ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Veterinary Parasitology 138 (2006) 61–74

veterinary parasitology

www.elsevier.com/locate/vetpar

Expressed sequence tag (EST) analysis of the erythrocytic stages of *Babesia bovis*

Erik de Vries [a,*], Craig Corton [b], Barbara Harris [b], Albert W.C.A. Cornelissen [a], Matthew Berriman [b]

[a] *Division of Infection Biology, Department of Infectious Diseases and Immunology, Utrecht University, P.O. Box 80165, 3508 TD Utrecht, The Netherlands*
[b] *Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK*

Abstract

Expressed sequence tags (ESTs) provide an efficient way to identify large numbers of genes expressed in a specific stage of the life cycle of an organism. Here we analysed ~13,000 ESTs derived from the erythrocytic stage of the apicomplexan parasite *Babesia bovis*. The ESTs were clustered in order to obtain information on the expression level of a gene and to increase sequence length and reliability. A total of 3522 clusters were obtained and annotated using BLAST algorithms. The clusters were estimated to represent ~2600 genes of which in total ~2.1 Mbp sequence information was obtained. Expression levels of the genes, as determined by the numbers of ESTs contained within a cluster, were compared to those of their closest homologs in the erythrocytic stage of *Plasmodium falciparum* and *Toxoplasma gondii* tachyzoites. Pathways that are represented relatively abundant in *B. bovis* are, amongst others, the purine salvage pathway (displaying characteristics not identified before in apicomplexans), isoprenoid biosynthesis in the apicoplast and many genes encoding mitochondrial proteins. Especially remarkable in the latter group are the F-type ATPases – which are hardly expressed in *P. falciparum* and *T. gondii* – and two highly expressed glycerol-3-phosphate dehydrogenases creating a shuttle possibly controlling the cytoplasmic $NADH/NAD^{+}$-ratio. A comparison of known antigenic proteins from Australian and American strains of *B. bovis* with the Israel strain used here identifies considerable sequence variation in the rhoptry associated protein-1 (RAP-1), merozoite surface proteins of the variable merozoite surface antigen (VMSA) family and spherical body proteins. Analysis of the EST clusters representing the variable erythocyte surface antigen family reveals many variant transcripts of which a few are dominant. Two putative pseudogenes also seem to be transcribed at high levels.
© 2006 Elsevier B.V. All rights reserved.

Keywords: *Babesia bovis*; Expressed sequence tag (EST); Erythrocytic stages

1. Introduction

Babesia bovis is a tick-transmitted apicomplexan hemoparasite of cattle. *Babesia* and the closely related genus *Theileria* are together classified as the piroplasmidae which in turn have been grouped

* Corresponding author. Tel.: +31 30 2532582; fax: +31 30 2540784.
E-mail address: E.devries@vet.uu.nl (E. de Vries).

0304-4017/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.vetpar.2006.01.040

together with the plasmodia within the hemosporidae. Despite its severe impact on the cattle industry in large parts of the tropical and sub-tropical world, sequence information on *B. bovis* is currently limited to ~75 genes, often encoding antigens that were studied for their potential role in induction of protective immunity (Brown and Palmer, 1999; Dalrymple, 1993). The use of expressed sequence tags (ESTs) obtained by single-pass sequencing of cDNAs provides a cost-efficient means of rapidly acquiring sequence information on a large number of expressed genes.

We have analysed ~13,000 ESTs of *B. bovis* that were sequenced from clones randomly obtained from a non-subtracted cDNA library representing the erythrocytic stages of an in vitro cultured clonal line. Asynchronous development seems to be an intrinsic property of *B. bovis* (Franssen et al., 2003). In vitro synchronization has never been achieved. A pre-erythrocytic cycle is lacking since *B. bovis* sporozoites directly infect erythrocytes. Thus the obtained EST set represents most of the developmental cycle within the vertebrate host. Gametocytes can be induced in in vitro *Babesia* cultures (Mosqueda et al., 2004) but, as far as can be determined by light-microscopic parasite morphology, do not occur in the clonal line used here.

Although ESTs are short sequences usually representing only part of a gene, multiple ESTs derived from overlapping regions of a gene can be assembled into clusters. This improves sequence reliability due to multiple coverage and sequence length because clusters often encode a full-length open reading frame. Also, in contrast to genome sequencing projects, a collection of randomly sequenced ESTs provides information on the expression level of a gene (Carulli et al., 1998) and the number of ESTs in a cluster have been shown to correspond to transcriptional levels as observed by microarray techniques although a length-related bias was observed (Munoz et al., 2004).

This article discusses the general summary statistics of the *B. bovis* EST collection. EST clusters were annotated employing the BLAST algorithms on general sequence databases as well as specific apicomplexan EST and genomic sequence databases, allowing comparison of *B. bovis* with parasites from related genera. Selected cases in which proteins or metabolic pathways of *B. bovis* display interesting features in comparison to other apicomplexans are discussed in some more detail. Finally, as the *B. bovis* strain analysed here originates from Israel, an opportunity is provided to assess the degree of sequence heterogeneity present in antigenic proteins or protein families that have been studied almost exclusively in isolates from Middle or South America and to a lesser extent from Australia.

2. Materials and methods

All ESTs have been derived from an in vitro cultured clonal line (C61411) derived from a *B. bovis* strain isolated in Israel. cDNA clones were randomly selected for sequencing from a cDNA library constructed as previously described (Gaffar et al., 2004a,b,c). In brief, mRNA was isolated from an asynchronous *B. bovis* C61411 in vitro culture. First-strand cDNA was synthesized on mRNA selected for sizes larger than 300 nucleotides using an oligo-dT primer containing a 5′-terminal Xho1 restriction site. After synthesis of the second strand EcoR1 adapters were ligated followed by digestion with Xho1 allowing uni-directional ligation into the EcoR1/Xho1 site of lamba-uniZAP-XR (Stratagene). Sequencing was performed from the 5′-end into the direction of the poly-A-tail using the standard double-strand reverse sequencing primer.

3. EST database generation, clustering and BLAST annotation

cDNA clones were randomly selected and sequenced from the 5′-end. Vector sequences were clipped from the resulting high quality sequences after which they were assembled into clusters. Clusters containing *E. coli* or ribosomal RNA sequences and clusters shorter than 150 bp were removed. A total number of 3522 clusters were obtained with a total length of 2,117,885 bp. Thus, maximally ~21% of the total genome sequence of *B. bovis* (~10 Mbp contained within four chromosomes) has been obtained. The clusters harboured 12,329 individual ESTs with still 2071 clusters (59%) being composed of a single EST. A recent analysis of EST clusters derived from non-subtracted libraries of *Plasmodium falciparum*, *Toxoplasma gondii*, *Neospora caninum*

Table 1
Summary statistics of EST clusters

No. of EST's in cluster	No. of clusters	No. of EST's in group	Total length (bp)	Average length	No. of clusters with BLAST SCORE			Average score
					>50	>75	>100	
1	2071	2071	923358	446	1077 (52%)	778 (38%)	504 (25%)	113
2	533	1066	342007	642	390 (73%)	314 (59%)	251 (47%)	138
3–5	508	1866	412910	813	421 (83%)	364 (72%)	303 (60%)	163
6–10	200	1482	201670	1008	175 (88%)	164 (82%)	140 (70%)	220
10–25	142	2312	156802	1104	124 (87%)	114 (80%)	105 (74%)	256
26 or more	68	3532	81138	1193	58 (85%)	56 (82%)	52 (76%)	334
All clusters	3522	12329	2117885	601	2245 (64%)	1790 (51%)	1356 (39%)	149

Numbers of blast hits were categorized in three groups (score: >50; >75; >100) according to the highest score obtained from the different databases that were used. Scores for clusters as obtained with the individual databases can been viewed in supplementary Table S1. All ESTs are available for download and BLAST searching at www.sanger.ac.uk/Projects/B_bovis/ and will be submitted to GenBank.

and *Eimeria tenella* (Li et al., 2003) revealed a relatively non-redundant EST dataset for *P. falciparum* and *T. gondii* (where >90% of ESTs occurred in assemblies that each constituted <0.1% of all ESTs). In contrast, *B. bovis* EST assemblies are more like those of *N. caninum* and *E. tenella* in having a group of highly abundant clusters – 210 clusters with more than 10 ESTs in *B. bovis* – that contain 47% of all ESTs (Table 1 summarizes cluster data). As shown below, this is primarily due to the high abundance of ESTs involved in translation and in spherical body protein encoding ESTs.

