

Draft genomes of two contemporary strains of Babesia divergens

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Apaa, T. T. ORCID: <https://orcid.org/0000-0001-7315-1262>, Jones, B. P., Blanchard, A. M. ORCID: <https://orcid.org/0000-0001-6991-7210> and Johnson, N. (2025) Draft genomes of two contemporary strains of Babesia divergens. Microbiology Resource Announcements, 14 (1). ISSN 2576-098X doi: 10.1128/mra.00898-24 Available at <https://centaur.reading.ac.uk/127597/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1128/mra.00898-24>

Publisher: American Society for Microbiology

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Draft genomes of two contemporary strains of *Babesia divergens*

Ternenge T. Apaa,^{1,2} Ben P. Jones,¹ Adam M. Blanchard,³ Nicholas Johnson^{1,4}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Babesia divergens*, a major cause of bovine babesiosis with zoonotic potential, was analyzed through genomes Bdiv23B and Bdiv24B following Illumina sequencing of DNA extracted from PCR-positive cattle blood. The genomes comprised 3888 and 4032 predicted coding sequences, respectively, comparable to the reference genome, Rouen 1987, highlighting genomic consistency across isolates.

KEYWORDS *Babesia divergens*, genomes, direct DNA sequencing, Great Britain

Babesia infection can lead to significant disease in both humans and domestic animals. Bovine babesiosis, commonly called redwater fever, causes febrile illness, anorexia, hemolytic anemia, dehydration, and diarrhea (1, 2). Disease results from high parasitic loads destroying erythrocytes, releasing hemoglobin into the urine that can be fatal if not treated promptly. Genomic analysis has emerged as a powerful tool for understanding pathogen biology, pathogenicity, and evolutionary relationships (3, 4). Previously, it was necessary to culture *Babesia divergens* prior to DNA preparation and sequencing, a process that is time-consuming. To avoid this limitation, we have derived *B. divergens* genomes directly from an infected cattle blood sample.

B. divergens was detected in two cattle blood samples by PCR (5). Prior to extraction, red blood cells were lysed with 0.15% saponin prepared in 1× phosphate-buffered saline (PBS) (150 mg: 99.85 mL 1× PBS) and the parasite preparation washed with PBS. Total DNA was extracted using the AllPrep DNA/RNA Kit (QIAgen, UK) following the manufacturer's instruction. Fragmented DNA libraries were prepared for both samples using the Nextera XT Kit (Illumina, Cambridge, UK) and sequenced using NovaSeq (Illumina) to generate 150 base pair-end reads. Reads were quality-filtered and trimmed using fastp v0.23.1 (6), then classified taxonomically with Kraken2 v2.1.2 protozoa and fungal reference database (accessed 10 March 2023) (7) to confirm *Babesia* presence. First, depletion of host reads from all data sets was conducted by mapping reads to a cattle reference genome, ARS-UCD2.0 (accession no. [GCF002263795.3](#)) and extracting unmapped reads. Second, recovered reads not mapped to cattle genome were then mapped to concatenated *B. divergens* reference genome assemblies available in the National Center for Biotechnology Information (NCBI) database ([1987:GCA001077455.2](#) and [1802A:GCA018398725](#)). Third, reads mapped to concatenated *B. divergens* genomes were extracted for *de novo* genome assembly using SPAdes Genome Assembler v3.15.5 (8). Mapping and extraction of both unmapped and mapped reads data sets (steps 1–3) were conducted using bowtie2 v2.5.2 (9) and samtools v1.9 (10). Assembled contigs were also classified against the NCBI Blastn Database (accessed 19 December 2023) using SprayNPray v1.0 (11). Filtering and extraction of contigs ≥ 500 bp from the resulting assemblies were conducted using SeqKit v2.7.0 (12). Quality evaluation in comparison with *B. divergens* reference genomes Rouen 1987 was performed on assembled genomes using QUAST v5.2.0 (13) (Table 1). All bioinformatics analyses were conducted using default parameters. This approach allowed genome derivation directly from infected cattle blood DNA without requiring passage in purified bovine red blood cells, enhancing our understanding of the parasite's genetic makeup.

