization sensitive. *Vet Parasitol.*167: 216-226.

Dowling SC, Perryman LE, Jasmer DP. 1996. A *Babesia bovis* 225-kilodalton spherical-body protein: localization to the cytoplasmic face of infected erythrocytes after merozoite invasion. *Infect Immun*. 64: 2618-2626.

Dubremetz JF, Garcia-Reguet N, Conseil V, Fourmaux MN. 1998. Apical organelles and host-cell invasion by Apicomplexa. *Int J Parasitol*. 28: 1007-1013.

Duzgun A, Wright IG, Waltisbuhl DJ, Gale KR, Goodger BV, Dargie JD, Alabay M, Cerci H. 1991. An ELISA for the diagnosis of *Babesia ovis* infection utilizing a synthetic, *Babesia bovis*-derived antigen. *Vet Parasitol*. 39:225-231.

Dvorak JA, Miller LH, Whitehouse WC, Shiroishi T. 1975. Invasion of erythrocytes by malaria merozoites. *Science*. 187: 748-750.

Ekici OD, Sevinc F, Isik N. 2012. Instability of ovine babesiosis in an endemic area in Turkey. *Vet Parasitol*. 188: 372-375.

Emre Z, Duzgun A, Iriadam M, Sert H. 2001. Seroprevalence of *Babesia ovis* in Awassi sheep in Urfa, Turkey. *Turk. J Vet Anim Sci*. 25: 759-762.

Fernandes EK, Bittencourt VR. 2008. *Entomopathogenic fungi* against South American tick species. *Exp Appl Acaro*l. 46:71-93.

Ferrer D, Castellà J, Gutiérrez JF, Lavín S, Marco I. 1998a. Seroprevalence of *Babesia ovis* in mouflon sheep in Spain. *J Wildl Dis*. 34:637-639.

Ferrer D, Castellà J, Gutiérrez JF, Lavín S, Marco I.1998b. Seroprevalence of *Babesia ovis* in Spanish ibex (Capra pyrenaica) in Catalonia, northeastern Spain. *Vet Parasitol*. 75: 93-98.

Ferrer D, Castellà J, Gutiérrez JF.1998c.Seroprevalence of *Babesia ovis* in sheep in Catalonia, northeastern Spain. Vet Parasitol. 79: 275-281.

Ferreri LM, Brayton KA, Sondgeroth KS, Lau AO, Suarez CE, McElwain TF. 2012. Expression and strain variation of the novel "small open reading frame" (*smorf*) multigene family in *Babesia bovis*. *Int J Parasitol*. 42:131-138.

Fish L, Leibovich B, Krigel Y, McElwain T, Shkap V. 2008. Vaccination of cattle against *B. bovis* infection with live attenuated parasites and non-viable immunogens. *Vaccine*. 26: 29-33.

Fisher TG, McElwain TF, Palmer GH. 2001. Molecular basis for variable expression of merozoite surface antigen gp45 among American isolates of *Babesia bigemina*. *Infect Immun*. 69: 3782-3790.

Florin-Christensen M, Suarez CE, Hines SA, Palmer GH, Brown WC, McElwain TF. 2002. The *Babesia bovis* merozoite surface antigen 2 locus contains four tandemly arranged and expressed genes encoding immunologically distinct proteins. *Infect Immun*. 70: 3566-3575.

Freeman RR, Trejdosiewicz AJ, Cross GA. 1980. Protective monoclonal antibodies recognizing stage-specific merozoite antigens of a rodent malaria parasite. *Nature*. 284: 366-368.

Freeman JM, Kappmeyer LS, Ueti MW, McElwain TF, Baszler TV, Echaide I, Nene VM, Knowles DP. 2010. A *Babesia bovis* gene syntenic to *Theileria parva* p67 is expressed in blood and tick stage parasites. *Vet Parasitol*. 173: 211-218.

Friedhoff KT. 1988. Transmission of *Babesia*. In: Ristic M. Babesiosis of domestic animals and man. Beca Raton: CRC Press, pp 23-52.

Friedhoff KT. 1997. Tick-borne disease of sheep and goats caused by *Babesia*, *Theileria* or *Anaplasma* spp.. *Parassitologia*. 39: 99-109.

Gaffar FR, Yatsuda AP, Franssen FF, de Vries E. 2004a. Erythrocyte Invasion by *Babesia bovis* Merozoites Is Inhibited by Polyclonal Antisera Directed against Peptides

Derived from a Homologue of *Plasmodium falciparum* Apical Membrane Antigen 1. *Infect Immun*. 72: 2947-2955.

Gaffar FR, Yatsuda AP, Franssen FF, de Vries E. 2004b. A *Babesia bovis* merozoite protein with a domain architecture highly similar to the thrombospondin-related anonymous protein (TRAP) present in *Plasmodium* sporozoites. *Mol Biochem Parasitol*. 136: 25-34.

GilsonPR, Crabb BS. 2009. Morphology and kinetics of the three distinct phases of red blood cell invasion by *Plasmodium falciparum* merozoites. *Inter J Parasitol*. 39: 91-96.

Gindin G, Samish M, Zangi G, Mishoutchenko A, Glazer I. 2002. The susceptibility of different species and stages of ticks to *entomopathogenic fungi*. *Exp Appl Acarol*. 28: 283-288.

Goff WL, Davis WC, Palmer GH, McElwain TF, Johnson WC, Bailey JF, McGuire TC. 1988. Identification of *Babesia bovis* merozoite surface antigens by using immune bovine sera and monoclonal antibodies. *Infect Immun*. 56: 2363-2368.

Gohil S, Kats LM, Sturm A, Cooke BM. 2010. Recent insights into alteration of red blood cells by *Babesia bovis*: Moovin' forward. *Trends Parasitol*. 26: 591-599.

Gohil S, Herrmann S, Günther S, Cooke BM. 2013. Bovine babesiosis in the 21st century: Advances in biology and functional genomics. *Int J Parasitol*. 43: 125-132.

Goodger BV, Waltisbuhl DJ, Commins MA, Wright IG. 1992. *Babesia bovis*: dextran sulfate as an adjuvant for and precipitant of protective immunogens. *Int J Parasitol*. 22: 465-469.

Goo YK, Terkawi MA, Jia H, Aboge GO, Ooka H, Nelson B, Kim S, Sunaga F, Namikawa K, Igarashi I, Nishikawa Y, Xuan X. 2010. Artesunate, a potential drug for treatment of *Babesia* infection. *Parasitolo int*. 59: 481-486.

Goo YK, Aboge GO, Terkawi MA, Jia H, Yamagishi J, Sunaga F, Namikawa K, Cha SY, Jang HK, Kim S, Nishikawa Y, Xuan X. 2012. Four promising antigens, BgP32, BgP45, BgP47, and BgP50, for serodiagnosis of *Babesia gibsoni* infection were classified as *B. gibsoni* merozoite surface protein family. *Parasitol Int*. 61: 364-368.

Goo YK, Ueno A, Terkawi MA, Aboge GO, Junya Y, Igarashi M, Kim JY, Hong YC, Chung DI, Nishikawa Y, Xuan X. 2013. Actin polymerization mediated by *Babesia gibsoni* aldolase is required for parasite invasion. *Exp Parasitol*. 135: 42-49.

Gou H, Guan G, Liu A, Ma M, Xu Z, Liu Z, Ren Q, Li Y, Yang J, Chen Z, Yin H, Luo J. 2012. A DNA barcode for Piroplasmea. *Acta Trop*. 124: 92-97.

Gou H, Guan G, Ma M, Liu A, Liu Z, Ren Q, Li Y, Yang J, Chen Z, Yin H, Luo J.

2013. Phylogenetic analysis based on 28S rRNA of *Babesia* spp. in ruminants in China. *Exp Appl Acarol*. 59: 463-472.

Graf JF, Gogolewski R, Leach-Bing N, Sabatini GA, Molento MB, Bordin EL, Arantes GJ. 2004. Tick control: an industry point of view. *Parasitology*. 129: S427-442.

Gray JS, Weiss LM. 2008. Taylor and Francis, Abingdon, UK, pp 303-349.

Guan GQ, Yin H, Luo JX, Lu WS, Zhang QC, Ma ML, Yuan GL, Lu BY, Wang YJ, Muhe TE. 2001. Isolation of a large ovine *Babesia* sp. in Xinjiang, China. *Chin J Vet Sci Technol*. 31: 35-36.

Guan GQ, Yin H, Luo JX, Lu WS, Zhang QC, Gao YL, Lu BY. 2002. Transmission of *Babesia* sp. to sheep with field-collected *Haemaphysalis qinghaiensis*. *Parasitol Res*. 88: S22-24.

Guan G, Chauvin A, Luo J, Inoue N, Moreau E, Liu Z, Gao J, Thekisoe OM, Ma M, Liu A, Dang Z, Liu J, Ren Q, Jin Y, Sugimoto C, Yin H. 2008.The development and evaluation of a loop-mediated isothermal amplification (LAMP) method for detection of *Babesia* spp. infective to sheep and goats in China. *Exp Parasito*l. 120: 39-44.

Guan G, Ma M, Moreau E, Liu J, Lu B, Bai Q, Luo J, Jorgensen W, Chauvin A, Yin H. 2009. A new ovine *Babesia* species transmitted by *Hyalomma anatolicum anatolicum*. *Exp Parasitol*. 122: 261-267.

Guan G, Moreau E, Liu J, Hao X, Ma M, Luo J, Chauvin A, Yin H. 2010a. *Babesia* sp. BQ1 (Lintan): molecular evidence of experimental transmission to sheep by *Haemaphysalis qinghaiensis* and *Haemaphysalis longicornis*. *Parasitol Int*. 59: 265-267.

Guan G, Moreau E, Brisseau N, Luo J, Yin H, Chauvin A. 2010b. Determination of erythrocyte susceptibility of Chinese sheep (Tan mutton breed) and French sheep (Vendéen breed) to *Babesia* sp. BQ1 (Lintan) by in vitro culture. *Vet Parasitol*. 170: 37-43.

Guan G, Chauvin A, Rogniaux H, Luo J, Yin H, Moreau E. 2010c. Merozoite proteins from *Babesia* sp. BQ1 (Lintan) as potential antigens for serodiagnosis by ELISA. *Parasitology*. 137: 927-938.

Guan G, Ma M, Liu A, Ren Q, Wang J, Yang J, Li A, Liu Z, Du P, Li Y, Liu Q, Zhu H, Yin H, Luo J. 2012a. A recently identified ovine *Babesia* in China: serology and sero-epidemiology. *Parasitol Int*. 61: 532-537.

Guan G, Ma M, Liu A, Du P, Ren Q, Li Y, Wang J, Liu Z, Yin H, Luo J.2012b. Continuous in vitro cultivation of a recently identified *Babesia* that infects small ruminants in China. *Vet Parasitol*. 187: 371-8.

Guan G, Moreau E, Liu J, Ma M, Rogniaux H, Liu A, Niu Q, Li Y, Ren Q, Luo J, Chauvin A, Yin H.2012c. BQP35 is a novel member of the intrinsically unstructured protein (IUP) family which is a potential antigen for the sero-diagnosis of *Babesia* sp. BQ1 (Lintan) infection. *Vet Parasitol*. 187: 421-30.

Gubbels MJ, de Vos AP, van der Weide M, Viseras J, Schouls LM, de Vries E, Jongejan F. 1999. Simultaneous detection of bovine *Theileria* and *Babesia* species by reverse line blot hybridization. *J Clin Microbiol*. 37: 1782-1789.

Guerrero FD, Miller RJ, Pérez de León AA. 2012. Cattle tick vaccines: many candidate antigens, but will a commercially viable product emerge? *Int J Parasitol*. 42: 421-427.

Guo S, Ma DL, Mou YJ, Yang SM, Wang WB, Ge GH, Ga DJ, Fang BQ. 2005. Investigations of ovine Theileriosis in north of China. *Chin J Vet Parasitol*. 13: 15-17.

Guo S, Mu Y, Liu Z, Ma D, Yang S, Ge G, Fang B, Ga D, Ma M, Luo J, Yin H, Seitzer U, Ahmed JS. 2007. Serological investigation of ovine theileriosis by ELISA in Gannan Tibet Region of Gansu Province in China. *Parasitol Res*. Suppl 2: S197-200.

Habela M, Reina D, Nieto C, Navarrete I. 1990. Isolation and identification of *Babesia ovis* in Extremadura (Spain). *Vet Parasitol*. 35:. 233-238.

Hadj-Kaddour K, Carcy B, Vallet A, Randazzo S, Delbecq S, Kleuskens J, Schetters T, Gorenflot A, Précigout E. 2007. Recombinant protein Bd37 protected gerbils against heterologous challenges with isolates of *Babesia divergens* polymorphic for the bd37 gene. *Parasitol*. 134: 187-96.

Hajdusek O, Almazán C, Loosova G, Villar M, Canales M, Grubhoffer L, Kopacek P, de la Fuente J. 2010. Characterization of ferritin 2 for the control of tick infestations. *Vaccine*. 28: 2993-2998.

Hashemi-Fesharki R. 1997. Tick-borne diseases of sheep and goats and their related vectors in Iran. *Parassitologia*. 39: 115-117.

Hashemi-Fesharki R, Uilenberg G. 1981. *Babesia crassa* of domestic sheep in Iran. *Vet Quarterly*. 3: 1-8.

He L, Feng HH, Zhang WJ, Zhang QL, Fang R, Wang LX, Tu P, Zhou YQ, Zhao JL, Oosthuizen MC. 2012. Occurrence of *Theileria* and *Babesia* species in water buffalo (Bubalus babalis, Linnaeus, 1758) in the Hubei province, South China. *Vet Parasitol*. 186: 490-496.

Hines SA, Palmer GH, Jasmer DP, McGuire TC, McElwain TF. 1992. Neutralization-sensitive merozoite surface antigens of *Babesia bovis* encoded by members of a polymorphic gene family. *Mol Biochem Parasitol*. 55: 85-94.

Hines SA, Palmer GH, Brown WC, McElwain TF, Suarez CE, Vidotto O, Rice-Ficht AC. 1995a. Genetic and antigenic characterization of *Babesia bovis* merozoite spherical body protein Bb-1. *Mol Biochem Parasitol*. 69: 149-159.

Hines SA, Palmer GH, Jasmer DP, Goff WL, McElwain TF. 1995b. Immunization of cattle with recombinant *Babesia bovis* merozoite surface antigen-1. *Infect Immun*. 63: 349-352.

Holder AA, Freeman RR, Uni S, Aikawa M. 1985. Isolation of a *Plasmodium falciparum* rhoptry antigen. *Mol Biochem Parasitol*. 14: 293-303.

Holman PJ, Bendele KG, Schoelkopf L, Jones-Witthuhn RL, Jones SO. 2003. Ribosomal RNA analysis of *Babesia odocoilei* isolates from farmed reindeer (*Rangifer tarandus tarandus*) and elk (*Cervus elaphu canadensis*) in Wisconsin. *Parasitol Res*. 91: 378-383.

Hope M, Riding G, Menzies M, Colditz I, Reverter A, Willadsen P. 2005. Potential for recombinant *Babesia bovis* antigens to protect against a highly virulent isolate. *Parasite Immunol*. 27: 439-445.

Hötzel I, Brown WC, McElwain TF, Rodríguez SD, Palmer GH. 1996. Dimorphic sequences of *rap-1* genes encode B and CD4+ T helper lymphocyte epitopes in the *Babesia bigemina* rhoptry associated protein-1. *Mol Biochem Parasitol*. 81: 89-99.

Hötzel I, Suarez CE, McElwain TF, Palmer GH. 1997. Genetic variation in the dimorphic regions of RAP-1 genes and *rap-1* loci of *Babesia bigemina*. *Mol Biochem Parasitol*. 90: 479-489.

Howell SA, Withers-Martinez C, Kocken CH, Thomas AW, Blackman MJ. 2001. Proteolytic processing and primary structure of *Plasmodium falciparum* apical membrane antigen-1. *J Biol Chem*. 276: 31311-31320.

Howell SA, Well I, Fleck SL, Kettleborough C, Collins CR, Blackman MJ. 2003. A single malaria merozoite serine protease mediates shedding of multiple surface proteins by juxtamembrane cleavage. *J Biol Chem*. 278: 23890-23898.

Igarashi I, Aikawa M, Kreier JP. 1988. Host cell–parasite interactions in babesiosis. In: Ristic, M. (Ed.), Babesiosis of Domestic Animals. CRC Press, Inc., Boca Raton, FL, pp 53-70.

Innan H. 2009. Population genetic models of duplicated genes. *Genetica* 137:19-37.

Iqbal F, Khattak R, Ozubek S, Khattak M, Rasul A, Aktas M. 2013. Application of the Reverse Line Blot Assay for the Molecular Detection of *Theileria* and *Babesia* sp. in Sheep and Goat Blood Samples from Pakistan. *Iran J Parasitol*.8: 289-295.

Jagannath MS, Hedge KS, Shivaram K, Nagaraja KV. 1974. An outbreak of babesiosis in sheep and goats and its control. *Mysor J Agric Sci.* 8: 441- 443.

Jasmer DP, Reduker DW, Perryman LE, McGuire TC. 1992. A *Babesia bovis* 225-kilodalton protein located on the cytoplasmic side of the erythrocyte membrane has sequence similarity with a region of glycogen phosphorylase. *Mol Biochem Parasitol*. 2: 263-269.

Jenkins MC. 2001. Advances and prospects for subunit vaccines against protozoa of veterinary importance. *Vet Parasitol*. 101: 291-310.

Jouglin M, Fernandez-de-Mera IG., De la Cotte N, Ruiz-Fons F, Gortazar C, et al 2014**.** Isolation and characterization of *Babesia pecorum* sp. nov. from farmed red deer (Cervus elaphus). *Vet Res* (in press)

Kappmeyer LS, Perryman LE, Hines SA, Baszler TV, Katz JB, Hennager SG, Knowles DP. 1999. Detection of equine antibodies to *babesia caballi* by recombinant *B. caballi* rhoptry-associated protein 1 in a competitive-inhibition enzyme-linked immunosorbent assay. *J Clin Microbiol*. 37: 2285-2290.

Kats LM, Cooke BM, Coppel RL, Black CG. 2008. Protein Trafficking to Apical Organelles of Malaria Parasites–Building an Invasion Machine. *Traffic*. 2: 176-186.

Kjemtrup AM, Thomford JW, Gardner IA, Conrad PA, Jessup DA, Boyce WM. 1995. Seroprevalence of two *Babesia* spp. isolates in selected bighorn sheep (Ovis canadensis) and mule deer (Odocoileus hemionus) populations in California. *J Wildl Dis*. 31:467-471.

Kumar S, Yokoyama N, Kim JY, Huang X, Inoue N, Xuan X, Igarashi I, Sugimoto C. 2004. Expression of *Babesia equi* EMA-1 and EMA-2 during merozoite developmental stages in erythrocyte and their interaction with erythrocytic membrane skeleton. *Mol Biochem. Parasitol*. 133: 221-227.

Kuttler KL. 1988a. Worldwide impact of Babesiosis. In: Ristic M. Babesiosis of domestic animals and man, Beca Raton: CRC Press, pp 1-10.

Kuttler KL. 1988b. Chemetherapy of babesiosis. In: Ristic M. Babesiosis of domestic animals and man, Beca Raton: CRC Press, pp 227-240.

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.

Leal AT, Seixas A, Pohl PC, Ferreira CA, Logullo C, Oliveira PL, Farias SE, Termignoni C, da Silva Vaz I Jr, Masuda A. 2006.Vaccination of bovines with recombinant Boophilus Yolk pro-Cathepsin. *Vet Immunol Immunopathol*. 114: 341-345.

LeRoith T, Berens SJ, Brayton KA, Hines SA, Brown WC, Norimine J, McElwain TF. 2006. The *Babesia bovis* merozoite surface antigen 1 hypervariable region induces surface-reactive antibodies that block merozoite invasion. *Infect Immun*. 74: 3663-3667.

Levine ND. 1971. Taxonomy of the piroplasms. *Trans Am Microsc Soc*.90: 2-33.

Levine ND, Corliss JO, Cox FE, Deroux G, Grain J, Honigberg BM, Leedale GF, Loeblich AR 3rd, Lom J, Lynn D, Merinfeld EG, Page FC, Poljansky G, Sprague V, Vavra J, Wallace FG.1980. A newly revised classification of the protozoa. *J Protozool*. 27: 37-58.

Levine ND. 1985. Veterinary Protozoology. Iowa State University Press, *Ames*, pp 291-328.

Levine ND. 1988. The protozoan phylum Apicomplexa. Volume I. Volume II. CRC Press Inc., Boca Raton, FL 33431; USA, pp 154 -203.

Lewis D, Holman MR, Purnell RE, Young ER, Herbert IV, Bevan WJ. 1981. Investigations on *Babesia motasi* isolated from Wales. *Res Vet Sci*. 31: 239-243.

Li Y, Guan G, Ma M, Liu J, Ren Q, Luo J, Yin H. 2011. *Theileria ovis* discovered in China. *Exp Parasitol*. 127: 304-307.

Li ZH, Zhang GW, Zhao DY. 2006. A case of ovine piroplasmoisis, diagnosis and treatment. *Shanxi Agri*. 11: 27.

Lian C, He HC, Bai Q, Han GF, Liu GY. 1997. A large new *Babesia* sp. found from infected sheep sheep in China. *Chin J Vet Sci*. 31: 239-243.

Liu AH, Yin H, Guan GQ, Schnittger L, Liu ZJ, Ma ML, Dang ZS, Liu JL, Ren QY, Bai Q, Ahmed JS, Luo JX. 2007. At least two genetically distinct large *Babesia* species infective to sheep and goats in China. *Vet Parasitol*. 147: 246-251.

Liu A, Guan G, Du P, Gou H, Liu Z, Liu J, Ma M, Yang J, Li Y, Niu Q, Ren Q, Bai Q, Yin H, Luo J. 2012. Loop-mediated isothermal amplification (LAMP) method based on two species-specific primer sets for the rapid identification of Chinese *Babesia bovis* and *B. bigemina*. *Parasitol Int*. 61: 658-663.

Liu A, Guan G, Du P, Gou H, Zhang J, Liu Z, Ma M, Ren Q, Liu J, Yang J, Li Y, Niu Q, Bai Q, Yin H, Luo J. 2013. Rapid identification and differentiation of *Theileria sergenti* and *Theileria sinensis* using a loop-mediated isothermal amplification (LAMP) assay. *Vet Parasitol*. 191: 15-22.