A number of sequence databases were employed for annotating the EST clusters by BLAST analysis (Altschul et al., 1990). These include the non-redundant (NR) protein database deposited at NCBI-genbank, the Swiss-Prot database, a database composed of all apicomplexan EST sequences (274,547 in total) available from Genbank and a database composed of the predicted proteins from the nearly complete genomic sequences of *P. falciparum*, *Theileria annulata*, *T. gondii* and *Cryptosporidium parvum* that recently have become available (Gardner et al., 2002; Abrahamsen et al., 2004; www.toxoDB.org; www.sanger.ac.uk/Pathogens/T_annulata/). The DNA sequences of the *B. bovis* clusters were analysed by BLASTX whereas a *B. bovis* protein database (3099 proteins; minimum length 40 amino acids) was derived from the clusters by conceptual translation of the longest open reading frames. The protein database was analysed using BLASTP on protein databases and TBLASTN on the EST database. To be able to compare blast results obtained with different blast-databases the normalized BLAST-score is reported here instead of its more often used derivative *P* (the probability score which is related to query and database size). As genes evolve at very different rates no single meaningful cut-off value can be determined above which sequences aligned by BLAST are guaranteed to represent homologous sequences. The great majority of BLAST hits with a score above 50 (comparable to probability score $P \sim e^{-6}$ when using the Swiss-Prot database, 100 is comparable to $P \sim e^{-20}$) are believed to represent alignments of homologous sequences (without the ability to discriminate between orthologous protein matches or, for instance, conserved domain matches within otherwise heterologous proteins). Table 1 lists the number of matches obtained at three cut-off values. A group of 2,245 clusters – 63.7% of all clusters, containing 75.3% of all ESTs – have a match to another protein of which 950 clusters only match proteins annotated as hypothetical. The group of 1277 clusters without matches, or scoring below 50, contains 423 clusters from which no protein sequence was predicted, expectedly because they are mainly composed of 5′- or 3′-untranslated regions. Groups harbouring the clusters with higher numbers of ESTs have more and better matches, up to 76.5% with a score above 100. Undoubtedly, this mainly relates to the longer, often complete, sequences present in these groups, as also suggested by the similar ratio between average score and average sequence length in the different groups displayed in Table 1.

How many genes are represented by the 3522 clusters? Inspection of the 2001 clusters with

similarity to *T. annulata* genes (see Table 3) shows that they match with 1496 different *T. annulata* genes. Many of these clusters map to different regions of the same protein or, to a small extent, are not correctly assembled into one cluster. Extrapolation to all 3522 clusters indicates the representation of approximately 2600 genes in the current *B. bovis* EST dataset which would correspond to 69% of all genes if the total number of genes equals the number found for *T. annulata* (3,789 genes according to the preliminary annotations available at www.sanger.ac.uk).

4. Comparison with genome and expressed protein complement of other apicomplexan genera

4.1. Overview

A comprehensive gene comparison with the genome sequences of *T. annulata*, *P. falciparum*, *T. gondii* and *C. parvum* awaits the sequencing of the complete *B. bovis* genome. Here, their sequences were used indirectly to compare the expression levels of *B. bovis* genes with their closest homologs in the comparable life cycle stages of *P. falciparum* and *T. gondii*. The ESTs from non-subtracted, randomly selected *P. falciparum* erythrocytic stage libraries (11,888 ESTs) and *T. gondii* in vitro cultured tachyzoite libraries (24,030 ESTs) were linked to their cognate genes using BLASTN which in turn were linked to the translated *B. bovis* EST clusters using BLASTP. *T. annulata* and *C. parvum* genomes were included in the analysis although insufficient ESTs of these species are available for comparison. A comparative analysis involving the indicated EST libraries directly provides a glimpse into the pathways and proteins active in the *B. bovis* erythrocytic cycle and allows to determine their potential presence – by comparison with genomes – and actual activity – by comparison with ESTs – in other apicomplexans.

Despite strong similarities between the erythrocytic cycle and associated pathogenesis of *B. bovis* and *P. falciparum* (Clark and Jacobson, 1998; O'Connor et al., 1999), phylogenetic analyses based on 18S rRNA alignments variably positioned the Piroplasms (*Babesia* and *Theileria*) in a clade with the coccidia (e.g. *T. gondii* and *Eimeria* sp.) or *Plasmodium* sp. (Escalante and Ayala, 1994; Morrison and Ellis, 1997; Li et al., 2003). This ambiguous evolutionary relationship is reflected in the comparative BLAST results (Table 2). The piroplasms display their close relationship by the highest number of BLAST hits (2001 *B. bovis* clusters matching a *T. annulata* gene with an average score of 147). *P. falciparum* and *T. gondii* have similar numbers of matches (1445 and

Table 2
Summary of blast results with group- or species-specific databases

	No. of proteins	No. of clusters with score			Exclusive to group	Average score[a]	Average score[b]
		>50	>75	>100			
T. annulata	3789	2001	1597	1241	499	147	199
P. falciparum	5365	1445	1035	727	17	128	155
T. gondii	8336	1315	904	670	9	130	152
C. parvum	3396	914	634	450	3	125	131
All four	20886	2069	1657	1277			
Eukaryota[c]		1052	703	490			
Viridiplantae		963	653	461			
Bacteria		473	260	167			
Archaea		396	232	154			

The first column indicates the number of proteins encoded by the four apicomplexan species. The next three columns indicate the number of *B. bovis* EST clusters with a blast hit scoring higher than 50, 75 or 100, respectively. The fifth column indicates the number of proteins that *B. bovis* shares uniquely with a specific species.

[a] Column six shows the average score of all clusters with a BLAST hit with a score higher than 50.

[b] Column seven shows the average score of the 788 clusters with a BLAST hit with a score higher than 50 in all four apicomplexan species.

[c] The Eukaryota database does not include apicomplexan and viridiplantae sequences.

1315) and average scores (128 and 130, respectively) that are hardly higher than those observed with the more distantly related *C. parvum* (125).

The average scores of the *B. bovis* clusters having BLAST hits with all four species (788 clusters, Table 2, last column) are, however, clearly lower for the 788 matches with *C. parvum* in comparison to those with *P. falciparum* and *T. gondii*. A group of 499 clusters exclusively match with *T. annulata*; 430 of these matches are with piroplasm-specific genes as they do not match with any other protein of organisms outside the apicomplexa. Another 508 clusters can be considered as apicomplexan-specific. It is of interest to note that such groups can never be strictly defined because homologous genes in other organisms may have evolved below levels of significant similarity, here arbitrarily determined as a score lower than 50. Rapid evolutionary divergence within the apicomplexan phylum is suggested by the fact that less BLAST hits are found with *C. parvum* as with plants or other eukaryotes (excluding apicomplexa).

