Genome-wide analysis of gene expression and protein secretion of *Babesia canis* during virulent infection identifies potential pathogenicity factors

Ramon M. Eichenberger\(^1\), Chandra Ramakrishnan\(^1\), Giancarlo Russo\(^2\), Peter Deplazes\(^1\) & Adrian B. Hehl\(^1\)

Infections of dogs with virulent strains of *Babesia canis* are characterized by rapid onset and high mortality, comparable to complicated human malaria. As in other apicomplexan parasites, most *Babesia* virulence factors responsible for survival and pathogenicity are secreted to the host cell surface and beyond where they remodel and biochemically modify the infected cell interacting with host proteins in a very specific manner. Here, we investigated factors secreted by *B. canis* during acute infections in dogs and report on *in silico* predictions and experimental analysis of the parasite’s exportome. As a backdrop, we generated a fully annotated *B. canis* genome sequence of a virulent Hungarian field isolate (strain BcH-CHIPZ) underpinned by extensive genome-wide RNA-seq analysis. We find evidence for conserved factors in apicomplexan hemoparasites involved in immune-evasion (e.g. VESA-protein family), proteins secreted across the iRBC membrane into the host bloodstream (e.g. SA- and Bc28 protein families), potential moonlighting proteins (e.g. profilin and histones), and uncharacterized antigens present during acute crisis in dogs. The combined data provides a first predicted and partially validated set of potential virulence factors exported during fatal infections, which can be exploited for urgently needed innovative intervention strategies aimed at facilitating diagnosis and management of canine babesiosis.

Apicomplexan parasites have evolved a plethora of targeted interactions with their host organisms and cellular machinery to elicit modification of gene expression and cellular functions. These modifications are critical for invasion and for establishment of intracellular niches for parasite replication, differentiation, and persistence, by securing nutrient import and mediating immune evasion\(^1\). As a result, host-parasite relationships in intracellular Apicomplexa can be viewed as co-adaptations driven by co-evolution of parasites with their hosts\(^2\). Most importantly, this complex interplay is directly linked to pathogenicity and virulence which characterize symptomatic infections with these protozoan parasites\(^3–8\).

Three genera of this phylum represent a triad of phylogenetically related hemoproteozoa: *Plasmodium* (human and animal malaria), *Theileria* (theileriosis in livestock), and *Babesia* (severe disease and death in cattle, horses, dogs, with some zoonotic potential). They are the causative agents of important human and/or animal diseases with high morbidity and mortality, responsible for significant economic burdens and healthcare challenges worldwide\(^6\). The clinically relevant part of the complex, heteroxenous *Babesia* life cycle takes place during asexual proliferation of parasites in red blood cells (RBC) of the mammalian intermediate host. This represents a distinct specialization of *Babesia* for growth in RBC, i.e. cells which lack organelles, as well as endocytic and exocytic membrane transport machinery\(^9–11\).

Canine babesiosis is an emerging tick-borne disease of dogs with global distribution caused by morphologically distinctive *Babesia* species\(^12–14\). The geographical distribution of the disease caused by *B. canis*, the species responsible for a virulent form of canine babesiosis in Europe and Western Asia, appears to extend, as
demonstrated by an increasing number of cases in areas outside of the established endemic regions. The clinical spectrum of *B. canis* babesiosis is broad, ranging from apparently silent, mild disease to fulminant and often fatal presentations. Indeed, in many respects canine babesiosis is highly comparable to clinical human malaria.

Analysis of correlated genomic, transcriptomic and proteomic data lays out the molecular underpinning for investigating the role of exported factors in disease manifestation and progression. However, experimental investigation of *B. canis* biology is limited by the current lack of cultivation techniques for propagation of this parasite in vitro. Here, we implemented short term in vitro cultivation of infected RBC in combination with whole genome sequencing, annotation, and comparative genomic characterization of a virulent Hungarian *B. canis* field isolate. By correlating RNA-seq and protein mass spectrometry data, we provide the first account of factors exported by *B. canis* blood stages with potential links to host-pathogen interaction and acute virulence.

**Results**

**Genome sequence of a virulent *Babesia canis* strain.** Until recently, only a few genes from canine *Babesia* species have been characterized and global genomic data were not available. To provide the molecular underpinning for targeted approaches to diagnosis, vaccination and clinical management of canine babesiosis we undertook complete sequencing, assembly, annotation and characterization of the genome of a virulent *B. canis* strain, designated BCh-CHIPZ. The BCh-CHIPZ isolate originated from the blood of a Swiss dog that was infected, presumably through the bite of an infected tick, during a trip to Hungary. The isolate was characterized by genotyping and identified as double B-type by restriction fragment length polymorphism (RFLP) analysis in the Bc28-multigene family and the 18S rRNA gene. We implemented a next-generation high throughput approach (PacBio® SMRT) for sequencing BCh-CHIPZ genomic DNA prepared after enrichment of infected RBCs (iRBCs). Assembly of the parasite-specific sequence reads after elimination of all reads from host DNA and gap-filling yielded 43 high-quality scaffolds at 55-fold average coverage representing a total of 7,002,180 bases with a median weighted contig size (N50) of 185 kb. Hence, the size of the *B. canis* genome lies between the phylogenetically related *B. microti* (6.5 Mb) and the *B. bovis* (8.2 Mb) genomes. The main features of the annotated *B. canis* genome and a comparison with other apicomplexan genomes are presented in Table 1. The quality of this assembly, based on the presence of core eukaryotic genes, indicates the *B. canis* genome sequence to be 92.6% complete as compared with the genome of the *T. gondii* ME49 type II reference strain and consistent with values of other *Babesia* genome sequences, which lie between 91.7% and 95.6% (Supplementary Document S1).

| Features                  | Species                  | *B. canis* [BCh-CHIPZ] | *B. bovis* [T2Bo] | *B. microti* [RI] | *Th. annulata* [Ankara] | *P. falciparum* [3D7] |
|---------------------------|--------------------------|------------------------|------------------|------------------|------------------------|-----------------------|
| Genome size (Mb)          | 7.0                      | 8.2                    | 6.5              | 8.4              | 23.3                   |
| # of chromosomes          | 4                        | 4                      | 4                | 4                | 14                     |
| # of scaffolds             | 43                       | 12                     | 3                | 7                | 14                     |
| G+C content (%)           | 45.8                     | 41.5                   | 36               | 32.5             | 19.4                   |
| Protein coding genes      | 3467                     | 3706                   | 3513             | 4082             | 5383                   |
| Mean gene size (bp)       | 1044                     | 1503                   | 1327             | 1602             | 2292                   |
| % of coding regions       | 51.7                     | 68                     | 73               | 73               | 53                     |
| Gene density              | 2020                     | 2194                   | 1816             | 2199             | 4374                   |

Table 1. Genome characteristics of related apicomplexan hemoprotozoa (more information is provided in Supplementary Document S2). 1Excluding introns and UTRs; 2Genome size/number of protein coding genes.

A robust annotation of species-specific gene models was based on the analysis of orthologues as well as on synteny in related species (described below). Furthermore, whole genome mapping approach allowed contiguous alignment of the *B. canis* scaffolds with the 4 *B. bovis* chromosomes. This indicated a high degree of synteny with *B. bovis* chromosomes.

