BMC Genomics

RESEARCH

Open Access

Genome assembly and annotation of *Babesia rossi*, a protozoan parasite for canine babesiosis

Neelam Redekar[1†], Xu Wang[2,3,4†], Luis Neves[5], Steven Brooks[6], Justin Lack[1], Andrew Leisewitz[7,8*] and Hans Ackerman[6*]

Abstract

Background Apicomplexan parasite, *Babesia rossi,* is an Ixodid tick-transmitted pathogen that causes the most severe form of canine babesiosis disease. Compared to other *Babesia* pathogens of dogs, *B. rossi* exhibits unique pathophysiology, virulence, and a responsiveness to drugs that differs from the small *Babesia* parasites.

Results Here we report the first near-complete chromosome-level assembly of *Babesia rossi* strain PMB – isolated from a sick dog from Pietermaritzburg, South Africa. Assembly with long-read HiFi data yielded 21.06 Mbp genome size, spanning across five gene-dense chromosome-level scaffolds, a single apicoplast scaffold, and a remaining 54 unplaced low gene density scaffolds with 1.32 Mb N50 and 96.6% BUSCO Apicomplexan completeness. The genome annotation identified a total of 3,098 protein-coding genes, 71 tRNA, and 16 rRNA genes. The mitochondrial genome (6.4 Kbp) was also identified. Genome assemblies of two additional field isolates of *B. rossi* were also reported. Comparative genomic analyses revealed four syntenic genomic inversions and multiple polymorphisms across three *B. rossi* isolates, although SNP and indel density was higher within the gene desserts of the genomes. Despite these differences, three *B. rossi* isolates' genome assemblies showed 99% conserved orthologous gene sets. About 76% of protein-coding genes of *Babesia rossi* isolate PMB were shared with four other *Babesia* species.

Conclusion This report provides valuable genomic information that is crucial to comprehend *B. rossi* evolution, virulence, and potential drug targets for canine babesiosis.

Keywords *Babesia*, Genome, *Babesia rossi*, Canine, Protozoan, Apicomplexa

†Neelam Redekar and Xu Wang contributed equally to this work.

*Correspondence:
Andrew Leisewitz
all0087@auburn.edu
Hans Ackerman
hans.ackerman@nih.gov
[1] Integrated Data Sciences Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA
[2] Department of Pathobiology, College of Veterinary Medicine, Auburn University, Auburn, AL 36849, USA
[3] HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA
[4] Scott-Ritchey Research Center, College of Veterinary Medicine, Auburn University, Auburn, AL 36849, USA
[5] Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Pretoria, South Africa
[6] Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Disease, National Institutes of Health, Rockville, MD 20852, USA
[7] Department of Clinical Sciences, College of Veterinary Medicine, Auburn University, Auburn, AL 36849, USA
[8] Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Pretoria, South Africa

Background

Parasites of the *Babesia* genus are highly successful Apicomplexan tick-transmitted organisms that infect the red blood cell in their vertebrate hosts [1]. Infections result in malaria-like disease characterized by fever and hemolytic anemia that can progress to multiple organ dysfunction in some cases resulting in a high morbidity and mortality. *Babesia* infections threaten all animal species important to man and human babesiosis is an emerging and increasingly important zoonosis in the USA and Europe [2]. Food-producing bovid species carry an enormous *Babesia* disease burden with over half of all cattle globally at risk of infection and this is of huge importance from both an economic and food security perspective [3]. The infection of domestic dogs is also a growing burden on the pet-owning population globally [4]. In human populations, the expanding tick vector distribution, increased interaction between humans and the natural environment, a growing population of immune-suppressed people, and emerging parasite drug resistance contribute to the increased number of cases of babesiosis [2].

It is quite typical for hemoprotozoan parasites of the same genus, but different species, to cause a wide spectrum of clinical disease depending on the specific parasite species responsible for the infection. A very typical and relevant example is provided by the important malaria parasite (genus *Plasmodium*) that infects humans. Infections with parasites of the same genus but different species result in very different clinical diseases [5–8]. Comparing the genomes of malaria parasites of widely differing pathogenicity has resulted in the identification of novel genes (with unknown function) that may well play a role in pathogenicity [9]. Another example is bovine *Babesia—B. bigemina* causes a relatively mild acute disease followed by persistent infection. *B. bovis* on the other hand causes a severe acute disease characterized by neurological signs and a high mortality [10].

The canine *Babesia* species further illustrates this biology. The disease caused by *B. rossi* infection is widely described as causing severe and complicated disease [11, 12]. The disease is acute, characterized by severe intravascular hemolysis, and, in around 10% of cases, death follows single or multiple organ failure, usually within 24 h of presentation for veterinary care [11] (Figure S1). The clinical disease caused by *B. rossi* is recognized as the most severe of all canid *Babesia spp*. infections [13]. *B. gibsoni* on the other hand has a lower mortality, is a far more chronic disease, with up to a third of infected dogs demonstrating parasitemia with no clinical signs [14], and the mortality rate is < 5% [15]. Multiple organ dysfunction has not been described with *B. gibsoni* infection. *B. rossi* is only known to be transmitted by an ixodid tick vector [16], whilst *B. gibsoni* can be transmitted by the bite of an infected dog during inter-dog aggression [15]. Diminazene aceturate is highly effective at sterilizing *B. rossi* infections, whilst it is ineffective against *B. gibsoni* where drug resistance to treatment is an ongoing problem [17].

The diseases caused by *B. rossi* and *B. gibsoni* have been studied in large case series [11, 18, 19] and, in the case of *B. rossi,* at an 'omics' level [20, 21]. The *B. gibsoni* genome has been published [22]. Here we present the *B. rossi* genome for the first time. Dogs have provided model systems to study numerous disease processes, including sepsis [23]. It is quite possible that canine babesiosis could provide a system in which to study parasite-associated disease mechanisms. This provides a basis for comparing the parasites that precipitate these diverse conditions at a genomics level as this could well help with discovering pathogenicity associated genes, as well as vaccine and drug targets.

Materials and methods

Parasite isolates

The *Babesia rossi* isolate used as the reference laboratory strain ('PMB') was first isolated in 1976 from a sick dog from Pietermaritzburg, South Africa [24]. This isolate has been used for experimental infections in dogs over several decades (for example [25]), maintained as a cryopreservate, and propagated in vitro in Dr. Leisewitz's laboratory at Auburn University College of Veterinary Medicine. The strains named 'K' and 'R' were collected from dogs presented to the Outpatients Clinic of the Veterinary Faculty of the University of Pretoria (Onderstpoort) for care because they were clinically ill. The infections were diagnosed by means of a stained peripheral thin blood smear, and the dogs were treated. A mono-infection with *B. rossi* was confirmed by means of PCR and Reverse Line Blot [26].

Parasite culture and sample collection

The R and K *Babesia rossi* clinical isolates were expanded by in vitro culture at the Onderstepoort (Gauteng, South Africa) laboratory by one of the authors (LN) while the PMB strain was cultured at the Auburn laboratory (Auburn, USA). Two hundred microlitres of cryopreserved isolate was thawed and placed into a culture system. Briefly, culture was performed in 25 mL culture flasks in a mixture of RPMI 1640 (with L-glutamine and 25 mM HEPES Corning 10–041-CV), with added dog serum (to 10%), reduced glutathione (Sigma #G4251-10 g) (2 mg/mL), insulin-transferrin-selenium (ITS, Gibco #51,500,056) (100 µL/10 mL) and gentamicin (10 mg/L). Cultures were incubated in 5% CO_2 at 37°C. Culture medium was changed daily, and parasite density was determined by standard light microscopy on

stained smears daily. Parasite density in canine red cells was allowed to reach approximately 30% at which time a portion of the red blood cell culture was harvested for DNA isolation and the remaining portion was further propagated until enough DNA had been generated for sequencing.