The majority of annotated clusters can be provisionally categorized in 23 functional groups. Assignment to unique groups is arbitrary in cases where proteins fit in multiple categories, e.g. a protein kinase can have a signalling function in cell cycle control. A supplementary Table (Table S1) is provided containing the cluster numbers, sequences, annotations and BLAST results of 1068 clusters composed of 6071 ESTs that are summarized in Table 3. Table S1 also includes the comparison of EST numbers between *B. bovis*, *P. falciparum* and *T. gondii*. EST numbers have not been normalized for EST database size as the *B. bovis* and *P. falciparum* EST collection are of equal size whereas the *T. gondii* collection is approximately two-fold larger.

4.2. Most abundantly expressed groups

The translational machinery represents the by far most abundant groups of ESTs (197 genes matching 19.1% of all ESTs). This large set includes the complete cytoplasmic ribosome gene complement (82 genes) as well as 18 putative organellar ribosomal proteins. This high relative abundance of translation related ESTs is also observed in *T. gondii* tachyzoites but is five-fold lower in *P. falciparum* merozoites. Spherical body proteins (discussed below) are remarkable in harbouring the most abundantly expressed individual genes (450 ESTs assembled in 13 clusters). ESTs related to transcription, mitochondrial function, purine salvage and glycolysis are also relatively abundant in *B. bovis*. The following sections briefly discuss some of the pathways or proteins for which relatively high numbers of ESTs have been identified in comparison to *P. falciparum* and/or *T. gondii*.

4.3. Organelles

Transcripts encoding mitochondrial proteins are six-fold more abundant in comparison to *P. falciparum* and *T. gondii*. This is not only caused by a few abundantly transcribed genes but consistent for a larger number of genes (see Table 3 and Table S1). For instance, ESTs encoding proteins of the TCA cycle (14 genes, 44 ESTs) and of respiratory chain complexes III and IV (14 genes, 86 ESTs) are two- to four-fold more abundant. Still, like in *P. falciparum*, a major role for mitochondrial oxidative phosphorylation in energy production is less likely considering the absence of ESTs encoding for the E1 subunit of pyruvate dehydrogenase, aconitase and nearly all components of complex I. Possibly, as suggested for malaria parasites (see Krungkai, 2004, for a recent review), the TCA cycle mainly serves to produce metabolites for other pathways whereas components of the electron transport system may be involved in maintaining a membrane potential required for transport of metabolites and proteins across the inner mitochondrial membrane.

Highly remarkable, considering the above mentioned facts, is the relatively high expression (76 ESTs in total, only four in *P. falciparum*) of the five subunits of the catalytic subunit CF(1) and subunit C of the proton channel CF(0) of the F-type ATP synthase. In *P. falciparum* a function of this ATP synthase has been questioned as genes encoding subunits A and B of the proton channel were not detected (Gardner et al., 2002), thus abolishing the link between a proton gradient and ATP generation or in reverse, proton-pumping by ATP hydrolysis. The observed high expression in *B. bovis* of the membrane-anchoring C subunit and the CF(1) catalytic subunit suggests a membrane-linked functional activity and presents a model for more detailed investigation. The most

Table 3
Comparison of EST abundance in functionally grouped EST clusters

Functional group	*B. bovis*			*P. falciparum*		*T. gondii*		*T. annulata*
	Genes	EST's	%	Genes	EST's	Genes	EST's	Genes
Translation machinery	197	2355	19.10	177	523	167	3258	188
Ribosomal proteins (cytoplasmic)	82	1800		72	308	77	2792	81
Ribosomal proteins (organellar)	18	34		16	7	13	45	17
Spherical body proteins	13	450	3.65					
Transcription machinery	82	386	3.13	70	165	68	162	81
Mitochondrial function	54	297	2.41	50	53	48	129	52
TCA cycle	14	44		14	23	14	30	14
ATPase F-type	6	76		5	4	5	14	
Carbohydrate metabolism	18	265	2.15	17	171	16	161	17
Glycolytic pathway	15	252		13	152	14	157	14
Cytoskeletal/motor proteins	26	193	1.57	22	107	21	234	23
Nucleotide metabolism	27	187	1.52	24	111	24	73	24
Purine salvage	16	151		12	52	13	48	
Variant erythrocyte surface antigen	42	172	1.40					
Protein trafficking	44	167	1.35	39	175	36	100	42
Chromatin structure/regulation	21	152	1.23	17	126	19	142	20
Protein folding/chaperones	35	151	1.22	30	151	27	275	33
Signalling	27	146	1.18	24	170	25	116	29
Protein degradation	47	145	1.18	45	220	42	294	46
Endoplasmic reticulum proteins	16	133	1.08	15	192	15	65	16
Transporters	30	128	1.04	22	36	23	59	29
Lipid metabolism	40	126	1.02	34	86	24	53	38
Mitosis/cell cycle regulation	27	114	0.92	22	114	22	35	27
B. bovis merozoite antigens	6	112	0.91					
Vacuolar	16	90	0.73	14	46	13	40	15
Protein kinases/phosphatases	25	83	0.67	23	74	23	29	25
Redox	14	61	0.49	12	58	10	124	13
DNA replication	22	54	0.44	20	81	18	25	20
DNA repair	20	43	0.35	14	22	10	19	19

The ESTs from non-subtracted, randomly selected *P. falciparum* erythrocytic stage libraries (11,888 ESTs) and *T. gondii* in vitro cultured tachyzoite libraries (24,030 ESTs) were used as described in the text. The total number of genes represented by the *B. bovis* clusters for each group as well as the total number of ESTs within a group is given. The number of genes present in the genomes of *P. falciparum*, *T. gondii* and *T. annulata* that are matching to a *B. bovis* cluster are indicated per group. The total numbers of ESTs matching to the *P. falciparum* or *T. gondii* genes within a group are listed. EST numbers and BLAST scores for individual genes are listed in supplementary Table S1.

prominent difference with *P. falciparum* and *T. gondii* however is presented by the high expression of two glycerol-3-phosphate (G-3-P) dehydrogenases. The simultaneous presence during the erythrocytic cycle of a mitochondrial inner membrane located FAD-dependent G-3-P dehydrogenase (27 ESTs, only 1 EST was found in both *P. falciparum* and *T. gondii*) and a cytosolic NAD-dependent G-3-P dehydrogenase (8 ESTs; 1 EST in *P. falciparum* and none in *T. gondii*) creates the possibility of a G-3-P shuttle. The shuttle allows controlling the cytosolic $NADH/NAD^+$ ratio (Rigoulet et al., 2004) by transferring reducing equivalents from cytosolic NADH to mitochondrial matrix $FADH_2$. NAD-dependent cytosolic G-3-P dehydrogenase reduces dehydroxy-acetonphosphate (DHAP) to G-3-P using NADH. G-3-P easily gains access to the inner mitochondrial membrane where FAD-dependent G-3-P dehydrogenase converts G-3-P back into DHAP by reducing FAD^+ to $FADH_2$.

The mitochondrial genome is actively transcribed as exemplified by ESTs for cytochrome *b* and cytochrome *c* oxidase subunits 1 and 3. An enormous

number of ESTs is derived from a short non-coding region of the mitochondrial genome (cluster 3823; 618 ESTs thus representing 5% of all ESTs). The cluster maps to a short region of the mitochondrial genome in between the genes for cytochrome *b* and cytochrome *c* oxidase subunit 3 and is homologous to domain IV of the 28S large subunit rRNA. Remarkably, this rRNA is encoded by a discontinuous gene which is scattered over both strands of the mitochondrial genome (Genbank Accession Nr. 2981472) like in *P. falciparum* (Feagin, 1992) and *T. parva* (Kairo et al., 1994). Short transcripts of these regions were observed in these species. ESTs all start in a 10 bp region and terminate at a stretch of seven A's which are encoded by the *B. bovis* mitochondrial genome and probably provided the annealing site for oligo-dT during cDNA library first-strand synthesis.