**Genome-wide comparison of *B. canis* gene models with those of other Apicomplexa reveals species-specific genes.** The *B. canis* gene models were grouped into orthologues in four representative apicomplexan parasites *B. bovis*, *Th. annulata*, *P. falciparum*, and *T. gondii* and the fully sequenced *Babesia* species *B. bovis*, *B. bigemina*, and *B. microti* using a Markov Cluster algorithm in OrthoMCL (Fig. 2A,B). The four *Babesia* species analysed here share a core set of 1216 pan-*Babesia* conserved orthologous groups (COG) (Supplementary Data S1). However, bovine and rodent *Babesia* genomes also share 1030 COGs which are not represented in the *B. canis* genome (Supplementary Data S2). The former and the pan-*Babesia* COGs together contain a high percentage of clusters with hypothetical proteins (46.9% and 38.1%, respectively), which is not surprising given the overall high proportions of proteins annotated as “hypothetical” (61.8% in *B. bigemina*, 49.9% in *B. bovis*, and 50.2% in *B. canis*). However, there was no specific enrichment with respect to protein domains and functions between the two groups. Functional classification of the 95 clusters unique to *B. canis* highlights potential species-specific innovations comprised of proteins annotated as “hypothetical” proteins (55.8%), surface proteins (proteins with VESA1-domains and merozoite surface antigen; 26.4%), a number of specified secreted proteins (9%), and various other proteins (8.6%) (Supplementary data S3). Analysis of the bovine *Babesia* shared
orthologues revealed hypothetical proteins in 76.2% of the clusters analysed and predicted membrane proteins in 8.2% of the gene models. Prima facie, we interpret the diverged set of annotated genes in the different Babesia species as a reflection of the narrow and highly specific host range and a high degree of specialization for their respective vertebrate- and arthropod niches. This is based on the assumption that the diversity of encoded proteins is inversely correlated with the degree of specialization for different environments, niches, and lifestyles of a parasite, thus reflecting lineage-specific adaptations. Nevertheless, with 347 genes/Mb, the B. canis genome has a similar gene density as the B. bovis (372 genes/Mb) and the B. microti (354 genes/Mb) genomes, albeit some regions appear to have higher density than others (Fig. 1).

Despite significant species-specific features, analysis of orthologues revealed a considerable degree of synteny between the genomes of B. canis and B. bovis compared with the genome of the rodent B. microti (Fig. 1). This is also reflected in the phylogenetic analysis based on a core set of 975 shared genes in representative apicomplexan parasites (Fig. 2B,C).

**Prediction of the B. canis exportome and analysis of mRNA expression levels.** Exported proteins of apicomplexan parasites include virulence factors, i.e. parasite-specific proteins which interfere directly with host cell functions. As mature RBC lack de novo protein synthesis or trafficking pathways, membrane transport machinery for targeting these proteins via the cytoplasm to the iRBC surface and beyond are provided by the parasite. These parasite-induced cytoplasmic alterations of iRBC are critical for its development and directly

---

**Figure 1.** Genome characteristics of Babesia species. Assembled B. canis contigs (top arc) were arranged based on synteny to related Babesia species. The predicted arrangement of B. canis contigs on four chromosomes is indicated by colours of the lines linking conserved genes. Blood stage mRNA expression data for B. canis genes is included (line plot); gene density is colour coded and represented as number of genes per 10 kb (heatmap). Chromosomal distribution of B. canis VESA gene models is indicated with red dots.
linked to the severity of babesiosis. As a first step to characterize parasite-secreted factors and surface proteins, we generated a predicted B. canis exportome based on the annotated BcH-CHIPZ genome. We used data mining tools to search for genes coding for predicted parasite surface- or iRBC-targeted factors. The criteria and prioritization for this selection were based on the presence of a canonical hydrophobic N-terminal signal peptide (cSP) sequence, an alternative non-classical secretion pathway (a/ncSP) signal, predicted transmembrane (TM) domain(s), a GPI anchor signal, or previously described domains of conserved apicomplexan secretory proteins and protein families (Supplementary Fig. S1; Supplementary Data S4). By this approach we generated a curated and parsed list of 509 genes coding for proteins that are predicted to be exported to the parasite surface and/or to the host, which corresponds to 14.6% of all B. canis gene models. This proportion is very similar to the 528 gene models (14%) representing the predicted B. bovis secretome.

All Apicomplexa express members of gene families that are exported to the surface of invasive stages or the plasma membrane of iRBCs where they interact with the host immune system. Members of the Babesia variant erythrocyte surface antigen (VESA) family on the surface of iRBCs comprise the largest protein family in B. canis (103 gene models) in the predicted exportome (Fig. 3A). Most VESA genes cluster together on few contigs, unlike other gene families of the predicted exportome which are distributed more randomly (Fig. 1). Furthermore, VESA genes are in regions of the genome with high gene density, suggesting frequent gene duplication events consistent with a driving selection pressure on these proteins associated with immune-evasion. B. canis merozoite surface antigens (MSA; 2 gene models in B. canis) as well as Bc28 family members (20 gene models) are highly abundant on the merozoite surface but are also shed and interact with the iRBC membrane and the host’s immune system. Currently, two members of the Bc28 gene family are characterized as major merozoite surface antigens playing a critical function in the interaction of merozoites with RBCs. The so-called “secreted antigens” (SA-1 and SA-3, 24 members) comprise yet another family of genes coding for exported proteins. Some SA genes have been identified in other Babesia species. The pBLAST analysis showed that the B. canis SAs are highly homologous (21 of 24 proteins with e-values < 10E-5) to those of B. gibsoni, which also propagates in dogs, and to a lesser extent to SAs of B. bigemina (3 of 24 proteins with e-values < 10E-5), which infects bovines. This supports the idea that host-specific factors drive expansion and diversification of these orthologous gene families.
this gene family appears to be expanded in Babesia species infecting canines. In contrast, genes coding for smORF and spherical body proteins (SBP) appear to be specific to B. bovis and are without orthologues in the B. canis genome\textsuperscript{27, 39}. Accordingly, these protein families are also only represented in the bovine specific COGs. Proteins of the "secreted antigen family", as the name indicates, are secreted beyond the iRBC and circulate in the blood of infected dogs but are otherwise not well characterized\textsuperscript{31}. The predicted B. canis exportome also includes a group of 23 protein kinases. Secreted protein kinases and pseudokinases are known virulence factors in toxoplasmosis\textsuperscript{40–43}. Kinases are also secreted into iRBCs in Plasmodium, where they are responsible inter alia for remodelling the host cell\textsuperscript{44, 45}. However, functional data related to secreted kinases in Babesia are not available. Nevertheless, kinase inhibitors have been identified as potential novel drugs for the treatment of babesiosis in cattle\textsuperscript{46, 47}.

To refine the predicted host-targeted B. canis exportome dataset we analysed the transcript levels of all 509 selected gene models in parasites harvested from a splenectomised dog. Evidence from a genome-wide RNAseq analysis showed transcription of 71.5% of the 509 exportome gene models (FPKM > 10; n = 364) (Fig. 3B). Remarkably, only a few members each of the five most abundant exported protein family genes (VESA-, DnaJ-, SA-, PK-, and Bc28- multigene families) with potential roles in antigenic variation and immune evasion were expressed at high levels (FPKM > 1000). The particular expression profile of the five gene families in this snapshot might be due to selection of a few specific phenotypes in the population at the peak of pathogenicity, but it should also be considered that the collected blood stage parasites do not represent a synchronously developing population. Nevertheless, this finding merits further investigation, in particular in the light of previous observations indicating that pathogenicity was associated with expression of a specific subset of VESA- and other exported genes in B. bovis infection\textsuperscript{26}.

