Phylogenetic analysis of *Ehrlichia canis* and *Rhipicephalus* spp. genes and subsequent primer and probe design.

Name: V.H. de Visser (3051684)

Supervisor: prof. dr. F. Jongejan

Division: Utrecht Centre for Tick-borne Diseases

Period: October 2011-January 2012

TABLE OF CONTENTS

TABLE OF CONTENTS 1
ABSTRACT 2
INTRODUCTION 3
MATERIALS AND METHODS 6
RESULTS 7
PRIMER AND PROBE DESIGN 15
REALTIME PCR PRIMER VALIDATION 19
DISCUSSION AND RECOMMENDATIONS 21
ACKNOWLEDGEMENTS 21
REFERENCES 22

ABSTRACT

A phylogenetic analysis was made targeting six different *Ehrlichia canis* genes (VIRB9, GP36, GP28, GLTA, GP200, and 16S rRNA). The same was done for the ITS2 gene of *Rhipicephalus* species including *R. sanguineus*, vector of *Ehrlichia canis*. Furthermore, multiple matching (q)PCR primer sets and RLB probes were designed. This will prove useful in distinguishing *Rhipicephalus* spp., since determination solely based on phenotype is not always reliable. A beginning was made in validating the designed GLTA primer sets.

A great deal of variation was found between the different *E. canis* genes and to a lesser degree between the different *Rhipicephalus* genes.

Asian (Taiwanese) strains of *E. canis* appeared as a stable group, branching off at an early stage. Brazilian and other South American strains tended to show the same characteristics within phylogenetic trees of several of the mentioned genes.

***Keywords**: Ehrlichia canis, Rhipicephalus, tick, phylogenetic analysis, realtime PCR, ITS2, GLTA.*

INTRODUCTION

Rhipicephalus sanguineus (*Acari, Ixodidae*) is a three-host species hard tick. Its appearance is medium-sized, pale yellowish brown or reddish brown. Another striking characteristic is the simus pattern covering the scutum.

Figure 1 R. sanguineus with characteristic simus pattern. [23]

R. sanguineus serves as a vector for *Ehrlichia canis.* Apart from *E. canis*, *R. sanguineus* is responsible for transmitting other pathogens, such as *Babesia* spp., *Haemobartonella canis*, *Hepatozoon canis*, *Dipetalonema dracunculoides* and various organisms pathogenic to humans, such as *Rickettsia* spp. *rickettsia* and *coroni*. [1]

Depending on environmental properties such as altitude, hosts, temperature and relative humidity, larvae and nymphs feed for about 4 days and females for 8. The number of eggs deposited varies from 1164 to almost 5000. Larvae can survive unfed for approximately 253 days, nymphs for 97 and adults for as long as 568 days.

All stages of development of *R. sanguineus* (also known as brown dog tick or kennel tick) prefer the domesticated dog as a host. In the Afrotropical region specimens have been taken from cattle, goats, wild animals (carnivores) and also humans. The latter must however be considered predominantly as accidental hosts.

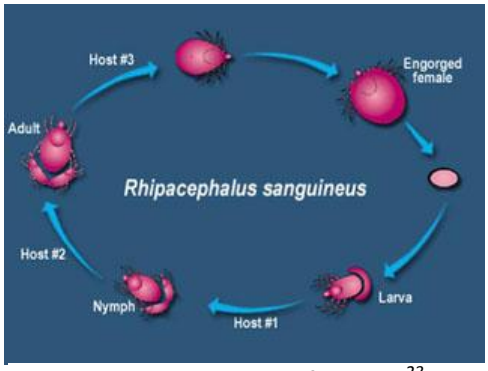

Figure 2: R. sanguineus life cycle. [23]

Immature stages of *R. sanguineus* are typically found on the dog's legs, chest and belly. Nymphs also attach to the ears. Adults are mainly found on the head and neck area. Once the nymphs and adult females are engorged, they will detach from their host and start moulting or laying eggs in cracks or crevices.

R. sanguineus is a widely distributed tick which is found circumglobally approximately between the latitudes of 50˚N and 30˚S. This worldwide distributing is due to the fact that the domestic dog serves as a primary host, as mentioned before. In the tropics and subtropics *R. sanguineus* can be found both indoors and outdoors. In colder climates it is mostly found indoors in homes, kennels and any other structures included in the habitats of dogs.

E. canis, the causative agent of canine monocytic ehrlichiosis (CME; tropical canine pancytopenia) was firstly described in Algeria. Nowadays it is distributed to numerous other areas on the globe. The disease is transmitted by *R. sanguineus* nymphs and adults. It is the most commonly reported canine infectious disease in the USA, and is present wherever the brown dog tick resides. *Ehrlichia spp.* are obligate intracellular bacteria that multiply in hematopoietic cells (i.e. monocytes). [2]

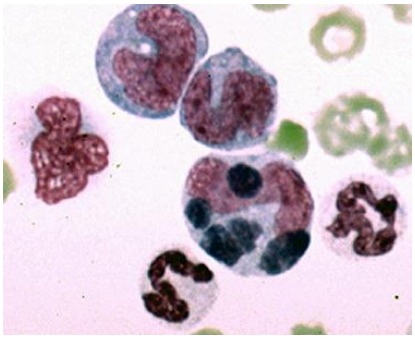

Figure 3: E. canis in a membrane-bound inclusion.[24]

CME generally progresses in three stages, including acute, subclinical and chronic. During the acute phase (8-20 days p.i.), symptoms include dyspnoe, anorexia and depression. Laboratory tests can show the presence of thrombocytopenia, leucopenia and hypergammaglobulinemia. Following this first (untreated) stage, a subclinical phase can commence, which is able to last from 40 days up to several years. During this symptomless period *E. canis* parasites are mainly found in the spleen. The last chronic phase manifests itself with hemorrhages, epistasis and edema. Similarly as in the acute phase, a thrombocytopenia usually develops. It is not uncommon for this stage of immunosuppression to facilitate secondary co-infections with other pathogens/micro-organisms. [2] In order to prevent CME from causing fatalities, a quick and accurate diagnosis is of the utmost importance, followed by adequate antibiotic therapy. (Note: although recovered, a 'CME dog' will remain seropositive.) [1, 2]

Realtime-PCR detection of Ehrlichia canis

Diagnosis of CME is based on hematological, biochemical and serologic findings. Detection of *E. canis* morulae in blood smears is a quick and inexpensive diagnostic method, although in some cases even less than 1% of infected cells are present in a CME positive animal which improves the chance of this method giving false negative results. [3] Serology tests, including *E. canis* specific IFAT and ELISA both have the advantage of a higher sensitivity, although these tests are incapable of distinguishing the current state of infection.

Because of its generally high sensitivity and specificity, the disadvantages of the methods mentioned above are less apparent when a (realtime-) PCR is used for determining the presence of CME. This technique can be used (additionally) as a way to confirm a current infection. [3]

The use of this diagnostic tool differs from other methods in multiple ways. Most commonly, a DNA-binding dye called SYBR-green is incorporated in a realtime PCR assay. It binds nonspecifically to double stranded DNA, increasing up to 1000 times in fluorescence whilst doing so. Therefore, the more amplicon present in the mixture, the more fluorescence is shown by SYBR-green. This type of dye makes the use of a probe unnecessary, only one set of primers is needed, and the amplification can be monitored during its development (hence the term 'realtime'). By using melt-curve analysis, clear assumptions can be made about the presence of the right fragment in the mixture. Each type of fragment melts at its own temperature, making it easy to distinguish amplicons from other fragments such as primer-dimers. [15]

Usually a (realtime-)PCR protocol targeting the 16S rRNA gene is used to detect canine ehrlichiosis. [5, 6, 7, 8, 9, 10, (...)] However, the presence of sequence variation of this gene between individual bacteria can cause amplification that is not specific enough. [3] Despite this, not many other genes have frequently been used to detect CME by means of PCR. [4] VIRB9 [14], DSBA [13], GP28 [3] and GP16/GP36 [4] are among the few exceptions. Therefore a selection of these genes has been phylogenetically analyzed in this study (see further).