B. bovis harbours a plastid genome (Lang-Unnasch et al., 1998). The apicoplast seems, as expected, to be a functional organelle by the presence of 109 clusters (331 ESTs in *B. bovis* versus 118 ESTs in *P. falciparum*, see Table S1) with similarity to *P. falciparum* proteins predicted to be targeted to the apicoplast (Foth et al., 2003). The group contains essential apicoplast proteins like gyrase subunits A and B (clusters 2710 and 2397) and a putative DNA polymerase (POM1, cluster 3671). The required presence of a predicted signal peptide at the putative N-terminus was confirmed for several abundantly expressed *B. bovis* proteins lending additional support for their potential apicoplast localization. These include a triose phosphate/phosphate translocator (clusters 1876 and 2374, 7 ESTs; no ESTs are found in *P. falciparum* and *T. gondii*), an *S*-adenosyl-methyltransferase (cluster 1800 and 1822, 13 ESTs; no ESTs in *P. falciparum* and *T. gondii*) with high similarity to the bacterial mraW family and located in the division cell wall cluster of *E. coli* (Carrion et al., 1999), and a chloroplast-like chaperone (HSP20, cluster 1621, 10 ESTs; 2 ESTs in *P. falciparum* and none in *T. gondii*). The most abundant signal peptide containing protein (cluster 1213, 17 ESTs) that matches a predicted apicoplast hypothetical protein of *P. falciparum* (no ESTs) appears to be a homologue of the intriguing polymorphic immunodominant protein (Toye et al., 1995) of *Theileria* species showing similarity to copper transporters. Possibly this protein gains extracellular access via the apicoplast targeting pathway as proposed for other proteins. Other relatively highly expressed potential apicoplast proteins, still incomplete at the N-terminus, are homologues of chaperones HSP83 (cluster 1671; 12 ESTs) and DnaJ-2 (clusters 1581 and 1583; 12 ESTs) and pyruvate kinase (cluster 1729; 7 ESTs). Considering the potential functions of the apicoplast (Wilson, 2002) the detection of many ESTs related to isoprenoid biosynthesis is apparent (24 ESTs in 11 clusters representing eight genes). In *P. falciparum* and *T. gondii* no ESTs are obtained while the genes are present. This suggests a potentially more important role for this pathway in *B. bovis*. In contrast, no ESTs related to type II fatty acid synthesis were detected in *B. bovis*.

4.4. Purine salvage

ESTs encoding proteins involved in purine metabolism are very abundant in comparison to *P. falciparum* and *T. gondii* and include genes that were not previously described in apicomplexan parasites. Apicomplexan genes involved in nucleotide biosynthesis pathways are of diverse phylogenetic origin with differences in between genera and their occurrence is related to multiple endosymbiotic events (Striepen et al., 2004; Chaudhary et al., 2004). A summary of pathways shown to be expressed in *B. bovis* adds further to this diversity (Fig. 1). Purine phosphoribosyl transferases catalyzing the direct conversion of purine bases into monophosphate nucleotides were not detected, fitting with the observation that the genomes of the related piroplasms *T. annulata* and *T. parva* do not appear to encode these enzymes (Striepen et al., 2004). High expression of two different adenosine kinases indicates that adenosine is the main, and probably sole, precursor in purine nucleotide salvage in *B. bovis*. In this respect *B. bovis* is most similar to *C. parvum* which fully depends on adenosine salvage and most different from *P. falciparum* which does not appear to encode an adenosine kinase. *T. gondii* does encode for both adenosine kinase and hypoxanthine-guanine phosphoribosyl transferase. In contrast to the eukaryotic-type adenosine kinase of *C. parvum* and *T. gondii* the two *B. bovis* adenosine kinases are highly similar to bacterial ribokinases. Recently a *B. canis* homologue was cloned and demonstrated to be an efficient adenosine kinase (Carret et al., 1999).

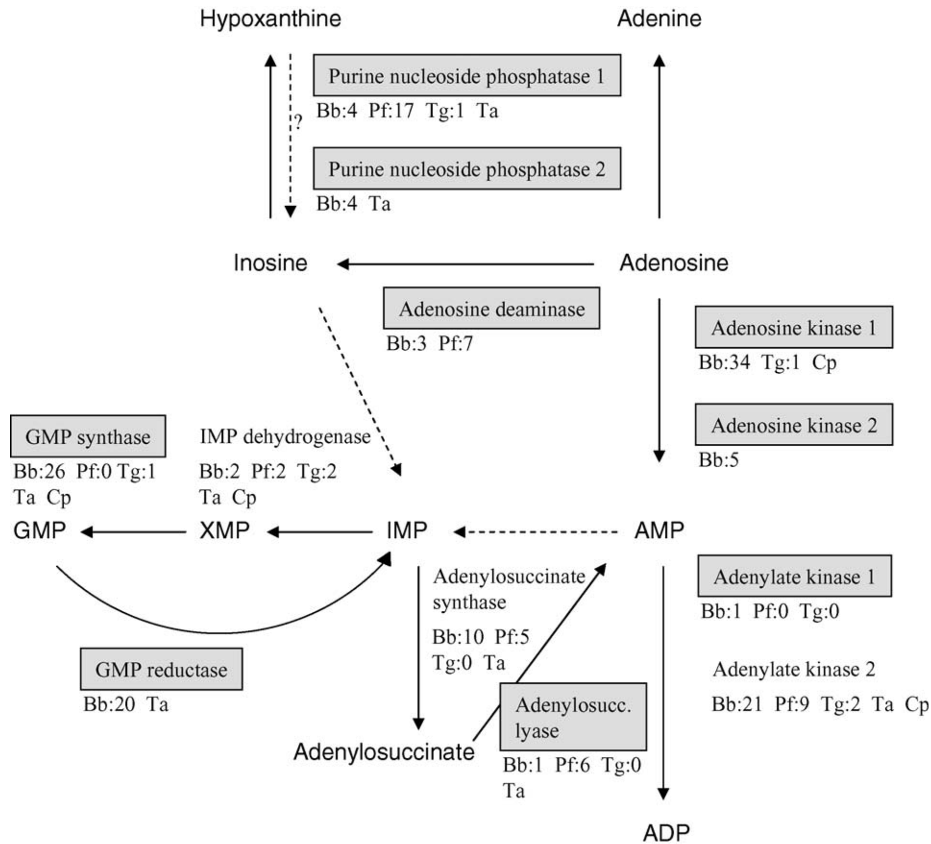

Fig. 1. Schematic representation of the putative purine salvage pathway of *B. bovis*. Enzymes indicated in a shaded box share high similarity with bacterial proteins whereas the others are more similar to eukaryotes. The presence of a homologous gene in *P. falciparum* (Pf), *T. gondii* (Tg), *T. annulata* (Ta) or *C. parvum* (Cp) is indicated as well as the numbers of ESTs present (not for Ta and Cp). *B. bovis* cluster numbers are as follows: purine nucleoside phosphatase 1 and 2, clusters 1979 and 2004; adenosine kinase 1 and 2, clusters 301 and 302; adenylate kinase 1 and 2, clusters 1177 and 1175; adenosine deaminase, cluster 1170; IMP dehydrogenase, clusters 3216 and 3680; GMP synthase, cluster 556; GMP reductase, cluster 84; adenylosuccinate synthase, clusters 1029, 1310 and 2098; adenylosuccinate lyase, cluster 3111. Sequences and BLAST results can be found in supplementary Table S1.