Figure 3. Exported protein families in Babesia canis. (A) In silico predicted exported protein families. Bar colours indicate the predicted route of secretion and/or type of membrane anchoring. cSP: canonical signal peptide; aSP: alternative (non-classical) secretory pathway; TM: transmembrane domain; GPI: predicted glycoprophosphatidylinositol anchor. (B) Transcript abundances of exported proteins. Each dot corresponds to the transcript abundance (FPKM) for each gene in the family during acute crisis in a splenectomised dog. Red bars indicate the median FPKM values in the presented protein family. FPKM: Fragments Per Kilobase of exon model per Million mapped reads.
Further in silico characterization of the predicted B. canis exportome revealed a significant number of proteins (210/509) lacking a canonical hydrophobic signal sequence for co-translational insertion into the lumen of the endoplasmic reticulum, which is suggestive of non-classical secretion. This export pathway is generally known as ‘leaderless protein secretion’ and was described recently also in P. falciparum. We screened the B. canis exportome for additional motifs associated with protein trafficking into the host cell and beyond. First, we searched for matches to Plasmodium PEXEL export elements (RxLxE/Q/D) for trafficking beyond the parasitophorous vacuolar membrane (PVM), which have been also identified in T. gondii and C. parvum. Although Babesia builds a PVM during invasion of RBC it disappears within minutes and the parasite resides directly in the host cell cytoplasm. A comprehensive comparative in silico analysis of protein export across 10 Apicomplexa species revealed PEXEL-like motifs (PLM) in B. bovis (but not in T. parva) with a role in retention of these proteins in so called spherical bodies and release in a cell-cycle dependent manner. We found hits for the B. bovis PEXEL-like motifs within the first 100 N-terminal amino acids in 64.3% of cSP-containing proteins and only in 7.6% of predicted alternatively secreted B. canis proteins (Supplementary Data S4). Using a hidden Markov model to identify alternative motifs by de novo pattern discovery did not yield any additional hits.

Expression levels of stage-specific genes identifies asexual and sexual Babesia blood stage parasites at acute crisis. We collected biological material from three experimentally infected dogs, which all showed very similar rapid disease progression consistent with fatal babesiosis (described in ref. 57). The first mild clinical signs (e.g. lethargy) preceded acute crisis with early symptoms of septic shock by a maximum of 27 hours (Supplementary Fig. S3). In spite of the fulminant disease manifestation, a hallmark of canine babesiosis is a remarkably low parasitemia with typically <2% infected erythrocytes observed in the peripheral blood. The asexual erythrocytic stage of the parasite is entirely responsible for the morbidity and mortality associated with canine babesiosis. Morphological characterization of blood stage parasites revealed pleomorphic forms in Giemsa-stained blood smears derived from a critically ill dog (Fig. 4A). Small and large ring stages were the most frequently detected forms, while only few pear-shaped forms could be identified. To characterize the parasite stages in iRBCs from experimentally infected animals on a molecular level in more detail, we used the high-throughput RNA-seq data to identify orthologous gene models with stage-specific expression in Plasmodium spp. We intersected the list of B. canis gene models with evidence for expression with a list of validated Plasmodium spp. stage-specifically expressed genes (e-value threshold of 10E-3) (Supplementary Data S5). We found evidence for RNA-seq reads (threshold FPKM > 10) mapping to 22 of 29 conserved merozoite, and 4 of 7 early gametocyte genes in a blood sample collected from a splenectomised dog (Fig. 4B). The higher mRNA abundance of merozoite- (proliferative stage) compared with early gametocyte-specific orthologues (sexual stage) during severe disease was consistent with the microscopy data and suggested only marginal numbers of circulating sexual stages in this peripheral blood sample. Nevertheless, some evidence for early gametocyte stages was found, which could indicate that acute clinical crisis in infected dogs might trigger development of sexual stages. For example in Plasmodium, gametocyte formation occurs in bone marrow and aggravates clinical symptoms and disease outcome. A potential bone marrow tropism could explain many typical clinical signs in canine babesiosis, e.g. leukopenia and thrombocytopenia, associated with poor outcome in acute infections. B. canis orthologues of the 18 identified invertebrate stage-specific proteins (flagellar “late” gametocyte-, ookinete- and sporozoite genes) described for Plasmodium had FPKM values between 0 and 10.8, indicating that no or only insignificant amounts of these mRNAs were present in the sampled parasite population. Based on the degree of
orthologue conservation and the microscopy data we interpret this as paucity of late gametocyte stages in the peripheral blood (Supplementary Data S5).

**Shotgun proteomics and transcriptomics identify potential *B. canis*-specific secreted virulence factors during severe disease.** Proteins involved in parasite-host interactions at acute crisis were identified by a comparative shotgun approach from *B. canis* infected and healthy host blood samples. Mass spectrometry datasets were generated from secreted soluble factors collected from pooled short-term cultures from three experimentally infected dogs at acute crisis and RBC cultures derived from healthy blood donors, as well as from RBC membrane fractions from experimentally infected animals and non-infected controls. We identified a set of abundant parasite-specific proteins in these fractions after filtering the data with the non-infected control sets. Based on their abundance in the iRBC and/or culture supernatant the identified secreted membrane-bound and soluble parasite proteins were considered candidate factors involved in host-pathogen interaction (Table 2; Supplementary Fig. S2).

Not surprisingly, considering the limited understanding of apicomplexan pathogenicity determinants, the resulting *B. canis* dataset of exported proteins comprises 47.8% (11/23) proteins annotated as "hypothetical". In addition, the following hits to functionally annotated proteins were identified: factors on cell surface membranes (from the parasite or the iRBC) which include 2 variant erythrocyte surface antigens (VESAs) located on the surface of iRBC, a *Babesia* spp. membrane protein (homologous to *B. bovis*, *B. microti*, *B. bigemina*, *Theileria annulata*, *T. equi*, *T. orientalis*, *T. parva*, and *Cyttauxzoon felis*), a spectrin repeat superfamily extracellular matrix binding protein with homology to *B. bovis* VESA, and 2 merozoite surface antigens of the Bc28 gene family. These proteins where shown to be involved in survival and virulence of several hemoproteozoa species. The dataset comprises SA1 and SA3, belonging to the above described third-largest exported family of proteins (secreted antigen family) suggesting a direct role in host-parasite interaction. The *B. canis* dataset also contains profilin, a secreted virulence factor described in *P. falciparum* and *T. gondii* with important roles in the process of crossing biological barriers during host cell invasion and egress, a SET-domain containing protein with putative function in lateral gene transfer in apicomplexans, and a SNARE protein mediating membrane fusion of secretory vesicles with the plasma membrane. Furthermore, a secreted histone protein (classified as DNA-binding protein) was detected. Histone proteins were described to have moonlighting functions outside of the nucleus in *P. falciparum*, specifically at the interface between parasite and erythrocyte cytoplasm, i.e. the parasitophorous vacuole membrane.

All hits in the proteomic dataset for which GO descriptions or blast homologies could be assigned were further validated using transcriptome evidence (threshold FPKM value > 10). Only Bc28.2, a known secreted *B. canis*-specific protein with a measured FPKM of 4.5, fell below this threshold. However, the correlation between transcript- and protein abundance is not linear and there is a complex relationship leading to notable discrepancies due to a range of factors such as protein- and mRNA turnover. Taken together with additional evidence from the literature, these proteins are good candidates for factors that are presented to the host on the parasite or iRBC surface, or released in to the bloodstream and potentially involved in parasite persistence and virulence.