Liu GY, Lian C, Han GF, Wang ZD, Wang MY, He HC, Bai Q. 1997. Experimental infection sheep and goats with two ovine *Babesia* isolated from Ningxian county of Gansu Province. *Chin J Vet Sci Technol*. 27: 10-11.

Liu GY, Lian C, He HC, Han GF, Wang MY, Bai Q. 1999. Isolation and characterizing infectivity of *Babesia ovis*. *Chin J Vet Sci Technol*. 29: 12-13.

Liu ZL, Zhao JL, Ma LH, Yao BA. 1997. *Babesia orientalis* sp. nov. parasitized in buddalo bubalus babalis in China (piroplasmida: babesiidae). *Acta Vet Zootechnica Sin*. 28: 84-89.

Lobo CA, Rodríguez M, Cursino-Santos JR. 2012. *Babesia* and red cell invasion. *Curr Opin Hematol*. 19:170-175.

Lu WS, Yin H, Luo JX, Lu WX, Zhang QC, Dou HF. 1992. Discovery of *Babesia ma*jor in Xinjiang. Symposium of the 3rd Meeting of Chinese Society of Veterinary Parasitology. p 220.

Ludwig W, Schleifer KH. 1994. Bacterial phylogeny based on16S and 23SrRNA sequence analysis. *FEMS Micr obiol Rev*. 15: 155-173.

Lund ED, Soudant P, Chu FL, Harvey E, Bolton S, Flowers A. 2005. Effects of triclosan on growth, viability and fatty acid synthesis of the oyster protozoan parasite *Perkinsus marinus*. *Dis Aquat Organ*. 67: 217-224.

Luo J, Yin H. 1997. Theileriosis of sheep and goats in China. *Trop Anim Health Prod*. 29: 8S-10S.

Luo JX, Yin H, Guan GQ, Yang DY, Liu AH, Ma ML, Liu ZJ, Dang ZS, Liu GY, Bai Q, Lu WS, Chen PY. 2005a. A comparison of small-subunit ribosomal RNA gene sequences of bovine *Babesia* species transmitted by *Haemaphysalis* spp. in China. *Parasitol Res.* 95:145-149.

Luo JX, Yin H, Liu ZJ, Yang DY, Guan GQ, Liu AH, Ma ML, Dang ZS, Lu BY, Sun CQ, Bai Q, Lu WS, Chen PY. 2005b. Molecular phylogenetic studies on an unnamed bovine *Babesia* sp. based on small subunit ribosomal RNA gene sequences. *Vet Parasitol*. 133: 1-6.

Machado RZ, McElwain TF, Pancracio HP, Freschi CR, Palmer GH.1999. *Babesia bigemina*: immunization with purified rhoptries induces protection against acute parasitemia. *Exp Parasitol*. 93:105-108.

Malandrin L, Jouglin M, Sun Y, Brisseau N, Chauvin A. 2010. Redescription of *Babesia capreoli* (Enigk and Friedhoff, 1962) from roe deer (Capreolus capreolus): isolation, cultivation, host specificity, molecular characterisation and differentiation from *Babesia divergens*. *Int J Parasitol*. 40: 277-284.

McCosker PJ. 1981. The global importance of babesiosis. In: Ristic M, Kreier JP. (Eds), Babesiosis. Academic Press, New York, pp 1-24.

McElwain TF, Perryman LE, Davis WC, McGuire TC. 1987. Antibodies define multiple proteins with epitopes exposed on the surface of live *Babesia bigemina* merozoites. *J Immunol*. 138: 2298-2304.

McElwain TF, Perryman LE, Musoke AJ, McGuire TC. 1991. Molecular characterization and immunogenicity of neutralization-sensitive *Babesia bigemina* merozoite surface proteins. *Mol Biochem Parasitol*. 47: 213-222.

Mehlhorn H, Schein E. 1984. The piroplasms: Life cycle and sexual stages. *Adv Parasitol*. 23: 37-103.

Ménard R. 2000. The journey of the malaria sporozoite through its hosts: two parasite proteins lead the way. *Microbes Infect*. 2: 633-642.

Mendes TA1, Lobo FP, Rodrígues TS, Rodrígues-Luiz GF, daRocha WD, Fujiwara RT, Teixeira SM, Bartholomeu DC. 2013. Repeat-enriched proteins are related to host cell invasion and immune evasion in parasitic protozoa. *Mol Biol Evol*. 30: 951-963.

Mesplet M, Echaide I, Dominguez M, Mosqueda JJ, Suarez CE, Schnittger L, Florin-Christensen M. 2010. Bovipain-2, the falcipain-2 ortholog, is expressed in intraerythrocytic stages of the tick-transmitted hemoparasite *Babesia bovis*. *Parasit Vectors*. 3: 113.

Mihalca AD. 2010. The quest for piroplasms: from Babeş and Smith to molecules. *Sci Parasitol*. 11: 14-19.

Mishra VS, Stephens EB, Dame JB, Perryman LE, McGuire TC, McElwain TF. 1991. Immunogenicity and sequence analysis of recombinant p58: a neutralization-sensitive, antigenically conserved *Babesia bigemina* merozoite surface protein. *Mol Biochem Parasitol*. 47: 207-212.

Mishra VS, McElwain TF, Dame JB, Stephens EB. 1992. Isolation, sequence and differential expression of the p58 gene family of *Babesia bigemina*. *Mol Biochem Parasitol*. 53: 149-158.

Montero E, Rodríguez M, OksovY, Cheryl A. Lobo. 2009. *Babesia divergens* Apical Membrane Antigen 1 and Its Interaction with the Human Red Blood Cell. *Infect Immun*. 77: 4783-4793.

Mordmüller B, Fendel R, Kreidenweiss A, Gille C, Hurwitz R, Metzger WG, Kun JF, Lamkemeyer T, Nordheim A, Kremsner PG. 2006. Plasmodia express two threonine-peptidase complexes during asexual development. *Mol Biochem Parasitol*. 148: 79-85.

Mosqueda J, McElwain TF, Stiller D, Palmer GH. 2002a. *Babesia bovis* Merozoite Surface Antigen 1 and Rhoptry-Associated Protein 1 Are Expressed in Sporozoites, and

Specific Antibodies Inhibit Sporozoite Attachment to Erythrocytes. *Infect Immun*. 70: 1599-1603.

Mosqueda J, McElwain TF, Palmer GH. 2002b. *Babesia bovis* Merozoite Surface Antigen 2 Proteins Are Expressed on the Merozoite and Sporozoite Surface, and Specific Antibodies Inhibit Attachment and Invasion of Erythrocytes. *Infect Immun*. 70: 6448-6455.

Mosqueda J, Olvera-Ramirez A, Aguilar-Tipacamu G, Canto GJ. 2012. Current advances in detection and treatment of babesiosis. *Curr Med Chem*. 19: 1504-1518.

Niu QL, Luo JX, Guan GQ, Liu ZJ, Ma ML, Liu AH, Gao JL, Ren QY, Li YQ, Qiu JX, Yin H. 2009a. Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. *Exp Parasitol*. 121: 64-68.

Niu Q, Luo J, Guan G, Ma M, Liu Z, Liu A, Dang Z, Gao J, Ren Q, Li Y, Liu J, Yin H. 2009b. Detection and differentiation of ovine *Theileria* and *Babesia* by reverse line blotting in China. *Parasitol Res*. 104:1417-1423.

Niu Q, Guan G, Liu Z, Ma M, Li Y, Liu A, Ren Q, Liu J, Luo J, Yin H. 2012. Simultaneous detection of piroplasma infections in field *Haemaphysalis qinghaiensis* ticks by reverse line blotting. *Exp Appl Acarol*. 56: 123-132.

Niu Q, Bonsergent C, Guan G, Yin H, Malandrin L. 2013. Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group). *Vet Parasitol*. 198: 24-38.

Norimine J, Suarez CE, McElwain TF, Florin-Christensen M, Brown WC. 2002. Immunodominant epitopes in *Babesia bovis* rhoptry-associated protein 1 that elicit memory CD4(+)-T-lymphocyte responses in *B. bovis*-immune individuals are located in the amino-terminal domain. *Infect Immun*. 70: 2039-2048.

Norimine J, Mosqueda J, Suarez C, Palmer GH, McElwain TF, Mbassa G, Brown WC. 2003. Stimulation of T-helper cell gamma interferon and immunoglobulin G responses specific for *Babesia bovis* rhoptry-associated protein 1 (RAP-1) or a RAP-1 protein lacking the carboxy-terminal repeat region is insufficient to provide protective immunity against virulent *B. bovis* challenge. *Infect Immun*. 71: 5021-5032.

Ogwang C, Afolabi M, Kimani D, Jagne YJ, Sheehy SH, Bliss CM, Duncan CJ, Collins KA, Garcia Knight MA, Kimani E, Anagnostou NA, Berrie E, Moyle S, Gilbert SC, Spencer AJ, Soipei P, Mueller J, Okebe J, Colloca S, Cortese R, Viebig NK, Roberts R, Gantlett K, Lawrie AM, Nicosia A, Imoukhuede EB, Bejon P, Urban BC, Flanagan KL, Ewer KJ, Chilengi R, Hill AV, Bojang K. 2013. Safety and immunogenicity of heterologous prime-boost immunisation with *Plasmodium falciparum* malaria candidate vaccines, ChAd63 ME-TRAP and MVA ME-TRAP, in healthy Gambian and Kenyan adults. *PLoS One*. 8: e57726.

Oka M, Aikawa M, Freeman RR, Holder AA, Fine E. 1984. Ultrastructural localization of protective antigens of *Plasmodium yoelii* merozoites by the use of monoclonal antibodies and ultrathin cryomicrotomy. *Am J Trop Med Hyg*. 33: 342-346.

Oosthuizen MC, Allsopp BA, Troskie M, Collins NE, Penzhorn BL. 2009. Identification of novel *Babesia* and *Theileria* species in South African giraffe (*Giraffa camelopardalis*, Linnaeus, 1758) and roan antelope (*Hippotragus equinus*, Desmarest 1804). *Vet Parasitol.* 163: 39-46.

Ozkuc U. 1979. Studies in the serological diagnosis of *Babesia ovis* infection in sheep by the fluorescent antibody technique. *Pendik Vet. Mikrobiyol. Enst. Derg*. 11: 70-83.

Palmer GH, McElwain TF, Perryman LE, Davis WC, Reduker DR, Jasmer DP, Shkap V, Pipano E, Goff WL, McGuire TC. 1991. Strain variation of *Babesia bovis* merozoite surface-exposed epitopes. *Infect Immun*. 59: 3340-3342.

Palmer GH, McElwain TF. 1995. Molecular basis for vaccine development against anaplasmosis and babesiosis. *Vet Parasitol*. 57: 233-253.

Papadopoulos B1, Perié NM, Uilenberg G. 1996a. Piroplasms of domestic animals in the Macedonia region of Greece. 1. Serological cross-reactions. *Vet Parasitol*. 63: 41-56.

Papadopoulos B, Brossard M, Perié NM. 1996b. Piroplasms of domestic animals in the Macedonia region of Greece. 3. Piroplasms of small ruminants. *Vet Parasitol*. 63: 67-74.

Parizi LF, Reck J Jr, Oldiges DP, Guizzo MG, Seixas A, Logullo C, de Oliveira PL, Termignoni C, Martins JR, Vaz Ida S Jr. 2012. Multi-antigenic vaccine against the cattle tick *Rhipicephalus* (*Boophilus*) *microplus*: a field evaluation. *Vaccin*e. 30: 6912-6917.

Pihlajamaa T, Kajander T, Knuuti J, Horkka K, Sharma A, Permi P. 2013. Structure of *Plasmodium falciparum* TRAP (thrombospondin-related anonymous protein) A domain highlights distinct features in apicomplexan von Willebrand factor A homologues. *Biochem J*. 450: 469-476.

Pipano E. 1995. Live vaccine against hemoparasitic disease in livestock. *Vet Parasitol*. 57: 213-231.

Polar P, de Muro MA, Kairo MT, Moore D, Pegram R, John SA, Roach-Benn C. 2005a. Thermal characteristics of *Metarhizium anisopliae* isolates important for the development of biological pesticides for the control of cattle ticks. *Vet Parasitol.* 134: 159-167.

Polar P, Kairo MT, Peterkin D, Moore D, Pegram R, John SA. 2005b. Assessment of fungal isolates for development of a myco-acaricide for cattle tick control. *Vector Borne Zoonotic*. 5: 276 -284.

Précigout E, Delbecq S, Vallet A, Carcy B, Camillieri S, Hadj-Kaddour K, Kleuskens J, Schetters T, Gorenflot A. 2004. Association between sequence polymorphism in an epitope of *Babesia divergens* Bd37 exoantigen and protection induced by passive transfer. *Int J Parasitol*. 34: 585-593.

Preiser P, Kaviratne M, Khan S, Bannister L, Jarra W. 2000. The apical organelles of malaria merozoites: host cell selection, invasion, host immunity, and immune evasion. *Microbes Infect*. 2: 1461-1477.

Pruett JH, Barrett CC, Fisher WF. 1987. Kinetic development of serum antibody to purified *Hypoderma lineatum* proteins in vaccinated and non-vaccinated cattle. *Southwestern Entomol*. 12: 79-88.

Qi C, Zhou D, Liu J, Cheng Z, Zhang L, Wang L, Wang Z, Yang D, Wang S, Chai T. 2011. Detection of *Babesia divergens* using molecular methods in anemic patients in Shandong Province, China. *Parasitol Res*. 109: 241-245**.**

Ramasamy R. 1998. Molecular basis for evasion of host immunity and pathogenesis in malaria. *Biochim Biophys Acta*. 1406: 10-27.

Ranjbar-Bahadori S, Eckert B, Omidian Z, Shirazi NS, Shayan P. 2012. *Babesia ovis* as the main causative agent of sheep babesiosis in Iran. *Parasitol Res*.110:1531-1536.

Razmi GR, Naghibi A, Aslani MR, Fathivand M, Dastjerdi K. 2002. An epidemiological study on ovine babesiosis in the Mashhad suburb area, province of Khorasan, Iran. *Vet Parasitol*. 108: 109-115.

Razmi GhR, Nouroozi E. 2010. Transovarial Transmission of *Babesia ovis* by *Rhipicephalus sanguineus* and *Hyalomma marginatum*. *Iranian J Parasitol*. 5: 35-39.

Razmi G, Pourhosseini M, Yaghfouri S, Rashidi A, Seidabadi M. 2013. Molecular detection of *Theileria* spp. and *Babesia* spp. in sheep and ixodid ticks from the northeast of Iran. *J Parasitol*. 99: 77-81.

Remarque EJ, Faber BW, Kocken CH, Thomas AW. 2008. Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends Parasitol*. 24: 74-84.

Ren QY, Sun M, Guan GQ, Li YQ, Liu ZJ, Liu AH, Ma ML, Niu QL, Liu JL, Yin H. 2011. Biological control of engorged female *Haemaphysalis qinghaiensis* (Acari: Ixodidae) ticks with different Chinese isolates of *Beauveria bassiana*. *Parasitol Res*. 4: 1059-1064.

Rijpkema SG, Molkenboer MJ, Schouls LM, Jongejan F, Schellekens JF. 1995. Simultaneous detection and genotyping of three genomic groups of *Borrelia burgdorferi sensu lato* in Dutch Ixodus ricinus ticks by characterization of the amplified intergenic spacer region between 5S and 23S rRNA genes. *J Clin Microbiol*. 33: 3091-3095.

Ristic M, Kreier JP. 1981. Babesiosis, Academic Press, New York.

Ristic M. 1988. Babesiosis of Domestic Animals and Man. CRC Press, Boca Raton, FL.

Robson KJ, Hall JR, Davies LC, Crisanti A, Hill AV, Wellems TE.1990. Polymorphism of the TRAP gene of *Plasmodium falciparum*. *Proc Biol Sci*. 242: 205-216**.**

Rodríguez SD, Palmer GH, McElwain TF, McGuire TC, Ruef BJ, Chitko-McKown MG, Brown WC. 1996. CD4+T helper lymphocyte responses against *Babesia bigemina* rhoptry-associated protein-1. *Infect Immun*. 64: 2079-2087.

Rodríguez LE, Curtidor H, Urquiza M, Cifuentes G, Reyes C, Patarroyo ME. 2008. Intimate molecular interactions of *P. falciparum* merozoite proteins involved in invasion of red blood cells and their implications for vaccine design. *Chem Rev*. 108: 3656-3705.

Rojas-Caraballo J, Mongui A, Giraldo MA, Delgado G, Granados D, Millan-Cortes D, Martinez P, Rodríguez R, Patarroyo MA. 2009. Immunogenicity and protection-inducing ability of recombinant *Plasmodium vivax* rhoptry-associated protein 2 in Aotus monkeys: a potential vaccine candidate. *Vaccine*. 27: 2870-2876.

Roma GC, Furquim KCS, Bechara GH, Camargo-Mathias MI. 2010. Permethrin-induced morphological changes in oocytes of *Rhipicephalus sanguineus* (Acari: Ixodidae) semi-engorged females. *Food Chem Toxicol*. 48: 825-830.

Ruef BJ, Dowling SC, Conley PG. 2000. A unique *Babesia bovis* spherical body protein is conserved among geographic isolates and localizes to the infected erythrocyte membrane. *Mol Biochem Parasitol*. 105: 1-12.

Salama AA, Terkawi MA, Kawai S, Aboulaila M, Nayel M, Mousa A, Zaghawa A, Yokoyama N, Igarashi I. 2013. Specific antibody to a conserved region of *Babesia* apical membrane antigen-1 inhibited the invasion of *B. bovis* into the erythrocyte. *Exp Parasitol*. 135: 623-628.

Sam-Yellowe TY. 1996. Rhoptry Organelles of the Apicomplexa: Their Role in Host Cell Invasion and Intracellular Survival. *Parasitol Today*. 12: 308-316.

Samish M, Rehacek J. 1999. Pathogens and predators of ticks and their potential in biological control. *Annu Rev Entomol*. 44:159-182.

Santangelo MP, McIntosh D, Bigi F, ArmôGR, Campos AS, Ruybal P, Dellagostin OA, McFadden J, Mendumd T, Gicquel B, Winter N, Farber M, Cataldi A. 2007. Mycobacterium bovis BCG as a delivery system for the RAP-1 antigen from *Babesia bovis*. *Vaccine*. 25: 1104-1113.

Sevinc F, Dik, B, 1996. The diagnosis of *Babesia ovis* in sheep in Konya around by ELISA. *Vet Bil Derg*. 12: 73-90.

Sevinc F, Turgut K, Sevinc M, Ekici OD, Coskun A, Koc Y, Erol M, Ica A. 2007. Therapeutic and prophylactic efficacy of imidocarb dipropionate on experimental *Babesia ovis* infection of lambs.*Vet Parasitol*. 149: 65-71.

Sevinc F, Sevinc M, Ekici OD, Yildiz R, Isik N, Aydogdu U. 2013a. *Babesia ovis* infections: detailed clinical and laboratory observations in the pre- and post-treatment periods of 97 field cases *Vet Parasitol*. 191:35-43.

Sevinc F, Guler L, Sevinc M, Ekici OD, Isik N. 2013b. Determination of immunoreactive proteins of *Babesia ovis*. *Vet Parasitol*. 198: 391-5. 5.

Schnittger L, Yin H, Gubbels MJ, Beyer D, Niemann S, Jongejan F, Ahmed JS. 2003. Phylogeny of sheep and goat *Theileria* and *Babesia* parasites. *Parasitol Res*. 91: 398-406.

Schnittger L, Yin H, Bai Q, Gubbels MJ, Beyer D, Nieman S, Jongejan F, Ahmed JS. 2004. Simultaneous detection and differentiation of *Theileria* and *Babesia* parasites infecting small ruminants by reverse line blotting. *Parasitol Res*. 92: 189-196.

Schnittger L, Rodríguez AE, Florin-Christensen M, Morrison DA. 2012. *Babesia*: a world emerging. *Infect Genet Evol*. 12: 1788-1809.

Schouls LM, Van De Pol I, Rijpkema SG, Schot CS. 1999. Detection and identification of *Ehrlichia*, *Borrelia burgdorferi sensu lato* and *Bartonella* species in Dutch Ixodes ricinus ticks. *J Clin Microbiol*. 37: 2215-2222.

Shayan P, Hooshmand E, Rahbari S, Nabian S. 2007. Determination of *Rhipicephalus* spp. as vectors for *Babesia ovis* in Iran. *Parasitol Res*. 101: 1029-1033.

Shayan P, Hooshmand E, Nabian S, Rahbari S. 2008. Biometrical and genetical characterization of large *Babesia ovis* in Iran. *Parasitol Res*. 103: 217-221.

Shkap V, Pipano E, McElwain TF, Herzberg U, Krigel Y, Fish L, Palmer GH. 1994. Cross-protective immunity induced by *Babesia bovis* clones with antigenically unrelated variable merozoite surface antigens. *Vet Immunol Immunopathol*. 41: 367-374.

Silva MG, Ueti MW, Norimine J, Florin-Christensen M, Bastos RG, Goff WL, Brown WC, Oliva A, Suarez CE. 2010. *Babesia bovis* expresses a neutralization-sensitive antigen that contains a microneme adhesive repeat (MAR) domain. *Parasitol Int*. 59:294-297.

Silva MG, Ueti MW, Norimine J, Florin-Christensen M, Bastos RG, Goff WL, Brown WC, Oliva A, Suarez CE. 2011. *Babesia bovis* expresses Bbo-6cys-E, a member of a novel gene family that is homologous to the 6-cys family of *Plasmodium*. *Parasitol Int*. 60:13-18.

Sivakumar T, Okubo K, Igarashi I, de Silva WK, Kothalawala H, Silva SS,Vimalakumar SC, Meewewa AS, Yokoyama N. 2013. Genetic diversity of merozoite surface antigens in *Babesia bovis* detected from Sri Lankan cattle. *Infect Genet Evol*. 19: 134-140.

Skuce PJ, Mallon TR, Taylor SM. 1996. Molecular cloning of a putative rhoptry associated protein homologue from *Babesia divergens*. *Mol Biochem Parasitol*. 77: 99-102.