Furthermore, a beginning has been made to detect *E. canis* by using the GLTA nucleotide sequence. This gene encodes for citrate synthase, which enables the first essential step in the citric acid cycle, the central metabolic pathway in all aerobic organisms taking place in mitochondria. [19] The degradation of amino acids, fatty acids and carbohydrates generates acetyl-CoA, which is oxidized

together with oxaloacetate by citrate synthase to form citrate. Outside of the mitochondrium, citrate is able to function as “an acetyl donor for acetyl-CoA synthesis by ATP-citrate-lyase after transport through the mitochondrial membrane”; however, mostly it acts as a substrate for subsequent steps of the citric acid cycle inside the mitochondrium. [19] In algae and gram-negative bacteria such as *E. canis* citrate synthase consists of 4-6 subunits, each with a molecular weight of about 50,000 (molar mass). It is dependent of NADH energy supply.

Reverse Line Blot

Following (q)PCR amplicon production of DNA samples, a Reverse Line Blot (RLB) can be performed. The essence of this molecular diagnostic tool is based upon the cross-wise hybridization of multiple PCR products and species-specific oligonucleotide probes on a membrane, enabling simultaneous detection of several spp. (ticks or pathogens). See figure 4) This hybridization causes a chemiluminescence which can be made visible using x-ray photography. After development of this film, positive signals will appear as a black dot with varying intensity depending on the amount of amplicons in the PCR product.

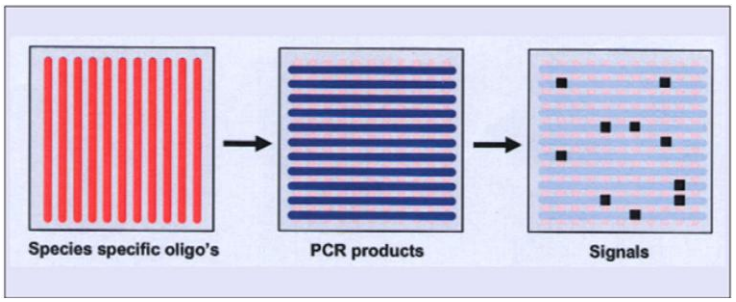

Figure 4: RLB detection.[20]

RLB assays combined with PCR amplification provides for a considerable increase in sensitivity compared to using PCR solely as a diagnostic tool.[20]

MATERIALS AND METHODS

Primers were designed using BioEdit Sequence Alignment Editor (ClustalW alignment) and following rules stated by the PCR manual leaflet (REF!). All nucleotide sequences were derived from the NCBI database (www.ncbi.nlm.nih.gov). Primer suitability was determined using Finnzymes Reagents Multiple Primer Analyzer(http://www.finnzymes.fi/java_applets/ multiple_primer_analyzer.html) and checked for specificity in NCBI's Nucleotide BLAST (www.ncbi/nlm/nih.gov).

Name	*GenBank accession numbers*
gltA	JN391409.1; JN391410.1; AY615901.1; AF304143.1; AY647155.1; EU078905.1
VirB9	JF706287.1; AY205343.1; AY205341.1; AY205339.1; AY205342.1; AY205340.1; AF546158.1
Gp36	EU139491.1; EF651794.1; EF560599.1; EF551366.1; HQ009756.1; HM188566.1; DQ146152.1; DQ146154.1; DQ146151.1; DQ146155.1; DQ146153.1; DQ085429.1; DQ085428.1; DQ085427.1
p28	EF014897.1; GU951532.1; AF082749.1; AF082747.1; AF082745.1; AF082750.1; AF082748.1; AF082746.1; DQ460713.1
16SrRNA	JN368080.1; JN121380.1; JN121379.1; JN187091.1; HQ718614.1; HQ718612.1; HQ718610.1; HQ718608.1; HQ718606.1; HQ718604.1; HQ718602.1;HQ718613.1; HQ718611.1; HQ718609.1; HQ718607.1; HQ718605.1; HQ718603.1; HQ718601.1; JF728840.1; JF429693.1; HQ908081.1; HQ844983.1; GU810149.1; EU567025.1; EU567023.1; EU567021.1; EU567024.1; EU567022.1; EU567020.1 ;EU439944.1; EF417993.1; EF424612.1; DQ401044.1; EU143637.1 EU123923.1; EU143636.1; HQ290362.1; EU106856.1; GQ857078.1; EF195134.1; EF195135.1; GU991633.1; EU781694.1; EU781692.1; EU781690.1; EU781688.1; EU781686.1; EU781695.1; EU781693.1; EU781691.1; EU781689.1; EU781687.1; DQ494536.1; DQ494537.1; GU386289.1; GU386287.1; GU386285.1 ;GU386288.1; GU386286.1; GU182114.1; GQ395380.1 GQ395381.1; GQ395378.1; AB287435.1; EF051166.1; M73226.1; M73221.1; DQ003032.1; EF139458.1; EU491504.1; EU376115.1; EU376113.1; EU376116.1; EU376114.1; EU376112.1; EU139493.1; AY621071.1; EU178797.1; EU263991.1; EF011110.1; EF011111.1; DQ915970.1; DQ460714.1; DQ648491.1; DQ228513.1; DQ228511.1; DQ228509.1; DQ228507.1; DQ228505.1; DQ228503.1; DQ228501.1; DQ228499.1; DQ228497.1; DQ228514.1; DQ228512.1; DQ228510.1; DQ228508.1; DQ228506.1; DQ228504.1; DQ228502.1; DQ228500.1; DQ228498.1; DQ228496.1; DQ206872.1; AY394465.1; AF536827.1; AF308455.1; AF373615.1; AF373613.1; AF373614.1;AF373612.1; U96437.1; AF162860.1; AF156785.1; AF156786.1; U26740.1

Figure 5: analyzed genes and their references.

Multiple alignment analysis, distance matrix calculation and construction of fylogenetic trees were performed using the MEGA (version 5.05) software program. Phylogenetic trees were made using the neighbor-joining method and the distance matrixes were made following Kimura's two parameters. Multiple alignment analysis was done by means of MUSCLE. Tree stability bootstrap values were calculated also by MEGA 5.05. See also figure below for details.

Analysis

Analysis ---------- Phylogeny Reconstruction
Scope ---------- All Selected Taxa
Statistical Method ---------- Neighbor-joining
Phylogeny Test
Test of Phylogeny ---------- Bootstrap method
No. of Bootstrap Replications ------- 500
Substitution Model
Substitutions Type ---------- Nucleotide
Model/Method ---------- Maximum Composite Likelihood
Substitutions to Include ---------- d: Transitions + Transversions
Rates and Patterns
Rates among Sites ---------- Uniform rates
Pattern among Lineages ---------- Same (Homogeneous)
Data Subset to Use
Gaps/Missing Data Treatment --------- Pairwise deletion
Codons Included ---------- 1st+2nd+3rd+Non-Coding

Figure 6: Specified settings of the MEGA 5.05 phylogenetic tree analysis program.

RESULTS

1) Phylogenetic analysis of *E. canis* genes

VIRB9

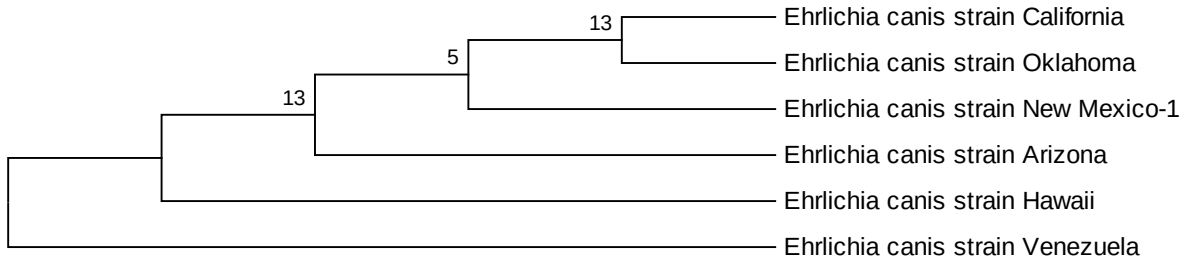

Figure 7: Phylogenetic tree of E. canis *VIRB9 gene.*

This phylogenetic tree suggests a strong relationship between both North and South American strains of the VIRB9 gene. Unfortunately, no European, Asian or African strains are available, making it impossible to analyze intercontinental relatedness of this gene.

Constructing a pairwise distance matrix for VIRB9 was unnecessary; overall distance average was 0. Hence, all these variances share no genetic differences and are classified in the same taxa. Bootstrap values are therefore considerably low. This close relatedness is easily visible when a VIRB9-like strain is added to the tree, as in the figure below.