**Discussion**

Secreted soluble and membrane-targeted parasite proteins have important roles at the host-parasite interface and are frequently virulence and pathogenesis factors. Systematic investigation of these interactions in infections with potentially fatal blood-borne parasites, such as *Babesia canis*, which modify the host cell, its microenvironment, as well as host physiology in multiple ways, is a challenge. Moreover, these questions cannot be addressed with biological material from regular patients from an animal clinic and the parasite is difficult to cultivate. Hence, in this study, we used experimental infections of dogs in a controlled setting to gather data on disease progression, clinical parameters, and parasite development. We report the first genome and gene expression data of the api- 
occlusion

74. Following RBC invasion, *B. bovis* exports VESA proteins to the surface of the host cell modifying the iRBC membrane to induce cytoadherence of iRBC at capillary and post-capillary venous endothelia. VESA expression is subject to rapid antigenic variation defined as the serial clonal replacement of the expressed VESA gene in daughter merozoites before invasion of a new RBC. This gives rise to antigenically distinct parasite populations, which complicates mounting of an effective host immune response. Antigenic variation, i.e. exclusive sequential expression of surface antigen family members is likely a sub-phyllum–specific immune-evasion strategy in hematozoa documented in several *Babesia* and *Plasmodium* species including *B. rodhaini*, *B. microti*, *B. bigemina*, *P. falciparum* and *P. vivax*. Not surprisingly, the largest exported protein family in *B. canis* consists of the variant erythrocyte surface antigen (VESAs) proteins. Indeed, a recent study using comparative transcriptomics and proteomics of attenuated and virulent *B. bovis* strains showed a significantly increased diversity of upregulated VESA genes in virulent strains. In the present study, we detected a high number (74/103) of transcribed VESA genes as well as their products in the *B. canis* secreted proteome. Hence, adherence of *B. canis* iRBC to endothelial cells is one explanation for the apparent discrepancy
| Identification       | GO-description\(^1\) | SP | TM | GPI | AS | Apicomplexan orthologue\(^2\) | Transcriptome evidence (FPKM) |
|----------------------|-----------------------|----|----|-----|----|-----------------------------|-------------------------------|
| **Specific B. canis secreted soluble proteins:** | | | | | | | |
| BcH000240            | Hypothetical protein, conserved | 1 | 0 | 0 | 0 | no (0.7) | |
| BcH000493            | Hypothetical protein [VESA] | 0 | 0 | 0 | 1 | VESA (B. bovis BBOV1004140) | yes (673.5) |
| BcH001113            | Hypothetical protein, conserved | 0 | 0 | 0 | 0 | yes (544) | |
| BcH002928            | Hypothetical protein, conserved | 1 | 0 | 0 | 0 | yes (64.2) | |
| BcH003445            | Secreted antigen 3 | 0 | 0 | 0 | 1 | Secreted antigen 3 (B. gibsoni BAH70325) | yes (26.4) |
| BcH003000            | Merozoite surface protein (Bc 28.2) | 1 | 0 | 0 | 0 | B. canis specific | no (4.5) |
| BcH002050            | SET domain-containing protein | 0 | 0 | 0 | 0 | SET- and MYND domain containing protein, putative (B. bigemina BBBOND0302680) | yes (147.6) |
| BcH002986            | hypothetical protein | 0 | 0 | 0 | 0 | yes (60.1) | |
| BcH003048            | Profilin | 0 | 0 | 0 | 0 | Polypeptide of profiling (B. bigemina BBBOND_0104160) | yes (100.6) |
| BcH002141            | Snare protein | 0 | 1 | 0 | 0 | Snare protein, putative (B. bigemina BBBOND0107960) | yes (236.6) |
| BcH000974            | Hypothetical protein, conserved | 1 | 0 | 0 | 0 | VESA (B. bovis BBOV1006920) | no (4.1) |
| BcH000821            | Spectrin repeat superfamily extracellular matrix binding | 0 | 1 | 0 | 0 | VESA (B. bovis BBOV1004140) | yes (123.2) |
| BcH000175            | Hypothetical protein, conserved | 0 | 0 | 0 | 0 | yes (57.9) | |
| BcH002958            | Hypothetical protein (membrane protein) | 0 | 0 | 0 | 1 | Membrane protein, putative (B. bigemina BBBOND_0105580) | yes (56.4) |
| **Specific B. canis host-RBC membrane presented proteins:** | | | | | | | |
| BcH000821            | Spectrin repeat superfamily extracellular matrix binding | 0 | 1 | 0 | 0 | VESA (B. bovis BBOV1004140) | yes (123.2) |
| BcH000300            | Merozoite surface protein (Bc 28.2) | 1 | 0 | 0 | 0 | B. canis specific | no (4.5) |
| BcH000299            | Merozoite surface protein (Bc 28.1) | 1 | 1 | 1 | 0 | B. canis specific | yes (18.3) |
| BcH002699            | Histone h2a | 0 | 0 | 0 | 0 | 41-2-protein antigen precursor, putative (B. bigemina BBBOND_0103700) | yes (31.9) |
| BcH001932            | Hypothetical protein, conserved | 0 | 0 | 0 | 0 | yes (12.5) | |
| BcH002268            | Hypothetical protein, conserved | 0 | 0 | 0 | 0 | yes (111.9) | |
| BcH001655            | VESA | 0 | 0 | 0 | 1 | VESA (B. bovis BBOV14002850) | yes (884.1) |
| BcH001448            | Hypothetical protein, conserved | 0 | 0 | 0 | 0 | no (0.9) | |
| BcH003376            | Secreted antigen 1 | 0 | 1 | 1 | 1 | Secreted antigen 1 (B. gibsoni ALE14558) | yes (142.5) |
| BcH000391            | Hypothetical protein, conserved | 0 | 0 | 0 | 1 | yes (331.2) | |
| BcH000895            | Hypothetical protein, conserved | 0 | 0 | 0 | 0 | yes (26.4) | |

Table 2. List *Babesia canis* specific candidates from culture supernatant derived and membrane-bound fractions. GO: gene ontology; SP: signal peptide; TM: transmembrane domain; GPI: AS: alternative secretion: GPI: glycosphatidylinositol; VESA: variant erythrocyte surface antigen. \(^1\)Automated annotation by Blast2GO; \([\text{if no Blast2GO annotation could be assigned, best pBlast hit was provided in brackets}]. \(^2\)Other than hypothetical.
between the low parasitemia observed in peripheral blood samples and the severity of the disease, consistent with sequestration of parasites and removal from the peripheral circulation. Correspondingly, using parasites isolated from a peripheral blood sample for the transcriptomic and proteomic studies might introduce a bias with respect to the detection of important pathogenicity factors. Direct comparison of proteomic and transcriptomic data has to be interpreted with caution, since the RNA-seq data was generated from parasites isolated from a splenectomized animal. The peripheral blood of splenectomized individuals would present a dominance of a sub-population of mature asexual and immature sexual parasites in iRBCs that failed to sequester, because they are not expressing appropriate ligands for cytoadherence, as repeatedly demonstrated in human \( P. falciparum\)\(^9,86 \).

Modulation of the adhesive properties of iRBCs by the spleen has been discussed, although no underlying mechanism has been identified\(^81,82 \). For example, in splenectomized rhesus monkeys infected with \( P. falciparum\) infections, the concentration of EVs in the blood is positively correlated with severe disease suggesting an important role in pathogenesis and immune modulation\(^83 \). Strikingly, the \( B. canis\) BcH-CHIP genome encodes homologues of important structural proteins and enzymes frequently found in exosomes, including heat shock proteins (e.g. HSP70 and 90), tetraspanins and ALIX protein, TSG101/ESCRT protein, 14-3-3 protein, thioredoxin peroxidase, histones, and the RAB GTPases 5, 7, 11, as well as a RAB GDI\(^85,88,89 \). Due to the mechanism for their biogenesis, exosomes often contain endosome-associated proteins, such as the SNARE proteins also detected in the \( B. canis\) culture supernatant. In \( P. falciparum\), a SNARE protein was localized proximal to the plasma membrane where it is possibly involved in membrane trafficking events associated with the parasite’s food vacuole\(^84 \). Hence, EVs in \( B. canis\) infections could be crucial vehicle carriers for the dissemination of pathogenicity factors that are linked to acute and severe disease, including coding and non-coding RNA which can interfere with host cell gene expression. In fact, secretion via EVs could account for many of the non-classically secreted proteins detected in the \( B. canis\) proteomic datasets.