Whereas synthesis of IMP and subsequently GMP takes place via adenosine monophosphate deaminase in *C. parvum*, an adenosine deaminase encoding EST cluster was found in *B. bovis*. Possibly the IMP required for GMP synthesis is generated from inosine by one of the adenosine kinases of *B. bovis* for which clear homologues have not been identified in any other apicomplexan. *B. bovis* is also unique so far in encoding two different adenylate kinases, one of which is abundantly expressed. Proteins involved in IMP metabolism are highly expressed in *B. bovis* and the presence of a cluster of 20 ESTs encoding a bacterial-type GMP reductase creates an apparently futile IMP-GMP cycle. This is also found in the *T. annulata* genome but not in other apicomplexans. Its presence suggests a tight control of GMP and IMP levels (note that GMP can be converted to the signalling molecule cyclic-GMP by GMP cyclase which is encoded by five ESTs in *B. bovis*). Clearly, *B. bovis* purine salvage pathways differ from those of other apicomplexans, the bacterial-type nature and high expression of many of the enzymes involved being the most remarkable features. Phylogenetic analysis will be required to determine their exact origin. One question that needs to be addressed experimentally is the role of hypoxanthine. *Babesia* and *Theileria* species incorporate radioactively labelled hypoxanthine (Nott et al., 1990; Kamiyama

et al., 1992). These significant observations are difficult to interpret in the absence of purine phosphoribosyl transferases in the *B. bovis* EST set and in the *T. annulata* and *T. parva* genomes. Of interest is the finding of two different purine nucleoside phosphorylases which, in theory, can also catalyze the reverse reaction, e.g. from hypoxanthine to inosine, at sufficient ribose-phosphate levels.

4.5. Highly expressed and conserved genes rarely occurring in P. falciparum and T. gondii EST collections

High abundance EST clusters of *B. bovis* that are absent or poorly represented in the EST collections of *P. falciparum* and *T. gondii* may help to identify processes critically important to this parasite. Many examples can be found in Table S1 of which the more apparent, e.g. highly abundant and high BLAST score, are summarized here. A ubiquitously conserved NTPase (cluster 380, ESTs: Bb, 48; Pf, 1; Tg, 1) of the MRP family of ATPases. This family was suggested to be involved in control of cell division/ chromosome partitioning (Leipe et al., 2002). However, recent results indicate an essential function in maturation of Fe/S proteins in bacteria as well as eukaryotes for other members (Skovrar and Downs, 2003; Hausmann et al., 2005); a poly-A tail binding protein (cluster 908; Bb, 47 ESTs; Pf, 2 ESTs; Tg, 2 ESTs) that has been implied in mRNA stabilization, localization as well as in translational initiation or silencing; myosin A tail interacting protein (cluster 218; Bb, 40 ESTs; Pf, 1 EST; Tg, 7 ESTs), possibly involved in the molecular motor driving host cell invasion (Bergman et al., 2003); a P-type ATPase cation transporter (cluster 915; Bb, 27 ESTs; Pf, no ESTs; Tg, 2 ESTs) classified as an endoplasmic reticulum-type calcium pump in plants (Wu et al., 2002) and highly similar to PfATP4 of *P. falciparum* (Krishna et al., 2001); a cyclin classified as negative regulator of transcription of the transcriptional activator NUC-1, which controls the expression of phosphorous acquisition enzymes (Kang and Metzenberg, 1993) (cluster 900; Bb, 17 ESTs; Pf, no ESTs; Tg, 2 ESTs); secretory pathway component Sec31 which is part of the coat of COPII coated vesicles travelling from ER to golgi (cluster 1149; Bb, 13 ESTs; Pf, no ESTs; Tg, no ESTs).

5. Antigenic proteins of *B. bovis*

The EST data show that nearly all described apical complex, merozoite surface or erythrocyte surface antigens, which were usually isolated by virtue of their immuno-dominance, belong to the most abundantly expressed genes. High EST numbers result in clusters mostly encoding full-length proteins thus enabling a reliable sequence comparison of the American and Australian *B. bovis* antigenic proteins with those from the Israel strain which mainly has been studied by immunological tools without further sequence knowledge (Shkap et al., 1994).

5.1. Merozoite surface antigens

Glycosyl-phosphatidylinositol (GPI) anchored merozoite membrane proteins of 42 and 44 kDa are the dominant surface proteins and have collectively been called variable merozoite surface antigens (VMSA) that are encoded by a gene family (Cowman et al., 1984) comprising a single *msa-1* gene (Jasmer et al., 1992a,b) and four tandemly arranged genes denoted *msa-2a*$_1$, *msa-2a*$_2$, *msa-2b* and *msa-2c* (Florin-Christensen et al., 2002) in a Mexican *B. bovis* isolate (Mo7 strain). *Msa-2c* is most closely related to the *babr* gene family (38% amino acid identity to BabR0.8) of an Australian isolate. A comparison of Mexican and Argentinean isolates (Suarez et al., 2000) revealed a better conservation of the MSA-2 family members (69–89% identity) than MSA-1 (52%). All *msa* genes, including the three EST clusters described below, encode for an extremely conserved C-terminus that constitutes the signal for cleavage and GPI-anchor addition.

Three full-length, abundantly expressed, VMSA-like protein encoding clusters were identified from the EST assembly. Cluster 1007 (nine ESTs) is most similar to MSA-2a2 but markedly different from both the Mexican and Argentinean isolate with a highest identity of 65% to Mo7 MSA-2a2. Cluster 669 (22 ESTs) belongs to the MSA-2c/BabR group but has a remarkable mosaic structure. The first 70 amino acids are 74% identical to Mo7 MSA-2c; however, due to a frame shift this part is missing from the BabR proteins. This is followed by a 122 amino acid domain that is 92% identical to BabR1.2 but only 51% to MSA-2c and subsequently by a 51 amino acid domain that is

41% identical to MSA-2c and 23% identical to BabR1.2 and a 31 amino acid domain that is absent from MSA-2c but 55% identical to BabR1.2. In addition, preceding the GPI-anchor signal, a 71 amino acid, proline-rich C-terminal repeat region similar to repeat domains of several adhesive proteins, like the thrombospondin related anonymous protein of *P. berghei*, is unique to cluster 669. No clear MSA-1 ortholog was identified. Instead, cluster 670 (17 ESTs) takes a unique position in between the currently identified *vmsa* genes as the encoded 283 amino acid protein is 29% identical to MSA-1 as well as to MSA-2a2 and MSA-2b. Sequence data on the VMSA proteins expressed by a *B. bovis* strain isolated in Israel thus enforce the image of a rapidly evolving gene family with even larger differences in between parasites from different geographical origin as previously reported. No significant similarity between EST clusters and any of the *Plasmodium* merozoite surface proteins – or MSP – was detected.

5.2. Apical organelle antigens

Rhoptries and micronemes are apically located organelles of apicomplexan parasites secreting a complex mixture of proteins directly involved in host cell binding and invasion (Soldati et al., 2001). They have been identified by electron microscopy and cell fractionation in *B. bovis* as well (Dowling et al., 1996) but their contents are yet poorly characterized. Two micronemal proteins (apical membrane antigen 1 (AMA1) and thrombospondin related anonymous protein (TRAP)), both involved in invasion and conserved in *Plasmodium* species and *T. gondii*, have been described for *B. bovis* Israel strain (Gaffar et al., 2004a,b). No clusters similar to any of the other known *P. falciparum* micronemal proteins, e.g. the host cell receptor binding ligand families containing Duffy binding ligand (DBL) domains or *T. gondii* micronemal adhesions, e.g. carrying Apple or EGF domains, were detected (Tomley and Soldati, 2001). Most likely, *B. bovis* host cell invasion is mainly facilitated by non-homologous or poorly conserved proteins when compared with *P. falciparum*.

The single known *B. bovis* rhoptry associated protein, RAP-1 (Suarez et al., 1998) appears to be one of the most abundantly expressed proteins (cluster 390, 60 ESTs) and was shown to be involved in host cell invasion in vitro (Yokoyama et al., 2002; Mosqueda et al., 2002). It has been studied extensively as a potential vaccine component (Norimine et al., 2003; Brown et al., 1998), also because of the apparently extreme sequence conservation (99% identity between strains from Mexico, Brazil, Argentina and Uruguay). In contrast, cluster 390 is only 90% identical to RAP-1 of any of these strains, having implications for potential RAP-1 antigenic diversity in the field and the extent of variation still permitting the proper functioning of RAP-1.