Slenning BD. 2010. Global climate change and implications for disease emergence. *Vet Pathol.* 47: 28-33**.**

Sossai S, Peconick AP, Sales-Junior PA, Marcelino FC, Vargas MI, Neves ES, Patarroyo JH. 2005. Polymorphism of the Bm86 gene in South American strains of the cattle tick *Boophilus microplus*. *Exp Appl Acarol*. 37: 199-214.

Starcovici C. 1893. Bemerkungen über den durch Babes entdeckten Blutparasiten und die durch denselben hervorgebrachten Krankheiten, die seuchenhafte Hämoglobinurie des Rindes (Babes), des Texasfieber (Th. Smith) und der Carceag der Schafe (Babes). *Zbl Bakt I Abt*. 14: 1-8.

Suarez CE, McElwain TF, Stephens EB, Mishra VS, Palmer GH. 1991a. Sequence conservation among merozoite apical complex proteins of *Babesia bovis*, *Babesia bigemina* and other apicomplexa. *Mol Biochem Parasitol*. 49: 329-332.

Suarez CE, Palmer GH, Jasmer DP, Hines SA, Perryman LE, McElwain TF. 1991b. Characterization of the gene encoding a 60-kilodalton *Babesia bovis* merozoite protein with conserved and surface exposed epitopes. *Mol Biochem Parasitol*. 46: 45-52.

Suarez CE, Palmer GH, Hines SA, McElwain TF. 1993. Immunogenic B-cell epitopes of *Babesia bovis* rhoptry-associated protein 1 are distinct from sequences conserved between species. *Infect Immun*. 61: 3511-3517.

Suarez CE, McElwain TF, Echaide I, Torioni de Echaide S, Palmer GH. 1994a. Interstrain conservation of babesial RAP-1 surface-exposed B-cell epitopes despite rap-1 genomic polymorphism. *Infect Immun*. 62: 3576-3579.

Suarez CE, Thompson SM, McElwain TF, Hines SA, Palmer GH. 1994b. Conservation of oligopeptide motifs in rhoptry proteins from different genera of erythroparasitic protozoa. *Exp Parasitol*. 78: 246-251.

Suarez CE, Palmer GH, Hötzel I, McElwain TF. 1998a. Structure, sequence, and transcriptional analysis of the *Babesia bovis rap-1* multigene locus. *Mol Biochem Parasitol*. 93: 215-224.

Suarez CE, Palmer GH, Hotzel I, Hines SA, McElwain TF. 1998b. Sequence and functional analysis of the intergenic regions separating babesial rhoptry-associated protein-1 (*rap-1*) genes. *Exp Parasitol*. 90: 189-194.

Suarez CE, Florin-Christensen M, Hines SA, Palmer GH, Brown WC, McElwain TF. 2000. Characterization of allelic variation in the *Babesia bovis* merozoite surface antigen-1 (MSA-1) locus and identification of a cross-reactive inhibition-sensitive MSA-1 epitope. *Infect Immun*. 68: 6865-6870.

Suarez CE, Palmer GH, Florin-Christensen M, Hines SA, Hötzel I, McElwain TF. 2003. Organization, transcription, and expression of rhoptry associated protein genes in the *Babesia bigemina rap-1* locus. *Mol Biochem Parasitol*. 127: 101-112.

Suarez CE, Palmer GH, LeRoith T, Florin-Christensen M, Crabb B, McElwain TF. 2004. Intergenic regions in the rhoptry associated protein-1 (*rap-1*) locus promote exogenous gene expression in *Babesia bovis*. *Int J Parasitol.* 34:1177-11184.

Suarez CE, Laughery JM, Bastos RG, Johnson WC, Norimine J, Asenzo G, Brown WC, Florin-Christensen M, Goff WL. 2011. A novel neutralization sensitive and subdominant RAP-1-related antigen (RRA) is expressed by *Babesia bovis* merozoites. *Parasitolog*y. 18:1-10.

Suarez CE, Noh S. 2011. Emerging perspectives in the research of bovine babesiosis and anaplasmosis. *Vet Parasitol*.180: 109-125.

Sultan AA, Thathy V, Frevert U, Robson KJ, Crisanti A, Nussenzweig V, Nussenzweig RS, Ménard R. 1997. TRAP is necessary for gliding motility and infectivity of *Plasmodium* sporozoites. *Cell*. 90: 511-522.

Sun C, Liu Z, Gao J, Guan G, Ma M, Luo J, Yin H. 2008. Investigations into the natural infection rate of *Haemaphysalis qinghaiensis* with *piroplasma* using a nested PCR. *Exp Appl Acarol*. 44: 107-114.

Sun M, Ren Q, Guan G, Liu Z, Ma M, Gou H, Chen Z, Li Y, Liu A, Niu Q, Yang J, Yin H, Luo J. 2011a.Virulence of *Beauveria bassiana*, *Metarhizium anisopliae* and *Paecilomyces lilacinus* to the engorged female *Hyalomma anatolicum anatolicum* tick (Acari: Ixodidae). *Vet Parasitol*. 180: 389-393.

Sun M, Ren Q, Liu Z, Guan G, Gou H, Ma M, Li Y, Liu A, Yang J, Yin H, Luo J. 2011b. *Beauveria bassiana*: Synergistic effect with acaricides against the tick *Hyalomma anatolicum anatolicum* (Acari: Ixodidae). *Exp Parasitol*. 128:192-195.

Sun Y, Moreau E, Chauvin A, Malandrin L. 2011. The invasion process of bovine erythrocyte by *Babesia divergens*: knowledge from an in vitro assay. *Vet Res*. 42: 62.

Suss-Toby E, Zimmerberg J, Ward GE. 1996. *Toxoplasma* invasion: the parasitophorous vacuole is formed from host cell plasma membrane and pinches off via afission pore. *Proc Natl Acad Sci U S A*. 93: 8413-8418.

Tattiyapong M, Sivakumar T, Ybanez AP, Ybanez RH, Perez ZO, Guswanto A, Igarashi I, Yokoyama N. 2014. Diversity of *Babesia bovis* merozoite surface antigen genes in the Philippines. *Parasitol Int*. 63: 57-63.

Telford 3rd SR, Gorenflot A, Brasseur P, Spielman A. 1993. Babesial infection in human and wildlife. In: Kreier, J.P., Baker, J.R. (Eds.), Parasitic Protozoa, vol. 5, 7. Academic Press, San Diego, pp 1-47.

Terkawi MA, Amornthep A, Ooka H, Aboge G, Jia H, Goo, YK, Nelson B, Yamagishi J, Nishikawa Y, Igarashi I, Kawazu SI, Fujisakik K, Xuan X. 2009. Molecular characterizations of three distinct *Babesia gibsoni* rhoptry-associated protein-1s (RAP-1s). *Parasitology*. 136: 1147-60.

Terkawi MA, Seuseu FJ, Eko-Wibowo P, Huyen NX, Minoda Y, AbouLaila M, Kawai S, Yokoyama N, Xuan X, Igarashi I. 2011a. Secretion of a new spherical body protein of *Babesia bovis* into the cytoplasm of infected erythrocytes. *Mol Biochem Parasitol*. 178:40-45.

Terkawi MA, Huyen NX, Wibowo PE, Seuseu FJ, Aboulaila M, Ueno A, Goo YK, Yokoyama N, Xuan X, Igarashi I. 2011b. Spherical body protein 4 is a new serological antigen for global detection of *Babesia bovis* infection in cattle. *Clin Vaccine Immunol*. 18: 337-342.

Terkawi MA, Ratthanophart J, Salama A, Aboulaila M, Asada M, Ueno A, Alhasan H, Guswanto A, Masatani T, Yokoyama N, Nishikawa Y, Xuan X, Igarashi I. 2013. Molecular Characterization of a New *Babesia bovis* Thrombospondin-Related Anonymous Protein (BbTRAP2). *PLoS One*. 8: e83305.

Theodoropoulos G, Gazouli M, Ikonomopoulos JA, Kantzoura V, Kominakis A. 2006. Determination of prevalence and risk factors of infection with *Babesia* in small ruminants from Greece by polymerase chain reaction amplification. *Vet Parasitol*. 135: 99-104.

Tian ZC, Liu GY, Yin H, Luo JX, Guan GQ, Luo J, Xie JR, Shen H, Tian MY, Zheng JF, Yuan XS, Wang FF. 2013a. RPS8--a new informative DNA marker for phylogeny of *Babesia* and *Theileria* parasites in China. PLoS One. 8:e79860.

Tian Z, Luo J, Zheng J, Xie J, Shen H, Yin H, Luo J, Tian M, Yuan X, Wang F, Liu G. 2013b. Phylogenetic analysis of *Babesia* species in China based on Cytochrome b (COB) gene. *Infect Genet Evol*. 13: 36-40.

Tian Z, Liu G, Yin H, Luo J, Guan G, Xie J, Luo J, Zheng J, Tian M, Yuan X, Wang F, Chen R, Wang H. 2013c. 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.

Tonkin ML, Crawford J, Lebrun ML, Boulanger MJ. 2013. *Babesia divergens* and *Neospora caninum* apical membrane antigen 1 structures reveal selectivity and plasticity in apicomplexan parasite host cell invasion. *Protein Sci*. 22: 114-127.

Torina A, Agnone A, Sireci G, Mosqueda JJ, Blanda V, Albanese I, La Farina M, Cerrone A, Cusumano F, Caracappa S. 2010. Characterization of the apical membrane antigen-1 in Italian strains of *Babesia bigemina*. *Transbound Emerg Dis*. 57: 52-56.

Uilenberg G, Rombach, MC, Perié NM, Zwart D. 1980. Blood parasites of sheep in the Netherlands. II. *Babesia motasi* (Sporozoa, Babesiidae). *Tijdschr Diergeneeske*. 105: 3-14.

Uilenberg G. 2006. *Babesia*-A historical overview. *Vet Parasitol*. 138: 3-10.

Vial HJ, Gorenflot A. 2006. Chemotherapy against babesiosis. *Vet parasitol*. 138: 147-160.

Vichido R1, Falcon A, Ramos JA, Alvarez A, Figueroa JV, Norimine J, Brown WC, Castro LA, Mosqueda J. 2008. Expression analysis of heat shock protein 20 and rhoptry-associated protein 1a in sexual stages and kinetes of *Babesia bigemina*. *Ann N Y Acad Sci* 1149: 136-140.

Wang CJ. 1993. Bovine piroplasmoisis. In: The diseases of Domestic Animals in China, edited by Division of Veterinary and Animal Husbandry, Minstry of Agriculture, Academic Press, 381-390.

Wang JM, Ma ML, Liu AH, Ren QY, Li AY, Liu ZJ, Li YQ, Yin H, Luo JX, Guan GQ. 2013. A sero-epidemiological survey of Chinese *Babesia motasi* for small ruminants in China. *Parasitol Res*. 112: 2387-2391.

Wengelnik K, Spaccapelo R, Naitza S, Robson KJ, Janse CJ, Bistoni F,Waters AP, Crisanti A. 1999. The A-domain and the thrombospondin-related motif of *Plasmodium falciparum* TRAP are implicated in the invasion process of mosquito salivary glands. *EMBO J*. 18: 5195-5204.

Wenyon CM. 1926. Protozoology. In: A Manual for Medical Men, Veterinarians and Zoologists, vol. II, Hafner Publish. Co., New Y ork, p. 986 (Reprinted in 1965).

Willadsen P, Bird P, Cobon GS, Hungerford J. 1995. Commercialisation of a recombinant vaccine against *Boophilus microplus*. *Parasitology*. 110: S43-S50.

Woese CR. 1987. Bacterial evolution. *Microbiol Rev*. 51: 221-271.

**Wright IG, Casu R, Commins MA, Dalrymple BP, Gale KR, Goodger BV, Riddles PW, Waltisbuhl DJ, Abetz I, Berrie DA, Bowles Y, Dimmock C, Hayes T, Kalnins H,

Leatch, G, McCrae T, Montague PE, Nisbet IT, Parrodi F, Peters JM, Scheiwe PC, Smith W, Rode-Bramanis K, White MA.1992.The development of a recombinant *Babesia* vaccine. *Vet Parasitol*. 44:3-13.

Wright IG, Goodger BV. 1988. Pathogenesis of Babesiosis, p.100–118. In M. Ristic (ed.), Babesiosis of domestic animals and man. CRC Press, Boca Raton, Fla.

Wright IG, Goodger BV, Clark IA. 1988. Immunopathophysiology of *Babesia bovis* and *Plasmodium falciparum* infections. *Parasitol Today*. 4: 214-8.

Xie J, Liu G, Tian Z, Luo J. 2013. Development of loop-mediated isothermal amplification (LAMP) for detection of *Theileria equi*. *Acta Trop*. 127: 245-50.

Yang YS, Murciano B, Moubri K, Cibrelus P, Schetters T, Gorenflot A, Delbecq S, Roumestand C. 2012. Structural and functional characterization of Bc28.1, major erythrocyte-binding protein from *Babesia canis* merozoite surface. *J Biol Chem*. 287: 9495-9508.

Yang ZJ, Long T, Jiao N, Wang BA, Zhao SK. 2003. Treatment of Small-tailed babesiosis. *Henan Agr Sci*. 12: 58-59.

Yeruham I, Hadani A, Galker F. 1998. Some epizootiological and clinical aspects of ovine babesiosis caused by *Babesia ovis* - a review. *Vet Parasitol*. 74: 153-163.

Yin H, Lu W, Luo J, Zhang Q, Dou H.1996. Experiments on the transmission of *Babesia major* and *Babesia bigemina* by *Haemaphysalis punctata* . *Vet Parasitol.* 67: 89-98.

Yin H, Lu WS, Luo JX. 1997a. Babesiosis in China. *Trop Anim Health Prod*. 29: 11S-15S.

Yin H, Lu WS, Luo JX, Zhang QC, Lu WX, Dou HF. 1997b. Isolation and morphological observation of *Babesia motasi* and *Babesia ovis* in China. *Chin J Vet Sci Technol*. 10: 7-9.

Yin H, Luo J, Guan G, Gao Y, Lu B, Zhang Q, Ma M, Lu W, Lu C, Yuan Z, Guo S, Wang B, Du H, Schnittger L, Ahmed JS, Jongejan F. 2002a. Transmission of an unidentified *Theileria* species to small ruminants by *Haemaphysalis qinghaiensis* tickscollected in the field. *Parasitol Res*. 88: S25-S27.

Yin H, Luo J, Guan G, Lu B, Ma M, Zhang Q, Lu W, Lu C, Ahmed J. 2002b. Experiments on transmission of an unidentified *Theileria* sp. to small ruminants with *Haemaphysalis qinghaiensis* and *Hyalomma anatolicum anatolicum*. *Vet Parasitol*. 108:21-30.

Yin H, Luo J, Schnittger L, Lu B, Beyer D, Ma M, Guan G, Bai Q, Lu C, Ahmed J. 2004. Phylogenetic analysis of *Theileria* species transmitted by *Haemaphysalis qinghaiensis*. *Parasitol Res*. 92: 36-42.

Yin H, Luo JX. 2007. Ticks of small ruminants in China. *Parasitol Res*. 101: S187-S189.

Yokoyama N, Suthisak B, Hirata H, Matsuo T, Inou, N, Sugimoto C, Igarashi I. 2002. Cellular localization of *Babesia bovis* merozoite rhoptry-associated protein 1 and its erythrocyte-binding activity. *Infect Immun*. 70: 5822-5826.

Yokoyama N, Bork S, Nishisaka M, Hirata H, Matsuo T, Inoue N, Xuan X, Suzuki H, Sugimoto C, Igarashi I. 2003. Roles of the Maltese cross form in the development of parasitemia and protection against *Babesia microti* infection in mice. *Infect Immun*. 71: 411-417.

Yokoyama N, Okamura M, Igarashi I. 2006. Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. *Vet Parasitol*. 138: 22-32.

Zhang CL, Liao HF, Han QP, Bao YL. 2010. Diagnosis and treatment of goat babesiosis. *Yuannan J Animal Sci Vet Medi*. 4: 19-20.

Zhao XR, Li CR, Min YG. 1986. Investigations on ovine babesiosis. *Chin J Vet Sci Technol* .1: 26-27.

Zhou JL, Yang J, Zhang GH, Nishikawa Y, Fujisaki K, Xuan XN. 2006a. *Babesia gibsoni*: An apical membrane antigen-1 homologue and its antibody response in the infected dogs. *Exp Parasitol*. 114: 329-333.

Zhou JL, Fukumoto S, Jia HL, Yokoyama N, Zhang GH, Fujisaki K, Lin JJ, Xuan XN. 2006b. Characterization of the *Babesia gibsoni* P18 as a homologue of thrombospondin related adhesive protein. *Mol Biochem Parasitol*.148: 190-198.

Zhou JL, Jia HL, Nishikawa Y, Fujisaki K, Xuan XX. 2007. *Babesia gibsoni* rhoptry-associated protein 1 and its potential use as a diagnostic antigen. *Vet Parasitol*. 145: 16-20.

Zintl A, Mulcahy G, Skerrett HE, Taylor SM, Gray JS. 2003. *Babesia divergens*, a Bovine Blood Parasite of Veterinary and Zoonotic Importance. *Clin Microbiol Rev*. 16: 622-636.

ANNEXES

Annexe 1. List of publications of Qingli Niu

Niu Q, Bonsergent C, Guan G, Yin H, Malandrin L. 2013. Sequence and organization of the rhoptry-associated-protein-1 (*rap-1*) locus for the sheep hemoprotozoan *Babesia* sp. BQ1 Lintan (*B. motasi* phylogenetic group). *Vet Parasitol*. 198: 24-38.

---

(1). **Niu Q**, Guan G, Liu Z, Ma M, Li Y, Liu A, Ren Q, Liu J, Luo J, Yin H. 2012. Simultaneous detection of piroplasma infections in field *Haemaphysalis qinghaiensis* ticks by reverse line blotting. *Exp Appl Acarol*. 56:123-132.

(2). **Niu Q**, Luo J, Guan G, Liu Z, Ma M, Liu A, Gao J, Ren Q, Li Y, Qiu J, Yin H. 2009. Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. *Exp Parasitol*. 121:64-68.

(3). **Niu Q**, Luo J, Guan G, Ma M, Liu Z, Liu A, Dang Z, Gao J, Ren Q, Li Y, Liu J, Yin H. 2009. Detection and differentiation of ovine *Theileria* and *Babesia* by reverse line blotting in China. *Parasitol Res*. 104:1417-1423.

(4). Guan G, **Niu Q**, Yang J, Li Y, Gao J, Luo J, Yin H. 2011. *Trypanosoma* (Herpetosoma) *grosi*: first isolation from Chinese striped field mouse (Apodemus agrarius). *Parasitol Int*. 60:101-104.

(5). Yang J, Guan G, **Niu Q**, Liu Z, Li Y, Liu J, Ma M, Ren Q, Liu A, Luo J, Yin H. 2013. Development and application of a loop-mediated isothermal amplification assay for rapid detection of *Borrelia burgdorferi s. l.* in ticks. *Transbound Emerg Dis*. 60: 238-44.

(6). Fu Y, Liu Z, Guan G, **Niu Q**, Li Y, Yang J, Ren Q, Ma M, Liu A, Peng Y, Luo J, Yin H. 2012. Development of Real-Time PCR for Detection of *Borrelia burgdorferi sensu l*ato in China. *Vector Borne and Zoonotic Dis*. 12: 341-345.

(7). Yang J, Liu Z, Guan G, Che R, **Niu Q**, Li Y, Liu J, Ma M, Ren Q, Liu A, Luo J, Yin H. 2012. Evaluation of molecular methods for detection of *Borrelia burgdorferi senso l*ato in ticks. *Diagn Microbiol Infect Di*s. 73: 80-83.

(8). Li Y, Liu Z, Yang J, Chen Z, Guan G, **Niu Q**, Zhang X, Luo J, Yin H. 2014. Infection of small ruminants and their red blood cells with *Theileria annulata* schizonts. *Exp Parasitol*. 137: 21-24.

(9). Ma ML, Guan GQ, Li YQ, Liu AH, Ren QY, **Niu QL**, Yin H, Luo JX.2009. Prokaryotic expression of Bm86 gene of Boophilus microplus and optimization of the expression condition. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 27:531-3.[Article in Chinese]

(10). Guan G, Moreau E, Liu J, Ma M, Rogniaux H, Liu A, **Niu Q**, Li Y, Ren Q, Luo J, Chauvin A, Yin H. 2012. BQP35 is a novel member of the intrinsically unstructured protein (IUP) family which is a potential antigen for the sero-diagnosis of *Babesia* sp. BQ1 (Lintan) infection. *Vet Parasitol*. 187:421-30.

(11). Ren Q, Sun M, Guan G, Li Y, Liu Z, Liu A, Ma M, **Niu Q**, Liu J, Yin H, Luo J. 2011. Biological control of engorged female *Haemaphysalis qinghaiensis* (Acari: Ixodidae) ticks

with different Chinese isolates of *Beauveria bassiana*. *Parasitol Res*. 109: 1059-64.

(12). Li Y, Luo J, Guan G, Ma M, Liu A, Liu J, Ren Q, **Niu Q**, Lu B, Gao J, Liu Z, Dang Z, Tian Z, Zhang B, He Z, Bai Q, Yin H. 2009. Experimental transmission of *Theileria uilenbergi* infective for small ruminants by *Haemaphysalis longicornis* and *Haemaphysalis qinghaiensis*. *Parasitol Res*. 104: 1227-1231.

(13). Ma M, Guan G, Chen Z, Liu Z, Liu A, Gou H, Ren Q, Li Y, **Niu Q**, Yang J, Yin H, Luo J. 2013. The life cycle of *Haemaphysalis qinghaiensis* (Acari: Ixodidae) ticks under laboratory conditions. *Exp Appl Acarol*. 59: 493-500.

(14). Li Y, Zhang X, Liu Z, Chen Z, Yang J, He H, Guan G, Liu A, Ren Q, **Niu Q**, Liu J, Luo J, Yin H. 2014. An epidemiological survey of *Theileria* infections in small ruminants in central China. *Vet Parasitol.* 200: 198-202.

(15). Liu A, Guan G, Du P, Gou H, Liu Z, Liu J, Ma M, Yang J, Li Y, **Niu Q**, Ren Q, Bai Q, Yin H, Luo J. 2012. Loop-mediated isothermal amplification (LAMP) method based on two species-specific primer sets for the rapid identification of Chinese *Babesia bovis* and *B. bigemina*. *Parasitol Int*. 61: 658-663.