We identified several proteins with moonlighting functions which have been identified as virulence factors in related Apicomplexa. For example profilin (pf) was shown to have important roles in the process of crossing biological barriers during host cell invasion and egress, in addition to its classical role as a regulator of actin dynamics\(^85 \). The structure of the \( P. falciparum\) pfh accommodates additional domains with as yet unknown functions, which are potentially involved in the parasite-host interplay\(^84 \). Another interesting finding was a histone protein in the iRBC membrane fraction. In several \( P. falciparum\) species, a histone protein methylated at position L9 was detected in the PV membrane\(^86 \). The conserved PVM localization in \( P. falciparum\) indicated that histones have evolved additional functions in invasion and interaction with the host. As \( Babesia\) spp. are not contained within a PV, the potential moonlighting function of this protein will have to be further investigated. Furthermore, histones are commonly found in EVs\(^80 \). We also detected a \( B. canis\) SET-domain containing protein (BcSET) in the secreted soluble fraction. SET proteins are involved in diverse mechanisms such as transcriptional regulation, enhancer function, mRNA splicing, DNA replication, and DNA damage response\(^86 \). An exported SET exerting an immune-suppressive function by transcriptional silencing of host gene expression was described in the pathogen \( B. anthraxis\)\(^82 \). Interestingly, deregulation of gene expression due to SET domain-containing proteins is correlated with unfavourable clinical outcomes in various forms of cancer\(^93 \).

Members of the Bc28 and SA gene families are among the most prominent \( B. canis\) secreted factors, and have recently gained attention as candidates for innovative vaccine strategies and in diagnostics\(^90,91 \). Bc28 gene products are highly represented on the merozoite surface, but are also shed and can interact with the iRBC membrane\(^80 \). Bc28 belong to a multigene family composed of polymorphic genes. Different \( B. canis\) strains in Europe are associated with different mortality rates and show genetic heterogeneity in Bc28.1, detectable by a PCR-RFLP test\(^92 \). It has been suggested that this variable repertoire has evolved to allow the parasite to evade host immune responses, although merozoite surface proteins of \( Babesia\) spp. are thought to be the main targets of the adaptive host immune response\(^93,94 \). However, whether all members are directly involved in immune evasion and virulence, or even show stage-specific behaviour, will have to be investigated in more detail.

In the absence of reverse genetic techniques for \( B. canis\), sequencing and annotation of a virulent \( B. canis\) strain together with the described parasite blood stage proteome and transcriptome provides a powerful underpinning for designing targeted intervention strategies such as the development of novel diagnostic tools, effective vaccines, and innovative drugs.

**Methods**

**Parasites.** Biological materials for the different experiments were collected from experimentally infected dogs. Animals were inoculated intravenously with \( 1 \times 10^9\) parasitized erythrocytes from a cryo-conserved field isolate, derived from a naturally infected Bernese mountain dog from Switzerland that had been exposed during a trip to Hungary. All experimentation with animals and protocols was in complete compliance with the strict Swiss animal welfare standards and regulations and approved by the responsible authorities (Veterinary Office of the Canton of Zurich; permission number 122/2012) previous to the study, including round the clock surveillance of the animals by a veterinarian. Whole blood was treated with citrate-phosphate-dextrose-adenine (12.3%
CPDA-1, v/v) for anticoagulation under sterile conditions and prepared immediately after euthanasia. Blood samples for genomic and transcriptomic studies were collected from an experimentally inoculated splenectomised dog (4 years old). Splenectomy increased the final amount of parasites in the peripheral blood to approximately 4% infected erythrocytes. The biological material for the transcriptome analysis was collected from three experimentally infected adult beagles with a parasitaemia of 1–1.75% (Supplementary Fig. S3). All samples were collected at the first signs of acute crisis detected by any clinical signs of acute shock or central nervous depression.

**Genome sequencing, assembly, and annotation pipeline.** Genomic DNA (gDNA) was prepared from purified and host leukocyte-depleted\(^{97}\) infected blood from an experimentally infected, splenectomised dog. High-throughput genome sequencing was performed on a PacBio Single Molecule Real Time (SMRT) sequencing platform\(^{98}\) run on 4 SMRT cells. 5 μg of input gDNA was used for library construction. Sequenced fragments were filtered to eliminate *Canis familiaris* reads and processed for assembly with PacBio read-alignment software BLASR\(^{97}\). The final draft of the genome was obtained using a two-step procedure. De novo assembly was initially performed using the HGAP2 algorithm\(^{98}\) and the outcome was further polished using the PacBio proprietary analysis package Quiver, a reference-guided consensus tool used as part of the SMRT analysis. The resulting draft was subjected to gap-closing and further scaffolding using PBJelly v14.1.15\(^{98}\). Genome annotation was performed using different sources of evidence-based prediction on the Maker2 pipeline\(^{100}\), on ab initio gene predictor algorithms\(^{100}\) with *P. falciparum* 3D7\(^{102}\), *B. bovis* T2Bo\(^{103}\), and *T. annulata* Ankara\(^{104}\) gene models as templates retrieved from EuPathDB release 26\(^{105}\), and on Trinity-assembled de novo transcriptome evidence\(^{106}\). Repetitive genomic elements were identified and masked from annotation with RepeatMasker using the full Repbase database\(^{106}\). Redundant transcripts potentially representing sequencing errors or genetic polymorphisms were clustered with the cd-hit software\(^{107}\). Non-overlapping sequences of the different evidence datasets were generated from the gff-files with BEDTools genome arithmetic software\(^{108}\). Functional annotation of the gene-predictions was achieved by gene ontology mapping through Blast2Go v3.1\(^{109}\) using an upper cut-off e-value of < 0.0005.

Full methodological details are provided in the Supplementary Material.

**RNA preparation and RNA-seq data analysis.** Total RNA was prepared from host leukocyte-depleted *B. canis* blood stages using the Zymo Direct-zol RNA MiniPrep kit (Zymo Research) including an on-column DNase I (QIAGEN) digest. Paired-end, stranded sequencing of the cDNA library was performed on 1 lane of an Illumina HiSeq 2000 sequencer (Illumina Inc.). Measurement of expression level was based on fragments per kilobase of exon model per million mapped reads (FPKM), normalized by the length of the gene. A gene was considered expressed if its FPKM was at least 10. Full methodological details are provided in the Supplementary Material.