Nucleic acid isolation and sequencing

Genomic DNA samples were extracted from the three *B. rossi* isolates using Quick DNA Miniprep Plus Kit (Zymo Research, CA) following the manufacturer's protocol. The concentration of genomic DNA was measured by a Qubit Fluorometer instrument (Thermo Fisher Scientific, MA). DNA size distributions were examined on a Bioanalyzer (Agilent Technologies, CA) to ensure a DNA integrity number (DIN) 7.2 or greater. For each isolate, ~ 200 ng genomic DNA was used as input for Illumina TruSeq Nano DNA Library Preparation following the manufacturer's protocol. The final libraries were sequenced on a MiSeq platform using 300-bp paired-end format. PacBio libraries were constructed using the same DNA samples (~ 300 ng input per sample), barcoded and pooled in a PacBio SMRT Cell for sequencing on PacBio Sequel II platform with standard PacBio protocol. Both sequencing runs were performed at CCR Sequencing Facility, NCI (Frederick, MD). The PacBio SMRT Cell sequencing resulted in 380 K to 425 K HiFi reads, with total yield ranging from 3.7 to 4.4 Gbp and mean read quality QV44.7 to QV47.9 for three libraries. The MiSeq sequencing resulted in 10 million short PE reads with base call quality of above 84% with Q30 and above.

Genome assembly and annotation

The genome assembly and annotation were performed separately for three strains of *B. rossi* using the following steps. The genome assembly was first created using PacBio HiFi long reads using Flye (v2.8–1) [27] in `trestle` mode allowing minimum 10,000 bp overlap between reads. Assembly was polished with both HiFi (long) and MiSeq (short) read sequences using NextPolish (v1.3.1) [28] with default parameters. Multiple iterations (> 3) of scaffolding were performed using long interval K-mer based scaffolder (LINKS, v1.8.7) [29] until the total number of assembled scaffolds was stabilized. Assembly quality and completeness were evaluated using QUAST (v5.0.2) [30], and BUSCO (v4.1.3) [31] respectively. HiFi reads were then mapped to assembly sequence using minimap2 (v.2.26) (Li 2018), and mean genome coverage was estimated using Samtools (v.1.17) [32]. The rRNA and tRNA prediction were performed using 'cmsearch' function in Infernal (v1.1.4) [33] and tRNAscanSE (v2.0.9) [34] respectively. GC content of assembled scaffolds was estimated using seqkit (v.2.7.0) [35]. Telomeric repeat detection was performed within 1 Kb ends of chromosomes and unplaced scaffolds using TelFinder [36]. Full-length mitochondrial genomic reads were identified by mapping HiFi data using minimap2 (v.2.28) to *Babesia gibsoni* mitochondrial genome.

Genome annotations were performed using MAKER (v2.31.10) [37, 38]. This included repeat masking with RepeatMasker (v4.1.0) [http://repeatmasker.org], ab-initio gene prediction with Augustus (v3.3.3) [39] and SNAP (v2013-11–29) [40], EST and protein alignment with Exonerate (v2.4.0) [41]. Predicted genes were BLAST against protein database to identify proteins (ncbi-blast-2.10.0 +). Genome sequences and gene annotations from other *Babesia* genomes such as *B. divergens*, *B. bigemina*, *B. bovis*, *B. microti*, *B. ovata* were used for EST and protein homology evidence. Gene prediction completeness was estimated against Apicomplexa lineage ('apicomplexa_odb10' dataset—creation date: 2024–01–08) using BUSCO (v.5.4.7) [31] in protein mode. Annotation files were summarized with AGAT (v.1.2.0) [42]. KEGG pathway annotation was performed for predicted protein sequences using eggnog-mapper 2.0.1 (http://eggnog-mapper.embl.de/). Three discontinued KEGG pathway IDs: map00072, map00281, and map01130 were merged to map00650, map00907, and map 01110 as indicated in KEGG pathway map update history (https://www.genome.jp/kegg/docs/upd_map.html) prior to summarizing the results.

RNA sequencing analysis and gene annotation

Bulk mRNA-Seq reads (GEO accession: GSEI67201) originating from *B. rossi* infected canine host blood [21] were used to verify protein predictions in *B. rossi* genome. RNA-Seq reads originating from *Babesia* were first extracted by mapping them against *Canis lupus familiaris* and *B. divergens* reference genomes together using Disambiguate (v1.0) [43]. *Babesia* RNA-Seq reads were then mapped to assembled *B. rossi* isolated genomes with STAR (v2.7.10b) [44] to validate gene predictions.

Genome variant analyses

To assess diversity between strains, we remapped the K and R strains back to the PMB reference strain and identified variants. For this workflow, CCS reads were mapped to the PMB strain assembly using pbmm2 (v1.13.1, https://github.com/PacificBiosciences/pbmm2) and variants were called using DeepVariant (v1.6.0) [45]. Joint genotyping was then performed using GLNexus (v1.4.1) [46], and resulting variants were counted using GATK4 (v4.5.0.0) CollectVariantCallingMetrics function [47]. Density SNPs in 10 Kb window was calculated using vcftools (v0.1.16) [48].

Table 1 Summary statistics of *Babesia rossi* genome assemblies

Genome assembly	Strain PMB (laboratory strain)	Strain K (field isolate K)	Strain R (field isolate R)
Sequencing data and coverage			
PacBio sequencing data	4.3 Gb PacBio HiFi reads	3.7 Gb PacBio HiFi reads	4.3 Gb PacBio HiFi reads
Illumina sequencing data	6.51 Gb Illumina reads	6.59 Gb Illumina reads	6.09 Gb Illumina reads
Genome coverage	HiFi: 210×; Illumina: 310×	HiFi: 173×; Illumina: 314×	HiFi: 208×; Illumina: 290×
Assembly statistics			
Assembled genome length	21,067,390 bp	22,437,258 bp	20,911,923 bp
Number of scaffolds	61	141	51
Number of contigs	62	157	58
Scaffold N50	1.32 Mbp	0.77 Mbp	1.27 Mbp
Contig N50	1.32 Mbp	0.70 Mbp	1.27 Mbp
Maximum scaffold length	4.7 Mbp	3.5 Mbp	4.1 Mbp
Maximum contig length	4.7 Mbp	3.5 Mbp	4.1 Mbp
GC content	44.79%	44.84%	44.86%
Completeness (Apicomplexa)			
BUSCO completeness	96.6%	96.8%	96.6%
single-copy BUSCO	96.4%	96.6%	96.4%
duplicated BUSCO	0.2%	0.2%	0.2%
fragmented BUSCO	0.2%	0.2%	0.2%
missing BUSCO	3.2%	3.0%	3.2%
Mapping metrics			
Illumina read mapping %	96.83%	96.83%	96.89%

Comparative gene analyses with species from Apicomplexa phylum

The comparative orthologous gene clusters were identified using OrthoFinder (v.2.5.4) [49] and compared with those in 7 species of *Babesia* (*B. rossi, B. gibsoni. B. divergens, B. bigemina, B. ovata, B. bovis, B. microti*), *Theileria parva, Plasmodium falciparum*, and *Toxoplasma gondii* species. Species Tree inference from All Genes (STAG) trees were constructed using single copy orthologous gene results from OrthoFinder and visualized in iTOL [50]. Orthogroup unassigned gene predictions were queried against translated nucleotide databases with tBLASTn (v2.15.0+) [51]. Syntenies were generated with bed positions of predicted protein from *B. rossi* and *B. gibsoni* using GENESPACE [52] with MCScanX [53], Diamond (v.2.0.15) [54], and OrthoFinder (v.2.5.4) [49] dependencies.