(16). Liu A, Guan G, Du P, Liu Z, Gou H, Liu J, Yang J, Li Y, Ma M, **Niu Q**, Ren Q, Bai Q, Yin H, Luo J. 2012. Loop-mediated isothermal amplification (LAMP) assays for the detection of *Theileria annulata* infection in China targeting the 18S rRNA and ITS sequences. *Exp Parasitol*. 131:125-9.

(17). Sun M, Ren Q, Guan G, Liu Z, Ma M, Gou H, Chen Z, Li Y, Liu A, **Niu Q**, Yang J, Yin H, Luo J. 2011. Virulence of *Beauveria bassiana*, *Metarhizium anisopliae* and *Paecilomyces lilacinus* to the engorged female *Hyalomma anatolicum anatolicum* tick (Acari: Ixodidae). *Vet Parasitol*. 180: 389-93.

(18). Liu AH, Guan GQ, Liu JL, Liu ZJ, Leblanc N, Li YQ, Gao JL, Ma ML, **Niu QL**, Ren QY, Bai Q, Yin H, Luo JX. 2011. Polymorphism analysis of Chinese *Theileria sergenti* using allele-specific polymerase chain reaction of the major piroplasm surface protein gene. *J Parasitol*.97: 116-21.

(19). Liu A, Guan G, Liu Z, Liu J, Leblanc N, Li Y, Gao J, Ma M, **Niu Q**, Ren Q, Bai Q, Yin H, Luo J. 2010. Detecting and differentiating *Theileria sergenti* and *Theileria sinensis* in cattle and yaks by PCR based on major piroplasm surface protein (MPSP). *Exp Parasitol*. 126: 476-481.

(20). Liu A, Guan G, Du P, Gou H, Zhang J, Liu Z, Ma M, Ren Q, Liu J, Yang J, Li Y, **Niu Q**, Bai Q, Yin H, Luo J. 2013. Rapid identification and differentiation of *Theileria sergenti* and *Theileria sinensis* using a loop-mediated isothermal amplification (LAMP) assay. *Vet Parasitol*. 191:15-22.

(21). **Niu Qingli**, Qiu Jiaxiang, Guan Guiquan, Ma Miling, Liu Zhijie, Dang Zhisheng, Liu Aihong, Gao Jinliang, Ren Qiaoyun, Li Youquan, Liu Junlong, Bi Qi, Luo Jianxun, Yin Hong. 2008. Sequencing and Phylogenetic analysis of ITS genes of *Theileria luwenshuni* and *T. uilenbergi* isolated in China. *Veterinary Science in China*. 5: 365-369. **(Article In Chinese)**

(22). **Niu Qingli**, Luo Jianxun, Yin Hong. 2008. Advances and application of rDNA-ITS in the molecular taxonomy of parasite. Chinese journal of veterinary parasitology. 16: 41-47. (**Article In Chinese**)

(23). **Niu Qingli**, Luo Jianxun, Yin Hong. 2009. Progresson Lyme Disease in China. Progress in Veterinary Medicine. 30: 89-93. (**Article In Chinese**)

(24). **Niu Qingli**, Guan Guiquan, Yang Jifei, Fu Yuguang, Ma Miling, Li Youquan, Liu Junlong, Hao Xuefeng, Luo Jianxun, Yinhong. 2010. Investigations of *Borrelia burgdorferi* infection in ticks by PCR. Chinese Journal of Preventive Veterinary Medicine. 12: 984-987. (**Article In Chinese**)

(25). **Niu Qingli**, Yang Jifei, Guan Guiquan, Fu Yuguang, Ma Miling, Li Youquan, Liu Junlong, Liu Zhijie, Liu Aihong, Ren Qiaoyun, Hao Xuefeng, Luo Jianxun, Yin Hong. 2010. Identification and phylogenetic analysis of Lyme disease Borrelia spp. isolated from Shangzhi Prefecture of Heilongjiang Province. *Veterinary Science in China*. 40: 661-666. (**Article In Chinese**)

Annexe 2: Article: **Niu Q**, Guan G, Liu Z, Ma M, Li Y, Liu A, Ren Q, Liu J, Luo J, Yin H. 2012.Simultaneous detection of piroplasma infections in field *Haemaphysalis qinghaiensis* ticks by reverse line blotting..*Exp Appl Acarol*. 56: 123-132.

Exp Appl Acarol (2012) 56:123–132
DOI 10.1007/s10493-011-9503-z

Simultaneous detection of piroplasma infections in field *Haemaphysalis qinghaiensis* ticks by reverse line blotting

Q. Niu · G. Guan · Z. Liu · M. Ma · Y. Li · A. Liu · Q. Ren · J. Liu · J. Luo · H. Yin

Received: 12 August 2011 / Accepted: 6 November 2011 / Published online: 6 December 2011
© Springer Science+Business Media B.V. 2011

Abstract Piroplasmosis is caused by tick-borne haemoprotozoan parasites in the genera of *Theileria* and *Babesia*, in which numbers of agents are highly pathogenic for cattle, sheep and goats. We developed a reverse line blot (RLB) assay for detection and differentiation of four different parasites, in which 18S ribosomal RNA gene sequence was amplified with a set of universal primers specific for all members in the genera of *Theileria* and *Babesia*; and the probes were designed on the basis of hypervariable region 4 (V4 region) of 18S rRNA gene. Three *Theileria* and one *Babesia* can be detected simultaneously on this system and it was sensitive to detect a parasitemia level between 10^{-5} and 10^{-8}%. A total of 149 *Haemaphysalis qinghaiensis* ticks collected from Lintan County of Gannan Tibetan Autonomous Region was tested by RLB. Among these, 136 tick samples were also tested by a nested PCR assay developed previously. After comparison of these results, it showed that more *T. luwenshuni* was detected in RLB assay, while more *T. uilenbergi* was detected in *H. qinghaiensis* ticks by nested PCR. The RLB has shown capability for simultaneous detection of four species of piroplasm in *H. qinghaiensis* ticks, indicating its usefulness for epidemiological studies of piroplasmosis.

Keywords Piroplasma · *Haemaphysalis qinghaiensis* · Reverse line blot (RLB) · Detection

Introduction

Piroplasmosis is also called “texas fever” or “haemosporidiasis”, it is caused by *Babesia* and *Theileria* (Apicomplexa, Piroplasmia, Babesiidae and Theileriidae). The disease is usually characterized by fever, hemolytic anemia, haemoglobinuria, and even death in severe case. The cattle and sheep/goat grazing areas are mainly located in northland China,

Q. Niu · G. Guan · Z. Liu · M. Ma · Y. Li · A. Liu · Q. Ren · J. Liu · J. Luo · H. Yin (✉)
State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Key Laboratory of Epizootic Diseases of Grazing Animals of Ministry of Agriculture, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou 730046, China
e-mail: yinhong@caas.net.cn

so the piroplasm of ruminants are common parasites in the north of China, and considered to be one of the most frequent and important tick-borne diseases of ruminants and responsible for vast economic losses.

The prevalence of piroplasmosis correlates with the geographic distribution and activity of vector ticks. In addition, environmental conditions changing, especially global warming, favors tick survival and reproduction, which correlates with a significant increase in the abundance of ticks (Slenning 2010).

Because there have all kinds of ticks in the world, of the 109 species of ticks identified to date in China, thus this disease is also widely distributed. *Haemaphysalis qinghaiensis* belongs to family of Ixodidae, which is a new species named by Teng Kuofan (Teng 1980). It distributes in Qinghai, Gansu, Ningxia, Sichuan, Yunnan provinces and at altitude of 1,600–4,200 meters areas, and is not recorded in the rest of the world. It is a three-host tick living in grass or bushes in mountain area (Yuan et al. 2002). Our previous studies have confirmed that *H. qinghaiensis* tick is the transmission vector of *Theileria luwenshuni*, *Theileria uilenbergi* and *Babesia motasi* infective to small ruminants, and *T. sinense* infective to cattle and yaks (Guan et al. 2002; Yin et al. 2002a, b, c, 2004, 2007).

Based on the 18S rRNA gene sequencing, PCR (Sun et al. 2008) was developed for distinguishing between these four species. However, PCR assays do not generally detect mixed infections, although there are some amplification protocols that can detect mixed piroplasm infections to some extent (Birkenheuer et al. 2003; Criado-Fornelio et al. 2003). but there is unsatisfactory. RLB hybridization assay was initially developed as a reverse dot blot assay for the diagnosis of sickle cell anemia (Saiki et al. 1988). The essence of both techniques is the hybridization of PCR products to a membrane where specific probes immobilized on in order to determine differences in the amplified sequences. In the "line" approach, multiple samples can be analyzed against multiple probes to enable simultaneous detection. The RLB is more sensitive and specific for identifying small piroplasms and it can simultaneously detect the genus and differentiate between the species of haemoparasites in blood, organs and ticks. It has been already used for the detection and characterization all known *Theileria* and *Babesia* species in previous studies successfully (Gubbels et al. 1999; Schnittger et al. 2004; Niu et al. 2009). In the present study, we aimed at developing a RLB assay which allows detection and differentiation of *T. luwenshuni*, *T. uilenbergi*, *T. sinensis* and *B. motasi*.

Materials and methods

Parasite stocks and DNA extraction

All parasite stocks used in this study were described in detail previously (Schnittger et al. 2004; Yin et al. 2004; Liu et al. 2007). *Babesia motasi* isolate was isolated from a sheep infested with adult *H. qinghaiensis* ticks from Lintan, Gansu Province (Guan et al. 2002). *Theileria luwenshuni* (Lintan), *T. uilenbergi* (Longde) and *T. sinensis* (Weiyuan) as defined previously (Bai et al. 2002a, b; Schnittger et al. 2004; Yin et al. 2004; Liu et al. 2010). All parasite isolates were stored as EDTA–blood stabilates in liquid nitrogen. 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 using spectrophotometer (Sedico Ltd., Biotech Engineering, UK). Control DNA was isolated from the venous blood of uninfected cattle/sheep and *H. qinghaiensis*.

Collection of *Haemaphysalis qinghaiensis* tick samples and DNA extraction

Total of 149 adult *H. qinghaiensis* ticks were collected from grass tips in Lintan County of the Gannan Tibetan Autonomous Region. One hundred thirty-six of them were separated as described previously (Sun et al. 2008). Each tick was soaked in 70% ethanol in 15 min, dried, and grounded in a separate 1.5 ml tube to avoid cross contamination. The sample was incubated with proteinase K for 2 h at 56°C, and then boiled at 100°C for 10 min to inactivate proteinase K. After centrifugation, the supernatant was transferred to a fresh sterile microtube, and DNA was extracted using a Genomic DNA Purifcation Kit (Gentra, USA) according to the manufacturer's instructions.

Primers and probes designation

General and species-specific RLB oligonucleotide probes were deduced from the hypervariable V4 region of the 18S rRNA gene sequences (Table 1). A pair general primers was: RLB-F (5′-GAGGTAGTGACAAGAAATAACAATA-3′), RLB-R (biotin-5′-TCTTCGA TCCCCTAACTTTC-3′) (Gubbels et al. 1999). The three *Theileria* species included in the assay were *T. luwenshuni T. uilenbergi* and *T. sinensis.* The *Babesia* specie was *B. motasi*. A catch-all *Theileria* and *Babesia* species control oligonucleotide was as shown in Table 2. All the probes were diluted to 50–800 pmol/150 μl in 500 mM NaHCO3 (pH 8.4). The sequences and optimum concentrations of the oligonucleotide probes used are summarized in Table 1. These sequences were aligned using MUTALIN online interface (http://www.ncbi.nlm.nih.gov/). The variable regions of these sequences were flanked by two PCR primers that were used for amplification.

PCR amplification and RLB hybridization

The amount of DNA was assessed to 100 ng by photometry. Negative control DNA was isolated from the venous blood of uninfected cattle, sheep/goat and *H. qinghaiensis*. The PCR reaction was performed as described in detail elsewhere. Amplification products were

Table 1 Sequence and concentration of specific oligonucleotide probes *Theileria* sp. and *Babesia* sp.

Oligonucleotides	Sequence 5′–3′	Optimised concentrations	Reference
Cat-all (ca 841–859)	CTGTCAGAGGTGAAATTCT	200 pmol	Niu et al. 2009
B-all (B-all 745–766)	CCTTGGTAATGGTTAATAGGAA	200 pmol	Niu et al. 2009
T-all (T-all 811–832)	TACCAAAGTAATGGTTAATAGG	50 pmol	Niu et al. 2009
B. motasi (B. m 466–487)	GAATGATGCCGACTTAAACCCT	200 pmol	Niu et al. 2009
T. sinensis (T. s 627–645)	TCGCATCTCTTGCTGAGTG	800 pmol	This study
T. luwenshuni (T. l 628–647)	ATCTTCTTTTTGATGAGTTG	400 pmol	Niu et al. 2009
T. uilenbergi (T. u 678–697)	TGCATTTTCCGAGTGTTACT	800 pmol	Niu et al. 2009

Table 2 Species and their recognition pattern by oligonuclotide probes

Isolates	Species	Oligonucleotide Species-specific		Oligonucleotide Group-specific
T. luwenshuni (Lintan)	*T. luwenshuni*	T. l	T-all	(Cat- All)
T. sinensis(Weiyuan)	*T. sinensi*	T. s	T-all	(Cat- All)
T. uilenbergi (Longde)	*T. uilenbergi*	T. u	T-all	(Cat- All)
B. motasi. (Lintan)	*B. motasi*	B. m	B-all	(Cat- All)

analyzed by electrophoresis in a 1% agarose gel. Preparation of RLB membrane and hybridization were carried out as described previously (Niu et al. 2009).The diluted probes were distributed into miniblotter slots, and incubated for 1 min for linking of oligonucleotides to the membrane. After aspiration of solutions, the membrane was incubated in 100 mM NaOH for 10 min, washed at 60°C for 5 min, then at 42°C for 5 min in 2 × SSPE, 0.1% (SDS). Subsequently, the membrane was placed perpendicular to its previous orientation into the miniblotter. The denatured PCR products were aliquoted into the slots of the miniblotter for 60 min at 42°C, then aspirated and the membrane washed at 42°C for 10 min in 2 × SSPE, 0.1% SDS. Subsequently, the membrane was treated at 42°C for 30 min with peroxidase-labeled streptavidin diluted 1:4,000 in 2 × SSPE/0.1% SDS, washed twice at 42°C for 10 min and twice at room temperature for 5 min in 2 × SSPE, 0.1% SDS. Finally chemiluminescence detection was performed according to standard procedures (Amersham).

Specificity and sensitivity of the RLB

The RLB was examined by 20 μl of 100 ng amplified product of *T. luwenshuni*, *T. uilenbergi*, *T. sinensis* and *B. motasi* extracted from infected blood. Meanwhile, the uninfected ovine/cattle, pathogen-free tick blood DNA and water were used as negative and blank controls. The sensitivity of the RLB was tested by hybridizing 20 μl of the tenfold diluted PCR product. The genomic DNA of the four pathogens was serially diluted from 100 ng/μl to 0.000001 pg/μl by tenfold dilution as templates for the PCR, and then examined by RLB.

Epidemiological study and analysis

A total of 149 *H. qinghaiensis* ticks were detected by the established RLB assay and the positive and negative samples were recorded to assess the 4 piroplasma infection rate in the detected ticks. Meanwhile, the result of 136 of the samples was compared with the nest-PCR. The differences between females and males were evaluated by using Chi-square test (χ^2).

Results

Specificity of RLB

Generated DNA fragments were hybridized to the oligonucleotide-linked membrane followed by chemiluminescence detection resulting in signals of equal intensity for each oligonucleotide. All probes bound to their respective target sequence only, resulting in clear recognition of individual strain, species, or group. No signal was observed when ovine/cattle genomic DNA, genomic DNA of uninfected pathogen tick and water were used as control (Fig. 1). Each *Theileria* species is identified by three oligonucleotide probes: the catch-all probe (ca 841–859), a probe recognizing *Theileria* parasites (T-all 811–832), and species-specific probes for either *T. luwenshuni* (T. u l628–647), or *T. uilenbergi* (T. u 678–697) or *T. sinensis* (T. s 627–645). *Babesia* is recognized by the following there oligonucleotides: a catch-all probe (ca 841–859), a probe recognizing *Babesia* parasites (B-all 745–766), and species-specific probes for either *B. motasi* species (B. m 466–487) (Fig. 1).

Sensitivity of RLB

Sensitivity of the assay was assessed by detection of serially diluted genomic DNA of *Babesia* and *Theileria*. Accordingly, the RLB is capable of amplification of about 1 pg of gDNA of *B. motasi*, 0.01 pg gDNA of *T. luwenshuni*, 0.01 pg gDNA of *T. uilenbergi*, 10 pg gDNA of *T. sinensis*. When a nest-PCR was used as control, it was restricted to detect about 100 pg, 1 ng, 1 pg, and 10 pg, respectively. The detection limit of the RLB was higher than that of the nest-PCR, obviously. The sensitivity of the RLB could be considerably enhanced to detect a parasitemia level of at least by amplification of PCR.

Detection of parasites in ticks

From 149 samples analysed, the prevalence of each hemoparasite species is summarized in Table 3. The positive samples of *T. luwenshuni*, *T. uilenbergi*, *T. sinensis* and *B. motasi* were 68 (45.6%), 12 (8.1%), 8 (5.4%) and 10 (6.7%) by RLB, respectively. The infection

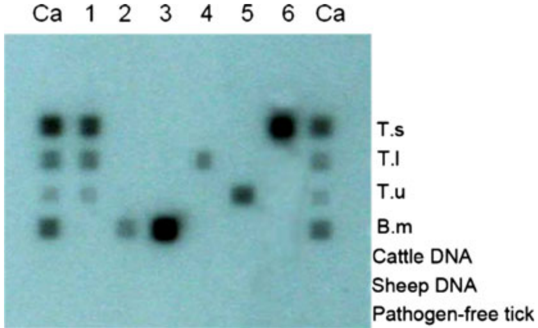


Fig. 1 Reverse line blot (RLB) assay of standard positive specificity test of 4 piroplasms. Ca: Probe of all of piroplasms; *Lane* 1–6: Probe of all of *Theileria*; Probe of all of *Babesia*; Probe of all of *B.motasi*; Probe of *B.* motasi (Lintan); Probe of all of *T. luwenshuni*; Probe of all of *T. uilenbergi*; Probe of all of *T. sinensis*; *row* 1–7: *T. sinensis* (Weiyuan strain); *T. luwenshuni* (Lintan strain); *T. uilenbergi* (Longde strain); *B. motasi* (Lintan strain); Cattle DNA; Sheep DNA; Pathogen-free tick

Table 3 Comparison of the results of RLB and nested PCR

Piroplasma	Female (66/55[a]) positive		Male (83/81[a]) positive		Total (149/136[a])		Infection rate (%)	
	RLB	PCR	RLB	PCR	RLB	PCR	RLB	PCR
T. luwenshuni	25	5	43	4	68	9	45.64	6.62
T. uilenbergi	5	27	7	35	12	62	8.1	45.59
T. sinensis	3	2	5	0	8	2	5.4	1.47
B. motasi	6	0	4	1	10	1	6.7	0.74
Total	25	27	43	35	68	62	45.64	45.59
Infection rate	37.9%	49.10%	51.8%	43.21%	–	–	–	–

[a] Number of sample examined with nested PCR, which was done by Sun et al. (2008)

rate of *T. uilenbergi* was 62 (45.6%) by nested PCR. Therefore the infection rate of *T. uilenbergi* was brought almost opposite from *T. luwenshuni* by using two methods. Most of the samples analyzed were negative with *T. uilenbergi* by RLB. It revealed that most of the ticks were infected by *T. luwenshuni* genotype, among all these positive samples, 30 were positive for more than one pathogen. The prevalence of *T. sinensis* and *B. motasi* were low with only eigth and ten positive samples respectively. Sixty eight *H. qinghaiensis* ticks (25 females and 43 males) were infected with *T. luwenshuni* and the infection rate was 37.9% and 51.8%. Five females and 7 males were infected with *T. uilenbergi*; 3 females and 5 males were infected with *T. sinensis* and 6 females and 4 males were infected with *B. motasi*. The high prevalence (45.6%) of *T. luwenshuni* infection indicated the severity of theileriosis; while the presence of *T. sinensis* and *B. motasi* implies the potential co-existence of these pathogens in the area studied.

Discussion

Theileriosis of sheep and goats was first reported in Sichuan province, 50 years ago (Yang et al. 1958), Babesiosis was first reported in small ruminants in China in 1982 (Chen 1982). Both ELISA and PCR have been developed for detection of the diseases (Gao et al. 2002, Guo et al. 2002, 2007; Yin et al. 2008; Sun et al. 2008). However, it has been demonstrated that the common vector of *T. luwenshuni*, *T. uilenbergi*, *T. sinensis* and *B. motasi* are *H. qinghaiensis*, which was found in the most endemic areas of these pathogens (Bai et al. 2002b; Guan et al. 2002; Yin et al. 2002a, 2008). Sheep and goats were infected by two or three of these species. The detection and discrimination of a parasite in the definitive host or the transmission vector are crucial for understanding the epidemiology of the disease. In the past, very limited information is available about the diagnosis of these *Theileria* and *Babesia* species by the same vector *H. qinghaiensis* transmitted simultaneously. In order to better understand theileriosis and babesiosis in China, more evidence about role of the *H. qinghaiensis* transmissed in the piroplasma is needed.

Many studies on the diagnosis of tick-borne hematozoa have been published during the last decade. Nucleic acid-based methods offer a variety of tools for the diagnosis of parasites. Generally, the detection and diagnosis of piroplasmosis consist of molecular methods, which are as following: (1) PCR (Almeria et al. 2001; Yin et al. 2008); (2) reverse line blotting (RLB) (Almeria et al. 2002; Niu et al. 2009; Silva et al. 2010); (3) real time PCR (Jaffer et al. 2010); (4) LAMP (Guan et al. 2008; Salih et al. 2008). None of

these methods could be considered better than another. PCR is a cheap alternative for small-scale handling few samples. On the contrary, large-scale laboratories prefer methods amenable to automation, like RLB, PCR-ELISA or qPCR. Based on V4 region of the 18S rRNA gene of *Theileria* and *Babesia* species, reverse line blot (RLB) hybridisation method has been developed for the detection and identification of *Theileria* and *Babesia* species (Nagore et al. 2004; Ica et al. 2007; Altar et al. 2008; Niu et al. 2009). Our research was focused on development of a very convenient and useful RLB method being able to detect all these four pathogens simultaneously and other piroplasms infective to small ruminants, which can be used for diagnosis, detection, quarantine, and epidemiological survey. It is quickly becoming a standard molecular tool for diagnostic and epidemiological studies in an increasing number of laboratories all over the world.