Chances of designing successful primer sets targeting this gene are good considering this high degree of relationship between the VIRB9 taxa, assuming eventual European varieties will find their place in this tree with equal straightforwardness.

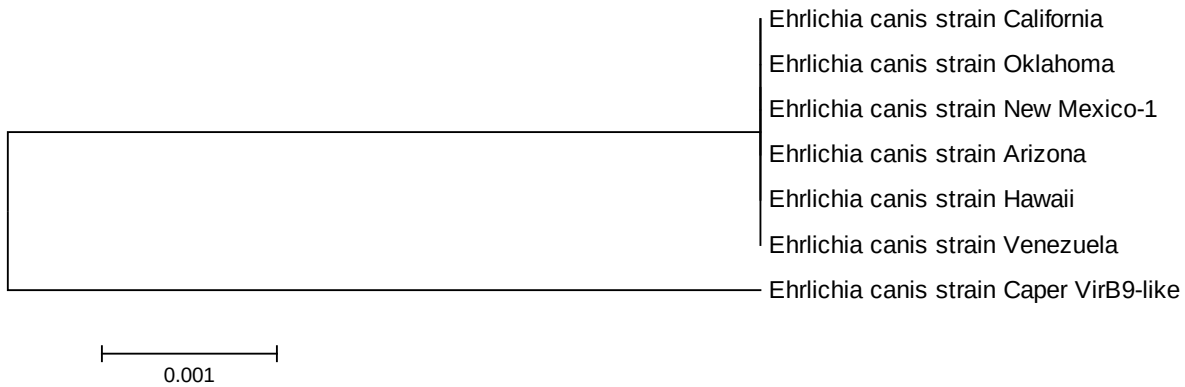

Figure 8: Phylogenetic tree of E. canis VIRB9 gene, including VIRB9-like gene.

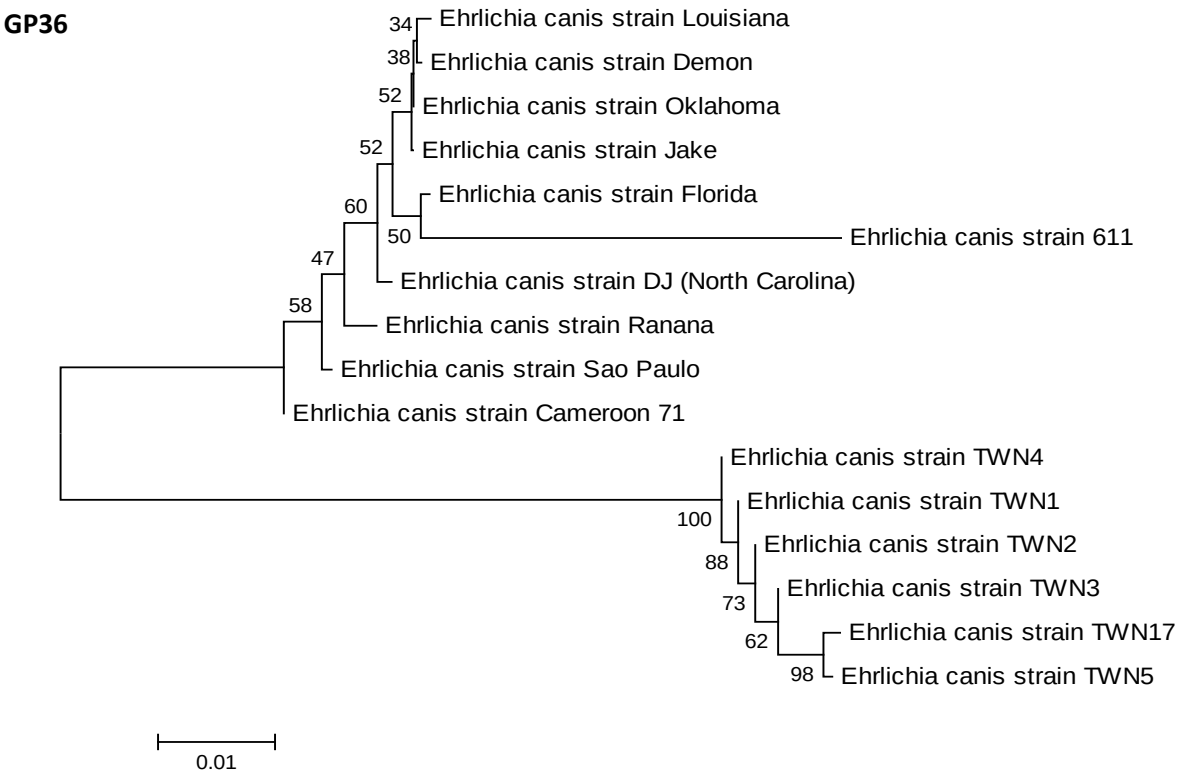

Figure 9: Phylogenetic tree of E. canis *GP36 gene.*

When characterizing the *E. canis* GP36 gene, a strong dichotomy appears to be present. A relatively late branch containing all Taiwanese strains (note: all to be referenced back to one study) is found, which seems to be practically unrelated to strains originating from both Africa and the American continents. The latter seemingly originate in an African-South American-North American sequence. Another prominent feature is the Israeli 611 strain, appearing in this tree as a novel 'brother' of the Florida strain. High bootstrap values are particularly found within the Taiwanese branch (av =88.6), whereas the North American strains share an average of 68.6. Widest distances were found between the North American (Florida, Louisiana and Oklahoma) and the six Taiwanese strains.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1)TWN4														
2)TWN3	0.002													
3)TWN2	0.000	0.002												
4)TWN1	0.001	0.004	0.001											
5)TWN17	0.001	0.004	0.001	0.003										
6)TWN5	0.001	0.004	0.001	0.003	0.002									
7)Florida	0.108	0.108	0.108	0.110	0.110	0.106								
8)Sao Paulo	0.037	0.084	0.087	0.077	0.096	0.096	0.011							
9)Louisiana	0.103	0.103	0.103	0.105	0.105	0.101	0.006	0.011						
10)Cameroon	0.070	0.081	0.075	0.074	0.092	0.093	0.010	0.005	0.009					
11)DJ	0.073	0.084	0.078	0.077	0.096	0.096	0.003	0.007	0.006	0.006				
12)Demon	0.072	0.082	0.077	0.075	0.094	0.095	0.005	0.007	0.002	0.006	0.004			
13)Oklahoma	0.101	0.101	0.101	0.103	0.103	0.099	0.005	0.009	0.002	0.008	0.005	0.000		
14)Jake	0.077	0.083	0.077	0.078	0.095	0.096	0.005	0.007	0.002	0.006	0.004	0.000	0.000	

Figure 10: Pairwise distance matrix of the GP36 gene; overall average 0.050.

GP28

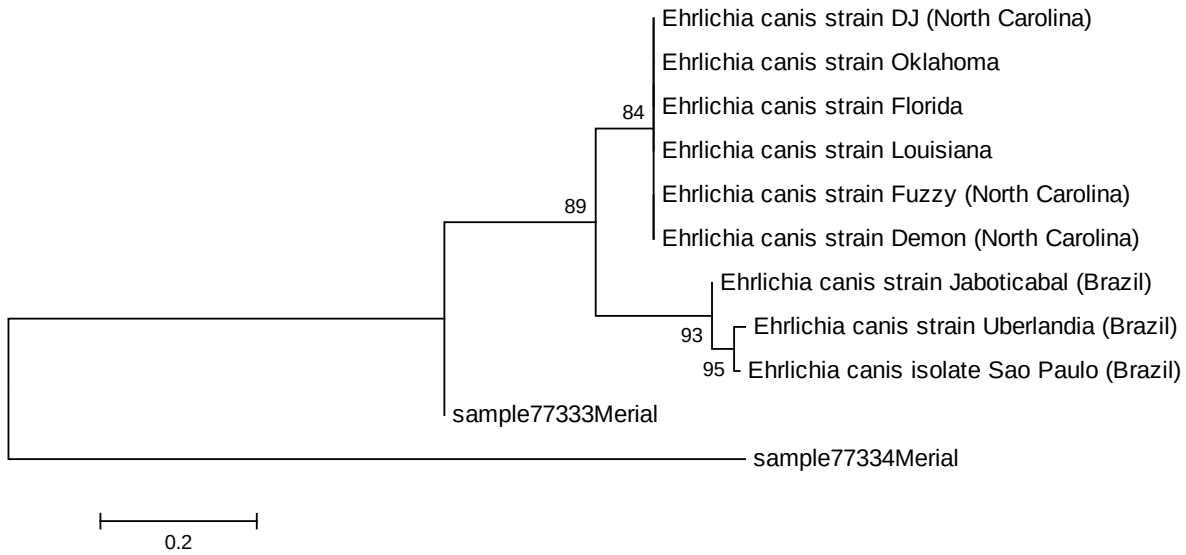

Figure 11: Phylogenetic tree of E. canis *GP36 gene, including Merial samples 77333 and 77334.*