Spherical bodies are large, heterogeneously structured, vesicles unique to the genus *Babesia* but may be functionally similar to the smaller dense granules of *Plasmodium* spp. and other apicomplexans. Three spherical body proteins have been described (SBP1–3) that were all shown to be secreted from spherical bodies and to localize to the cytoplasmic face of the erythrocyte membrane. A fourth protein (SBP4) has been deposited in GenBank. High antibody titres to SBP1–3 are measured in infected animals and some correlation to protection of disease has been reported (Goodger et al., 1992; Tetzlaff et al., 1992; Dalrymple et al., 1993; Ruef et al., 2000). Spherical body proteins are abundantly transcribed (440 ESTs) and include the two most abundant clusters. Extensive sequence variation and the presence of additional paralogs are detected in the EST database for SBP1 and 2.

SBP1, also called Bb1 (Mexico strain) or Bv80 (Australia strain), has been described as an 80 kDa protein (Tripp et al., 1989; Dalrymple et al., 1993; Hines et al., 1995) composed of highly conserved N- and C-terminal regions separated by a short repeat region displaying considerable variation in between strains. Two clusters (7 and 765; 37 ESTs) nearly cover the complete protein and display extreme sequence conservation to Mexican strain SBP1 (99% identity in terminal domains, 80% in the 120 amino acid repeat region) in contrast to 85% conservation to the terminal domains of the Australian strain. Two additional SBP1-like genes are present (cluster 248, six ESTs; cluster 249, two ESTs) of which only the N-terminal regions were obtained. These display 59% and 55% identity to Mexican strain SBP1.

SBP2 has been described as a 225 kDa protein of which only the N-terminal sequence has been determined from the Mexico Mo7 strain (Jasmer et al., 1992a,b; Dowling et al., 1996). Pieces of

homologous genes (named BvVA1) were cloned from an Australian isolate (Dalrymple et al., 1993) but no full-length cDNAs have been reported. Genomic analyses indicated that SBP2 is a member of a gene family with conserved N- and C-terminal domains and size variation within the central region. The terminal domains are separated by complex repeat regions created by multiple rearrangements (Dalrymple et al., 1993). Six EST clusters encoding members of this family were detected. Cluster 1319 (17 ESTs) encodes a protein of 521 amino acids still lacking a probably small part of the conserved N-terminus (43 aa in comparison to SBP2 and BvVA1). The N- and C-terminal domains are 88% to 90% identical to BvVA1 and the central repeat region displays complicated repeats in a slightly altered arrangement as in BvVA1. Cluster 1320 (four ESTs) encodes the central repeat region of a similar protein. The other four SBP-2 like clusters are more abundantly expressed and encode relatively small full-length proteins for which the repeat region is lacking. Of these, cluster 24 (150 ESTs) is the most abundant of all clusters encoding a 388 amino acid protein with a highly conserved N-terminus and a moderately conserved C-terminus in comparison to SBP2/BvVA1. Clusters 27 (23 ESTs), 28 (17 ESTs) and 29 (17 ESTs) are only well conserved at the N-terminus, especially the first 30 amino acids, showing 70 to 94% identity. SBP3 (cluster 383; 142 ESTs) and SBP4 (cluster 668; 41 ESTs) appear to be unique and highly conserved proteins that are abundantly expressed. SBP3 is 99% identical to the reported (Ruef et al., 2000) Mexican strain Mo7 sequence whereas SBP4 is 92% identical to the sequence reported in Genbank.

5.3. *Erythrocyte surface antigens*

As in the case of *P. falciparum*, infected erythrocytes sequester in the microvasculature by adherence to endothelial cells (Allred and Al-Khedery, 2006, this issue; Allred and Al-Khedery, 2004; O'Connor and Allred, 2000). A heterodimeric, parasite encoded, variable erythrocyte surface antigen (VESA1) involved in antigenic variation is thought to act as a cytoadherence ligand (Allred et al., 1993, 1994). The VESA1a subunit is encoded by genes (*ves1α*) that are part of a multigene family (*ves*) undergoing rapid variation, seemingly by a segmental gene conversion mechanism (Allred et al., 2000). Large scale EST sequencing provides an unbiased view on the expressed complement of mRNA of this family as it is independent of the potentially biased primer selection required for micro-array and PCR-based approaches. A total of 170 ESTs, grouped in 42 clusters, were found to match the *ves* gene family. The clonal line used here has been in continuous culture for several years prior to cDNA library construction. As a consequence, clonality may not have been maintained within many genes of the actively evolving *ves* gene family. Similarly, the ability of this line to vary has not been established, nor the mechanisms it may employ if it does vary. Nevertheless, the majority of ESTs assemble into three homogeneous clusters in which no variation, except for some occasional sequencing errors, occurs. One of these clusters (Cluster 845; 22 ESTs) is most similar (66% identity) to the C-terminal part of the VESA1a subunit (the N-terminal part is not covered). The other two abundant clusters more closely resemble a sequence annotated as *vesD* pseudo gene (accession AY279553) as well as five other, partially sequenced, *ves* pseudogenes. Cluster 957 (27 ESTs) encodes for a 123 kDa full-length protein. The C-terminal region (177 aa), containing the single predicted transmembrane region, is not encoded by the *vesD* pseudo gene but is highly similar to the VESA1a C-terminus. Possibly, cluster 957 encodes for the b subunit of the VESA1 heterodimer, which is size-polymorphic (Allred et al., 1994; O'Connor et al., 1997) and may be subject to antigenic variation as well. VESA1b was suggested to be related to VESA1a, but a genetic relationship has not yet been directly established (D.R. Allred, B. Al-Khedery, personal communication). Remarkably, cluster 3841 (37 ESTs) has the appearance of a transcribed pseudogene. It encodes a 37 kDa polypeptide, highly similar to the N-terminus of the polypeptides encoded by the *vesD* pseudogene and cluster 957, but is lacking a methionine for translation initiation. The premature stop codon at the 3′-end is followed by a poly-A tail. Cluster 71 (9 ESTs), containing an open reading frame terminated before the transmembrane region, may represent a transcribed pseudogene of the *ves1α* group to which it is highly similar. Of all other 38 clusters only two are most similar to the cluster 957/pseudogene group whereas all others belong to the *ves1α* group. These clusters were assembled from 1 to

10 ESTs and, as some of the shorter clusters match to different sections of *ves1α*-like genes without overlap, a number of them may derive from the same gene. In summary, dominantly transcribed versions of a *ves1α*-like gene and another *ves* gene (possibly encoding VESA1B) are present with the additional transcription, at lower levels, of at least 25 different *ves1α*-like sequences. Whether these represent transcription of multiple different alleles or modified forms of *ves1α* at the same allelic locus is unknown. Of both groups a pseudogene seems to be transcribed at high levels.

6. Discussion

ESTs provide a rapid and cost-efficient means of exploring the expressed gene complement during a specific stage of the life cycle of *B. bovis*. BLAST analysis has allowed assigning a putative function to many of the assembled gene clusters. By linking the clusters to EST collections of *P. falciparum* and *T. gondii* via their completely sequenced genomes, a first comparison of expression levels of homologous genes has been performed. This approach will provide new leads for further research into the biology of *B. bovis* and related apicomplexans. The identification of processes expressed at a comparatively high level may point at crucial events in the erythrocytic lifecycle of *B. bovis*. These may provide attractive targets for chemotherapeutic intervention, for example, several of the highly expressed bacterium-like enzymes involved in purine salvage. Also, increased knowledge on the sequence variability and expression of related family members of antigenic proteins is necessary for further studies on their potential use as vaccine components. It should be stressed however that all homology assignments are based on *in silico* sequence comparisons. Although, based on the acquired experiences with the BLAST algorithms, the great majority of assignments will eventually turn out to be correct, further experimental proof will be required to extend any individual case.