In this study, four piroplasmas species were detected with RLB by species-specific oligonucleotides, and no cross-reactions were observed. The assay allows a remarkable higher specificity than other methods through comparison. The *Theileria* genus-specific probe specifically detected all *Theileria* species (T-all), and the *Babesia* genus-specific probe specifically detected all *Babesia* species (B-all). Meanwhile, four species-specific probe detected distinct species (*T. luwenshuni*; *T. uilenbergi*; *T. sinensis*; B. *motasi*). This assay can be used to identify piroplasms in *H. qinghaiensis* ticks.

To determine the detection limit of the RLB, genomic DNA (100 ng/ul) of *T. luwenshuni*; *T. uilenbergi*; *T. sinensis* and *B. motasi* was serially tenfold diluted from 10^{-1} to 10^{-12}. we found that the detection limit of the RLB were 0.01, 0.01, 10 and 1 pg of *T. luwenshuni*; *T. uilenbergi*; *T. sinensis* and *B. motasi* in this study, while the detection level of a nest-PCR previous, carried out in parallel, was restricted to about 10 pg, 1 ng, 10 and 100 pg, respectively (Sun et al. 2008). The sensitivity of RLB is relatively higher than nest PCR. The results showed that RLB could be used for detecting *Theileria* and *Babesia* species, even in very low parasitemias in ticks. Many RLB studies were employed for the detection of various piroplasma species and similar senitivity of RLB for *Theileria* and *Babesia* species (Gubbels et al. 1999; Schnittger et al. 2004; Altay et al. 2008; Niu et al. 2009). But no report on use of RLB for detecton of piroplasma pathogens in vector *H. qinghaiensis* ticks was reported.

RLB is a powerful tool in field studies for detection of piroplasma infections, particularly to determine vector and carrier animals. As there was very limited data on the distribution of tick-borne pathogens in potential vector ticks and epidemiological situation in China. To assess the sensitivity and specificity of RLB and the epidemiological situation of these parasites' infection are very important work.

Haemaphysalis qinghaiensis is a specific tick species in China. Piroplasma were transmissed with *H. qinghaiensis* to show in China previously (Guan et al. 2002; Yin et al. 2002a, b, c; Li et al. 2007, 2009; Guan et al. 2010). In the present study, we determined prevalence of piroplasm infection (*Theileria* + *Babesia*) by RLB in ticks. Our data suggests that 68 out of 149 samples were positive for the presence of *Theileria* and *Babesia* parasites, and the sensitivity was higher than nest-PCR (62/136) (Table 3). All 68 positive ticks were infected with *T. luwenshuni* by RLB, while 62 positive ticks were infected with *T. uilenbergi* with nest-PCR. The infection rate of *T. luwenshuni* and *T. uilenbergi* were opposite through two methods on the whole. It has been still unknown yet why the positive rate by nest-PCR was reverse that by RLB with the *T. luwenshuni* and *T. uilenbergi* species. It would be of interest to collect more samples from that area and evaluate the infection of *T. luwenshuni* and *T. uilenbergi* in *H. qinghaiensis* ticks in the future study. Nevertheless, our result showed that Theileria. sp is widely distributed in China, and *T. luwenshuni* was the dominant species, which is in agreement with the pervious data by

PCR and RLB (Yin et al. 2008; Niu et al. 2009). According to the existing piroplasm prevalence information in China, the opposite result found in two *Theileria* was not surprising, most of theileriaosis cause due to mixed infection of *T. luwenshuni* and *T. uilenbergi* in prevalent area, single case also occurs, but this is very few (Li 2007). Eigth and ten positive tick samples had infected with *T. sinensis* and *B. motasi*. The general results were in concordance with previous reports that the positive rate of *T. sinensis* and *B. motasi* are detected with RLB and nest-PCR and all of them the mixed infection to parallel with *T. luwenshuni*. Analysis of prevalence by χ^2 test indicated that there was no significant difference ($P > 0.05$) in piroplasmida infection levels between females (25/66, prevalence 0.378) and males (43/83, prevalence 0.518).The infection rate of males detected by using RLB was greater than by nest-PCR (35/81, prevalence 0.432), while the infection rate of females was lower than in nest-PCR (27/55, prevalence 0.491). But the difference of ticks of either sex was not statistically significant.

In conclusion, RLB is a potential diagnostic tool for piroplasm infection not only for animals showing clinical signs and suffering acute infections, but for those apparently healthy animals with potential subclinical infections and vector ticks, which can be an optimal approach for the detection and discrimination of these important parasites. There is increasing evidence that piroplasma are present in wild-collected *H. qinghaiensis* from Lintan County of the Gannan Tibetan Autonomous Region with high to moderate prevalence. Further studies are necessary to detect four piroplasma pathogens. RLB could be a useful detection method of piroplasma and for discrimination among *T. luwenshuni*, *T. uilenbergi*, *T. sinensis* and *B. motasi* infections in *H. qinghaiensis*, and determination of them in adult of *H. qinghaiensis* has exhibited the vector potential of these ticks. The enhanced sensitivity of RLB-based arrays allows direct identification of species and mixed infections and provides more accurate prevalence data. This information would prove a valuable contribution to the successful implementation of control strategies for parasitic hematozoa of domestic animals. We recommend its use for integrated epidemiological monitoring of tick-borne disease, since RLB can also be used for screening ticks and can easily be expanded to include additional haemoparasite species.

Acknowledgments This study was financially supported by the Natural Sciences Foundation of China (no. 30800820), Supporting Plan (2007BAD40B00), "863" Project (2006AA10A207), the '973' Project, Specific Fund for Sino-Europe Cooperation, MOST, China; Key Project of Gansu Province (0801NKDA033), Lanzhou, Gansu, China. State Key Laboratory of Veterinary Etiological Biology Project (SKLVEB2008ZZKT019) and National Public Interests Research Institute Basic Scientific Research Expenses Special Fund Project. The research was also facilitated by ARBOZOONET, ASFRISK, ICTTD 3, EPIZONE and PIROVAC of European Commission, Brussels, Belgium.

References

Almeria S, Castellà J, Ferrer D, Ortuño A, Estrada-Peña A, Gutiérrez JF (2001) Bovine piroplasms in Minorca (Balearic Islands, Spain): a comparison of PCR-based and light microscopy detection. Vet Parasitol 99:249–259

Almeria S, Castella J, Ferrer D, Gutierrez JF, Estrad-Pena A, Sparagano O (2002) Reverse line blot hybridization used to identify hemoprotozoa in Minorcan cattle. Ann N Y Acad Sci 969:78–82

Altay K, Aktas M, Dumanli N, Aydin MF (2008) Evaluation of a PCR and comparison with RLB for detection and differentiation of *Theileria* sp. MK and other Theileria and Babesia species of small ruminants. Parasitol Res 103:319–323

Bai Q, Liu GY, Yin H, Zhao QZ, Liu DK, Ren JX, Li X (2002a) *Theileria sinensis* sp nov: a new species of bovine *Theileria*—classical taxonomic studies. Acta Veterinaria et Zootechnica Sinica 33:73–77

Bai Q, Liu G, Yin H, Zhao Q, Liu D, Ren J, Li X (2002b) *Theileria sinensis* sp. nov: a new species of bovine Theileria, molecular taxonomic studies. Acta Anim Husb Vet Sci 33:185–190

Birkenheuer AJ, Levy MG, Breitschwerdt EB (2003) Development and evaluation of a seminested PCR for detection and differentiation of *Babesia gibsoni* (Asian genotype) and *B.* canis DNA in canine blood samples. J Clin Microbiol 41:4172–4177

Chen DM (1982) Investigations on ovine piroplasmosis. Chin J Vet Sci Technol 3:31–32

Criado-Fornelio A, Martinez-Marcos A, Buling-Sarana A, Barba-Carretero JC (2003) Molecular studies on *Babesia*, *Theileria* and *Hepatozoon* in southern Europe. Part I. Epizootiological aspects. Vet Parasitol 113:189–201

Gao YL, Yin H, Luo J, Ouyang W, Bao H, Guan G, Zhang Q, Lu W, Ma M (2002) Development of an enzyme-linked immunosorbent assay for the diagnosis of *Theileria* sp infection in sheep. Parasitol Res 13:S8–S10

Guan GQ, Yin H, Luo J, Lu W, Zhang Q, Gao Y, Lu B (2002) Transmission of *Babesia* sp to sheep with field-collected *Heamaphysalis qinghaiensis*. Parasitol Res 88:S22–S24

Guan GQ, Chauvin A, Luo J, Inoue N, Moreau E, Liu Z, Gao J, Thekisoe OMM, Ma M, Liu A, Dang Z, Liu J, Ren Q, Jin Y, Sugimoto C, Yin H (2008) The development and evaluation of a loop-mediated isothermal amplification (LAMP) method for detection of Babesia spp. infective to sheep and goats in China. Exp Parasitol 120:39–44

Guan GQ, Moreau E, Liu J, Hao X, Ma M, Luo J, Chauvin A, Yin H (2010) *Babesia* sp. BQ1 (Lintan): Molecular evidence of experimental transmission to sheep by *Haemaphysalis qinghaiensis* and *Haemaphysalis longicornis*. Parasitol Int 59:265–267

Gubbels MJ, de Vos AP, van der Weide M, Viseras J, Schouls LM, de Vries E, Jongejan F (1999) Simultaneous detection of bovine *Theileria* and *Babesia* species by reverse line blot hybridization. J Clin Microbiol 37:1782–1789

Guo S, Yuan Z, Wu G, Wang W, Ma D, Du H (2002) Epidemiology of ovine theileriosis in Ganan region, Gansu Province, China. Parasitol Res 88:S36–S37

Guo SZ, Mu Y, Liu Z, Ma D, Yang S, Ge G, Fang B, Ga D, Ma M, Luo J, Yin H, Seitzer U, Ahmed J (2007) Serological investigation of ovine theileriosis by ELISA in Gannan Tibet Region of Gansu Province in China. Parasitol Res 101:S197–S200

Ica A, Vatansever Z, Yildirima A, Duzlu O, Inci A (2007) Detection of *Theileria* and *Babesia* species in ticks collected from cattle. Vet Parasitol 148:156–160

Jaffer O, Abdishakur F, Hakimuddin F, Riya A, Wernery U, Schuster R (2010) A comparative study of serological tests and PCR for the diagnosis of equine piroplasmosis. Parasitol Res 106:709–713

Li YQ (2007) Biological Characterization of *Theileria* spp. Infective for sheep and goats in China. PhD. Dissertation. p 6

Li YQ, Luo J, Liu Z, Guan G, Gao J, Ma M, Dang Z, Liu A, Ren Q, Lu B, Liu J, Zhao H, Li J, Liu G, Bai Q, Yin H (2007) Experimental transmission of sp. *Theileria* sp. (China 1) infective for small ruminants by *Haemaphysalis longicornis* and *Haemaphysalis qinghaiensin*. Parasitol Res 101:533–538

Li YQ, Luo J, Guan G, Ma M, Liu A, Liu J, Ren Q, Niu Q, Lu B, Gao J, Liu Z, Dang Z, Tian Z, Zhang B, He Z, Bai Q, Yin H (2009) Experimental transmission of *T. uilenbergi* infective for small ruminants by Haemaphysalis longicornis and Haemaphysalis qinghaiensin. Parasitol Res 104:1227–1231

Liu A, Guan G, Liu Z, Liu J, Leblanc N, Li Y, Gao J, Ma M, Niu Q, Ren Q, Bai Q, Yin H, Luo J (2010) Detecting and differentiating *Theileria sergenti* and *Theileria sinensis* in cattle and yaks by PCR based on major piroplasm surface protein (MPSP). Exp Parasitol 126(4):476–481

Liu AH, Yin H, Guan G, Schnittger L, Liu Z, Ma M, Dang Z, Liu J, Ren Q, Bai Q (2007) At least two genetically distinct large *Babesia* species infective to sheep and goats in China. Vet Parasitol 147:246–251

Nagore D, García-Sanmartín J, García-Pérez A, Juste R, Hurtado A (2004) Identification, genetic diversity and prevalence of *Theileria* and *Babesia* species in a sheep population from Northern Spain. Int J Parasitol 34:1059–1067

Niu Q, Luo J, Guan G, Ma M, Liu Z, Liu A, Dang Z, Gao J, Ren Q, Li Y, Liu J, Yin H (2009) Detection and differentiation of ovine *Theileria* and *Babesia* by reverse line blotting in China. Parasitol Res 104:1417–1423

Saiki RK, Chang CA, Levenson CH, Warren TC, Boehm CD, Kazazian HH Jr, Ehrlich HA (1988) Diagnosis of sickle cell anemia and beta-thalassemia with enzymatically amplified DNA and nonradioactive allele specific oligonucleotide probes. N Engl J Med 319:537–541

Salih DA, Liu Z, Bakheit MA, Ali AM, El Hussein AM, Unger H, Viljoen G, Seitzer U, Ahmed JS (2008) Development and evaluation of a loop-mediated isothermal amplification method for diagnosis of tropical theileriosis. Transbound Emerg Dis 55(5–6):238–243

Schnittger L, Yin H, Qi B, Gubbels MJ, Beyer D, Jongejan SNF, Ahmed JS (2004) Simultaneous detection and differentiation of *Theileria* and *Babesia* parasites infecting small ruminants by reverse line blotting. Parasitol Res 92:189–196

Slenning BD (2010) Global climate change and implications for disease emergence. Vet Pathol 47:28–33

Silva MG, Marques PX, Oliva A (2010) Detection of *Babesia* and *Theileria* species infection in cattle from Portugal using a reverse line blotting method. Vet parasitol 174:199–205

Sun CQ, Liu Z, Gao J, Guan G, Ma M, Luo J, Yin H (2008) Investigations into the natural infection rate of *Haemaphysalis qinghaiensis* with Piroplasma using a nested PCR. Exp Appl Acarol 44:107–114

Teng KF (1980) Two new species of *Haemaphysalis* from China (acarina: Ixodidae). Acta Zootaxonomica Sinica 5:144–149

Yang FG, Feng Z, Yu G, Liu J, Wei Z, He X (1958) A report on ovine theileriosis in Ganning animal farm station in Ganzi District. Chin J Vet Med 2:33–37

Yin H, Guan G, Ma M, Luo J, Lu B, Yuan G, Bai Q, Lu C, Yuan Z, Preston P (2002a) *Haemaphysalis qinghaiensis* ticks transmit at least two different *Theileria* species: one is infective to yaks, one is infective to sheep. Vet Parasitol 107:29–35

Yin H, Luo J, Guan G, Gao Y, Lu B, Zhang Q, Ma M, Lu W, Lu C, Yuan Z, Guo S, Wang B, Du H, Schnittger L, Ahmed J, Jongejan F (2002b) Transmission of an unidentified *Theileria* species to small ruminants by *Haemaphysalis qinghaiensis* ticks collected in the field. Parasitol Res 88:S25–S27

Yin H, Luo J, Guan G, Lu B, Ma M, Zhang Q, Lu W, Lu C, Ahmed J (2002c) Experiments on transmission of an unidentified *Theileria* sp. to small ruminants with Haemaphysalis qinghaiensis and Hyalomma anatolicum anatolicum. Vet Parasitol 108:21–30

Yin H, Luo J, Schnittger L, Lu B, Beyer D, Ma M, Guan G, Bai Q, Lu C, Ahmed J (2004) Phylogenetic analysis of *Theileria* species thransmitted by *Haemaphysalis qinghaiensis*. Parasitol Res 92:36–42

Yin H, Schnittger L, Luo J, Seitzer U, Ahmed J (2007) Ovine theileriosis in China: a new look at an old story. Parasitol Res 101:S191–S195

Yin H, Liu Z, Guan G, Liu A, Ma M, Ren Q, Luo J (2008) Detection and differentiation of *Theileria luwenshuni* and *T. uilenbergi* in small ruminants by PCR. Transbound Emerg Dis 55:233–237

Yuan GL, Yin H, Luo J, Guo Y, Guan G, Ma M, Lu B, Lv W, Zhang Q (2002) Studies on the bionomics of *Haemaphysalis qinghaiensis*. Chin J Vet Sci Technol 32:10

Annexe 3. Article: **Niu Q**, Luo J, Guan G, Liu Z, Ma M, Liu A, Gao J, Ren Q, Li Y, Qiu J, Yin H. 2009. Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences. *Exp Parasitol*. 121: 64-68.

Experimental Parasitology 121 (2009) 64–68

Contents lists available at ScienceDirect

Experimental Parasitology

journal homepage: www.elsevier.com/locate/yexpr

Differentiation of two ovine *Babesia* based on the ribosomal DNA internal transcribed spacer (ITS) sequences

Qingli Niu [a,b], Jianxun Luo [a], Guiquan Guan [a], Zhijie Liu [a], Miling Ma [a], Aihong Liu [a], Jinliang Gao [a], Qiaoyun Ren [a], Youquan Li [a], Jiaxiang Qiu [b], H. Yin [a,*]

[a] *State Key Laboratory of Veterinary Etiological Biology, Gansu Provincal Key Laboratory of Veterinary Parasitology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences (CAAS), Xujiaping1, Lanzhou, Gansu 730046, People's Republic of China*

[b] *College of Veterinary Medicine, Xinjiang Agricultural University, Urumqi 830052, People's Republic of China*

ARTICLE INFO

Article history:
Received 5 July 2008
Received in revised form 29 September 2008
Accepted 30 September 2008
Available online 17 October 2008

Keywords:
Differentiation
Babesia sp.
Internal transcribed spacer
Ovine
Babesia moatsi

ABSTRACT

The first and second internal transcribed spacers (ITS1, ITS2) as well as the intervening 5.8S coding region of the rRNA gene for six *Babesia* spp. isolated from different geographic origins were characterized. Varying degrees of ITS1 and ITS2 intra- and inter-species sequence polymorphism were found among these isolates. Phylogenetic analysis of the ITS1-5.8S gene-ITS2 region clearly separated the isolates into two clusters. One held an unidentified *Babesia* sp. transmitted by *Hyalomma anatolicum anatolicum*. The second held five other isolates, which were considered to be *Babesia motasi*. Each *Babesia* species cluster possessed ITS1 and ITS2 of unique size(s) and species specific nucleotide sequences. The results showed that ITS1, ITS2 and the complete ITS1-5.8S-ITS2 region could be used to discriminate these ovine *Babesia* spp. effectively.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Babesia are a tick-borne intracellular erythrocytic haemoprotozoan parasites of mammals, birds and reptiles (Bush et al., 2001; Levine, 1985). Clinical symptoms of babesiosis include hemolytic anemia which may be fatal if not treated. The disease can be of great economic importance to production animals where the specific tick vectors occur. Five ovine *Babesia* species have been described, *B. ovis*, *B. motasi*, *B. crassa*, *B. foliata* and *B. taylori* from sheep and goats (Hashemi-Fesharki and Uilenberg, 1981). The disease caused by *B. motasi* may be acute and/or chronic. Conversely, *B. ovis* infections are usually less severe than *B. motasi* (Morel, 1989). *B. crassa* appears to cause little or no pathogenicity (Hashemi-Fesharki, 1997). The taxonomy of these organisms is based on the classical methods of morphology, pathogenicity, antigenicity, host specificity, tick vectors, transmission mode and epidemiological data.

In China, ovine babesiosis is endemic in the Sichuan and Heilongjiang provinces and is believed to be caused by *B. motasi* due to the severe clinical symptoms observed (Chen, 1982; Zhao et al., 1986; Yin et al., 1997a,b). Later, several ovine *Babesia* isolates were collected from distinct geographical areas (Yin et al., 1997a,b; Guan et al., 2001). The isolate from Xinjiang region showed different characteristics; the other isolates originated from Lintan, Ningxian, Tianzhu, Madang in Gansu province and Chengde in Hebei province. The vector for the Xinjiang isolate was *Hyalomma anatolicum anatolicum*, whereas *Haemaphysalis* spp. transmitted the other isolates. Infections with the Xinjiang isolate often produced clinically mild infections, while the other isolates could lead to a moderate or severe clinical disease (Guan et al., 2002; Liu et al., 2007). Liu et al (2007) compared the six isolates of *Babesia* species based on 18S rRNA gene and primarily determined that there are two distinct ovine *Babesia* species by the observation from the phylogenetic tree. In the phylogenetic tree, the five isolates of *Babesia* sp. from different areas of China grouped with *B. motasi*, while *Babesia* sp. Xinjiang grouped in a separate clade between *B. ovis* and *B. crassa*. (Liu et al., 2007). Although the 18S rRNA gene is widely used as a genetic marker for taxonomic study, it is highly conserved and may not always reliably differentiate closely related species (Aktas et al., 2007). The rRNA internal transcribed spacers (ITS) that include internal transcribed spacer 1 (ITS1), 5.8S rRNA gene and internal transcribed spacer 2 (ITS2) are subject to higher evolutionary rates leading to greater variability in both nucleotide sequences and length (Hillis and Dixon, 1991). As a consequence, the ITS regions have been used as genetic markers for phylogenetic separation of closely related species, recognition of new species, determination of piroplasm species and/or isolates and has been widely applied for evolutionary and taxonomic studies for *Theileria* and *Babesia*

* Corresponding author. Fax: +86 931 8340977.
E-mail address: ttbdcn@public.lz.gs.cn (H. Yin).

0014-4894/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2008.09.021

(Aktas et al., 2007; Zahler et al., 1998; Collins and Allsopp, 1999; Baneth et al., 2004).

The aim of this work is to confirm the relationship between *Babesia* species Xinjiang isolate and the other 5 isolates by using the ITS regions as a genetic marker.