**Summary of computational genome analysis.** The identification of potential homologues of *B. canis* genes to selected apicomplexan parasites (*B. bovis* T2Bo\(^{102}\), *B. bigemina* BOND\(^{103}\), *B. microti* RI 20, *T. annulata* Ankara\(^{104}\), *P. falciparum* 3D7\(^{102}\), and *T. gondii* ME49\(^{105}\)) was carried out by protein BLAST (BLAST + executables v2.2.31) of each individual protein sequence. A e-value < 0.0001 cut-off was considered appropriate. For comparative genomics analysis of species-specific genes and clustering of orthologues, OrthoMCL (v2.0.9) was used\(^{110,111}\). Synteny analysis was performed in MCScanX\(^{112}\) by the pairwise determination of syntenic regions between species based on the order of orthologues from the OrthoMCL output with default values. *B. canis* contigs were arranged according to the syntenic relationship to *B. bovis* with ‘orderchr’ execution in Circos utilities. The graphical representation was made with Circos\(^{113}\). Whole genome phylogeny was based on the 975 identified OrthoMCL-orthologues shared among Apicomplexan species. Each of these gene models in 4 *Babesia* sp. (*B. canis*, *B. bovis*, *B. bigemina*, and *B. microti*), 2 *Theileria* sp. (*T. annulata* and *T. equi*), 2 *Plasmodium* sp. (*P. falciparum* and *P. vivax*) and *Toxoplasma gondii* was aligned using MAFFT v7\(^{114}\). Highly variable sites were trimmed using trimal\(^{115}\). A supermatrix of the alignments was concatenated using FASconCAT\(^{116}\). Based on this alignment a maximum likelihood phylogenetic tree was constructed using RAxML\(^{117,118}\) based on amino acid input, gamma model, LG substitution matrix and 100 initial bootstrap (bootstrap percentage support is shown along the branching nodes), whereas the tree was rooted using *T. gondii* as an outgroup. Prediction of the *B. canis* exportome was based on several in silico analysis tools, databases, and manual curation (Supplementary Fig. S1). Full methodological details are provided in the Supplementary Material.

**Proteomic studies.** Proteomic analyses were based on secreted soluble *Babesia* antigens and membrane-bound *B. canis* proteins. Culture supernatant was collected after 18 h of short term in vitro parasite cultures of *B. canis* RBC in serum-free medium prepared from anticoagulated whole blood from experimentally inoculated dogs. Similarly, blood from a healthy canine blood donor was collected and supernatant was processed as a control. Isolation of membrane bound proteins was based on anticoagulated, leucocyte-depleted blood. iRBC were concentrated by Percoll (GE Healthcare) gradient centrifugation\(^{119}\) and parasites were released from the iRBC by streptolysin-O mediated lysis\(^{120}\) and removed by an additional Percoll gradient centrifugation step\(^{121}\). Membrane proteins were collected by homogenization and repeated ultracentrifugation steps, mild sonication and acetone precipitation. The same protocol was used to collect proteins from uninfected erythrocyte membranes. All samples were collected and stored at −80 °C until further use.

Protein preparations were separated on one dimensional SDS-PAGE under sterile and reducing conditions. Each stained gel was cut into 8 equal sections and in-gel trypsin digestion was carried out for subsequent mass spectrometry analysis. Liquid chromatography-MS/MS was performed on a Q-exactive mass spectrometer (Thermo Scientific) equipped with a nanoAcuity UPLC (Waters Corporation).

Following peptide data acquisition, searches were performed using the MASCOT search program against a database built on the *B. canis* annotated genes with a concatenated decoy database supplemented with commonly
observed contaminants and the Swissprot database to increase database size, or against NCBI database for Canis lupus familiaris. Identified hits were filtered in Scaffold Viewer V4 (Proteome Software, Portland, US) based on stringent parameters to reach false discovery rates of <5%. Furthermore, all datasets were corrected for obvious host- and environmental contaminants by an additional pBLAST search of the NCBI non-redundant protein sequences database with the single peptide hits.

**Data Availability.** The annotated B. canis genome will be made accessible via PiroplasmaDB hosted by EuPathDB (http://piroplasmadb.org/piro/).