Two samples submitted by Merial from Abidjan (Ivory Coast) were co-mapped with other strains of ehrlichial 28-kDA immunodominant outer membrane protein, also known as the GP28 gene. In contrast to sample 77334, sample 77333 appears to be most closely related to the strains found in the GenBank register. Moreover, this African sample is most likely ancestral to all three Brazilian strains and North American strains. Bootstrap values were found to be particularly high within the Brazilian taxa. The overall distance matrix average was considerably high compared to the other analyzed *E. canis* genes. The most influential factor in this value is the erratic Merial 77334 sample, although the North and South American strains also show a significant divergence.

	1	2	3	4	5	6	7	8	9	10	11
1)Jaboticabal											
2)Uberlandia	0.0027										
3)Fuzzy	0.206	0.266									
4)Demon	0.206	0.266	0.000								
5)Louisiana	0.206	0.266	0.000	0.000							
6)sample77333	0.134	0.161	0.118	0.118	0.118						
7)sample77334	1.861	1.934	1.752	1.752	1.333						
8)Florida	0.206	0.266	0.000	0.000	0.118	1.752					
9)DJ	0.206	0.266	0.000	0.000	0.118	1.752	0.000				
10)Oklahoma	0.206	0.266	0.000	0.000	0.118	1.752	0.000	0.000			
11)Sao Paulo	0.017	0.021	0.260	0.260	0.260	0.171	1.905	0.260	0.260	0.260	

Figure 12: Pairwise distance matrix of the GP28 gene; overall average 0.421.

GP200

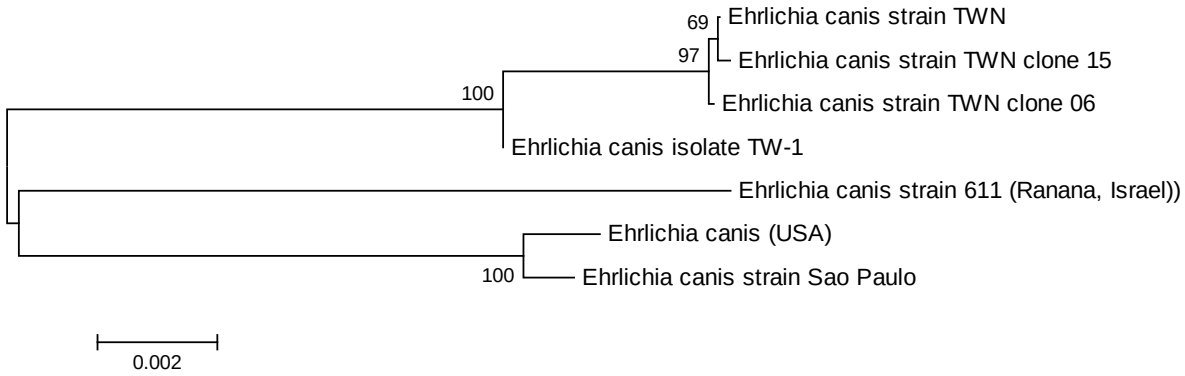

Figure 13: Phylogenetic tree of E. canis *GP200 gene.*

Next analysis was of the GP200 gene. As shown above, American and Brazilian strains appear to be closely related, while Taiwanese variances of this gene share a separate branch in which the TW-1 isolate shows ancestral properties. [21]
Widest distances were found between Brazilian and Taiwanese strains (as seen below in distance matrix). The USA and Sao Paulan strains were considered as siblings in all (100%) of bootstrap replications. Moreover, the sequences TW-1, TWN, TWN15 and TWN06 were grouped together in a monophyletic clade (i.e. a group containing the most common ancestor of a set of taxa and all the descendants of that most recent common ancestor) in all bootstrap replications as well. The Ranana 611 Israeli strain appears to be fairly recent compared to its GP200 neighbors, which is in accordance to the previously shown GP36 analysis.

	1	2	3	4	5	6	7
1) TWN 6							
2) TWN	0.000						
3) TWN 15	0.000	0.000					
4) TW-1	0.001	0.001	0.001				
5) USA	0.022	0.022	0.022	0.015			
6) 611	0.021	0.021	0.021	0.016	0.002		
7) Sao Paulo	0.024	0.024	0.024	0.018	0.022	0.021	

Figure 14: Pairwise distance matrix of the GP200 gene; overall average 0.014.

16S rRNA

As many articles have used this gene to both categorize and amplify *E. canis*, a phylogenetic analysis of this gene should be included.
In the bottom part of this tree, many strains and isolates appear as unbranched, including the Taiwanese strain TWN1-6 that were analyzed before. Their early branching-out is on the other hand in accordance with the GP36 tree.
Next several clusters appear in which strains from practically all continents and geographically unrelated countries coincide. Therefore, drawing consistent conclusions from this tree is again a difficult task. On the other hand, some branches containing mostly Asian strains can be recognized, including Malaysia, China and Taiwan. Overall, the balance of sequenced samples tips over to the Asian side and European and American variances are insufficiently represented, causing the tree to be imbalanced.

Abbreviations:

TW(N)x > Taiwan
Hdx > Cape Verde
UFVx, GOx, S22 > Brasil
xB,10Mx, CNx, YNEx, GX > China
VTE > Venezuela
EDx > Tunesia
EC > Mexico
MSIA > Malaysia
ECAN > Thailand
95E10 > USA