2. Materials and methods

2.1. Parasite isolates

The unidentified *Babesia* Xinjiang isolate was collected from Kashi of Xinjiang region, while *B. motasi* isolates were collected Chengde of Hebei Province, Tianzhu, Madang, Ningxian and Lintan of Gansu Provine by either ticks or blood from sheep. These isolates were designated as *Babesia* sp. Kashi isolate, *B. motasi* (Chengde), *B. motasi* (Lintan), *B. motasi* (Ningxian), *B. motasi* (Madang), and *B. motasi* (Tianzhu), respectively. The Chengde, Tianzhu, Madang and Ningxian isolates were obtained by inoculating field collected blood of asymptomatic sheep into splenectomised sheep. The Lintan isolate was isolated from a sheep infested with adult *Haemaphysalis qinghaiensis* ticks from Lintan, Gansu Province. The Kashi isolate was obtained from a sheep infested with *H. anatolicum anatolicum* from Kashi, Xinjiang Region.

2.2. Sheep

Sheep, 6–12 months of age, were purchased from a haemoprotozoa-free area and maintained in an isolated stable. All sheep used in these experiments were splenectomized thirty days before the study began. Ten days prior to the experiments, blood films taken from the ears of the sheep were fixed with methanol, stained with Giemsa and examined for the presence of haemoparasites. All those sheep were negative for haemoparasites and were used for the experiment.

2.3. DNA extraction

Each of six sheep were inoculated i.v. with 15 ml blood infected with one of the different *Babesia* isolates that had been cryopreserved in liquid nitrogen. When the parasitemia was more than 5%, venous 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 photometrically. Control DNA was isolated from the venous blood of uninfected sheep.

2.4. PCR amplification

A pair of primers used for amplifying the target segments was designed as described by previous publication (Zahler et al., 1998). The upstream PCR primer, Uni-up (5′-CCG AAT TCT TTG TGA ACC TTA TCA-3′), was specific to target the 3′ end of the 18S rRNA coding region. The region has been confirmed to be conserved among sequences of Chinese *Babesia* sp. isolates (GenBank Accession No DQ159074, AY260181 and DQ159073). Specific 28S sequence was not available, therefore a universal primer Uni-down (5′-CGG GAT CCT TCA CTC GCC GCT ACT-3′) that can specifically bind to a highly conserved 28S rRNA gene region, was used as the downstream primer.

The PCR amplification was performed in a final volume of 50 μl containing 1 μl of (100 pmol) each primer, 3 μl of 10 mM deoxynucleotide triphosphates, 5.0 μl of 10× PCR buffer (Mg^{2+} free), 0.2 μl of 5 units/μl r*Taq* DNA polymerase and 1 μl of specimen containing genomic DNA 100 ng. All the above reagents were purchased from TaKaRa, China. Amplification cycles were carried out with a DNA Thermal Cycler 2400 (Perkin-Elmer Life Sciences Inc., USA). The reaction mixture was incubated at 94 °C 5 min to denature the specimen DNA; followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 1 min. The final extension step was at 72 °C for 10 min and then the reaction was held at 4 °C. PCR products were separated in 1.5% TAE agarose gel containing 0.5 μg/ml ethidium bromide together with DNA markers (BD2000 Ladder Marker, TaKaRa) and visualized under UV illumination.

2.5. Cloning and sequencing of amplified DNA fragments

The PCR products were extracted from agarose gel with Agarose Gel DNA Purification Kit Ver. 2.0 (TaKaRa, China), ligated into pGEM-T Easy vectors (Promega, USA) before transformed into *Escherichia coli* JM 109 competent cells and plated onto LB plates containing 100 μg/ml ampicillin and 32 μg X-gal/400 μg IPTG (isopropyl-β-D-thiogalactoside). The plates were cultured overnight at 37 °C and only white colonies were picked and transferred into LB medium containing 100 μg/ml ampicillin and cultured at 37 °C overnight. The plasmids with inserts were extracted by using MiniBEST Plasmid Purification Kit Ver. 2.0 (TaKaRa, China) and identified by PCR (program and reaction mixtures same as above) and *EcoR* I (10 U) enzyme restriction. Three independent positive clones for each *Babesia* isolate were sequenced using the BigDye Terminator Mix (TaKaRa Company, China).

2.6. Sequence alignment and analysis

The ITS sequences of the six isolates were used as the template sequence to do blast searches from the NCBI database, respectively. The other sequences of *Babesia* and *Theileria* species available in GenBank were selected, and the alignment was done using ClustalW method of Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0 (Tamura et al., 2007). A phylogenetic tree was constructed using neighbor-joining algorithm of the phylogeny program of MEGA 4.0 (Saitou and Nei, 1987), which included all of the sequences from 6 Chinese isolates and the downloaded sequences of *B. caballi* (AF394536), *B. bigemina* (DQ301778, DQ301775, DQ301771 and DQ301769), *B. canis* (EF180055, AF394534, AY272048 and AF394535), *Babesia* sp. California (AY339745), *B. odocoilei* (AY345122, AY339759), *T. buffeli* (AY661534), *T. sergenti* (AY661531), *T. annulata* (AY684843), *T. mutans* (AY663655) and *T. parva* (U03602). The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. All of the Chinese isolate sequences were aligned by using ClustalW (http://www.ebi.ac.uk/clustalw) (Thompson et al., 1994). The percent identity value between Chinese isolates was calculated after alignment of Clustal W by using DNAstar (Version 4.01, Madision, Wis.).

3. Results

3.1. Amplification of the ITS1-5.8S-ITS2 rRNA gene

The ITS1-5.8S-ITS2rRNA gene amplification products were generated only from the DNA of *Babesia* species and no amplification was seen in either the sheep genome DNA or the water controls. The length of the PCR products that contained part of the 18S rRNA and 28S rRNA gene sequences was about 1000 bp. After removing the 18S rRNA and 28S rRNA gene sequence, the length of ITS gene regions of the six ovine *Babesia* species varied from 718 bp to 769 bp. The length of ITS1, ITS2 and 5.8S rRNA

Table 1
Percent identity of ITS1, ITS2 and 5.8S rDNA gene sequences of *Babesia* sp. Chinese isolates deduced after CLUSTAL W alignment.

Isolates	Percent identity of ITS1					Percent identity of ITS2					Percent identity of 5.8 S rDNA				
	B. motasi Chengde	NX	MD	TZh	LT	*B. motasi* Chengde de	NX	MD	TZh	LT	*B. motasi* Chengde	NX	MD	TZh	LT
B. motasi Ningxian	92.3					92.1					100.0				
B. motasi Madang	71.5	71.7				77.5	66.9				98.7	98.7			
B. motasi Tianzhu	71.2	71.2	98.4			77.5	65.7	100.0			99.4	99.4	99.4		
B. motasi Lintan	71.5	71.5	98.4	99.5		78.0	66.7	98.9	98.9		99.4	99.4	99.4	100.0	
Babesia sp. Kashi	2.5	2.8	2.6	2.6	2.6	12.1	12.2	98.9	9.0	9.0	93.7	93.7	91.9	93.7	93.7

sequences was 382–400 bp, 177–199 bp and 159–179 bp, respectively. The sequence data have been deposited at the GenBank database and the accession numbers for Chengde isolate, Ningxian isolate, Lintan isolate, Tianzhu isolate, Madang isolate and Kashi isolate are EF564710, EF564711, EF564712, EF564713, EF564714, EF564715, respectively.

3.2. Sequence alignment and phylogenetic analysis

The percent identity among the 6 Chinese *Babesia* sp. isolates based on the ITS1, ITS2 and 5.8S rDNA gene was determined (Table 1). The minimum identity among *B. motasi* Ningxian, Madang, Tianzhu, Lintan and Chengde isolates was 71.2% in ITS1, whereas the maximum identity of these isolates to the Kashi isolate was only 2.8%. For ITS2, Kashi isolate showed the maximum identity to the other isolates was 12.2%, while the other the isolates had the minimum identity of 65.7%. For the 5.8S rDNA, the Kashi isolate had the maximum identity of 93.7% to other isolates of China, which is lower that the minimum identity of 98.7% amongst the other isolates from China. By aligning the ITS sequences of all of 6 Chinese isolates, the sequence length and the higher variability of ITS of Kashi isolate were shown (Fig. 1.). A phylogenetic tree was inferred with the Chinese *Babesia* sp. isolates and other *Babesia* and *Theileria* species available in GenBank (Fig. 2.). The tree topology separated the 6 Chinese isolates into two far clades, while all of the *Theileria* species, *B. bigemina*, *B. caballi*, *B. canis* and *B. odocoilei* were classified into separate branches.

4. Discussion

Babesia sp. Kashi isolate was considered to be a new *Babesia* species because it showed different morphology, transmission vector and molecular taxonomy position based on 18S rRNA gene sequences (Guan et al., 2002; Bai et al., 2002; Liu et al., 2007) from the other *Babesia* sp. isolates originated from Chengde, Ningxian, Tianzhu, Lintan and Madang. That the remaining isolates could be *B. motasi* has been primarily inferred because they showed similar morphological shapes, transmission vector and pathogenicity (Yin et al., 1997a,b; Bai et al., 2002; Alani and Herbert, 1988). Furthermore, these isolates were placed onto *B. motasi* branch according to recent phylogenetic research based on 18S rRNA gene (Liu et al., 2007).

In this study, we confirm that there are two different ovine *Babesia* species co-existing in China using ITS phylogenetic analysis. The Kashi isolate has a very low percent identity with the other Chinese isolates in either ITS1 or ITS2 or 5.8S rDNA gene sequences, which have maximum percent identity of 2.8, 12.2 and 93.7, respectively. However, the other isolates from China showed the minimum percent identity in ITS1, ITS2 and 5.8S rDNA was 71.2, 65.7 and 99.4, respectively (Table 1). Through alignment of the sequences of all of the Chinese isolates, the high variation in regions of ITS1 and ITS2 between the Kashi isolate and the other isolates were clearly showed (Fig. 1.). Although the 5.8S gene region was only 160 bp and was shown to be a conserved region amongst the isolates and closely related species in study on *Theileria* species (Aktas et al., 2007). However, there were still seven mutations in the sequence of the Kashi isolate in comparison with sequences of the other isolates. Furthermore, the phylogenetic tree based on the whole ITS sequences placed Kashi isolate in a separate clade from the other Chinese isolates as well as from other *Babesia* species, such as *B. bigemina*, *B. caballi, B. canis and B. odocoilei*. The other five Chinese isolates were grouped together in an independent clade at the top of the tree (Fig. 2.). The tree also suggested that the isolates of Chengde, Madang, Tianzhu, Ningxian and Lintan may have evolved from the same ancestor. *Theileria* species were an out-group separated independently, which indicates the reliability of this phylogenetic work.

The transmission vector of *B. motasi* was considered to be *Heamaphysalis punctata* and its morphological shape was characterized as budding-form parasites (Alani and Herbert., 1988). The transmission vector of the Lintan isolate has proved to be *H. qinghaiensis* (Guan et al., 2002). Though we still have no evidence of the tick vectors for the isolates from Chengde, Madang, Tianzhu and Ningxian, it is well known that *H. qinghaiensis* and *H. longicornis* are widely distributed in Madang and Tianzhu County and *H. longicornis* is found Ningxian County and Chengde Hebei province (Yin et al., 1997a,b). Moreover, the budding-form parasites were also observed in Lintan and Ningxian isolates (Guan et al., 2002; Bai et al., 2002). Interestingly, it can be seen that there are two sub-groups at the top clade of the phylogenetic tree in this study, where isolates from Lintan, Madang and Tianzhu formed one sub-group, and isolates of Chengde and Ningxian formed another. Previously it has been demonstrated that the *Babesia* sp. Lintan isolate was a non-pathogenic organism (Guan et al., 2002), while the isolate from Ningxian was considered as a pathogenic organism (Yin et al., 1997a,b), and this may imply that the strains from different regions may differ in their virulence.

The Kashi isolate was originally isolated from *H. anatolicum anatolicum*, which suggests that the Kashi isolate may have different vector ticks from the other isolates collected. Morphologically, the parasites of the Kashi isolate were larger than other isolates (Guan et al., 2001,2002; Bai et al., 2002). Genetically, the Kashi isolate was separated from the other Chinese isolates with respect to 18S rRNA gene sequences. In this study, it was shown that there is a marked variance in ITS sequences between Kashi isolate and the other Chinese *B. motasi* isolates (Fig. 1). In the phylogenetic

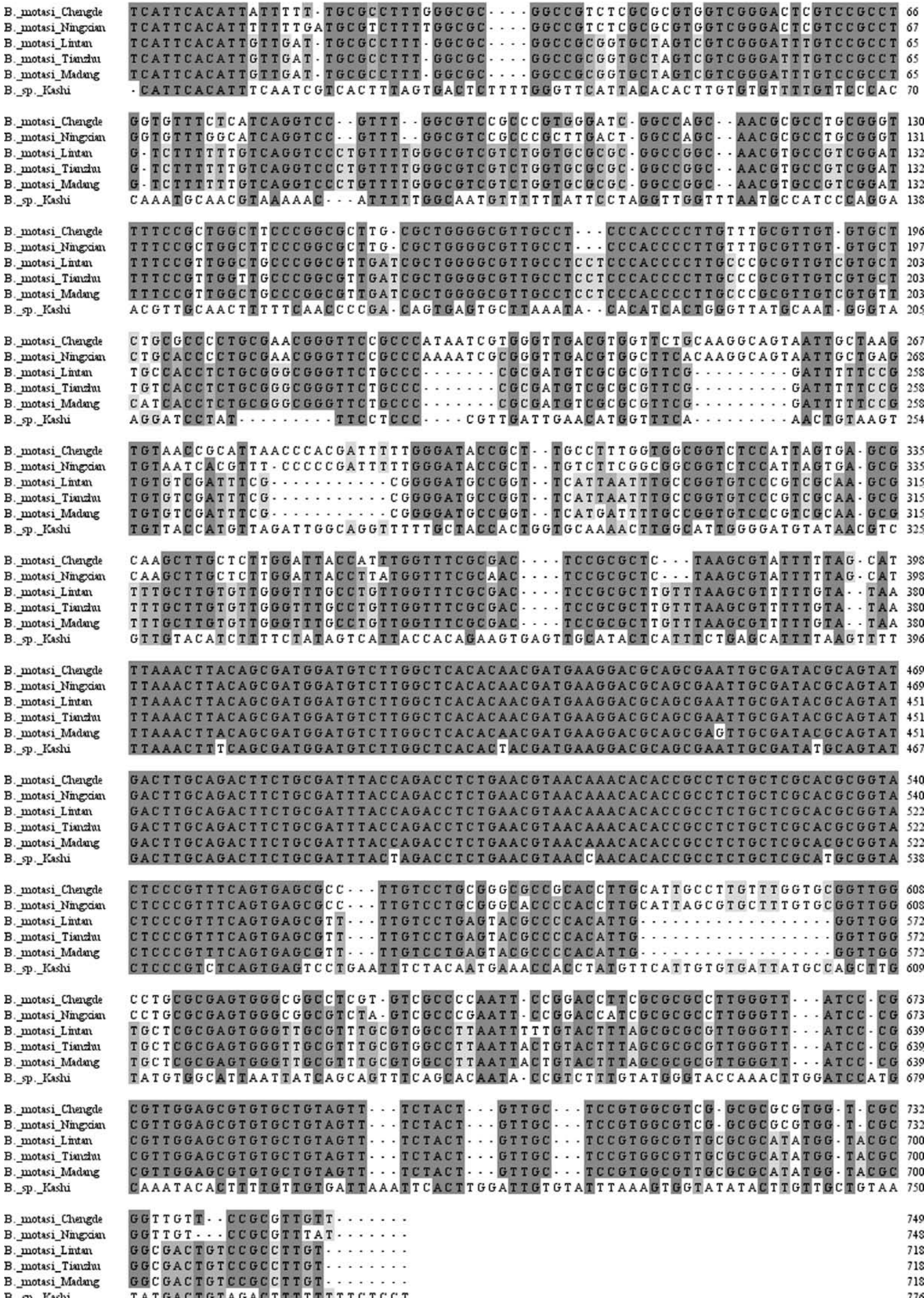


Fig. 1. ClustalW alignment of ITS sequences from 6 *Babesia* sp. Chinese isolates: Chengde, Ningxian, Madang, Tianzhu, Lintan and Kashi. The numbers of nucleotides are indicated at the right end of the sequences. From 1 to 426 is ITS1gene region, from 427 to 586 is 5.8S rRNA gene region and from 587 to 780 is ITS2 gene region. The degree of darkness of background indicates the level of the nucleotide sequence conservation among of different isolates. The darker it is, the more conservative the nucleotide sequences are.

tree Kashi isolate was in a clade highly separated from the other Chinese *B. motasi* isolates (Fig. 2.). There appear to be two ovine *Babesia* species in China, of which one could be *B. motasi* and the other is *Babesia* sp. Kashi.

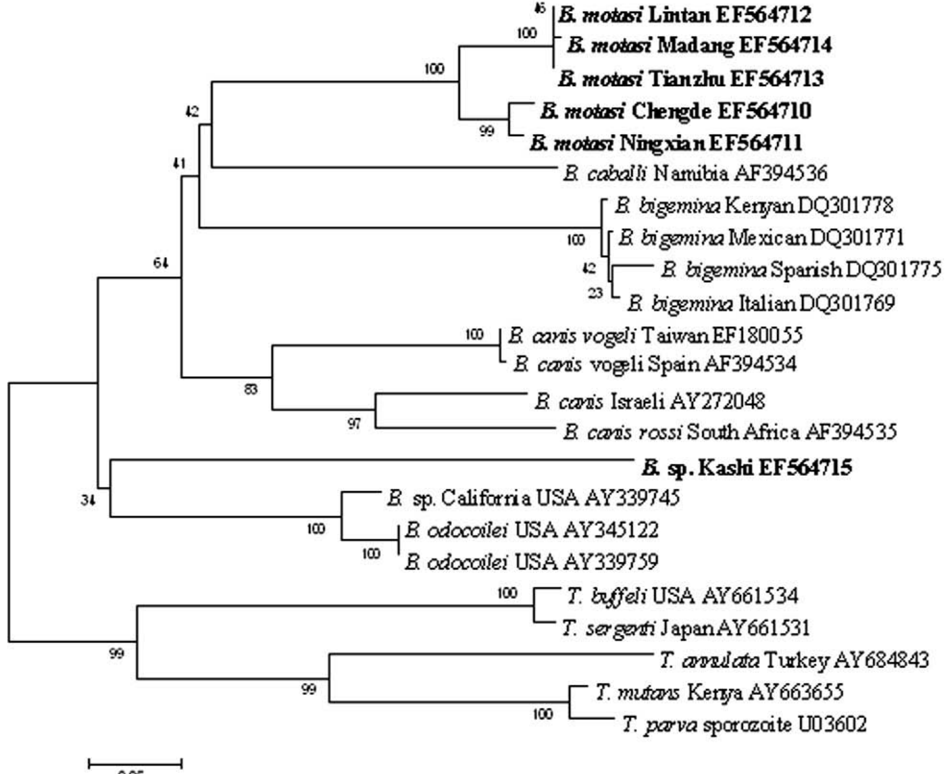


Fig. 2. Phylogenetic tree depicting the relationship among *Babesia* sp. Chinese isolates and other *Babesia* species and *Theileria* species based on ITS region sequences. The tree was inferred using the Neighbor-Joining method. Bootstrap values are shown as percentages at each node based on 500.

Acknowledgements

This study was financially supported by the National Natural Sciences Foundation of China (No 30571397), International Cooperation Fund of NSFC, the Natural Resource Platform Project (No 2005DKA21104), Supporting Plan (2007BAD40B00), “863” Project (2006AA10A207), Specific Fund for Sino-Europe Cooperation. MOST, China; and Outstanding Research Fellowship of CAAS, Beijing, China. The research was also facilitated by ICT-TD-3, SSA-income and EPIZONE projects of European Commission, Brussels, Belgium. The authors are grateful for Dr. Wayne Jorgensen, Focus Team Leader, Applied Biotechnology (Animals) Emerging Technologies Delivery, Department of Primary Industries and Fisheries, Australia, for his critical correction and comments of this article.

References

Aktas, M., Bendele, K.G., Altay, K., Dumanli, N., Tsuji, M., Holman, P.J., 2007. Sequence polymorphism in the ribosomal DNA internal transcribed spacers differs among *Theileria* species. Veterinary Parasitology 147, 221–230.

Alani, A.J., Herbert, I.V., 1988. The morphometrics of *Babesia motasi* (Wales) and its transmission by *Haemaphysalis punctata* (Canestrini and Fanzago 1877) to sheep. Veterinary Parasitology 30, 87–95.

Bai, Q., Guan, G.Q., Liu, D., Liu, J.X., Ren, X.L., 2002. Isolation and preliminary characterization of a large *Babesia* sp., from sheep and goats in the eastern part of Gansu Province. China Parasitology Research 88, S16–S21, [in Chinese].

Baneth, G., Kenny, M.J., Tasker, S., 2004. Infection with a proposed new subspecies of *Babesia canis*, *Babesia canis* subsp. Presentii, in domestic cat. Journal of Clinical Microbiology 42, 99–105.

Bush, A.O., Fernandes, J.C., Esch, G.W., Seed, R.J., 2001. Parasitism: The Diversity and Ecology of Animal Parasites. Cambridge University Press, Cambridge, pp. 89–91.

Chen, D.M., 1982. Investigations on ovine piroplasmosis. Chinese Journal of Veterinary Science and Technology 3, 31–32, [in Chinese].

Collins, N.E., Allsopp, B.A., 1999. *Theileria parva* ribosomal internal transcribed spacer sequences exhibit extensive polymorphism and mosaic evolution: application to the characterization of parasites from cattle and buffalo. Parasitologia 118, 541–551.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Guan, G.Q., Yin, H., Luo, J.X., Lu, W.S., Zhang, Q.C., Ma, M.L., Yuan, G.L., Lu, B.Y., Wang, Y., Muhe, T., 2001. Morphology and pathogenicity initial investigate of *Babesia* sp.. Chinese Journal Veterinary Science and Technology 31, 35–36, [in Chinese].

Guan, G.Q., Yin, H., Luo, J.X., Lu, W.S., Zhang, Q.C., Gao, Y.L., Lu, B.Y., 2002. Transmission of *Babesia* sp. to sheep with field collected *Haemaphysalis qinghaiensis*. Parasitology Research 88, S22–S24.

Hashemi-Fesharki, R., Uilenberg, G., 1981. *Babesia crassa* of domestic sheep in Iran. Veterinary Quarterly 3, 1–8.