Sources of genomic data used in this study

The following Apicomplexa genomes used in this study were downloaded from NCBI Genome database: *Babesia bovis* T2Bo (GCA_000165395.2), *Babesia bigemina* BOND (GCA_000723445.1), *Babesia divergens* 1802A (GCA_018398725.1), *Babesia microti* RI (GCA_000691945.2), *Toxoplasma gondii* GT1 (GCA_000149715.2), *Plasmodium falciparum* 3D7 (GCF_000002765.5), and *Theileria parva* Muguga (GCA_000165365.1). *Babesia gibsoni* WH58 (GWHBJTY00000000) nuclear assembly data was downloaded from National Genomics Data Center, China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences. *Babesia gibsoni* apicoplast genome information was downloaded from GenBank accession MN481613.1.

Results

Genome assembly statistics

We generated on average 4.1 Gb of Pacific Biosciences (PacBio) HiFi reads for each of the three *B. rossi* strains: the laboratory PMB strain, and two clinical isolates K and R (Table 1). All three assemblies had excellent continuity and genome completeness. The PMB strain was selected as the reference genome assembled due to fewer scaffolds and the lowest proportion of missing bases within them (Table 1). The final assembly size was 21,067,390 bp in 61 scaffolds, with a scaffold N50 of 1.32 Mb and a contig N50 of 1.32 Mb. Five gene-dense scaffolds (scaffold1, scaffold2, scaffold3, scaffold5, and scaffold6) were assigned to five chromosomes based on their size (Table 2). Collectively, they account for 11.7 Mb or 55.5% of the assembled genome. Our assembly was 0.5 Mb longer than the previously estimated *B. rossi* genome size of 20.5 Mb and in agreement with the five chromosomes previously shown in a pulsed-field

Table 2 Chromosome-level assembly and gene annotation statistics of *Babesia rossi* genome

	Chr1	Chr2	Chr3	Chr4	Chr5	Apicoplast	Unplaced
B. rossi PMB							
Scaffold ID	1	2	3	5	6	43	
Protein coding genes	1258	774	493	256	268	20	29
Total tRNA genes	20	11	7	1	8	24	
tRNA types	12	7	6	1	5	20	
% mRNA reads mapped	50.37	5.32	42.96	0.89	0.45	0	0
% GC content	44.73	44.55	43.89	43.98	44.30	20.20	45.03*
Telomeric repeats	ND	left	both	left	both	ND	
B. rossi K							
Scaffold ID	1	2	3	4	5	91	
Protein coding genes	1260	771	491	258	272	18	28
Total tRNA genes	20	11	7	1	8	24	
tRNA types	12	7	6	1	5	20	
% mRNA reads mapped	54.01	5.50	39.12	0.91	0.46	0	0
% GC content	44.44	44.17	43.97	44.32	44.27	20.21	45.19*
Telomeric repeats	ND	left	both	left	both	ND	
B. rossi R							
Scaffold ID	1	2	3	4	7	41	
Protein coding genes	1254	778	489	259	275	19	23
Total tRNA genes	20	11	7	1	8	24	
tRNA types	12	7	6	1	5	20	
% mRNA reads mapped	60.31	4.61	33.92	0.77	0.39	0	0
% GC content	44.68	44.29	44.02	44.67	44.58	20.20	45.38*
Telomeric repeats	ND	left	both	left	both	ND	

* Average GC content across all unplaced scaffolds is reported

Abbreviations: *ND* Not detected

gel electrophoresis study [55]. Mitochondrial genome for *B. rossi* PMB isolate was 6,412 bp long. Annotation of mitochondrial genome detected three genes: cytochrome B (1,092 bp), cytochrome oxidase 3 (654 bp), and cytochrome oxidase 1 (1,434 bp) genes.

Genome completeness and quality assessment

The genome completeness of the PMB assembly was assessed using Benchmarking Universal Single-Copy Orthologs (BUSCO). The overall completeness score against Apicomplexa was 96.6%, which is comparable to other *Babesia* species genomes, *B. gibsoni* (97%), *B. divergens* (97.5%), *B. bigemina* (96%), *B. bovis* (96.7%), *B.ovata* (96.8%), and *B. microti* (94.8%). The proportions of duplicated (0.2%) or fragmented BUSCO genes (0.2%) are relatively low in the *B. rossi* genome, suggesting a largely complete assembly (Table 1). The isolate K and isolate R assemblies had similar BUSCO statistics (Table 1). To assess the mapping quality of the reference genome, we aligned the Illumina short reads generated from all three isolates, and the mapping rate was ~ 97% (Table 1), consistent with excellent genome representation.

Gene annotation

A total of 3,098 protein-coding genes were annotated by the MAKER pipeline in the *B. rossi* PMB genome. Of these, 3,049 (98.4%) were located on the five chromosomes and 20 were located on the apicoplast (Table 2). Only 29 predicted genes were identified within the 55 unplaced scaffolds (Table 2). This low gene density suggests that these unplaced scaffolds represent gene deserts. De novo gene annotations were also performed for the K (3,098 genes) and R isolate assemblies with a similar number of genes identified (3,097) (Table 2). A total of 99% (3,068 genes) of predicted genes were assigned to shared orthogroups (N = 2882) among the three *B. rossi* isolates (Fig. 1A). The average gene length was 1,447 bp with an N50 of 1,821. The average number of exons per gene model was 2.9. To validate these predicted gene models, we utilized our previous RNA-seq data from *B. rossi*-infected canine blood samples [21], which contain both host and parasite sequences. A substantial amount of RNA-seq reads mapped to *B. rossi* PMB, confirming that they are truly expressed in vivo. Interestingly, the majority (over 90%) of the *Babesia* reads were mapped to