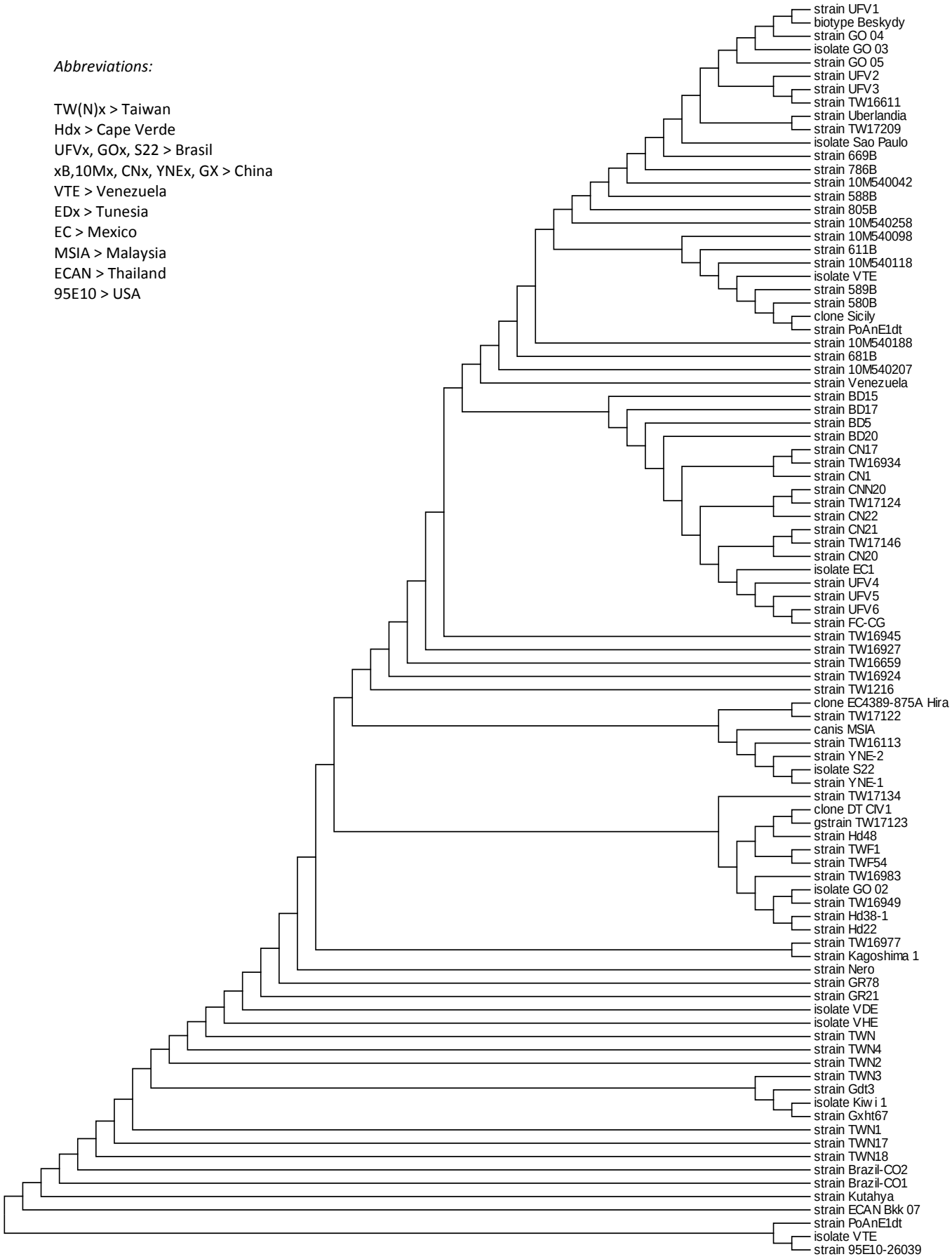

Figure 15: Phylogenetic tree of E. canis *16S rRNA gene.*

GLTA

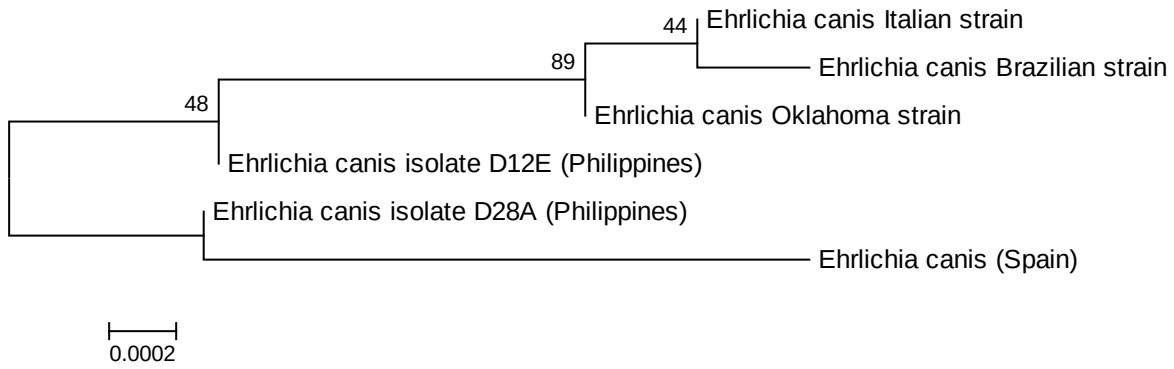

Figure 16: Phylogenetic tree of E. canis *GLTA gene.*

The above phylogenetic tree shows the genetic correspondence of the citrate synthase gene, also noted as GLTA. The matrix below shows the widest distance between strain Brazil and Spain. Remarkably, there is no logical relatedness between the strains from different continents, as to be expected. For instance, both Philippine strains appear in different taxa, and so do the Spanish and Italian varieties. Moreover, the scale bar represents only a 0.02% estimated sequence divergence, indicating a very low number of mutations. Therefore no logical phylogenetic relationship can be deduced from this tree.

	1	2	3	4	5	6
1)D12						
2)Spain	0.000					
3)D28	0.000	0.000				
4)Oklahoma	0.000	0.006	0.000			
5)Italy	0.000	0.006	0.000	0.006		
6)Brazil	0.000	0.008	0.000	0.000	0.000	

Figure 17: Pairwise distance matrix of the GLTA gene; overall average 0.002

2) Phylogenetic analysis of *Rhipicephalus* spp.

Based upon the Internal Transcribed Spacer (ITS) 2 gene, the following phylogenetic tree was made in order to show the ancestry and relative context of most relevant *Rhipicephalus* tick species, all vector of *E. canis*. The subsequent primers and probes will help to distinguish between these species, since phenotypical factors are not always reliable in this respect.

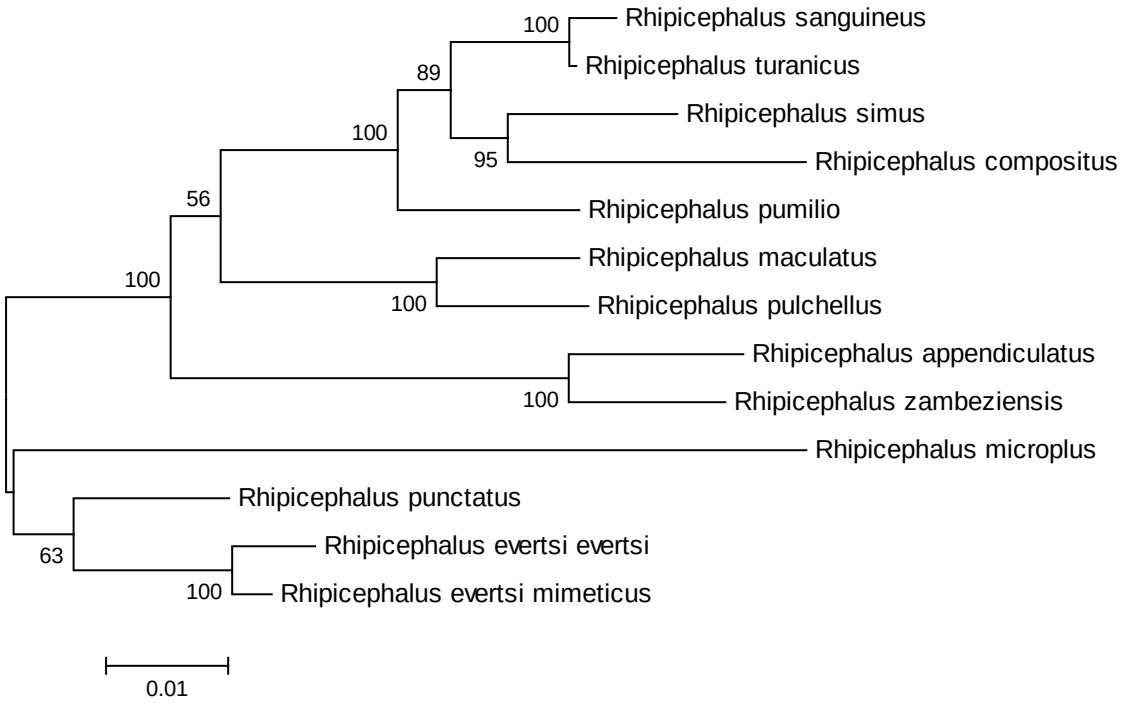

Figure 18: Phylogenetic tree of R. sanguineus *ITS2 gene.*

Apparently, *R. punctatus* and *R. evertsi* are the earliest subspecies. A very close relatedness is seen between *R. sanguineus* and *R. turanicus, R. maculatus and R. pulchellus, R. appendiculatus and R. zambeziensis,* respectively. The used sequence varied in country of origin: most were from Australia, except for R. sanguineus and *R. simus* (China), *R. appendiculatus* and *R. turanicus* and *R. evertsi evertsi* (Zambia), and *R. zambeziensis* was from South Africa. These differences appear to have no influence on their distribution in this tree. The high degree of similarity (distance matrix value 0.004, lowest overall value) between *R. sanguineus* and *R. turanicus* will have a challenging influence on primer design. The relevance of the discussion on classifying *Boophilus (R.) microplus* as a subgenus of *Rhipicephalus* is made visible both in its high distance values as well as its phylogenetic tree position. [17]