Hashemi-Fesharki, R., 1997. Tick-borne disease of sheep and goats and their related vectors in Iran. Parassitologia 39, 115–117.

Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. The Quarterly Review of Biology 66, 411–453.

Levine, N.D., 1985. Veterinary Protozoology. Iowa State University Press, Ames, pp. 291–328.

Liu, A.H., Yin, H., Guan, G.Q., Schnittger, L., Liu, Z.J., Ma, M.L., Dang, Z.S., Liu, J.L., Ren, Q.Y., Bai, Q., 2007. At least two genetically distinct large *Babesia* species infective to sheep and goats in China. Veterinary Parasitology 147, 246–251.

Morel, P., 1989. Tick-Borne diseases of livestock in Africa. In: Fischer, M.S., Ralph, S. (Eds.), Manual of Tropical Veterinary Parasitology. CAB International, Wallingford, p. 473.

Saitou, N., Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4, 406–425.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.

Thompson, J.D., Higgins, D.G., Gibson, T.J., Clustal, W., 1994. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

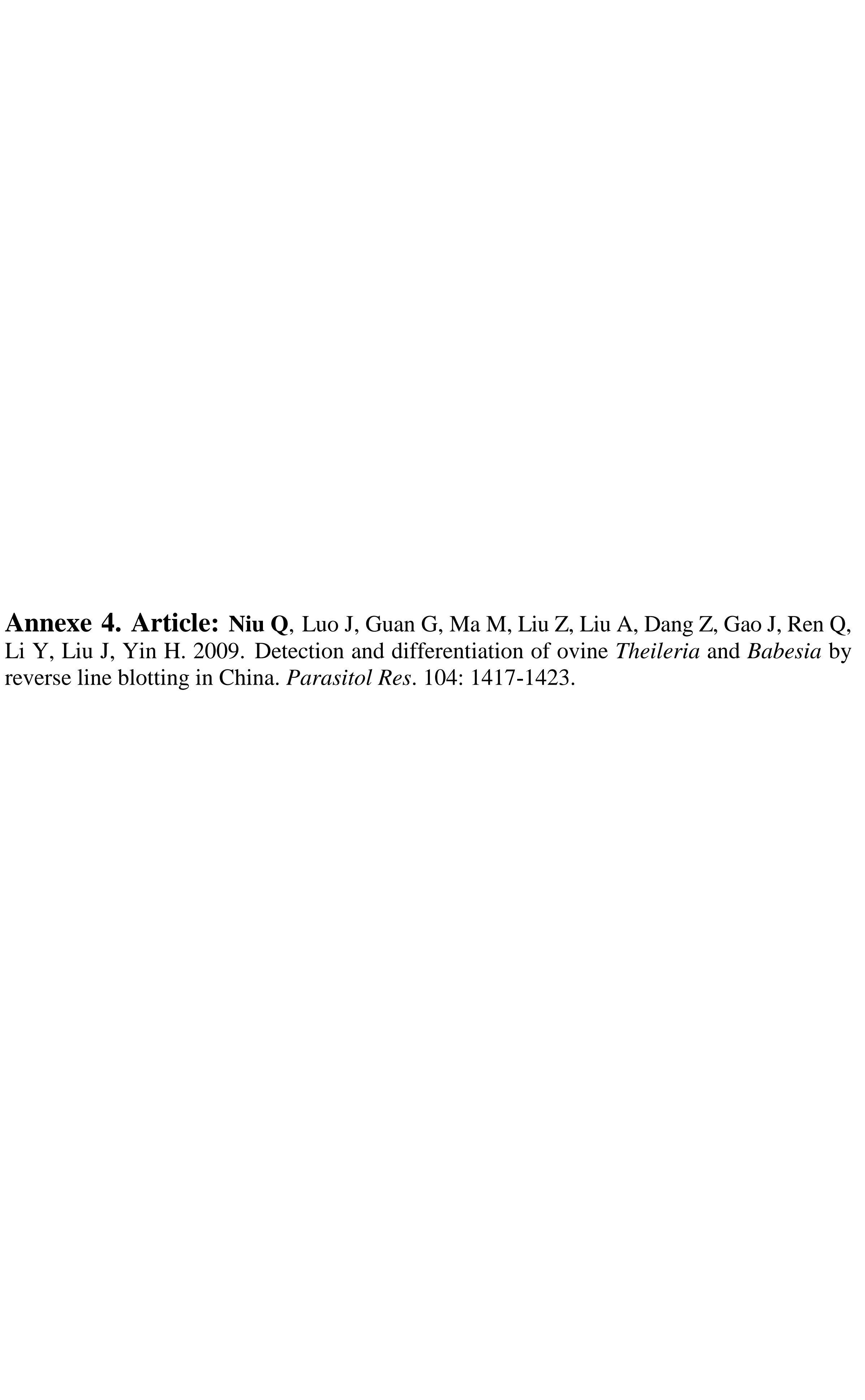
Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of National Academy of Sciences 101, 11030–11035.

Yin, H., Luo, J.X., Lu, W.S., Lu, W.X., Zhang, Q.C., Wang, Y., Liu, T., Lu, W., Li, Z., Dou, Y., Zhao, P., Chang, S., 1997a. Isolation and Morphology survey of the *Babesia motasi* and *Babesia ovis* in China. Chinese Journal Veterinary Science and Technology 27, 7–9, [in Chinese].

Yin, H., Lu, W., Luo, J.X., 1997b. Babesiosis in China. Tropical Animal Health and Production 29, 11S–15S.

Zhao, X.R., Li, C.R., Min, Y.G., 1986. Investigations on ovine babesiosis. Chinese Journal Veterinary Science and Technology 1, 26–27, [in Chinese].

Zahler, M., Schein, E., Rinder, H., Gothe, R., 1998. Characteristic genotypes discriminate between *Babesia canis* isolates of differing vector specificity and pathogenicity to dogs. Parasitology 84, 544–554.



Annexe 4. Article: Niu Q, Luo J, Guan G, Ma M, Liu Z, Liu A, Dang Z, Gao J, Ren Q, Li Y, Liu J, Yin H. 2009. Detection and differentiation of ovine *Theileria* and *Babesia* by reverse line blotting in China. *Parasitol Res*. 104: 1417-1423.

Parasitol Res (2009) 104:1417–1423
DOI 10.1007/s00436-009-1344-x

ORIGINAL PAPER

Detection and differentiation of ovine *Theileria* and *Babesia* by reverse line blotting in China

Qingli Niu · Jianxun Luo · Guiquan Guan · Miling Ma · Zhijie Liu · Aihong Liu · Zhisheng Dang · Jinliang Gao · Qiaoyun Ren · Youquan Li · Junlong Liu · Hong Yin

Received: 8 August 2008 / Accepted: 13 January 2009 / Published online: 11 February 2009
© Springer-Verlag 2009

Abstract A reverse line blot (RLB) assay was developed for detection and specific identification of the different ovine *Theileria* and *Babesia* parasites. In a polymerase chain reaction (PCR), the hypervariable region 4 (V4 region) of the 18S ribosomal DNA gene was amplified with a set of general primers specific for members of the genera *Theileria* and *Babesia*. Meanwhile, specific oligonucleotide probes were designed and bound on membrane. After one single-PCR amplification, the amplified fragment was hybridized against different generic and species-specific probes. It was able to detect four species, i.e., *Babesia motasi* (Chengde, Lintan, Ningxian, Tianzhu), *Babesia* sp. (Kashi), *Theileria luwenshuni* (Lintan, Madang, Ningxian), *Theileria uilenbergi* (Longde, Zhangjiachuan) as defined previously. All probes bound to their respective target sequence only; therefore, no cross-reaction was observed, resulting in clear recognition of either individual strains, species, or groups in normal positive tests. Meanwhile, no signal was observed when ovine genomic DNA and water were used as a control, demonstrating that the signals are due to the presence of parasite DNA in the samples. Furthermore, the sensitivity of RLB could be considerably enhanced to detect a parasitemia level between10^{-3}% and 10^{-8}%. Finally, 117 samples from field were tested with RLB, PCR, and enzyme-linked immunosorbent assay (ELISA). The positive rate of RLB was higher than that of PCR and ELISA, and furthermore, RLB could determinate the species of piroplasms, the samples were infected with. Samples, 1,117, from five areas in Gannan Tibet Autonomous Region have been examined with RLB assay and compared with ELISA assay for corresponding samples. The results showed that the positive rate of RLB was higher than that of ELISA test obviously, and both *T. luwenshuni* and *T. uilenbergi* were widely distributed in these areas. RLB developed here could be used for differentiation of *Babesia* and *Theileria* infection and for epidemiological survey, which was difficult to achieve by classical methods. In conclusion, the RLB is a versatile technique for simultaneous detection and identification of all ovine piroplasms.

Q. Niu · J. Luo · G. Guan · M. Ma · Z. Liu · A. Liu · Z. Dang · J. Gao · Q. Ren · Y. Li · J. Liu · H. Yin (✉)
State Key Laboratory of Veterinary Etiological Biology and Gansu Provincal Key Laboratory of Veterinary Parasitology, Lanzhou Veterinary Research Institute, CAAS, Lanzhou, Gansu 730046, China
e-mail: yinhong@public.lz.gs.cn

Introduction

Piroplasma species are tick-borne parasitic protozoa which are differentiated into the genera *Theileria* and *Babesia*. A number of these parasites are highly pathogenic for cattle, sheep, and goats; and the diseases emerging from these infections are referred to as theileriosis and babesiosis, respectively. The economic losses due to theileriosis and babesiosis are enormous in tropical and subtropical areas (Mehlhorn and Schein 1984; Mehlhorn et al. 1994). *Theileria ovis*, *Theileria lestoquardi* (formerly *T. hirci*), and *Theileria separata* are recognized as the species that cause ovine theileriosis (Preston 2001), whereas ovine babesiosis is caused by *Babesia ovis*, *Babesia motasi*, and *Babesia crassa* (Uilenberg 2001).

In China, it was first recorded that the ovine babesiosis spreads in the Sichuan and Heilongjiang province, and its pathogen was suspected to be *B. ovis* due to its relatively high pathogenicity (Chen 1982; Zhao et al. 1986). Later, several ovine *Babesia* isolates were isolated from different geographical distributions (Yin et al. 1997; Guan et al.

2001). At present, different characters between *Babesia* sp. Kashi isolate and others were observed, and molecular taxonomy based on 18S rRNA gene had been introduced to make more accurate taxonomic relationship between ovine *Babesia* isolates (Liu et al. 2007), demonstrating that *B. motasi* and an unidentified *Babesia* sp. (Kashi) existed in China. Ovine *Theileria* was found in the Sichuan province firstly, and was found latterly in the Qinghai, Gansu, Liaoning, Inner Mongolia, Ningxia, and Shaanxi province in China. Based on 18S rRNA gene sequencing of the Chinese isolates and other species of *Theileria* available in Genbank, it was found that ovine *Theileria* spp. infective to small ruminants was composed of two species of *Theileria*, one was closely related to benign *Theileria*, such as *Theileria buffeli*, *Theileria sergenti*, while the other one was more close to *Theileria mutans*. These two species isolated in China were named *Theileria luwenshuni* and *Theileria uilenbergi* respectively. (Yin et al. 2007).

Generally, the detection and diagnosis of sheep piroplasmosis consist of three methods: (1) Traditional methods to detect and identify these parasites rely on light-microscopy examination of thin blood smears. However, this can be difficult in the case of carrier animals where presence of parasites is scant and even in acute cases at the onset of the disease. In addition, species identification based only on morphology is not easy, being particularly difficult if mixed infections occur. (2) Some of serological tests like the complement fixation test (CFT) and the indirect fluorescent antibody test (IFAT) can be useful to detect past infections, but cross-reactions between species have been reported (Bruning 1996; Papadopoulos et al. 1996). Furthermore, false positive and negative results are commonly observed in these tests. (3) Polymerase chain reaction (PCR) is the most commonly used molecular technique to detect piroplasms, and this technique is more sensitive and specific than conventional methods. (Nagore et al. 2004a; Schnittger et al. 2004; Alhassan et al. 2005; Altay et al. 2005), but PCR assays do not generally detect mixed infections, although there are some amplification protocols that can detect mixed piroplasm infections to some extent (Birkenheuer et al. 2003; Criado-Fornelio et al. 2003b).

In order to overcome these limitations, a reverse line blot (RLB) assay was developed to detect virtually all the species of piroplasms infecting our sheep population. RLB was initially developed as a reverse dot blot assay for the diagnosis of sickle cell anemia (Saiki RK et al. 1988), but the essence of both techniques is the hybridization of PCR products to specific probes immobilized on a membrane in order to identify differences in the amplified sequences. In the "line" approach, multiple samples can be analyzed against multiple probes to enable simultaneous detection. This approach was initially developed for the identification of *Streptococcus* serotypes (Kaufhold et al. 1994). The first application of RLB for the detection and differentiation of pathogens in ticks was developed for *Borrelia* spirochetes by Rijpkema et al (1995) and followed by an RLB for *Mycobacterium tuberculosis* strain differentiation (Kamerbeek et al. 1997) and was subsequently combined with *Ehrlichia* spp (Schouls et al. 1999). Successful application of RLB to detect and differentiate all known *Theileria* and *Babesia* species was carried out by Gubbels et al. (1999) and Schnittger et al. (2004).

In this study, we used RLB to detect and differentiate *Theileria* and *Babesia* species of importance in small ruminants on the basis of their differences in 18S rRNA gene sequences in China.

Materials and methods

Primers and probes

Species-specific RLB oligonucleotide probes were deduced from the hypervariable V4 region of the 18S rRNA gene sequences (Table 1). These probes were shown to bind only to their respective target sequences. The two *Theileria* species included in the assay are *T. luwenshuni* and *T. uilenbergi*. The two *Babesia* species included are *B. motasi* and *Babesia* sp. (Kashi). A catch-all *Theileria* and *Babesia* species control oligonucleotide is also included. All the specific oligonucleotide probes containing a *N*-(trifluoroacetamidohexyl-cyanoethyl, *N*,*N*-diisopropyl phosphoramidite [TFA])-C6 amino linker, were designed and diluted to give

Table 1 Sequence and concentration of specific oligonucleotide probes *Theileria* sp. and *Babesia* sp are as following

Oligonucleotides	Sequence(5′-3′)	Dilute concentrations (pmol)
Cat-all	CTGTCAGAGGTGAAATTCT	200
B-all	CCTTGGTAATGGTTAATAGGAA	200
T-all	TACCAAAGTAATGGTTAATAGG	50
B. m	GAATGATGCCGACTTAAACCCT	200
B. sp. (Kashi)	CGGGTTTCGTCTACTTCGC	800
T. l	ATCTTCTTTTGATGAGTTG	400
T. u	TGCATTTTCCGAGTGTTACT	800
Cat –All	CTGTCAGAGGTGAAATTCT	200

50–1,200 pmol/150 ul in 500 mM NaHCO3 (pH 8.4). The sequence and optimized concentrations of the oligonucleotide probes used are summarized in Table 1. These sequences were aligned using MUTALIN online interface (http://www.ncbi.nlm.nih.gov/). The variable regions of these sequences were flanked by the sequences of the two PCR primers that were used for amplification.

Parasites isolates

All parasite stocks used in this study were described in detail previously (Schnittger et al. 2004; Yin et al 2004; Liu et al 2007). Five *Babesia* isolates originated from Kashi of Xinjiang region, Chengde of Hebei Province, Tianzhu, Ningxian, and Lintan of Gansu Provine, were used in this research. Among these isolates, Chengde, Tianzhu, and Ningxian isolates were obtained by inoculating field-collected blood of the asymptomatic sheep to the splenectomized sheep. Lintan isolate was isolated from a sheep infested with adult *Haemaphysalis qinghaiensis* ticks from Lintan, Gansu Province. The Kashi isolate was obtained from a sheep infested with a batch of mixed *Rhipicephalus sanguineus* and *Hyalomma anatolicum anatolicum* from Kashi, Xinjiang Region. These isolates were designated as *B. motasi* (Chengde), *B. motasi* (Lintan), *B. motasi* (Ningxian), *B. motasi* (Tianzhu), and *Babesia* sp. (Kashi) respectively; *T. luwenshuni* (Lintan, Madang, Ningxian), *T. uilenbergi* (Longde, Zhangjiachuan) as defined previously. All parasite isolates were stored as EDTA–blood stabilates in liquid nitrogen.

Field samples collection

Anticoagulated blood samples and sera, all of 117 samples, were collected from sheep and goats in Gansu (Tianzhu, Xiahe), Sichuan (Qianning), Liaoyang (Liaoning), Xinlong (Hebei), and Inner Mongolia (Tongliao) regions. All blood samples were collected in EDTA and stored at −20°C. Sera were detected by enzyme-linked immunosorbent assay (ELISA) and are described in detail (Gao et al. 2002). Anticoagulated blood samples were handled using Gentra Kit following the protocol of the manufacturer and amplified PCR.

Samples, 1,117, were collected from Biandu and Yangyong districts of Lintan County, Kache and Azitan districts of Zhuoni County and Ganjia districts of Xiahe County from July 2004 to June 2005 in Gannan Tibet Autonomous Region, Gansu Province by Guo et al. (2007).

DNA isolation and PCR amplification

DNA was isolated using a genomic DNA Purfication Kit (Gentra, USA) according to the manufacturer's instructions. The amount of DNA isolated was assessed to 100 ng by photometry. Negative control DNA was isolated from the venous blood of uninfected sheep/goat. Genomic DNA (100 ng) was added to a reaction mixture (final volume of 67 ul). For each sample, the PCR mixer was prepared as follows: H_2O 56.0 µl, 10× reaction buffer (200 mM Tris–HCl (pH 8.55), 160 mM $(NH_4)_2SO_4$ and 20 mM $MgCl_2$) 7.2 µl, 10 mM dNTP 1.6 µl, 10 pM sense primer (RLB-F) 3.6 µl, 10 pM antisense primer (RLB-R [labelled with biotin]) 3.6 µl, Taq polymerase 1.8 U (0.36 µl if the concentration is 5 U/µl). Prepare the PCR mixer according to number of samples and aliquot 66 µl to each tube. Add 1 µl of genomic DNA from filed samples. For blank control, nothing was added. PCR amplification was performed in an automatic DNA thermocycler (Eppendorf). The reaction was incubated at 94°C for 3 min to denature genomic DNA and the thermal cycle reaction program was 1 min at 94°C, 90 s at 55°C, and 90 s at 72°C for 40 cycles with a final extension step at 72°C for 5 min. Samples were held at 4°C until analysis. The PCR products were verified using agarose gel electrophoresis before it was analyzed by RLB hybridization.

Reverse line blot hybridization

Preparation of the membrane

Preparation of RLB membrane and hybridization were carried out as previously described (Gubbels et al. 1999) with the following adaptations: Cut Biodyne C membrane into 14.5 cm×14.5 cm size without removing protective paper. Mark the membrane with a ball pen for identifying the direction late. Activate the Biodyne C membrane for 10 min by incubation in 10 ml freshly prepared 16% (*w*/*v*) 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDAC) in demineralized water, in a rolling bottle at room temperature. Place the membrane in a plastic container and shake with demineralized water for 2 min and place it on a supporting cushion in a clean miniblotter system.

Hybridization

The diluted probes were aliquoted into the miniblotter slots, and an incubation of 1 min for linking of oligonucleotides to the membrane occurred. After aspiration of solutions, the membrane was incubated in 100 mM NaOH for 10 min, washed at 60°C for 5 min and then at 42°C for 5 min in 2× SSPE, 0.1% (SDS). Subsequently, the membrane was placed perpendicular to its previous orientation into the miniblotter. Twenty microliters of PCR products was diluted with 2× SSPE, 0.1% SDS to a final volume of 150 ul, heated to 99°C for 5 min and then cooled on ice. The denatured samples were aliquoted into the slots of the

miniblotter for 60 min at 42°C, then aspirated and the membrane washed at 42°C for 10 min in 2× SSPE, 0.1% SDS. Subsequently, the membrane was treated at 42°C for 30 min with peroxidase-labeled streptavidin diluted 1:4,000 in 2× SSPE/0.1% SDS, washed twice at 42°C for 10 min and twice at room temperature for 5 min in 2 × SSPE, 0.1% SDS. Finally chemiluminescence detection was performed according to standard procedures (Amersham). After examination and documentation, membranes were stripped as described by Gubbels et al. (1999) and could be reused about 15 times.

Results

Specificity of RLB of different isolates

Amplification of the V4 region of the 18S rRNA gene was performed by PCR on all isolates. Generated DNA fragments were hybridized to the oligonucleotide-linked membrane followed by chemiluminescence detection resulting in signals of equal intensity for each oligonucleotide. All probes bound to their respective target sequence only; therefore, no cross reaction was observed, resulting in the clear recognition of either individual strains, species, or groups. No signal was observed when ovine genomic DNA and water were used as the control, demonstrating that observed signals are due to the presence of parasite DNA in investigated samples. Each *Theileria* species is identified by three oligonucleotide probes: the catch-all probe (ca841–859), a probe recognizing *Theileria* parasites (T-all811–832), and the species-specific probes for either *T. luwenshuni* (T-l628-647) or *T. uilenbergi* (T-u678-697). *Babesia* piroplasms are recognized by the following and there oligonucleotides as well: a catch-all probe (ca841–859), a probe recognizing *Babesia* parasites (B-all745-766), and the species-specific probes for either *B. motasi* species (Bm466–487) or *Babesia* sp. (Kashi).

Sensitivity of RLB for ovine piroplasms

Sensitivity of the hybridization assay was assessed by RLB processing the serially diluted genomic DNA of *Babesia* and *Theileria*, accordingly; the presented RLB system is capable of identifying a parasitemia of about 10^{-6}% (*B. motasi*), 10^{-3}% (*Babesia* sp. Kashi), 10^{-8}% (*T. luwenshuni*), 10^{-8}% (*T. uilenbergi*), while the detection level of a traditional PCR, carried out as a control, was restricted to about 10^{-3}%, 10^{-2}%, 10^{-5}%, 10^{-3}%, respectively. The detection rate of RLB was higher than the PCR, obviously. For all species, just the sensitivity of *Babesia* sp. (Kashi) is relatively lower when compared with others. The sensitivity of RLB could be considerably enhanced to detect a parasitemia level of at least by amplification of PCR, respectively. The established RLB system was tested for its sensitivity, and observed detection limits were equal to those of a corresponding RLB assay designed for the detection of cattle-infecting piroplasms (Gubbels et al. 1999).