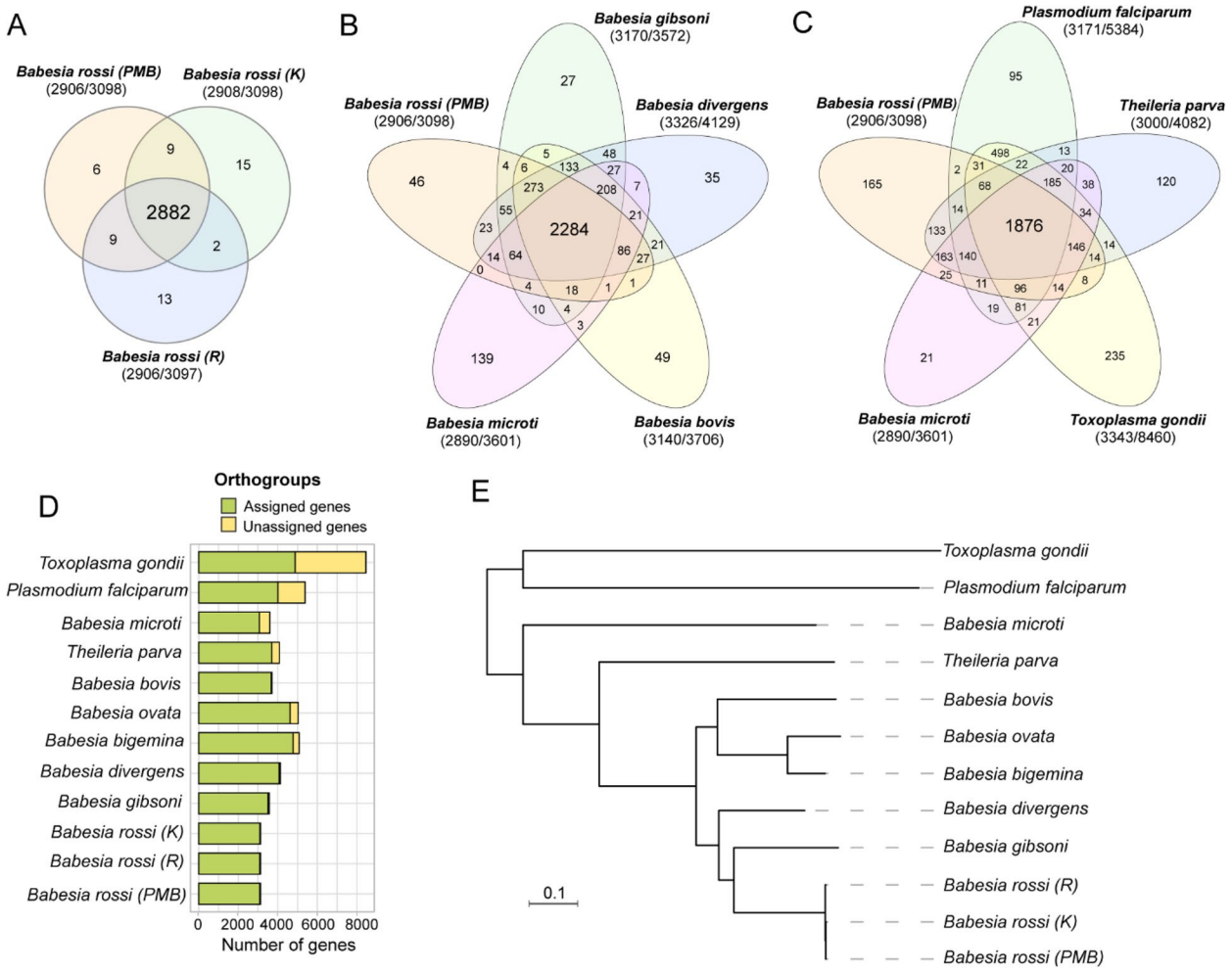

Fig. 1 Comparison of orthologous genes across multiple Apicomplexan parasite genomes. **A**–**C** Venn diagram showing the comparisons of orthoclusters across **A** three *Babesia rossi* clinical isolates (PMB, K, and R), **B** multiple species of*Babesia* genus (*B. rossi PMB, B. gibsoni, B. divergens, B. bovis, and B. microti*), and **C** multiple Apicomplexan parasite species (*B. rossi PMB, B. microti, Toxoplasma gondii, Theileria Parva, Plasmodium falciparum*). The total number of orthogroups and predicted gene models in each species are shown in parentheses. **D** Number of predicted genes assigned and unassigned to orthogroups for each species. **E** Species trees based on all orthogroups

Chr1 and Chr3, suggesting that these two chromosomes harbour the most highly expressed genes.

Annotation of non-coding RNAs

A total of 47 tRNA genes were identified in the nuclear genome of *B. rossi* PMB across five chromosomes (Table 3). These tRNAs decode the 20 standard amino acids with redundancy. All predicted tRNAs are single exon genes, except for the tRNA gene encoding Tyr on chromosome 5, which has one intron in its gene model – similar to the *B. gibsoni* genome assembly (a closely related canine *Babesia* species). Two copies each of 5.8S, 18S rRNA, and 28S rRNA genes were annotated on Chr3, and 5 copies of 5S rRNA gene and one copy each of 5.8S, 18S, and 28S rRNA genes were found on Chr1 (Table 3). *B. rossi* 5S, 18S, and 28S rDNA sequences are 99%, 95%, and 91% similar to that of *B. gibsoni*. The tRNA and rRNA gene prediction were a perfect match across the three *Babesia rossi* genome isolate assemblies.

Assembly and annotation of the apicoplast

As a specialized organelle in protozoan parasites, apicoplast is essential for key metabolic functions for parasite survival and propagation. In *B. rossi* PMB strain, the apicoplast genome was assembled into a single circularized contig (scaffold43). The apicoplast genome is 29,103 bp in size, which is slightly longer than that of *B. gibsoni* (28,386 bp). A total of 20 protein-coding genes, 24 tRNA genes decoding all 20 standard codons, and one copy of each 18S and 28S rRNA genes were annotated in the apicoplast genome (Table 2). The annotation of *B. gibsoni* apicoplast (GenBank: MN481613.1), on other hand, has

Table 3 Annotation of non-coding RNAs in *Babesia rossi*

B. rossi PMB	RNA types (count)	# genes
tRNA genes		
Chr1	Ala, Arg, Asn, Gln, Gly (2), Ile (2), Leu (3), Lys (2), Met, Ser (2), Thr, Val (3)	20
Chr2	Ala (2), Arg (3), Cys, Glu (2), Gly, Leu, Met	11
Chr3	Arg, Gln, His, Phe, Thr (2), Trp	7
Chr4	Leu	1
Chr5*	Asp, Lys, Pro (3), Ser (2), Tyr	8
Apicoplast	All 20 tRNAs; 2 copies of Arg, Met, Phe, Ser	24
rRNA genes		
Chr1	5s (5), 5.8s, 18s, 28s	8
Chr2	-	0
Chr3	5.8s (2), 18s (2), 28s (2)	6
Chr4	-	0
Chr5	-	0
Apicoplast	18s (1), 28s (1)	2

* One tRNA gene encoding Tyr on Chr5 is with one intron; while all other tRNA genes are intronless

30 protein coding genes, 23 tRNA genes decoding all 20 standard codons, and one copy each of 18S and 28S rRNA genes.

Comparative phylogenomic analysis of orthologous genes among *Babesia* species

To reveal the phylogenetic and evolutionary relationship between *B. rossi* and other *Babesia* parasites, we identified orthologous gene clusters (or orthogroups) between three *B. rossi* isolates and five additional *Babesia* parasites, as well as four Apicomplexan outgroup species. Within the *Babesia* genus, 76% of *B. rossi* PMB predicted genes (2,378 genes) are assigned to shared orthogroups (N = 2,284) in all four other *Babesia* species (Fig. 1B), suggesting a high level of conservation. If we go further out of the *Babesi*a genus, *B. rossi* PMB shares 75% (2,332 genes, 2238 orthogroups) of protein-coding genes with *Plasmodium falciparum*, the human malaria parasite, and 85% (2,657 genes, 2554 orthogroups) genes with *Theileria parva,* the theileriosis pathogen, and 76% (2,343 genes, 2253 orthogroups) genes with *Toxoplasma gondii,* the toxoplasmosis pathogen (Fig. 1C). Over 99% of our predicted genes in *B. rossi* PMB strain were assigned to an orthogroup (Fig. 1D). The phylogenomic analysis using orthologous genes demonstrated that all three *B. rossi* isolates fall into the same clade with no visible branch length, confirming that they are different strains within the same species (Fig. 1E). The immediate outgroup of *B. rossi* is *B. gibsoni*, which is another canine *Babesia* parasite. Interestingly, despite belonging to the same genus, the seven *Babesia* species did not form a monophyletic group in our analysis.

Babesia microti, a blood-borne parasite transmitted by the deer tick *Ixodes scapularis* typically classified as a *Babesia*, was found to be closer to the sister genus *Theileria* (Fig. 1E), suggesting potential historical classification issues or strong convergent evolution. This relationship was suggested by previous ribosomal RNA sequence comparisons, and here we demonstrate that at the genomic level [56, 57].