**References**

1. Goldberg, D. E. & Cowman, A. F. Moving in and renovating: exporting proteins from Plasmodium into host erythrocytes. *Nat Rev Microbiol* **8**, 617–621 (2010).
2. Morrison, D. A. Evolution of the Apicomplexa: where are we now? *Trends Parasitol* **25**, 375–382 (2009).
3. Bradley, R. J. & Sibley, L. D. Rhoopies: an arsenal of secreted virulence factors. *Curr Opin Microbiol* **10**, 582–587 (2007).
4. Haldar, K. & Mohandas, N. Erythrocyte remodeling by malaria parasites. *Curr Opin Hematol* **14**, 203–209 (2007).
5. Hutchings, C. L. et al. New insights into the altered adhesive and mechanical properties of red blood cells parasitized by Babesia bovis. *Mol Microbiol* **65**, 1092–1105 (2007).
6. Maier, A. G. et al. Exported proteins required for virulence and rigidity of Plasmodium falciaparum-infected human erythrocytes. *Cell* **134**, 48–61 (2008).
7. Lobo, C. A., Rodriguez, M. & Cursino-Santos, J. R. Babesia and red cell invasion. *Curr Opin Hematol* **19**, 170–175 (2012).
8. Singh, V., Gupta, P. & Pande, V. Revisiting the multigene families: *Plasmodium var* and vir genes. *J Vector Born Dis* **51**, 75–81 (2014).
9. Lau, A. O. An overview of the Babesia, Plasmodium and Theileria genomes: a comparative perspective. *Mol Biochem Parasitol* **164**, 1–8 (2009).
10. Gohil, S., Kats, L. M., Sturm, A. & Cooke, B. M. Recent insights into alteration of red blood cells by Babesia bovis: moovin’ forward. *Trends Parasitol* **26**, 591–599 (2010).
11. Spielmann, T. & Gilberger, T. W. Critical Steps in Protein Export of Plasmodium falciaparum Blood Stages. *Trends Parasitol* **31**, 514–525 (2015).
12. Schetters, T. P. et al. Different Babesia canis isolates, different diseases. *Parasitology* **115**, 485–493 (1997).
13. Boozer, A. L. & Macintire, D. K. Canine babesiosis. *Vet Clin North Am Small Anim Pract* **33**, 885–904, viii (2003).
14. Irwin, P. J. Canine babesiosis: from molecular taxonomy to control. *Parasit Vectors* **2**, S4 (2009).
15. Halos, L. et al. Questionnaire-based survey on the distribution and incidence of canine babesiosis in countries of Western Europe. *Parasite* **21**, 13 (2014).
16. Reyes, F. et al. Canine babesiosis in South Africa: more than one disease. Does this serve as a model for falciparum malaria? *Ann Trop Med Parasitol* **92**, 503–511 (1998).
17. Schetters, T. P. & Eling, W. M. Can Babesia infections be used as a model for cerebral malaria? *Parasitol Today* **15**, 492–497 (1999).
18. Carcy, B. et al. Classification of Babesia canis strains in Europe based on polymorphism of the Bc28.1-gene from the Babesia canis Bc28 multigene family. *Vet Parasitol* **211**, 111–123 (2015).
19. Brayton, K. A. et al. Genome sequence of Babesia bovis and comparative analysis of apicomplexan hemoproteozoa. *PLoS Pathog* **3**, 1401–1413 (2007).
20. Cornillot, E. et al. Sequencing of the smallest Apicomplexan genome from the human pathogen *Babesia microti*. *Nucleic Acids Res* **40**, 9102–9114 (2012).
21. Parra, G., Bradnam, K. & Korf, I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**, 1061–1067 (2007).
22. Reid, A. J. et al. Genomic analysis of the causative agents of coccidiosis in domestic chickens. *Genome Res* **24**, 1676–1685 (2014).
23. Gajria, B. et al. ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic Acids Res* **36**, D553–556 (2008).
24. Templeton, T. J. et al. The genome of *Eimeria falciformis*: reduction and specialization in a single host apicomplexan parasite. *BMC Genomics* **15**, 696 (2014).
25. Heitlinger, E., Spork, S., Lucius, R. & Dieterich, C. The genome of *Eimeria falciformis*-reduction and specialization in a single host apicomplexan parasite. *BMC Genomics* **15**, 696 (2014).
26. Pedroni, M. I., Sondgeroth, K. S., Gallego-Lopez, G. M., Echaide, I. & Lau, A. O. Comparative transcriptome analysis of geographically distinct virulent and attenuated *Babesia bovis* strains reveals similar gene expression changes through attenuation. *BMC Genomics* **14**, 763 (2013).
27. Pelle, K. G. et al. Shared elements of host-targeting pathways among apicomplexan parasites of differing lifestyles. *Cell Microbiol* **17**, 1618–1639 (2015).
28. Spence, P. J. et al. Vector transmission regulates immune control of *Plasmodium* virulence. *Nature* **498**, 228–231 (2013).
29. Allred, D. R., Hines, S. A. & Ahrens, K. P. Isolate-specific parasitic antigens of the *Babesia bovis*-infected erythrocyte surface. *Mol Biochem Parasitol* **60**, 121–132 (1993).
30. Yang, Y. S. et al. Structural and functional characterization of Bc28.1, major erythrocyte-binding protein from *Babesia canis* merozoite surface. *J Biol Chem* **287**, 9495–9508 (2012).
31. Zhou, M. et al. Molecular identification and antigenic characterization of a merozoite surface antigen and a secreted antigen of *Babesia canis* (BcMSA1 and BcSA1). *Parasit Vectors* **9**, 257 (2016).
32. Carcy, B., Precigout, E., Schetters, T. & Gorenflot, A. Genetic basis for GPI-anchor merozoite surface antigen polymorphism of Babesia and resulting antigenic diversity. *Vet Parasitol* **138**, 33–49 (2006).
33. Rug, M. & Maier, A. G. The heat shock protein 40 family of the malaria parasite *Plasmodium falciaparum*. *IUBMB Life* **63**, 1081–1086 (2011).
34. Njunje, I. et al. Babesia gibsoni. *Clin Vaccine Immunol* **16**, 944–948 (2009).
35. Jia, H. et al. Identification of a novel gene encoding a secreted antigen 1 of *Babesia gibsoni* and evaluation of its use in serodiagnosis. *Ann Trop Med Hyg* **75**, 843–850 (2006).
36. Luo, Y. et al. Identification and characterization of a novel secreted antigen 1 of *Babesia microti* and evaluation of its potential use in enzyme-linked immunosorbent assay and immunochromatographic test. *Parasitol Int* **60**, 119–125 (2011).
37. Sevinc, F., Cao, S., Xuan, X., Sevinc, M. & Ceylan, O. Identification and expression of *Babesia ovis* secreted antigen 1 and evaluation of its diagnostic potential in an enzyme-linked immunosorbent assay. *J Clin Microbiol* **73**, 1531–1536 (2015).
38. Jackson, A. P. et al. The evolutionary dynamics of variant antigen genes in Babesia reveal a history of genomic innovation underlying host-parasite interaction. *Nucleic Acids Res* **42**, 7113–7131 (2014).
39. Saeij, J. P. et al. *Toxoplasma* co-opts host gene expression by injection of a polymorphic kinase homologue. *Nature* **445**, 324–327 (2007).
41. Steinfeldt, T. et al. Phosphorylation of mouse immunity-related GTPase (IRG) resistance proteins is an evasion strategy for virulent Toxoplasma gondii. PLoS Biol 8, e1000576 (2010).
42. Fleckenstein, M. C. et al. A Toxoplasma gondii pseudokinase inhibits host IRG resistance proteins. PLoS Biol 10, e1001358 (2012).
43. Etheridge, R. D. et al. The Toxoplasma gondii pseudokinase ROP5 forms complexes with ROP18 and ROP17 kinases that synergize to control acute virulence in mice. Cell Host Microbe 15, 537–550 (2014).
44. Saei, J. P. et al. Polymorphic secreted kinases are key virulence factors in toxoplasmosis. Science 314, 1780–1783 (2006).
45. Lim, D. C., Cooke, B. M., Doering, C. & Saei, J. P. Toxoplasma and Plasmodium protein kinase roles: in invasion and host cell remodelling. Int J Parasitol 42, 21–32 (2012).
46. Bork, S., Das, S., Okubo, K., Yokoyama, N. & Igarashi, I. Effects of protein kinase inhibitors on the in vitro growth of Babesia bovis. Parasitology 132, 775–779 (2006).
47. Pedroni, M. et al. Bumped kinase inhibitor prohibits egression in Babesia bovis. Vet Parasitol 215, 22–28 (2016).
48. Nickel, W. The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. Eur J Biochem 270, 2109–2119 (2003).
49. Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G. & Brunak, S. Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel 17, 349–356 (2004).
50. Lingebach, K. & Przyborski, J. M. The long and winding road: protein trafficking mechanisms in the Plasmodium falciparum infected erythrocyte. Mol Biochem Parasitol 147, 1–8 (2006).
51. Thavayogarajah, T. et al. Alternative Protein Secretion in the Malaria Parasite Plasmodium falciparum. PLoS One 10, e0125191 (2015).
52. Sudarsan, R., Chopra, R. K., Khan, M. A. & Sharma, S. Ribosomal protein P2 localizes to the parasite zoite-surface and is a target for invasion inhibitory antibodies in Toxoplasma gondii and Plasmodium falciparum. Parasitol Int 64, 43–49 (2015).
53. Marti, M., Good, R. T., Rug, M., Knuepfer, E. & Cowman, A. F. Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science 306, 1930–1933 (2004).
54. Hiller, N. L. et al. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science 306, 1934–1937 (2004).
55. Hsiao, C. H., Luisa Hiller, N., Haldar, K. & Knoll, L. J. A HT/PEXEL motif in Toxoplasma dense granule proteins is a signal for protein cleavage but not export into the host cell. Traffic 14, 519–531 (2013).
56. Asada, M. et al. Gliding motility of Babesia bovis merozoites visualized by time-lapse video microscopy. PLoS One 7, e35227 (2012).
57. Eichenberger, R. M., Riond, B., Hofmann-Lehmann, R. & Deplazes, P. Prognostic markers in acute Babesia canis infections. Vet Parasitol 170, 168–175 (2010).