	1	2	3	4	5	6	7	8	9	10	11	12	13
1)R. sanguineus													
2)R.(B.) microplus	*0.098*												
3)R. appendiculatus	*0.087*	*0.132*											
4)R.zambeziensis	*0.085*	*0.120*	*0.027*										
5)R.turanicus	*0.004*	*0.108*	*0.082*	*0.080*									
6)R. evertsi evertsi	*0.077*	*0.084*	*0.086*	*0.078*	*0.073*								
7)R. maculatus	*0.066*	*0.128*	*0.073*	*0.078*	*0.064*	*0.068*							
8)R. evertsi mimeticus	*0.075*	*0.082*	*0.082*	*0.078*	*0.073*	*0.010*	*0.065*						
9)R. pulchellus	*0.065*	*0.030*	*0.070*	*0.074*	*0.062*	*0.070*	*0.024*	*0.065*					
10)R. simus	*0.032*	*0.121*	*0.088*	*0.090*	*0.028*	*0.085*	*0.065*	*0.080*	*0.066*				
11)R. pumilio	*0.032*	*0.115*	*0.082*	*0.080*	*0.028*	*0.075*	*0.059*	*0.070*	*0.057*	*0.039*			
12)R. punctatus	*0.068*	*0.088*	*0.082*	*0.076*	*0.066*	*0.033*	*0.016*	*0.029*	*0.066*	*0.069*	*0.065*		
13)R. compositus	*0.043*	*0.135*	*0.095*	*0.102*	*0.040*	*0.094*	*0.070*	*0.087*	*0.078*	*0.038*	*0.049*	*0.085*	

Figure 19: Pairwise distance matrix of Rhipicephalus spp., overall average 0.072.

PRIMER AND PROBE DESIGN

Both for *Rhipicephalus* spp. and *E. canis,* several primersets and matching probes were computed. After being validated, these sequences can be used in detecting both *E. canis* and its tick vector by means of PCR or RLB.

PRIMERS

1.1) Rhipicephalus catch-all primer sets

In order to distinguish with accuracy between *Rhipicephalus* spp. and other seemingly equal tick species, a catch-all RLB primer set was designed using the *Rhipicephalus* internal transcribed spacer (ITS) gene (Genbank accession numbers: JF758643.1, AF271283.1, FJ416321.1). After validation, these primers can be used in realtime PCR assays.

Name	**Sequence (5' to 3')**	**Tm˚C**	**CG%**	**nt**
RHIPISforw1	GAT CAC ATA TCA AGA GAG ACT TCG G	64.3	44.0	25
RHIPISrev1	GTT TAT CAC GCA ACT GCT CG	63.8	60.0	20
RHIPISforw2	CTA CAC GAG ACG ATG CCT CTC	64.3	57.1	21
RHIPISrev2	CTT GGT AGG GCG TCG TAC TC	63.9	60.0	20

Both sets will amplify *R. sanguineus, R. turanicus, R. bursa, R. appendiculatus, R. zambeziensis, R. evertsi* and also *R.(Boophilus) annulatus/microplus*. (Note: the two latter are referred to as both *Rhipicephalus* and *Boophilus* species.)

1.2) Differentiated Rhipicephalus primer sets.

Three primer sets were designed to replicate *R. sanguineus, R. turanicus* and *R. bursa*, respectively.

Name	**Sequence (5' to 3')**	**Tm˚C**	**CG%**	**nt**
RSANGforw	GAG ACT CGG ACG TGC AAC TG	66.6	60.0	20
RSANGrev	CCT GAA GCT TTC CGT CGT AGT C	66.3	54.5	22

The above forward primer replicates *R. sanguineus* exclusively. Reverse primer also replicates *R. turanicus* and *R. compositus*, but these two combined will be usable in reacting to *R. sanguineus* only. Moreover, other criteria (Tm˚C, CG% and nt) are met so that combining these two primers could very well work out.

Name	**Sequence (5' to 3')**	**Tm˚C**	**CG%**	**nt**
RBURSAforw	TGG TTN GCG GAC TCC TCT TTG	67.8	55.0	20
RBURSArev	TAT TGC GGT TCG CTG CGT AC	68.3	55.0	20

N=degenerate position A/C/T/G (gap).

RBURSAforw catches *R. bursa* exclusively.

Name	Sequence (5' to 3')	Tm˚C	CG%	nt
RTURforw	CTT GTT GCC TTC CGA ATA AGC	64.9	47.6	21
RTURrev	--C CGA AAC GGA AAA ATG TCT C	63.7	45.0	20

Only possible replication of *R. turanicus* was of one strain (GenBank number FJ416320.1).

2) Differentiated realtime PCR primer sets of E. canis genes

GLTA

Name	Sequence (5' to 3')	Tm˚C	CG%	nt
GLTAforw1	GCT ATT GGA ATA CCA GTG AG	*56.8*	*45.0*	*20*
GLTArev1	GGT GCA GTC AAT ATA TGA CCA G	*58.1*	*45.5*	*22*
GLTAforw2	CTC AGG AGT ATA TGC CTC CTG	*60.8*	*52.4*	*21*
GLTArev2	CAT GAG GAA GCA GTT GAT AAA G	*61.6*	*40.9*	*22*
GLTAforw3	GAT CCA CGT GCT AAG ATA ATT TGT G	*64.6*	*40.0*	*25*
GLTArev3	GAA CGT AGC TTG TAT CCT AAT GTG G	*63.8*	*44.0*	*25*

A beginning was made in validating the above GLTA primer sets, including the set found in an article by Marsilio et al. [22]

VIRB9

Name	Sequence (5' to 3')	Tm˚C	CG%	nt
VIRB9forw1	GCA ATG CAC ACT CCA TAA GC	*63.7*	*50*	*20*
VIRB9rev1	CCT AAC ATT ACA AGC GAC AAC TAC C	*63.8*	*44.0*	*25*
VIRB9forw2	GAC CTA AGA GAT CTA GCA TAC ATA GTA CGG	*64.3*	*43.3*	*30*
VIRB9rev2	CCTGTGGAGTTATTTGATGATGG	*64.5*	*43.5*	*23*

Using BLAST, both primer sets were found to amplify the *E. canis* VirB9 gene exclusively (100%).

GP36

Name	Sequence (5' to 3')	Tm˚C	CG%	nt
GP36forw1	CTG CTC CAG CTA CTG AAG ATT C	62.5	50.0	22
GP36rev1	AGC TAC TGA AGA TTC TGT TTC TGC	62.2	41.7	24

This gene mostly consist of multiple 'ctgctccagctactgaagattctgttt'- repeats, making it virtually impossible to produce a consistent primer pair. GP36 is therefore more or less unsuitable for use in realtime PCR and RLB diagnostics.

GP28

Name	Sequence (5’ to 3’)	Tm˚C	CG%	nt
GP28forw1	TAT CGG TTA CTC AAT GGG TGG C	67.0	50.0	22
GP28rev1	CAC AGG TAC TGC GCT CTA TCT CAT C	67.2	52.0	25

GP200

Name	Sequence (5’ to 3’)	Tm˚C	CG%	nt
GP200forw1	GGC ATA TAC CAA CCT CCC ATA G	63.4	50.0	22
GP200rev1	TGA TAC TCC TTG GCC AGA TG	63.1	50.0	20
GP200forw2	TCT CCT ACT CCT GAA CCG AAA G	62.9	50.0	22
GP200rev2	GCC AGA TGT TAG GAA TGC AG	62.7	50.0	20
GP200forw3	GGC ATA TAC CAA CCT CCC ATA G	63.4	50.0	22
GP200rev3	GGT TAA CTA CTT CTG GGC CAG TAC	63.1	50.0	24

More primer sets could be deduced with relative ease, since the GP200 gene is overall rich in CG content and has low nucleotide variability among various strains.

PROBES

On the 5’ side of the following sequences, a C6-aminolinker is attached.

1.1) Rhipicephalu*s catch-all RLB probes*

Coupled with primer pair	Sequence (5’ to 3’)	Rhipicephalus species
1	GAG TAA GCC GGG TRG CCC GCA GAC CG	sanguineus, turanicus, bursa, evertsi, punctatus and compositus.
	CGG GGA GCG AAA GCC GGC CAW CGA	sanguineus, turanicus, bursa, evertsi, punctatus, maculatus, simus and pumilio.
	CGA ACA GGG AAC GTT CGR GCG CGG AG	sanguineus, turanicus, bursa, evertsi, punctatus and maculatus.
2	ACG CAC GCG TGC AGC GGG ATA CCG CT	*sanguineus, turanicus, bursa, evertsi, punctatus, maculatus, compositus, appendiculatus, simus, pumilio, pulchellus, zambezienzis, microplus.*

R= degenerate position C/T in sequence 1 and 3; W= degenerate position G/T in sequence 2.