Potential species-specific protein-coding genes in the *Babesia rossi* genome

To identify *B. rossi*-specific genes that are not present in any other apicomplexan species, we examined the list of protein-coding genes that cannot be assigned to any orthogroups. In total 6, 5, and 4 *B. rossi*-specific unassigned gene candidates were found in isolates PMB, R, and K, respectively. Four of the unassigned genes were identical across two or three *B. rossi* isolate assemblies, and they encoded for proteins less than 15 amino acids in length. Despite their annotation edit distance score (AED) of 0.5 or less, these genes are likely to be incomplete gene models or gene prediction artifact (Table S1). The remaining two unassigned genes from PMB assembly encoded for the following proteins: PMB_chr1_P1 (75aa, AED = 0.20) and PMB_chr4_P1 (155aa, AED = 0.43) (Table S1). Translated nucleotide databases search (tblastn) using protein sequences of PMB_chr1_P1 gene showed 46% similarity to putative integral membrane protein mRNA in *B. bovis* T2Bo (XM_001611607.2), while the PMB_chr4_P1 gene showed 57% similarity to putative reverse transcriptase homolog partial mRNA in *B. bigemina* (XM_012911626.1). One of the unassigned genes from R isolate assembly: R_chr2_P1 (207aa, AED = 0.05) showed 65% similarity to putative multidrug resistance ribosomal protein RPL4 partial mRNA in *B. ovata* (XM_029011130.1); while another gene, R_chr2_P2 (36aa, AED = 0.47) showed no BLAST hit. These *B. rossi*-specific genes warrant further investigation, as these could be informative for exploring the functional divergence of this species.

Syntenic analysis and chromosomal rearrangements

Through comparative genomic analysis using gene order-based whole-genome alignment, large syntenic blocks were identified between *B. rossi* and *B. gibsoni* homologous chromosomes (Fig. 2). Notably, there are a dozen major inter-chromosomal rearrangements between the two species, indicating potential translocation and inversion events in the evolutionary history (Fig. 2). Nevertheless, most of the gene-dense chromosomal regions in *B. rossi* are represented in *B. gibsoni* assembly, and vice

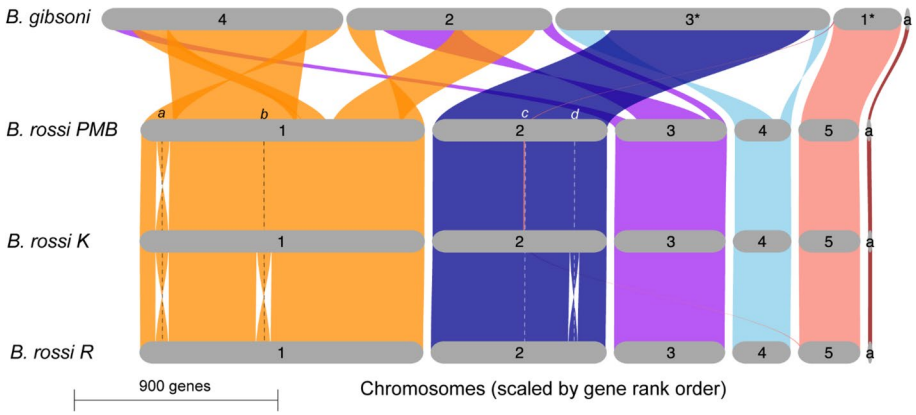

Fig. 2 Syntenic protein-coding regions between *Babesia rossi* isolates and *Babesia gibsoni*. Chromosomes are scaled by gene rank order. Chromosomal inversions relative to *B. gibsoni* are marked with asterisk. Syntenic inversion breakpoints between *Babesia rossi* clinical isolates are indicated as *a-d* and are further described in Table S2. Apicoplast is indicated as 'a'

versa (Fig. 2). The results suggest a high level of conservation in terms of gene content, which is consistent with our ortholog group analysis (Fig. 1). In addition to interspecies rearrangements, we also discovered four minor genomic rearrangements among the *B. rossi* lab strain (PMB) and two clinical isolates (K and R; Fig. 2). Genome neighbourhood analyses identified the exact genomic breakpoints for these events (Table S2), and the inversion events (a) are shared between K-PMB and R-PMB comparisons (Fig. 2).

SNP and indel density among three *B. rossi* strains

The *B. rossi* PMB strain was first isolated in 1976 and has been maintained as a laboratory strain for several decades, whereas isolates K and R were collected from recent clinical samples. By comparing each clinical isolate with the lab reference PMB strain, we identified 41,137 biallelic SNPs in isolate K and 29,070 SNPs in isolate R, with 7,200 SNPs shared between the two (Table 4 and Table S3). A total of 2,994 biallelic indels were called in isolate K, while isolate R had 1,815 indels (Table 4 and Table S3). Multiallelic SNPs and indels were extremely rare, which is as expected (Table 4). The overall SNP density is 2.4 SNPs per 10 Kb for field isolate K, and 2.2 SNPs per 10 Kb for isolate R, which represents the level of polymorphism within species. Along the five chromosomes, there is an inverse relationship between SNP/indel density and gene density (Fig. 3), indicating low levels of polymorphisms within a gene region, which is presumably due to selective constraint. Large gene-poor regions (ranging from 200 Kb to 1 Mb), including the beginning of Chr1, end of Chr2, and the beginning of Chr4 and Chr5, harbour the majority of the identified SNPs and indels (Fig. 3).

Table 4 Polymorphisms in *Babesia rossi* field isolates compared to assembled chromosomes of reference laboratory strain

	Private variants		Shared variants
	***B. rossi* field isolate K**	***B. rossi* field isolate R**	
Assembled 5 chromosomes[$]			
Total SNPs	34,094	21,931	7380
Total INDELs	2534	1358	535
SNP Ti/Tv ratio	0.9243	0.9644	0.9154
SNP density (per 10Kb)	2.4130	2.2077	NA

[$] PF variants that passed 'strict' validation stringency criteria in 'CollectVariantCallingMetrics' function in Picard is reported here

Prediction of metabolic pathways in the *Babesia rossi* genome

A total of 1386 putative proteins of the *B. rossi* PMB genome were annotated with 307 KEGG pathways, which were comprised of 616 unique pathway map identifiers. These annotated KEGG pathways belonged to major pathway groups such as metabolism (810 protein), genetic information processing (661 proteins), human diseases (636 proteins), organismal systems (472 proteins), cellular processes (317 proteins) and environmental information processing (225 proteins) (Fig. 4, Table S4). About 210 proteins were mapped to metabolic pathways (map01100) within global and overview maps (384 proteins) group for metabolism.

Discussion

Babesia rossi is a protozoan parasite, which infects red blood cells in dogs. It is predominantly found in sub-Saharan Africa and is known to cause some of the most severe forms of canine babesiosis. Isolates used in this study were all collected from dogs that presented with