58. Schetters, T. P. et al. Systemic inflammatory response in dogs experimentally infected with Babesia canis; a haematological study. Vet Parasitol 162, 7–15 (2009).
59. Furlanello, T., Fierro, F., Caldin, M., Labas, G. & Solano-Gallego, L. Clinicopathological findings in naturally occurring cases of babesiosis caused by large form Babesia from dogs of northeastern Italy. Vet Parasitol 134, 77–85 (2005).
60. Joice, R. et al. Plasmodium falciparum transmission stages accumulate in the human bone marrow. Sci Transl Med 6, 244re245 (2014).
61. Aguilar, R. et al. Severity of anaemia is associated with bone marrow haemozoin in children exposed to Plasmodium falciparum. Br J Haematol 164, 877–887 (2014).
62. Foley, M., Corcoran, L., Tilley, L. & Anders, R. Plasmodium falciparum: mapping the membrane-binding domain in the ring-infected erythrocyte surface antigen. Exp Parasitol 79, 340–350 (1994).
63. Lax, A. O. et al. Attenuation of virulence in an apicomplexan hemoparasite results in reduced genome diversity at the population level. BMC Genomics 12, 410 (2011).
64. Kursula, I. et al. Structural basis for parasite-specific functions of the divergent profilin of Plasmodium falciparum. Structure 16, 1638–1648 (2008).
65. Plattner, F. et al. Toxoplasma profilin is essential for host cell invasion and TLR11-dependent induction of an interleukin-12 response. Cell Host Microbe 3, 77–87 (2008).
66. Kishore, S. P., Metha, B. & Deitsch, K. W. Horizontal gene transfer of epigenetic machinery and evolution of parasitism in the malaria parasite Plasmodium falciparum and other apicomplexans. BMC Evol Biol 13, 37 (2013).
67. Ayong, L., Pagnotti, G., Tobon, A. B. & Chalraborti, D. Identification of Plasmodium falciparum family of SNAReS. Mol Biochem Parasitol 152, 113–122 (2007).
68. Parish, L. A. & Rayner, J. C. Plasmodium falciparum secretory pathway: characterization of PfStx1, a plasma membrane Qa-SNARE. Mol Biochem Parasitol 164, 153–156 (2009).
69. Luah, Y. H., Chaal, B. K., Ong, E. Z. & Bozdech, Z. A moonlighting function of Babesia bovis encodes components of rapid antigenic variation at the infected erythrocyte surface. PLoS One 5, e102530 (2010).
70. Wastling, J. M., Armstrong, S. D., Krishna, R. & Xia, D. Parasites, proteomes and systems: has Descartes’ clock run out of time? Parasitology 139, 1103–1118 (2012).
71. Udomsangpetch, R. et al. Involvement of cytokines in the histopathology of cerebral malaria. Am J Trop Med Hyg 57, 501–506 (1997).
72. Clark, I. A., Budd, A. C., Alleya, L. M. & Cowden, W. B. Human malarial disease: a consequence of inflammatory cytokine release. Malar J 5, 85 (2006).
73. Matijakko, V., Torti, M. & Schetters, T. P. Canine babesiosis in Europe: how many diseases? Trends Parasitol 28, 99–105 (2012).
74. Allred, D. R. & Al-Khedery, B. Antigenic variation and cytoadhesion in Babesia bovis and Plasmodium falciparum: different logic achieves the same goal. Mol Biochem Parasitol 134, 27–35 (2004).
75. Allred, D. R., Exline, R. M., Lane, T. J. & Ahrens, K. P. Antigenic variation of parasite-derived antigens on the surface of Babesia bovis-infected erythrocytes. Infect Immun 62, 91–98 (1994).
76. O’Connor, R. M., Lane, T. J., Stroup, S. E. & Allred, D. R. Characterization of a variant erythrocyte surface antigen (VESA1) expressed by Babesia bovis during antigenic variation. Mol Biochem Parasitol 89, 259–270 (1997).
77. Allred, D. R. et al. The ves multigene family of B. bovis encodes components of rapid antigenic variation at the infected erythrocyte surface. Mol Cell 5, 153–162 (2000).
78. Allred, D. R. & Al-Khedery, B. Antigenic variation as an exploitable weakness of babesial parasites. Vet Parasitol 138, 50–60 (2006).
79. Saul, A. The role of variant surface antigens on malaria-infected red blood cells. Parasitol Today 15, 455–457 (1999).
80. Bachmann, A. et al. Absence of erythrocyte sequestration and lack of multicopy gene family expression in Plasmodium falciparum from a spleenectomized malaria patient. PLoS One 4, e7459 (2009).
81. Ho, M., Bannister, L. H., Looareesuwan, S. & Suntharasamai, P. Cytoadherence and ultrastructure of Plasmodium falciparum-infected erythrocytes from a spleenectomized patient. Infect Immun 60, 2225–2228 (1992).
82. Pongpornratn, E. et al. Absence of knobs on parasitized red blood cells in a spleenectomized patient in fatal falciparum malaria. Southeast Asian J Trop Med Public Health 31, 829–835 (2000).
83. Lapp, S. A. et al. Spleen-dependent regulation of antigenic variation in malaria parasites: Plasmodium knowlesi SICAvar expression profiles in splenic and asplenic hosts. PLoS One 8, e78014 (2013).
84. Al-Khedery, B. & Allred, D. R. Antigenic variation in Babesia bovis occurs through segmental gene conversion of the ves multigene family, within a bidirectional locus of active transcription. Mol Microbiol 59, 402–414 (2006).
85. Raposo, G. & Stoorvogel, W. Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol 200, 373–383 (2013).
86. Yanez-Mo, M. et al. Biological properties of extracellular vesicles and their physiological functions. J Extracell Vesicles 4, 27066 (2015).
87. Nantakomol, D. et al. Circulating red cell-derived microparticles in human malaria. J Infect Dis 203, 700–706 (2011).
88. Thery, C., Zitvogel, L. & Amigorena, S. Exosomes: composition, biogenesis and function. Nat Rev Immunol 2, 569–579 (2002).
89. Silverman, J. M. & Reiner, N. E. Exosomes and other microvesicles in infection biology: organelles with unanticipated phenotypes. Cell Microbiol 13, 1–9 (2011).
90. Colombo, M., Raposo, G. & Thery, C. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. Annu Rev Cell Dev Biol 30, 255–289 (2014).
91. Herz, H. M., Garruss, A. & Shalitafard, A. SET for life: biochemical activities and biological functions of SET domain-containing proteins. Trends Biochem Sci 38, 621–639 (2013).
92. Mmetic, J. V., Buening, G. M. & Kinden, D. A. Purification of the erythrocytic stages of Babesia gibsoni. Vet Parasitol 190, 595–598 (2012).
93. Srivastava, K. et al. Effective and cheap removal of leukocytes and platelets from Plasmodium vivax infected blood. Malar J 8, 115 (2009).
94. Eid, J. et al. Real-time DNA sequencing from single polymerase molecules. Science 323, 133–138 (2009).
95. Chaisson, M. J. & Tesler, G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. BMC Bioinformatics 13, 238 (2012).
96. Chin, C. S. et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10, 563–569 (2013).
97. English, A. C. et al. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. PLoS One 7, e47768 (2012).
98. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics 19(Suppl 2), ii215–225 (2003).
99. Koet, I. Gene finding in novel genomes. BMC Bioinformatics 5, 59 (2004).
100. Gardner, M. J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419, 498–511 (2002).
101. Pain, A. et al. Genome of the host-cell transforming parasite Theileria annulata compared with T. parva. Science 309, 131–133 (2005).
102. Aurorecoecha, C. et al. EuPathDB: a portal to eukaryotic pathogen databases. Nucleic Acids Res 38, D415–419 (2010).
103. Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8, 1494–1512 (2013).
104. Jurka, J. et al. RepBase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110, 462–467 (2005).
105. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152 (2012).
106. Quinlan, A. R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics 47, 11 12 11–11 12 34 (2014).
107. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676 (2005).
108. Li, L., Stoeckert, C. I., Jr. & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13, 2178–2189 (2003).
109. Fischer, S. et al. Using OrthoMCL to assign proteins to OrthoMCL–DB groups or to cluster proteomes into new ortholog groups. Curr Protoc Bioinformatics Chapter 6, Unit 6 12 11–19 (2011).
110. Wang, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40, e49 (2012).
111. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. Genome Res 19, 1639–1645 (2009).
112. Kohr, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30, 772–780 (2013).
113. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic studies. Bioinformatics 25, 1972–1973 (2009).
114. Kuck, P. & Meusemann, K. FASconCAT: Convenient handling of the RNA Seq. G.R. coordinated gDNA Seq, performed genome assembly and annotation, and assisted in bioinformatics approaches. P.D. conceived the study and assisted in designing and planning of the study and supervised the infection experiments. ABH conceived the study, supervised data analyses and drafted the manuscript. All authors have read and approved the final manuscript.
**Additional Information**

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-03445-x

**Accession Codes:** Sequences and data have been deposited for public access on the European Nucleotide Archive (http://www.ebi.ac.uk/ena; PRJEB15861) and on the ProteomeXchange platform via PRIDE (http://www.proteomexchange.org; PXD005117).

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[Open Access] This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017