1.2) Differentiated R. sanguineus *probes*

Species	**Sequence (5' to 3')**
R. sanguineus	CAG GAG GTG CGG CGA CTC GTC GCG AA AGG AGG TGC GGC GAC TCG TCG CGA A
R. bursa	GNN ACG GTG CCG NTG CCC CTC CGG GCT CTT
R. turanicus	ACA CCC CAG ACG GTG ACG GCG CCC CTC CG

N= degenerate position A/C/G/T (sequence gap).

2) E. canis *probes*

GLTA

Coupled with primer set	**Sequence (5' to 3')**
1	TTG CGT TGG CTA GAA CAT CTG GTT G
2	AAA TGG TTT GTG ATG TTA TAA AGT CA
3	GTG TAA ATG ATC CAT TAT TAG AAA TA

VIRB9

Coupled with primer set	**Sequence (5' to 3')**
1	GGA GTT ATG CTT TTG ACC TGA TAT GC
2	TAG TAA TAA AAC CTA ACA TTA CAA GC

GP28

Coupled with primer set	**Sequence (5' to 3')**
1	TA{T,C}GA{A,G}{G,A}CATTCGA{C,T}GT{A,G}AAAA{G,A}TC {C,T}{T,G}TA{T,C}GA{A,G}{G,A}CATTCGA{C,T}GT{A,G}AAAA{G,A}T

Degenerate nucleotide positions could not be avoided in designing the GP28 probes, since the amplicon is overall irregular.

GP200

Coupled with primer set	**Sequence (5' to 3')**
1	CAA GAT AGT AAT CTA TAT TCA AGT ATT G TCA AGT ATT GGT GGC GTA CCA CAA GA
2	ATA CAT TCA GTA ATG GTC AGG AAA TT CAA TAT GAT GCA GCA GCT CGG GCT GG
3	TAA TCT ATA TTC AAG TAT TGG TGG CG CAA TAT GAT GCA GCA GCT CGG GCT GG GCC ATA TAC ATT CAG TAA TGG TCA GGA GAA TTT GAT ACT CCT TGG CCA GAT GT

REALTIME PCR PRIMER VALIDATION

As mentioned before, a beginning was made in validating the designed GltA primer sets.

All three primer sets (GltA1-3), including a set derived from Marilio et al. [22], were tested on positive control by conventional PCR and run on an agarose gel (see below).

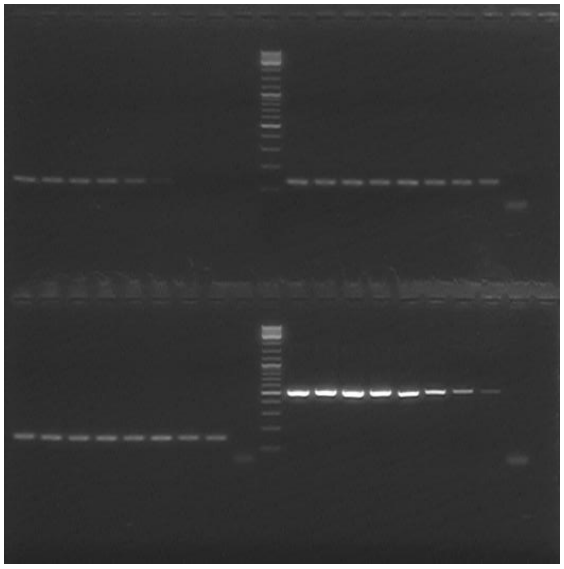

Figure 20: clockwise from upper left:GLTA1, GLTA3, GLTA2, GLTA from article[22].

This was followed by a gradient test, which proved GltA2 to be unsuitable due to a low plateau value. The other two sets had top values at 57.5˚C (GltA1) at 59.5˚C (GltA3).

Next a dilution series was made with GltA3 as seen in the figure below, followed by the corresponding meltcurve which correctly showed one peak i.e. no nonspecific products.

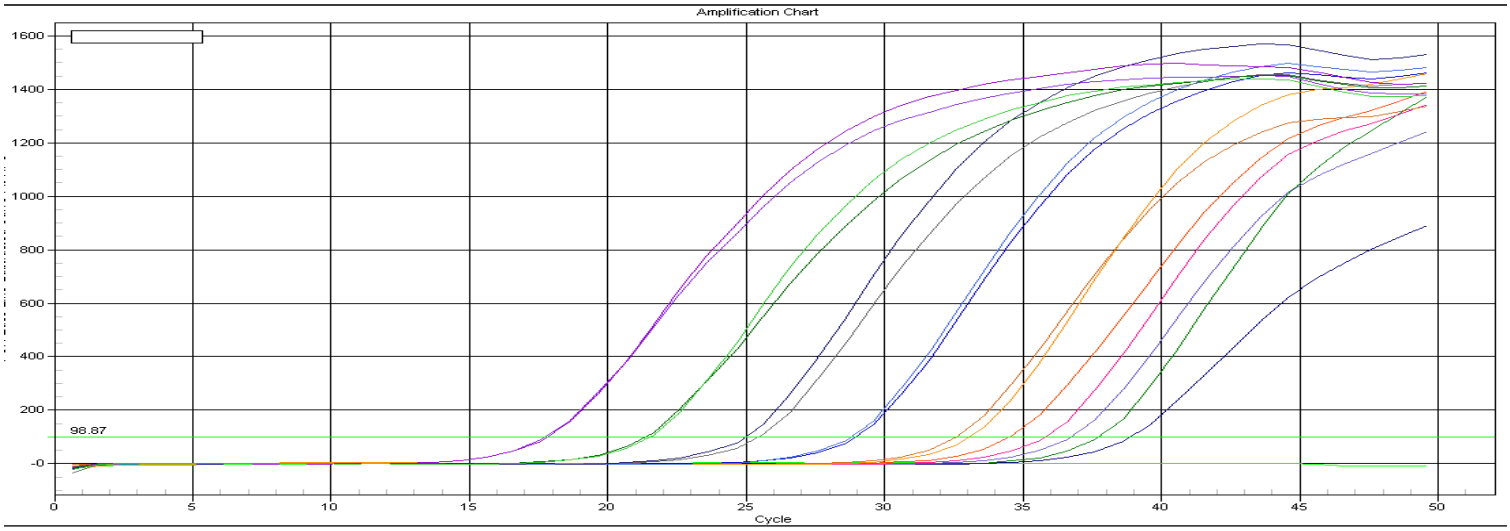

Figure 21: GLTA3 dilution series.

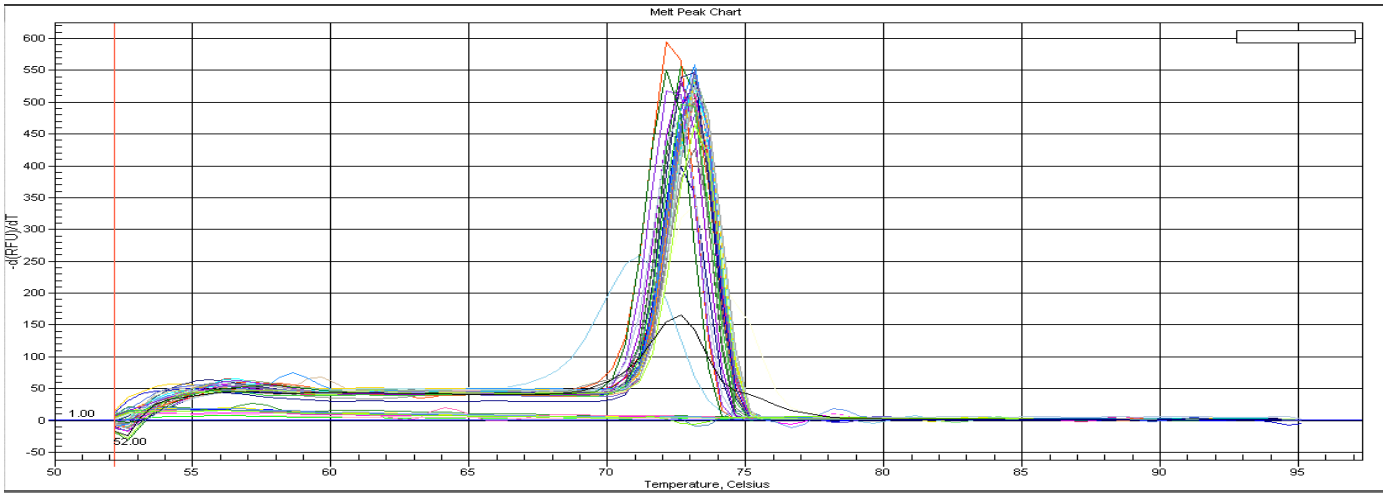

Figure 22: GLTA3 meltcurve.