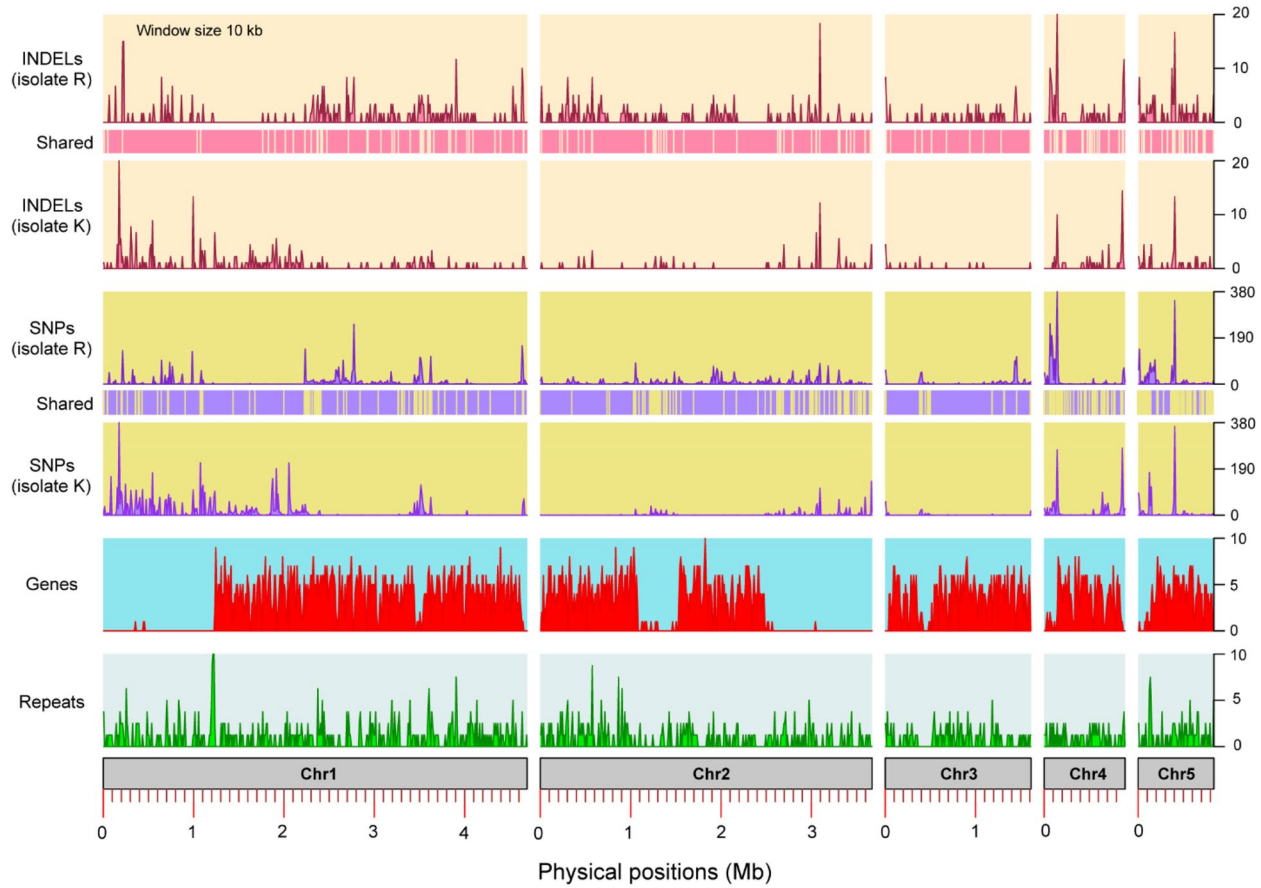

Fig. 3 Characteristics of assembled chromosomes of *Babesia rossi* PMB isolate. Polymorphisms in *B. rossi* clinical isolates (K and R) with respect to PMB isolate are plotted separately as single nucleotide polymorphisms (SNPs) and insertion-deletion (INDELs) for each isolate. Genomic bed positions of polymorphisms shared between two field isolates are represented within 'Shared' panel with yellow bars. Protein-coding gene predictions and simple repeats in PMB isolate are represented in two lower panels. All feature densities are plotted at 10 Kb window size

severe anemia and were all similarly affected. Understanding the genome of *B. rossi* is crucial for developing effective treatments, vaccines, and diagnostic tools specific to this highly virulent species. There are significant differences in virulence between the various *Babesia* species that infect dogs. *B. rossi* is widely accepted as the most virulent parasite frequently resulting in organ dysfunction and/or failure resulting in the highest mortality of all species. The next most virulent is *B. canis*. Infections with this parasite may result in severe disease but less frequently than *B. rossi*. This is followed by *B. gibsoni,* which often results in chronic infections that are usually subclinical or only mildly clinical. Infections with *B. vogeli* are typically asymptomatic [13]. Our study presents the first genome assembly reported for *Babesia rossi*, demonstrating high continuity (a total of 61 scaffolds with an N50 of 1.3 Mb) and excellent completeness (BUSCO completeness score 96.6%). Five gene-dense scaffolds comprise over 55% of the total assembled length and encompass > 99% of protein-coding genes, which correspond to the five *B. rossi* chromosomes. Despite the presence of several unplaced scaffolds, our assembly is nearly chromosome-level and 99% of gene models reside in five chromosomes.

A set of high-quality protein-coding genes was annotated in this reference *B. rossi* genome, with RNA-seq transcriptome data support. The majority of *B. rossi* genes have orthologs in *B. gibsoni,* a closely related canine *Babesia* species. Comparative genomic analysis revealed *B. rossi*'s evolutionary relationship with other apicomplexans, aligning well with the established phylogeny. In comparison with the other *Babesia* species, a unique feature of the *B. rossi* genome is a notable expansion in its genome size. While the typical *Babesia* genome ranges from 8 to 9 Mb, K-mer estimation indicated that the *B. rossi* genome is approximately double this size. The size of the *B. rossi* assembly in our study is approximately 21 Mb, with a slight inflation compared to the estimated size, which is likely due to some minor redundancy in repetitive contigs. The five chromosomes in our assembly exhibit strong