Acknowledgements

B. bovis clonal line C61411 was kindly provided by Dr. E. Pipano (Kimron Veterinary Institute, Israel). We thank Drs D.R. Allred and B. Al-Khedery for discussion and comments. *T. annulata* genomic sequence data were produced by the pathogen Sequencing Group at the Sanger Institute and can be obtained from ftp://ftp.sanger.ac.uk/pub/Pathogens/T_annulata/. Preliminary *T. gondii* genomic sequence data was accessed via http://ToxoDB.org and/or http://www.tigr.org/tdb/t_gondii/. Genomic data were provided by The Institute for Genomic Research (supported by the NIH grant #AI05093), and by the Sanger Center (Wellcome Trust). This study was supported in part by the Netherlands Foundation for Advancement of Tropical Research (NWO-WOTRO).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.vetpar.2006.01.040.

References

Abrahamsen, M.S., Templeton, T.J., Enomoto, S., et al., 2004. Complete genome sequence of the apicomplexan *Cryptosporidium parvum*. Science 304, 441–445.

Allred, D.R., Al-Khedery, B., 2004. Antigenic variation and cytoadhesion in *Babesia bovis* and *Plasmodium falciparum*: different logics achieve the same goal. Mol. Biochem. Parasitol. 134, 27–35.

Allred, D.R., Carlton, J.M.R., Satcher, R.L., Long, J.A., Brown, W.C., Patterson, P.A., 2000. The *ves* multigene family of *B. bovis* encodes components of rapid antigenic variation at the infected erythrocyte surface. Mol. Cell 5, 153–162.

Allred, D.R., Cinque, D.M., Lane, T.J., Ahrens, K.P., 1994. Antigenic variation of parasite-derived antigens on the surface of *Babesia bovis*-infected erythrocytes. Infect. Immun. 62, 91–98.

Allred, D.R., Hines, S.A., Ahrens, K.P., 1993. Isolate-specific parasite antigens of the *Babesia bovis*-infected erythrocyte surface. Mol. Biochem. Parasitol. 60, 121–132.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Bergman, L.W., Kaiser, K., Fujioka, H., Coppens, I., Daly, T.M., Fox, S., Matuschewski, K., Nussenzweig, V., Kappe, S.H., 2003. Myosin A tail-domain interacting protein (MTIP) localizes to the inner membrane complex of *Plasmodium* sporozoites. J. Cell Sci. 116, 39–49.

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

Brown, W.C., McElwain, T.F., Hotzel, 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.

Carret, C., Delbecq, S., Labesse, G., Carcy, B., Precigout, E., Moubri, K., Schetters, T.P., Gorenflot, A., 1999. Characterization and molecular cloning of an adenosine kinase from *Babesia canis rossi*. Eur. J. Biochem. 265, 1015–1021.

Carrion, M., Gomez, M.J., Merchante-Schubert, R., Dongarra, S., Ayala, J.A., 1999. mraW, an essential gene at the dcw cluster of *Escherichia coli* encodes for a cytoplasmic protein with methyltransferase activity. Biochimie 81, 879–888.

Carulli, J.P., Artinger, M., Swain, P.M., Root, C.D., Chee, L., Tulig, C., Guerin, J., Osborne, M., Stein, G., Lian, J., Lomedico, P.T., 1998. High throughput analysis of differential gene expression. J. Cell. Biochem. Suppl. 30, 286–296.

Chaudhary, K., Darling, J.A., Fohl, L.M., Sullivan Jr., W.J., Donald, R.G., Pfefferkorn, E.R., Ullman, B., Roos, D.S., 2004. Purine salvage pathways in the apicomplexan parasite *Toxoplasma gondii*. J. Biol. Chem. 279, 31221–31227.

Clark, I.A., Jacobson, L.S., 1998. Do babesiosis and malaria share a common disease process? Ann. Trop. Med. Parasitol. 92, 483–488.

Cowman, A.F., Bernard, O., Stewart, N., Kemp, D.J., 1984. Genes of the protozoan parasite *Babesia bovis* that rearrange to produce RNA species with different sequences. Cell 37, 653–660.

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

Dalrymple, B.P., Peters, J.M., Goodger, B.V., Bushell, G.R., Waltisbuhl, D.J., Wright, I.G., 1993. Cloning and characterisation of cDNA clones encoding two *Babesia bovis* proteins with homologous amino- and carboxy-terminal domains. Mol. Biochem. Parasitol. 59, 181–190.

Dowling, S.C., Perryman, L.E., Jasmer, D.P., 1996. A *Babesia bovis* 225 kDa spherical body protein: localization to the cytoplasmic face of infected erythrocytes after merozoite invasion. Infect. Immun. 64, 2618–2626.

Escalante, A.A., Ayala, F.J., 1994. Phylogeny of the malarial genus *Plasmodium*, derived from rRNA gene sequences. Proc. Natl. Acad. Sci. U.S.A. 91, 11373–11377.

Feagin, J.E., 1992. The 6 kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. Mol. Biochem. Parasitol. 52, 145–148.

Florin-Christensen, M., Suarez, C.E., Hines, S.A., Palmer, G.H., Brown, W.C., McElwain, T.F., 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.

Foth, B.J., Ralph, S.A., Tonkin, C.J., Struck, N.S., Fraunholz, M., Roos, D.S., Cowman, A.F., McFadden, G.I., 2003. Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. Science 299, 705–708.

Franssen, F.F., Gaffar, F.R., Yatsuda, A.P., De Vries, E., 2003. Characterization of erythrocyte invasion by *Babesia bovis* merozoites efficiently released from their host cell after high voltage pulsing. Microb. Infect. 23, 641–653.

Gaffar, F.R., Yatsuda, A.P., Franssen, F.F., 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 (AMA-1). Infect. Immun. 72, 2947–2955.

Gaffar, F.R., Yatsuda, A.P., Franssen, F.F., 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.

Gaffar, F.R., Wilschut, K., Franssen, F.F., De Vries, E., 2004c. An amino acid substitution in *Babesia bovis* dihydrofolate reductase-thymidilate synthase is correlated to cross-resistance agains pyrimethamine and WR99210. Mol. Biochem. Parasitol. 133, 209–219.

Gardner, M.J., Hall, N., Fung, E., et al., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419, 498–511.

Goodger, B.V., Waltisbuhl, D.J., Wright, I.G., White, M., 1992. *Babesia bovis*: analysis of and preliminary vaccination studies with a defined infected erythrocyte membrane binding antigen. Int. J. Parasitol. 22, 533–535.

Hausmann, A., Netz, D.A., Balk, J., Pierik, A.J., Mühlenhoff, U., Lill, R., 2005. The eukaryotic P loop NTPase Nbp35: an essential component of the cytosolic and nuclear iron–sulfur protein assembly machinery. Proc. Natl. Acad. Sci. U.S.A. 102, 3266–3271.

Hines, S.A., Palmer, G.H., Brown, W.C., McElwain, T.F., Suarez, C.E., Vidotto, O., Rice-Ficht, A.C., 1995. Genetic and antigenic characterization of *Babesia bovis* merozoite spherical body protein Bb-1. Mol. Biochem. Parasitol. 69, 149–159.

Jasmer, D.P., Reduker, D.W., Hines, S.A., Perryman, L.E., McGuire, T.C., 1992a. Surface epitope localization and gene structure of a *Babesia bovis* 44-kilodalton variable merozoite surface antigen. Mol. Biochem. Parasitol. 55, 75–83.