DISCUSSION AND RECOMMENDATIONS

Primers and probes were made for the three most relevant *Rhipicephalus* spp. Based on the results so far, the first catch-all primer pair (RHIPISforw1/RHIPISrev1) appears to be most promising, since the concerning criteria are best met and multiple options for matching RLB probes are present. All of these probes should be capable of catching the three most relevant *Rhipicephalus* species (*R. turanicus, R. bursa* and *R. sanguineus*).

Yet, sets are still to be designed for the remaining spp., including *R. evertsi, R. punctatus, R. maculatus* and so forth. This will increase the degree of usefulness in creating a novel (q)PCR/ RLB assay.

The GP genes (36, 28 and 200) provided a good starting point for phylogenetic analysis of *E. canis*, as opposed to the VIRB9, GLTA and 16S rRNA genes. The latter is particularly extensive, meaning that the number of used strains should be restricted in order to make a clearer ancestral analysis. The GLTA and VIRB9 assessments would need more sequences from other (European) studies to improve the degree of accuracy of the analyses.

The GP200 gene showed many possibilities for primer design, so that the sets proposed in this study could be increased in number with relative ease, if necessary. Matching probe design for this gene showed enough opportunities as well. GLTA primers and probes also appear useable, but to a lesser degree. GP36 on the other hand proved to be highly unsuitable for primer design due to multiple sequence repeats.

Most importantly, primer sets and probes still need to be fully validated. A start has been made in validating the GLTA3 primer set, with promising prospects.

ACKNOWLEDGEMENTS

I would like to thank prof. dr. Frans Jongejan and Michiel Wijnveld for their guidance and Marleen for our daily gezelligheid.

REFERENCES

1) WALKER J.B., KEIRANS J.E., HORAK I.G. 'The genus Rhipicephalus (Acari, Ixodidae): a guide to the brown ticks of the world' Cambridge University Press, 2000.
2) SKOTARCZAK B. Canine ehrlichiosis. *Ann Agric Environ Med* **10**, p137–141, 2003.
3) NAKAGHI A.C.H., MACHADO R.Z., FERRO J.A., LABRUNA M.B., CHRYSSAFIDIS A.L., ANDRÉ M.R., BALDANI C.D.: Sensitivity evaluation of a single step PCR assay using Ehrlichia canis p28 gene as a target and its application in diagnosis of canine ehrlichiosis. *Rev. Bras. Parasitol. Jaboticabal.* **19(2)** p. 75-79, 2010.
4) HSIEH Y.-C., LEE C.-C., TSANG C.-L., CHUNG Y.-T. Detection and characterization of four novel genotypes of Ehrlichia canis from dogs. *Veterinary Microbiology* **146** p. 70-75, 2010.
5) CARDOSO L., TUNA J., VIEIRA L., YISASSHAR-MEKUZAS Y., BANETH G. Molecular detection of Anaplasma platys and Ehrlichia canis in dogs from the north of Portugal. *The Veterinary Journal* **183** p. 232-233, 2010.
6) BANETH G., HARRUS S., OHNONA F.S., SCHLESINGER Y. Longitudinal quantification of Ehrlichia canis in experimental infection with comparison to natural infection. *Veterinary Microbiology* **136** p. 321-325, 2009.
7) AGUIRRE E., TESOURO M.A., AMUSATEGUI I., RODRIGUEZ-FRANCO F., SAINZ A. Comparison between different polymerase chain reaction methods for the diagnosis of Ehrlichia canis infection. *Animal Biodiversity and Emerging Diseases* **1149** p.118-120, 2008.
8) ALEXANDRE N., SANTOS A.S., NÚNCIO M.S., DE SOUSA R., BOINAS F., BACELLAR F. Detection of Ehrlichia canis by polymerase chain reaction in dogs from Portugal. *The Veterinary Journal **181*** p. 343-344, 2009.
9) GAL A., LOEB E., YISASCHAR-MEKUZASY., BANETH G. Detection of Ehrlichia canis by PCR in different tissues obtained during necropsy from dogs surveyed for naturally occurring canine monocytic ehrlichiosis. *The Veterinary Journal* **175** p. 212-217, 2008.
10) VINASCO J., LI O., ALVARADO A., DIAZ D., HOYOS L., TABACHI L., SIRIGIREDDY K., FERGUSON C., MORO M.H. Molecular evidence of a new strain of Ehrlichia canis from South America. *Journal of Clinical Microbiology* **45**(8) p.2716-2719, 2007.
11) PELEG O., BANETH G., EYEL O., INBAR J., HARRUS S. Multiplex real-time qPCR for the detection of Ehrlichia canis and Babesia vogeli. *Veterinary Parasitology* **173** p. 292-299, 2010.
12) SASHIKA M., ABE G., MATSUMOTO K., INOKUMA H. Molecular survey of Anaplasma and Ehrlichia infections of feral raccoons (Procyon lotor) in Hokkaido, Japan. *Vector-Borne and Zoonotic Diseases* **11**(4) p. 349-354, 2011.
13) NDIP L.M., NDIP R.N., NDIVE V.E., AWUH J.A., WALKER D.H., McBRIDE J.W. Ehrlichia species in Rhipicephalus sanguineus ticks in Cameroon. Vector-Borne and Zoonotic Diseases **7**(2) p. 221-227, 2007.
14) KLEDMANEE K., SUWANPAKDEE S., KRAJANGWONG S., CHATSIRIWECH J., SUKSAI P., SUWANNACHAT P., SARIYA L., BUDDHIRONGAWATR R., CHAROONRUT P., CHAICHOUN K. Development of multiplex polymerase chain reaction for detection of Ehrlichia canis, Babesia spp. and Hepatozoon canis in canine blood. *Southeast Asian J. Trop Med Public Health* **40** p.35-39, 2009.
15) Realtime PCR applications guide. Bio-Rad Laboratories, Inc.
16) YBANEZ A.P., MATSUMOTO K., INOKUMA H. First molecular detection of Ehrlichia canis and Anaplasma platys in ticks from Dogs in Cebu, Philippines. *Obihiro University of Agriculture and Veterinary Medicine, Japan.* Unpublished, submitted 22-07-2011.

17) JONGEJAN F., KAUFMAN W.R. 'Ticks and Tick-borne pathogens-proceedings of the 4th international conference on ticks and tick-borne pathogens.' Academic Publishers, 2003.
18) NIU Q., GUAN G., LIU Z., MA M., LI Y., LIU A., REN Q., LIU J., LUO J., YIN H. Simultaneous detection of piroplasma infections in field *Haemaphysalis qinghaiensis* ticks by reverse line blotting. *Exp. Appl. Acarol.* **56** (2) p. 123-132, 2012.
19) BEARD D.A., VINNAKOTA K.C., WU F. Detailed enzyme kinetics in terms of biochemical species: study of citrate synthase. *PLoS ONE* **3** (3), 2008.
20) Reverse line blot hybridization in the detection of tick-borne diseases. BTi, 2004.
21) ZHANG X., LUO T., KEYSARY A., BANETH G. Genetic and antigenetic diversities of major immunoreactive proteins in globally distributed *Ehrlichia canis* strains. *Clin Vaccine Immunol.* **15** (7) p.1080-1088, 2008.
22) MARSILIO F., DI MARTINO B., MERIDIANI I., BIANCIARDI P. Direct identification of *Ehrlichia canis by* a novel polymerase chain method and molecular analysis of the citrate synthase (gltA) gene from various Italian strains. *J Vet Diagn Invest.* **18**, p. 215-217, 2006.
23) www.vetpda.ucdavis.edu
24) www.vet.uga.edu