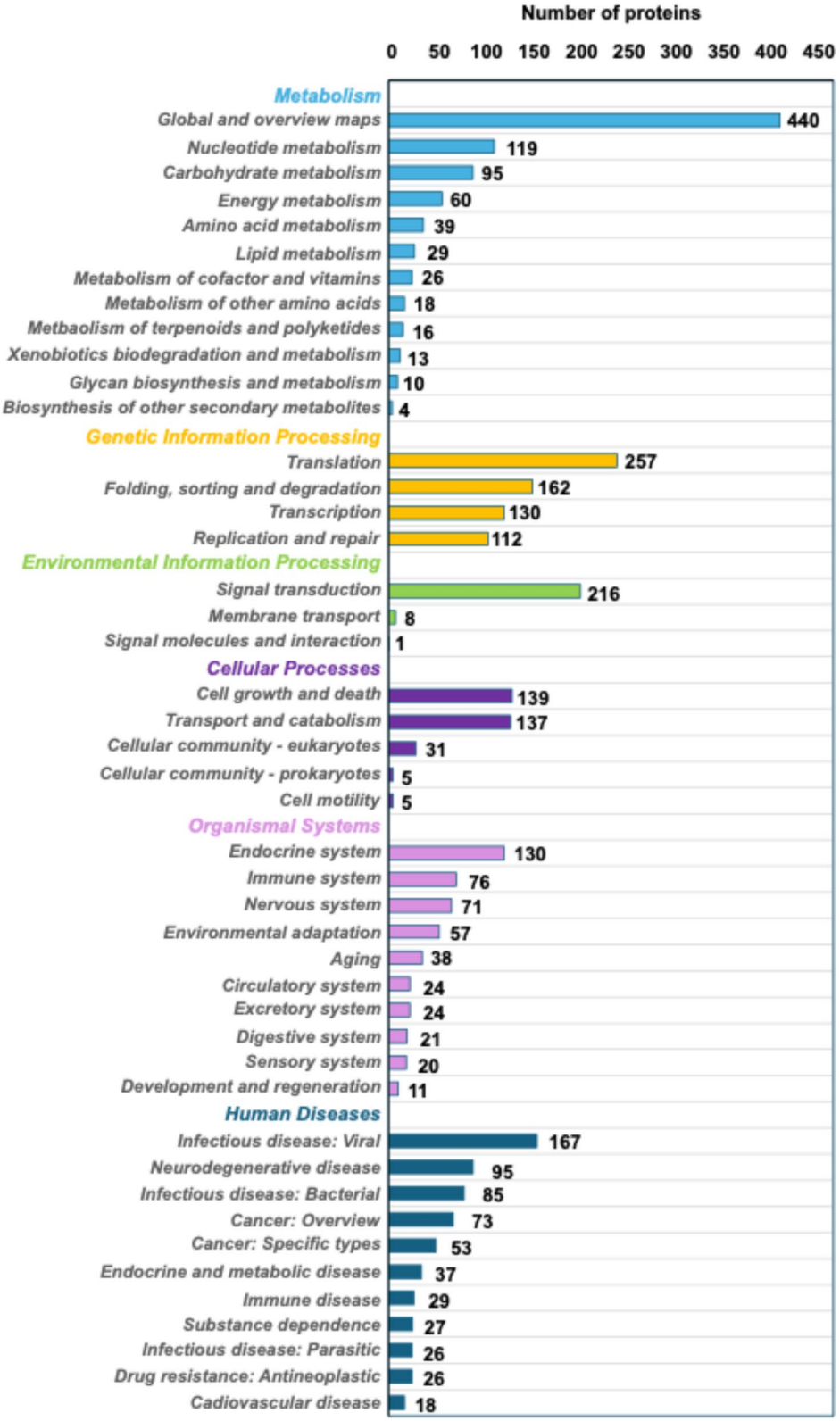

Fig. 4 Pathway annotation for *Babesia rossi* PMB isolate assembly. KEGG pathway annotations for protein-coding genes predicted in *B. rossi* PMB genome assembly are summarized

synteny with *B. gibsoni* gene models, spanning all four *B. gibsoni* chromosomes without apparent gaps. Our results suggest that the expanded regions in *B. rossi* are primarily non-coding or repetitive in nature. Whether these expanded genomic regions contribute to the increased virulence of *B. rossi* warrants further investigation.

Apicomplexan parasites exhibit significant genetic diversity, which can affect disease severity, immune response, and treatment outcomes. Studies in *Toxoplasma*, *Eimeria*, and *Plasmodium* have identified single nucleotide changes with significant effects on drug resistance and antigenicity [58–60]. Our reference genome was built on the PMB strain, a *B. rossi* strain isolated in 1970s and maintained in culture for almost 50 years. To determine the genetic diversity among laboratory and clinical strains, we independently assembled and annotated genomes from two clinical isolates, named as strain K and strain R. The assembled genome size of strain K is 6.5% larger than the reference, and strain R is almost identical to the reference in size. The observed size differences are likely attributable to variations in intergenic regions, as all three strains exhibited a similar number of genes. We identified three major inversions in *B. rossi* clinical isolates relative to the reference PMB strain: one inversion is shared between strains K and R, while two are private to strain R. Compared to the reference, we detected 33,937 biallelic SNPs and 2,526 indels in strain K, which translates to 2.6 polymorphic markers per 10 Kb. Strain R exhibits 55% fewer SNPs and 88% fewer indels than strain K, with one-third of its SNPs shared with strain K. These identified SNPs and indels are enriched in the intergenic regions. Although only two clinical isolates were analysed, our results serve as the first catalogue of *B. rossi* genetic diversity, facilitating the understanding of its population structure.

Conclusions

We report a high-quality genome assembly of *Babesia rossi*, an apicomplexan parasite responsible for the most virulent form of canine babesiosis. The *B. rossi* genome demonstrates a high degree of intraspecific sequence polymorphism and is most closely related to another canine *Babesia*, *B. gibsoni*. *B. rossi* possesses vast gene deserts that contribute to a substantial expansion of the genome size compared to other apicomplexa; however, the functional significance of this expansion remains to be determined.

Abbreviations

BUSCO	Benchmarking Universal Single-Copy Orthologs
AED	Annotation Edit Distance Score
STAG	Species Tree Inference from All Genes

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11495-z.

Supplementary Material 1. Figure S1.

Supplementary Material 2. Table S1, S2, S3.

Supplementary Material 3. Table S4.

Supplementary Material 4. Table S5.

Acknowledgements

This work utilized the computational resources of the NIH HPC Biowulf cluster (https://hpc.nih.gov).

Authors' contributions

NR: bioinformatics, data analysis, manuscript preparation (figures and tables) XW: data analysis and manuscript preparation LN: parasite culture, provision of parasite DNA, manuscript preparation SB: biospecimen processing and DNA extraction JL: bioinformatics, data analysis AL: Parasite collection, parasite culture, manuscript preparation HA: Overall project coordination, data analysis and manuscript preparation.

Funding

This research was supported in part by the Intramural Research Program of the NIH, including project AI001150 to HCA. Contributions from XW and AL were supported by the Scott Fund from the Scott-Ritchey Research Center, Auburn University.

Data availability

Babesia rossi PMB isolate genome was deposited to NCBI Genomes under accession number JAYQWK000000000. (BioProject accession: PRJNA1063654). Mitochondrial genome of *Babesia rossi* PMB isolate was deposited separately under PV059163 accession.

Declarations

Consent for publication

Consent has been provided by the NIH/NIAID. No consent is required from Auburn University or the University of Pretoria.

Competing interests

The authors declare no competing interests.

Received: 29 October 2024 Accepted: 17 March 2025
Published online: 01 July 2025

References

1. Mehlhorn H, Schein E. The piroplasms: life cycle and sexual stages. Adv Parasitol. 1985;23:37–103.
2. Vannier E, Krause PJ. Human babesiosis. N Engl J Med. 2012;366(25):2397–407.
3. Bock R, Jackson L, de Vos A, Jorgensen W. Babesiosis of cattle. Parasitology. 2004;129(Suppl):S247-269.
4. Irwin PJ. Canine babesiosis. Vet Clin North Am Small Anim Pract. 2010;40(6):1141–56.
5. Krause PJ, Daily J, Telford SR, Vannier E, Lantos P, Spielman A. Shared features in the pathobiology of babesiosis and malaria. Trends Parasitol. 2007;23(12):605–10.
6. Anstey NM, Douglas NM, Poespoprodjo JR, Price RN. *Plasmodium vivax*: clinical spectrum, risk factors and pathogenesis. Adv Parasitol. 2012;80:151–201.
7. Wassmer SC, Grau GE. Severe malaria: what's new on the pathogenesis front? Int J Parasitol. 2017;47(2–3):145–52.