Jasmer, D.P., Reduker, D.W., Perryman, L.E., McGuire, T.C., 1992b. 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. 52, 263–270.

Kairo, A., Fairlamb, A.H., Gobright, E., Nene, V., 1994. A 7.1 kb linear DNA molecule of *Theileria parva* has scrambled rDNA sequences and open reading frames for mitochondrially encoded proteins. EMBO J. 13, 898–905.

Kamiyama, T., Niinuma, S., Ikeda, I., Matsubara, J., 1992. Incorporation of [3H]hypoxanthine by short-term cultured *Theileria sergenti* and its inhibition by drugs. J. Parasitol. 78, 856–860.

Kang, S., Metzenberg, R.L., 1993. Insertional mutagenesis in *Neurospora crassa*: cloning and molecular analysis of the preg(+) gene controlling the activity of the transcriptional activator NUC-1. Genetics 133, 193–202.

Krishna, S., Woodrow, C., Webb, R., Penny, J., Takeyasu, K., Kimura, M., East, J.M., 2001. Expression and functional characterization of a *Plasmodium falciparum* Ca2+-ATPase (PfATP4) belonging to a subclass unique to apicomplexan organisms. J. Biol. Chem. 276, 10782–10787.

Krungkai, J., 2004. The multiple roles of the mitochondrion of the malarial parasite. Parasitology 129, 511–524.

Lang-Unnasch, N., Reith, M.E., Munholland, J., Barta, J.R., 1998. Plastids are widespread and ancient in parasites of the phylum apicomplexa. Int. J. Parasitol. 28, 1743–1754.

Leipe, D.D., Wolf, Y.I., Koonin, E.V., Aravind, L., 2002. Classification and evolution of P-loop GTPases and related ATPases. J. Mol. Biol. 317, 41–72.

Li, L., Brunk, B.P., Kissinger, J.C., Pape, D., Tang, K., Cole, R.H., Martin, J., Wylie, T., Dante, M., Fogarty, S.J., Howe, D.K., Liberator, P., Diaz, C., Anderson, J., White, M., Jerome, M.E., Johnson, E.A., Radke, J.A., Stoeckert Jr., C.J., Waterston, R.H., Clifton, S.W., Roos, D.S., Sibley, L.D., 2003. Gene discovery in the apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. Genome Res. 13, 443–454.

Morrison, D.A., Ellis, J.T., 1997. Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of apicomplexa. Mol. Biol. Evol. 14, 428–481.

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.

Mosqueda, J., Falcon, A., Antonio Alvarez, J., Alberto Ramos, J., Oropeza-Hernandez, L.F., Figueroa, J.V., 2004. *Babesia bigemina* sexual stages are induced in vitro and are specifically recognized by antibodies in the midgut of infected *Boophilus microplus* ticks. Int. J. Parasitol. 34, 1229–1236.

Munoz, E.T., Bogarad, L.D., Deem, M.W., 2004. Microarray and EST database estimates of mRNA expression levels differ: the protein length versus expression curve for *C. elegans*. BMC Genomics 5, 30–38.

Nott, M., Bagnara, A.S., O'Sullivan, W.J., Gero, A.M., 1990. Purine salvage and metabolism in *Babesia bovis*. Parasitol. Res. 76, 207–213.

Norimine, J., Mosqueda, J., Suarez, C., Palmer, G.H., McElwain, T.F., Mbassa, G., Brown, W.C., 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.

O'Connor, R.M., Lane, T.J., Stroup, S.E., Allred, D.R., 1997. Characterization of a variant erythrocyte surface antigen (VESA1) expressed by *Babesia bovis* during antigenic variation. Mol. Biochem. Parasitol. 89, 259–270.

O'Connor, R.M., Allred, D.R., 2000. Selection of *Babesia bovis*-infected erythrocytes for adhesion to endothelial cells co-selects for altered variant erythrocyte surface antigen isoforms. J. Immunol. 164, 2037–2045.

O'Connor, R.M., Long, J.A., Allred, D.R., 1999. Cytoadherence of *Babesia bovis* infected erythrocytes to bovine capillary endothelial cells provides an in vitro model for sequestration. Infect. Immun. 67, 3921–3928.

Rigoulet, M., Aguilaniu, H., Averet, N., Bunoust, O., Camougrand, N., Grandier-Vazeille, X., Larrson, C., Pahlman, I.L., Manon, S., Gustafsson, L., 2004. Organization and regulation of the cytosolic NADH metabolism in the yeast *Saccharomyces cerevisiae*. Mol. Cell Biochem. 256, 73–81.

Ruef, B.J., Dowling, S.C., Conley, P.G., Perryman, L.E., Brown, W.C., Jasmer, D.P., Rice-Ficht, A.C., 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.

Shkap, V., Pipano, E., McElwain, T.F., Herzberg, U., Krigel, Y., Fish, L., Palmer, G.H., 1994. Cross-protective immunity induced by *Babesia bovis* clones with antigenically unrelated variable merozoite surface antigens. Vet. Immunol. Immunopathol. 41, 367–374.

Skovrar, E., Downs, D.M., 2003. Lack of the ApbC or ApbE protein results in a defect in Fe-S cluster metabolism in *Salmonella enterica* Serovar typhimurium. J. Bacteriol. 185, 98–106.

Soldati, D., Dubremetz, J.F., Lebrun, M., 2001. Microneme proteins: structural and functional requirements to promote adhesion and invasion by the apicomplexan parasite *Toxoplasma gondii*. Int. J. Parasitol. 31, 1293–1302.

Striepen, B., Pruijssers, A.J.P., Huang, J., Li, C., Gubbels, M.J., Umejiego, N.N., Hedstrom, L., Kissinger, J.C., 2004. Gene transfer in the evolution of parasite nucleotide biosynthesis. Proc. Natl. Acad. Sci. U.S.A. 101, 3154–3159.

Suarez, C.E., Florin-Christensen, M., Hines, S.A., Palmer, G.H., Brown, W.C., McElwain, T.F., 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, C.E., Palmer, G.H., Hotzel, I., McElwain, T.F., 1998. Structure, sequence, and transcriptional analysis of the *Babesia bovis* rap-1 multigene locus. Mol. Biochem. Parasitol. 93, 215–224.

Tetzlaff, C.L., Rice-Ficht, A.C., Woods, V.M., Brown, W.C., 1992. Induction of proliferative responses of T cells from *Babesia bovis*-immune cattle with a recombinant 77-Kilodalton merozoite protein (Bb-1). Infect. Immun. 60, 644–652.

Tomley, F.M., Soldati, D.S., 2001. Mix and match modules: structure and function of microneme proteins in apicomplexan parasites. Trends Parasitol. 17, 81–88.

Toye, P.G., Metzelaar, M.J., Wijngaard, P.L., Nene, V., Iams, K., Roose, J., Nyanjui, J.K., Gobright, E., Musoke, A.J., Clevers, H.C., 1995. Characterization of the gene encoding the polymorphic immunodominant molecule, a neutralizing antigen of *Theileria parva*. J. Immunol. 155, 1370–1381.

Tripp, C.A., Wagner, G.G., Rice-Ficht, A.C., 1989. *Babesia bovis*: gene isolation and characterization using a mung bean nuclease-derived expression library. Exp. Parasitol. 69, 211–225.

Wilson, R.J., 2002. Progress with parasite plastids. J. Mol. Biol. 319, 257–274.

Wu, Z., Liang, F., Hong, B., Young, J.C., Sussman, M.R., Harper, J.F., Sze, H., 2002. An endoplasmic reticulum-bound Ca(2+)/Mn(2+) pump, ECA1, supports plant growth and confers tolerance to Mn(2+) stress. Plant Physiol. 130, 128–137.

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.