Detection of parasites in field samples collected from sheep in China

The field samples were collected from six regions, Liaoning, Sichuan, Hebei, Inner Mongolia, and Gansu (Tianzhu and Xiahe). The prevalence of each hemoparasite species identified by RLB, PCR, and ELISA is summarized in Table 2. RLB was more sensitive compared to the results obtained after PCR and ELISA examination. The number of positive samples with RLB was at least two times or even higher than those detected by other methods. Most of the samples analyzed were negative by PCR and ELISA. The number of positive samples to *T. luwenshuni* by RLB of these six regions were 17 (94.1%), 20 (100%), 24 (91.6%), 20 (60%), 20 (40%), 16 (62.5%) (Table 2), respectively. It has been revealed that most of the positive samples were infected by *T. luwenshuni* and/or *T. uilenbergi* genotypes, and prevalence of *Babesia* spp. was low with only three, three, and one sample positive to *B. motasi* out of the samples from Madang, Tongliao, and Xinglong, respectively, and no samples positive to *Babesia* sp. (Kashi). Most of the samples were negative to hemoparasites by PCR and/or ELISA (mainly samples from Sichuan). Furthermore, 1,117 samples from five areas in Gannan Tibet Autonomous Region were examined with RLB assay and compared with ELISA assay for corresponding samples. The average detection rate with ELISA was 71.3% for the whole samples, while mean detection rate with RLB was 80.3% (data not shown). It was found that the positive rate of the samples from Biandu area by RLB was much higher than that by ELISA test (Table 2). Only two samples, one is from Kache and the other one is from Biandu, were found to be positive to *B. motasi*. No sample was positive to *Babesia* sp. (Kashi).

Discussion

Babesiosis was first reported in small ruminants in China in 1982, and the pathogen was identified as *B. ovis* (Chen 1982). In 1985, an outbreak of babesiosis in sheep occurred in Heilongjiang, and the causative agent was again identified as *B. ovis* (Zhao et al. 1986). Although more than two decades passed, the data on the distribution and epidemiology of ovine babesiosis is very limited. Theileriosis of sheep and goats was first reported in Sichuan province,

Table 2 Comparison of the results of RLB, PCR, and ELISA

Origin	Samples amount	Positive amount							Positive rate (%)						
		RLB			PCR			ELISA	RLB			PCR			ELISA
		Tl	Tu	Mix	Tl	Tu	Mix		Tl	Tu	Mix	Tl	Tu	Mix	
Liaoyang	17	16	0	0	0	0	0	0	94.1	0	94	0	0	0	0
Tianzhu	20	20	10	9	17	0	0	13	100	50	45	85	0	0	65
Madang	24	22	15	15	4	11	1	14	91.6	62.5	62.5	16.6	45.8	0.04	58
Qianning	20	12	7	7	6	7	5	9	60	35	35	30	35	25	45
Xinglong	20	8	0	0	0	0	0	0	40	0	0	0	0	0	0
Tongliao	16	10	0	0	0	0	0	0	62.5	0	0	0	0	0	0
Azitan	311/311[a]	233	84	57	–	–	–	230	74.9	27	18.3	–	–	–	74
Kache	325/327[a]	280	224	201	–	–	–	254	86.2	68.9	61.3	–	–	–	77.4
Yangyong	89/83[a]	70	29	14	–	–	–	55	78.6	27	21.3	–	–	–	66.3
Biandu	197/197[a]	150	114	81	–	–	–	97	76	57.8	41.1	–	–	–	49.2
Ganjia	195/199[a]	165	69	43	–	–	–	161	84.6	35	22.1	–	–	–	80.9
Total	1234/1234	986	547	432	27	18	1	824	79.9	44.2	35.0	23.1	15.4	0.009	23.1

[a] Number of sample examined with ELISA, which was done by Guo et al. (2007)

50 years ago (Yang 1958), and an ELISA and PCR were developed for detection of the diseases (Gao et al. 2002; Yin et al. 2008). However, it has been demonstrated that *T. luwenshuni*, *T. uilenbergi*, and *B. motasi* are sharing the same vector as *H. qinghaiensis*, and in most endemic areas (Yin et al. 2007), the sheep and goats were infected by two or three of the species of piroplasms mentioned above. It would be very convenient and useful to develop a method being able to detect simultaneously all these three pathogens and other piroplasms infective to small ruminants for diagnosis, detection, quarantine, and epidemiological survey.

The RLB hybridization assay is a versatile diagnostic tool to sensitively and simultaneously detect and differentiate hemoparasites in blood, tissue, or ticks. RLB is based on simultaneous PCR amplification of related species. Each species can be identified by a species-specific oligonucleotide probe using a line-blotter apparatus. It is quickly becoming a standard molecular tool for diagnostic and epidemiological studies in an increasing number of laboratories all over the world. Centeno-Lima et al. (2003) used RLB for the characterization of *Babesia divergens* in a human case, whereas Nijhof et al. (2003) and Altay et al. (2007) discovered novel *Theileria* and *Babesia* species through the application of RLB. Furthermore, detection and differentiation by using RLB for many *Babesia* spp. and *Theileria* spp. occurring in small ruminants have been reported by Schnittger et al. (2004).

In this study, an RLB assay was developed based on the V4 region of the 18S rRNA gene of *Theileria* and *Babesia* species infecting sheep in China. All species included were detected by single species-specific oligonucleotides, and no cross-reactions were observed. The assay allows a remarkable higher specificity than other methods through comparison. The generic oligonucleotide probe (catch-all) and group-specific probes designed specifically to detect the 18S rRNA genes of the genera *Theileria* and *Babesia* were found to hybridize in all the cases where one or more species/genotypes were present, and the catch-all probe would indicate that a new species or genotype is present when the PCR products were hybridized with probes and no species-specific signal. Similarly, the *Theileria* genus-specific probe designed specifically detected any *Theileria* species present (T-all), and the *Babesia* genus-specific probe designed specifically detected any *Babesia* species (B-all) in this study. Meanwhile, four species-specific probe also designed used to detect distinct species (*B. motasi*; *Babesia* sp. Kashi; *T. luwenshuni*; *T. uilenbergi*). This assay can be used to identify the different ovine-infecting piroplasms species at present.

As there was very limited data on the distribution and epidemiological situation of *Babesia* spp. and *Theileria* spp. infection in China, sera and bloods samples were collected from different areas in China to evaluate the sensitivity and specificity of RLB and to assess the epidemiological situation of these parasites' infection. The RLB technique revealed that 88 out of 117 samples were positive for the presence of ovine *Theileria* parasites, and the sensitivity was higher than PCR and ELISA obviously, and further confirmed the sensitivity of the RLB. On the other hand, as the difference of results of the samples from Liaoning region by RLB was obviously with other two assays, in order to certify whether it was caused by infection of parasites or the false positive, the sequence analysis was conducted for amplicons (PCR product) from these samples by random sampling. It was found that the fragment was about 500 bp, and the homology was 99.6% with the hypervariable region 4 (V4 region) of the 18S

ribosomal DNA gene of *T. luwenshuni* by DNAstar analysis, suggesting that the sheep were infected by *T. luwenshuni*, and the RLB was a reliable diagnostic tool. The RLB has been used as a standard molecular tool for discrimination of piroplasm from nonpathogenic piroplasm species and the simultaneous detection and identification of *Theileria* and *Babesia* species when they occurred in the same animal (Nagore et al. 2004b; Georges et al. 2001; Brígido et al. 2004; García-Sanmartín et al. 2006).

From the results, it could be seen that *Theileria* sp. is widely distributed in China, and *T. luwenshuni* was the dominant species, which agreed with the previous data by PCR (Yin et al. 2008). Only nine out of 1,234 (1,117+117) samples were positive for *B. motasi*, suggesting that the prevalence of *B. motasi* was much less than *T. luwenshuni* or *T. uilenbergi.* However, the distribution of *B. motasi* is relatively large as the positive samples were from three provinces. In this study, no positive samples were detected for *Babesia* sp. Kashi. The possible reason is that all samples are from areas where the vector of *Babesia* sp. Kashi, *Hyalomma anatolicum anatolicum*, is free. As *H. anatolicum anatolicum* is only distributed in Xinjiang, it would be interesting to collect some samples from these areas and evaluate the situation of *Babesia* sp. Kashi infection in sheep and goats in the further study.

In conclusion, RLB should be considered as a piroplasm infection diagnostic tool not only for animals showing clinical signs and suffering acute infections, but for those apparently healthy animals with potential subclinical infections and seems the optimal approach for the detection and discrimination of these important parasites. We recommend its use for integrated epidemiological monitoring of tick-borne disease, since RLB can also be used for screening ticks and can easily be expanded to include additional haemoparasite species.

Acknowledgements This study was financially supported by the Natural Sciences Foundation of China (no. 30571397, 30800820), Supporting Plan (2007BAD40B00), “863” Project (2006AA10A207), the National Natural Resource Platform Project(2005DKA21100), Specific Fund for Sino-Europe Cooperation, MOST, China; Key Project of Gansu Province (0801NKDA033), Lanzhou, Gansu, China. State Key Laboratory of Veterinary Etiological Biology Project (SKLVEB2008ZZKT019) and National Public Interests Research Institute Basic Scientific Research Expenses Special Fund Project. The research was also facilitated by SSA-Income Projects, ICTTD 3 and EPIZONE of European Commission, Brussels, Belgium.

References

Alhassan A, Pumidonming W, Okamura M, Hirita H, Battsetseg B, Fujisaki F, Yokoyama N, Igarashi I (2005) Development of a single-round and multiplex PCR method for the simultaneous detection of *Babesia caballi* and *Babesia equi* in horse blood. Vet Parasitol 129:43–49

Altay K, Dumanli N, Holman PJ, Aktas M (2005) Detection of *Theileria* ovis infected sheep by nested PCR. Vet Parasitol 127:99–104

Altay K, Dumanli N, Aktas M (2007) Molecular identification, genetic diversity and distribution of *Theileria* and *Babesia* species infecting small ruminants. Vet Parasitol l47:161–165

Birkenheuer AJ, Levy MG, Breitschwerdt EB (2003) Development and evaluation of a seminested PCR for detection and differentiation of *Babesia gibsoni* (Asian genotype) and *B. canis* DNA in canine blood samples. J Clin Microbiol 41:4172–4177

Brígido C, da Fonseca IP, Parreira R, Fazendeiro I, do Rosário VE, Centeno-Lima S (2004) Molecular and phylogenetic characterization of *Theileria* spp. parasites in autochthonous bovines (Mirandesa breed) in Portugal. Vet Parasitol 123:17–23

Bruning A (1996) Equine piroplasmosis an update on diagnosis. Br Vet J 152:139–151

Centeno-Lima S, do Rosário V, Parreira R, Maia AJ, Freudenthal AM, Nijhof AM, Jongejan F (2003) A fatal case of human babesiosis in Portugal: molecular and phylogenetic analysis. Trop Med Int Health 8:760–764

Chen DM (1982) Investigations on ovine piroplasmosis. Chin J Vet Sci Technol 3:31–32

Criado-Fornelio A, Martinez-Marcos A, Buling-Sarana A, Barba-Carretero JC (2003b) Molecular studies on *Babesia*, *Theileria* and *Hepatozoon* in southern Europe. Part

García-Sanmartín J, Nagore D, García-Pérez AL, Juste RA, Hurtado A (2006) Molecular diagnosis of *Theileria* and *Babesia* species infecting cattle in Northern Spain using reverse line blot macro-arrays. BMC Vet Res 16:3–7

Gao YL, Yin H, Luo JX, Ouyang WQ, Bao HM, Guan GQ, Zhang QC, Lu WS, Ma ML (2002) Development of an enzyme-linked immunosorbent assay for the diagnosis of *Theileria* sp infection in sheep. Parasitol Res 13:S8–S10

Georges K, Lori GR, Riili S, Greco A, Caracappa S, Jongejan F, Sparagano O (2001) Detection of haemoparasites in cattle by reverse line blot hybridisation with a note on the distribution of ticks in Sicily. Vet Parasitol 99:273–286

Gubbels MJ, de Vos AP, van der Weide M, Viseras J, Schouls LM, de Vries E, Jongejan F (1999) Simultaneous detection of bovine *Theileria* and *Babesia* species by reverse line blot hybridization. J Clin Microbiol 37:1782–1789

Guan GQ, Yin H, Luo JX, Lu W, Zhang Q, Ma ML, Yuan GL, Lu BY, Wang Y, Muhe T (2001) Morphology and pathogenicity initial investigate of *Babesia* sp. Chin J Vet Sci Tech 31:35–36

Guo SZ, Mu YJ, Liu ZJ, Ma DL, Yang SM, Ge GH, Fang BQ, Ga DJ, Ma ML, Luo JX, Yin H, Seitzer U, Ahmed JS (2007) Serological investigation of ovine theileriosis by ELISA in Gannan Tibet Region of Gansu Province in China. Parasitol Res 101:S197–S200

Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J (1997) Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 35:907–914

Kaufhold A, Podbieldski A, Baumgarten G, Blokpoel M, Top J, Schouls L (1994) Rapid typing of group A streptococci by the use of DNA amplification and non-radioactive allele-specific oligonucleotide probes. FEMS Microbiol Lett 119:19–25

Liu AH, Yin H, Guan GQ, Schnittger L, Liu ZJ, Ma ML, Dang ZS, Liu JL, Ren QY, Bai Q (2007) At least two genetically distinct large *Babesia* species infective to sheep and goats in China. Vet Parasitol 147:246–251

Mehlhorn H, Schein E (1984) The piroplasms: life cycle and sexual stages. Adv Parasitol 23:37–103

Mehlhorn H, Schein E, Ahmed JS (1994) Theileria. In: Kreier JP (ed) Parasitic protozoa, vol 7. Academic Press, San Diego, pp 217–304

Nagore D, García-Sanmartín J, García-Pérez AL, Juste RA, Hurtado A (2004a) Detection and identification of equine *Theileria* and *Babesia* species by reverse line blotting: epidemiological survey and phylogenetic analysis Veterinary assistant. Vet Parasitol 123:41–54

Nagore D, García-Sanmartín J, García-Pérez AL, Juste RA, Hurtado A (2004b) Identification, genetic diversity and prevalence of *Theileria* and *Babesia* species in a sheep population from Northern Spain. Int J Parasitol 34:1059–1067

Nijhof AM, Penzhorn BL, Lynen G, Mollel JO, Morkel P, Bekker CP, Jongejan F (2003) *Babesia* bicornis sp. nov. and *Theileria bicornis* sp. nov.: tick-borne parasites associated with mortality in the black rhinoceros (*Diceros bicornis*). J Clin Microbiol 41:2249–2254

Papadopoulos B, Brossard M, Perie NM (1996) Piroplasms of domestic animals in the Macedonia region of Greece. 3. Piroplasms of small ruminants. Vet Parasitol 63:67–74

Preston PM (2001) Theilerioses. In: Service MW (ed) Encyclopedia of Arthropod-transmitted infections of man and domesticated animals. CABI, Wallingford, pp 487–502

Rijpkema SG, Molkenboer MJ, Schouls LM, Jongejan F, Schellekens JF (1995) Simultaneous detection and genotyping of three genomic groups of *Borrelia burgdorferi sensu lato* in Dutch *Ixodus ricinus* ticks by characterization of the amplified intergenic spacer region between 5S and 23S rRNA genes. J Clin Microbiol 33:3091–3095

Saiki RK, Chang CA, Levenson CH, Warren TC, Boehm CD, Kazazian HH Jr, Ehrlich HA (1988) Diagnosis of sickle cell anemia and beta-thalassemia with enzymatically amplified DNA and nonradioactive allele specific oligonucleotide probes. N Engl J Med 319:537–541

Schnittger L, Yin H, Bai Q, Gubbels MJ, Beyer D, Nieman S, Jongejan F, Ahmed JS (2004) Simultaneous detection and differentiation of *Theileria* and *Babesia* parasites infecting small ruminants by reverse line blotting. Parasitol Res 92:189–196

Schouls LM, Van De Pol I, Rijpkema SG, Schot CS (1999) Detection and identification of *Ehrlichia*, *Borrelia burgdorferi sensu lato* and *Bartonella* species in Dutch *Ixodes ricinus* ticks. J Clin Microbiol 37:2215–2222

Uilenberg G (2001) Babesiosis. In: Service MW (ed) Encyclopedia of Arthropod-transmitted infections of man and domesticated animals. CABI, Wallingford, pp 53–60

Yang FG, Feng ZG, Yu GH, Liu JX, Wei ZQ, He XZ (1958) A report on ovine theileriosis in Ganning animal farm station in Ganzi District (in Chinese). Chin J Vet Med 2:33–37

Yin H, Luo JX, Lu WS, Lu WX, Zhang Q, Wang Y, Liu T, Lu W, Li Z, Dou Y, Zhao P, Chang S (1997) Segregation and morphology survey of the *Babesia motasi* and *Babesia ovis* in China. Chin J Vet Sci Tech 27:7–9

Yin H, Luo JX, Schnittger L, Lu BY, Beyer D, Ma ML, Guan GQ, Bai Q, Lu CP, Ahmed J (2004) Phylogenetic analysis of *Theileria* species transmitted by *Haemaphysalis qinghaiensis*. Parasitol Res 92:36–42

Yin H, Schnittger L, Luo JX, Seitzer U, Ahmed JS (2007) Ovine theileriosis in China: a new look at an old story. Parasitol Res 101:S191–S195

Yin H, Liu Z, Guan G, Liu A, Ma M, Ren Q, Luo J (2008) Detection and differentiation of *Theileria luwenshuni* and *T. uilenbergi* in small ruminants by PCR. Transbound Emerg Dis 55:233–237

Zhao XR, Li CR, Min YG (1986) Investigations on ovine babesiosis. Chin J Vet Sci Technol 1:26–27

Thèse de Doctorat

Qingli NIU

Comparison of sequence, organisation and expression of the rhoptry-associated-protein-1 genes in *Babesia* isolates responsible for ovine Babesiosis in China

Comparaison des séquences, de l'organisation, et de l'expression des gènes *rap-1* (rhoptry-associated-protein-1) chez les isolats du genre *Babesia* responsables de babésiose ovine en Chine

Résumé

Les babésioses ovines chinoises sont dues à plusieurs parasites intra-érythrocytaires du genre *Babesia* répartis en 2 groupes phylogénétiques. L'un comporte *Babesia* sp. Xinjiang; l'autre, proche de *B. motasi*, regroupe plusieurs agents au statut d'espèce ou de genre encore à définir. Dans l'objectif de développer une stratégie vaccinale visant à bloquer la multiplication du parasite chez son hôte, la caractérisation de la famille multigénique *rap-1* (rhoptry-associated-protein-1), codant des protéines impliquées dans l'invasion du globule rouge par le parasite, a été menée.
Dans le groupe *B. motasi*-like, le locus *rap-1* chez 4 agents étudiés (*Babesia* sp. Tianzhu, *Babesia* sp. BQ1 Lintan et BQ1 Ningxian, *Babesia* sp. Hebei) est complexe, avec 3 types de gènes, *rap-1a* (6 copies), *rap-1b* (5 copies identiques intercalées entre les gènes *rap-1a*) et *rap-1c* (1 copie en 3' du locus). Si les séquences sont très proches pour 3 des agents (99,9% d'identité sur le locus entier), celles de Hebei divergent plus (78%). Les 3 types de gènes sont transcrits chez *Babesia* sp. BQ1 Lintan. En culture in vitro, seule une partie des copies de *rap-1a* est traduite. In vivo, l'ensemble des gènes semble traduit avec des variations de cinétique.
Chez *Babesia* sp. Xinjiang, 7 gènes *rap-1a* sont caractérisés, et séparés en types α et β en fonction de leurs séquences en 3'. Ces gènes en tandem sont identiques sur les 936 nt en 5', puis diffèrent en 3' à la fois en longueur et en séquence, avec un nombre variable de répétitions de 36 nt plus ou moins conservés. Au moins 5 gènes sont transcrits par ce parasite cultivé in vitro.
Ces résultats sont discutés en terme de taxonomie, d'évolution, de vaccination.

Mots clés
***Babesia*, babésiose ovine, rhoptry-associated-protein, expression des protéines, invasion de l'érythrocyte, taxonomie, vaccin, motifs répétés ------**

Abstract

Several intra-erythrocytic parasites of the genus *Babesia* are responsible of ovine babesiosis in China. They belong to two phylogenetic groups. One includes only *Babesia* sp. Xinjiang ; the other, close to *B. motasi*, gathers several different parasites whose status as genus or species remains to be clarified. To develop a vaccination strategy aiming at blocking the parasite multiplication within its host, the characterization of the multigene family *rap-1* (rhoptry-associated-protein-1) has been undertaken, as these proteins are known to be involved in the process of the erythrocyte invasion by the parasite.
In the *B. motasi*-like group, the *rap-1* locus is complex in the 4 parasites studied (*Babesia* sp. Tianzhu, *Babesia* sp. BQ1 Lintan and BQ1 Ningxian, *Babesia* sp. Hebei) with 3 types of genes : *rap-1a* (6 polymorphic copies), *rap-1b* (5 identical copies intercalated with *rap-1a* genes) and *rap-1c* (1 copy at the 3' end of the locus). The sequences are similar for 3 of these parasites (99.9% over the whole locus), but *Babesia* sp. Hebei *rap-1* sequences are more divergent (78%). The 3 types of genes are transcribed in vitro by *Babesia* sp. BQ1 Lintan. In vitro, only some copies of *rap-1a* are translated. In vivo, all genes seem to be translated with different kinetics.
Seven *rap-1a* genes have been characterized in *Babesia* sp. Xinjiang, belonging to two types, α and β, according to their 3' end sequences. These tandemly arranged genes are identical over 936 nt in 5', and differ in 3' by a variable number of more or less conserved 36 nt repeats. At least 5 of these genes are transcribed in in vitro culture.
These results are discussed in terms of taxonomy, evolution and vaccination.

Key Words
***Babesia*, ovine babesiosis, rhoptry-associated-protein, protein expression, red blood cell invasion, taxonomy, vaccine development, protein repeats ----**

L'UNIVERSITÉ NANTES ANGERS LE MANS