8. Garcia LS. Malaria. Clin Lab Med. 2010;30(1):93–129.
9. Frech C, Chen N. Genome comparison of human and non-human malaria parasites reveals species subset-specific genes potentially linked to human disease. PLoS Comput Biol. 2011;7(12): e1002320.
10. Suarez CE, Alzan HF, Silva MG, Rathinasamy V, Poole WA, Cooke BM. Unravelling the cellular and molecular pathogenesis of bovine babesiosis: is the sky the limit? Int J Parasitol. 2019;49(2):183–97.
11. Leisewitz AL, Goddard A, Clift S, Thompson PN, de Gier J, Van Engelshoven J, Schoeman JP. A clinical and pathological description of 320 cases of naturally acquired *Babesia rossi* infection in dogs. Vet Parasitol. 2019;271:22–30.
12. Penzhorn BL. Why is Southern African canine babesiosis so virulent? An evolutionary perspective. Parasit Vectors. 2011;4:51.
13. Leisewitz AL, Mrljak V, Dear JD, Birkenheuer A. The Diverse Pathogenicity of Various *Babesia* Parasite Species That Infect Dogs. Pathogens. 2023;12(12):1437.
14. Lee MJ, Yu DH, Yoon JS, Li YH, Lee JH, Chae JS, Park J. Epidemiologic and clinical surveys in dogs infected with *Babesia gibsoni* in South Korea. Vector Borne Zoonotic Dis. 2009;9(6):681–6.
15. Birkenheuer AJ, Correa MT, Levy MG, Breitschwerdt EB. Geographic distribution of babesiosis among dogs in the United States and association with dog bites: 150 cases (2000–2003). J Am Vet Med Assoc. 2005;227(6):942–7.
16. Lewis B, Penzhorn B, Lopez-Rebollar L, De Waal D. Isolation of a South African vector-specific strain of *Babesia canis*. Vet Parasitol. 1996;63(1–2):9–16.
17. Baneth G. Antiprotozoal treatment of canine babesiosis. Vet Parasitol. 2018;254:58–63.
18. Liu P-C, Lin C-N, Su B-L. Clinical characteristics of naturally *Babesia gibsoni* infected dogs: a study of 60 dogs. Veterinary Parasitology: Regional Studies and Reports. 2022;28: 100675.
19. Macintire DK, Boudreaux MK, West GD, Bourne C, Wright JC, Conrad PA. *Babesia gibsoni* infection among dogs in the southeastern United States. J Am Vet Med Assoc. 2002;220(3):325–9.
20. Kules J, Rubic I, Farkas V, Baric Rafaj R, Gotic J, Crnogaj M, Burchmore R, Eckersall D, Mrljak V, Leisewitz AL. Serum proteome profiling of naturally acquired *Babesia rossi* infection in dogs. Sci Rep. 2023;13(1):10249.
21. Smith RL, Goddard A, Boddapati A, Brooks S, Schoeman JP, Lack J, Leisewitz A, Ackerman H. Experimental *Babesia rossi* infection induces hemolytic, metabolic, and viral response pathways in the canine host. BMC Genomics. 2021;22(1):1–16.
22. Liu Q, Guan X-A, Li D-F, Zheng Y-X, Wang S, Xuan X-N, Zhao J-L, He L. Babesia gibsoni whole-genome sequencing, assembling, annotation, and comparative analysis. Microbiology Spectrum. 2023;11:e00721-00723.
23. Minneci PC, Deans KJ, Hansen B, Parent C, Romines C, Gonzales DA, Ying SX, Munson P, Suffredini AF, Feng J, et al. A canine model of septic shock: balancing animal welfare and scientific relevance. Am J Physiol Heart Circ Physiol. 2007;293(4):H2487-2500.
24. Uilenberg G, Franssen FF, Perie NM, Spanjer AA. Three groups of *Babesia canis* distinguished and a proposal for nomenclature. Vet Q. 1989;11(1):33–40.
25. Schetters TP, Strydom T, Crafford D, Kleuskens JA, van de Crommert J, Vermeulen AN. Immunity against *Babesia rossi* infection in dogs vaccinated with antigens from culture supernatants. Vet Parasitol. 2007;144(1–2):10–9.
26. Matjila PT, Leisewitz AL, Jongejan F, Penzhorn BL. Molecular detection of tick-borne protozoal and ehrlichial infections in domestic dogs in South Africa. Vet Parasitol. 2008;155(1–2):152–7.
27. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–6.
28. Hu J, Fan J, Sun Z, Liu S. NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics. 2020;36(7):2253–5.
29. Warren RL, Yang C, Vandervalk BP, Behsaz B, Lagman A, Jones SJM, Birol I. LINKS: Scalable, alignment-free scaffolding of draft genomes with long reads. GigaScience. 2015;4(1):35.
30. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics. 2018;34(13):i142–50.
31. Manni M, Berkeley MR, Seppey M, Zdobnov EM. BUSCO: assessing genomic data quality and beyond. Current Protocols. 2021;1(12): e323.
32. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10(2):giab008.
33. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics (Oxford, England). 2013;29(22):2933–5.
34. Chan Patricia P, Lin Brian Y, Mak Allysia J, Lowe Todd M. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Research. 2021;49(16):9077–96.
35. Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One. 2016;11(10):e0163962.
36. Sun Q, Wang H, Tao S, Xi X. Large-scale detection of telomeric motif sequences in genomic data using TelFinder. Microbiol Spectr. 2023;11(2):e03928-03922.
37. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 2008;18(1):188–96.
38. Campbell MS, Holt C, Moore B, Yandell M. Genome annotation and curation using MAKER and MAKER-P. Curr Protoc Bioinformatics. 2014;48(1):4–11.
39. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics. 2008;24(5):637–44.
40. Korf I. Gene finding in novel genomes. BMC Bioinformatics. 2004;5(1): 59.
41. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics. 2005;6(1): 31.
42. Dainat J, Hereñú D, Dr KDM, Davis E, Crouch K, LucileSol, Agostinho N, Pascal G, Zollman Z, Tayyrov: NBISweden/AGAT: AGAT-v1.2.0. In.: Zenodo; 2023.
43. Ahdesmäki MJ, Gray SR, Johnson JH, Lai Z. Disambiguate: an open-source application for disambiguating two species in next generation sequencing data from grafted samples. F1000Research. 2016;5:2741.
44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
45. Poplin R, Chang P-C, Alexander D, Schwartz S, Colthurst T, Ku A, Newburger D, Dijamco J, Nguyen N, Afshar PT, et al. A universal SNP and small-indel variant caller using deep neural networks. Nat Biotechnol. 2018;36(10):983–7.
46. Yun T, Li H, Chang P-C, Lin MF, Carroll A, McLean CY. Accurate, scalable cohort variant calls using DeepVariant and GLnexus. Bioinformatics (Oxford, England). 2021;36(24):5582–9.
47. Van der Auwera GA, O'Connor BD. Genomics in the cloud: using Docker, GATK, and WDL in Terra. O'Reilly Media. 2020.
48. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8.
49. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20(1):238.
50. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6.
51. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402.
52. Lovell JT, Sreedasyam A, Schranz ME, Wilson M, Carlson JW, Harkess A, Emms D, Goodstein DM, Schmutz J. GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. eLife. 2022;11:e78526.
53. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee Th, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49–e49.
54. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59–60.
55. Depoix D, Carcy B, Jumas-Bilak E, Pages M, Precigout E, Schetters T, Ravel C, Gorenflot A. Chromosome number, genome size and polymorphism of European and South African isolates of large *Babesia* parasites that infect dogs. Parasitology. 2002;125(4):313–21.
56. Homer MJ, Aguilar-Delfin I, Telford SR III, Krause PJ, Persing DH. Babesiosis. Clin Microbiol Rev. 2000;13(3):451–69.

57. Baneth G, Florin-Christensen M, Cardoso L, Schnittger L. Reclassification of Theileria annae as Babesia vulpes sp. nov. Parasites and vectors. 2015;8:1–7.
58. Dautu G, Munyaka B, Carmen G, Zhang G, Omata Y, Xuenan X, Igarashi M. Toxoplasma gondii: DNA vaccination with genes encoding antigens MIC2, M2AP, AMA1 and BAG1 and evaluation of their immunogenic potential. Exp Parasitol. 2007;116(3):273–82.
59. Healer J, Murphy V, Hodder AN, Masciantonio R, Gemmill AW, Anders RF, Cowman AF, Batchelor A. Allelic polymorphisms in apical membrane antigen-1 are responsible for evasion of antibody-mediated inhibition in Plasmodium falciparum. Mol Microbiol. 2004;52(1):159–68.
60. Blake DP, Clark EL, Macdonald SE, Thenmozhi V, Kundu K, Garg R, Jatau ID, Ayoade S, Kawahara F, Moftah A. Population, genetic, and antigenic diversity of the apicomplexan Eimeria tenella and their relevance to vaccine development. Proc Natl Acad Sci. 2015;112(38):E5343–50.

Publisher's Note