Editor Jason E. Stajich, University of California
Riverside, Riverside, California, USA

Address correspondence to Ternenge T. Apaa,
Ter.apaa@apha.gov.uk.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 18 September 2024

Accepted 22 November 2024

Published 13 December 2024

© Crown copyright 2024. This is an open-access
article distributed under the terms of the [Creative
Commons Attribution 4.0 International license](#).

TABLE 1 Genome comparison of *Babesia divergens* genomes generated by this study and the reference Rouen 1987^a

Parameter	Bdiv23B	Bdiv24B	Rouen 1987 ^a
Genome size (Mb)	8.15	8.30	9.7
GC% content	45.7	45.6	45.5
No. of contigs	530	725	459
Contig N50 (kb)/L50	146.42/17	148.82/19	123.5/22
Total reads sequenced (M)	59.8	64.8	NA
Total reads (M) used in genome assembly	28.82	28.96	NA
Reads mapped to reference (M)	28.61	28.61	NA
Avg. coverage depth	464	459	353
GenBank accession number	JBDMXZ000000000	JBDMXY000000000	GCA001077455.2

^a*B. divergens* (14); NA: not applicable; M: total sequences in millions.

ACKNOWLEDGMENTS

This study was supported by the Department for Environment, Food and Rural Affairs (Defra) and the Scottish and Welsh governments through project SE0566 and the Biotechnology and Biological Sciences Research Council through project BB/X018008/1. We also acknowledge teams within Animal and Plant Health Agency that contributed to sample collection, shipment, and Illumina sequencing.

AUTHOR AFFILIATIONS

- ¹Vector Borne Diseases, Virology Department, Animal Plant and Health Agency, Addlestone, Surrey, United Kingdom
²Rabies and Viral Zoonoses, Virology Department, Animal Plant and Health Agency, Addlestone, Surrey, United Kingdom
³School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington, United Kingdom
⁴Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, United Kingdom

AUTHOR ORCIDs

Ternenge T. Apaa <http://orcid.org/0000-0001-7315-1262>
Adam M. Blanchard <http://orcid.org/0000-0001-6991-7210>

FUNDING

Funder	Grant(s)	Author(s)
Department for Environment, Food and Rural Affairs, UK Government (Defra)	SE0566	Nicholas Johnson
UKRI Biotechnology and Biological Sciences Research Council (BBSRC)	BB/X018008/1	Nicholas Johnson

AUTHOR CONTRIBUTIONS

Ternenge T. Apaa, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Ben P. Jones, Data curation, Investigation, Methodology, Visualization, Writing – review and editing | Adam M. Blanchard, Conceptualization, Data curation, Methodology, Supervision, Validation, Visualization, Writing – review and editing | Nicholas Johnson, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

This Whole Genome Shotgun (WGS) project has been deposited at DDBJ/ENA/GenBank under the Bioproject [PRJNA1091155](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1091155), accession numbers [JBDMXZ000000000](https://www.ncbi.nlm.nih.gov/assembly/JBDMXZ000000000) (version [JBDMXZ000000000.1](https://www.ncbi.nlm.nih.gov/assembly/JBDMXZ000000000.1); SRA: [SRR28427025](https://www.ncbi.nlm.nih.gov/sra/SRR28427025)) and [JBDMY000000000](https://www.ncbi.nlm.nih.gov/assembly/JBDMY000000000) (version [JBDMY000000000.1](https://www.ncbi.nlm.nih.gov/assembly/JBDMY000000000.1); SRA: [SRR28427024](https://www.ncbi.nlm.nih.gov/sra/SRR28427024)). A set of bioinformatic scripts used in this analysis has been deposited in GitHub to ensure reproducibility and consistency (<https://github.com/Ter-lab/Babesia-divergens-genome-analysis/tree/main>).

ETHICAL APPROVAL

The animal study protocol was approved by the Veterinary 428 Ethical Review Committee of The Royal (Dick) School of Veterinary Studies (Ref. 59/21, 18 May 429, 2021).

REFERENCES

1. Singh K, Kumar S, Sharma AK, Jacob SS, RamVerma M, Singh NK, Shakya M, Sankar M, Ghosh S. 2022. Economic impact of predominant ticks and tick-borne diseases on Indian dairy production systems. *Exp Parasitol* 243:108408. <https://doi.org/10.1016/j.exppara.2022.108408>
2. Johnson N, Paul Phipps L, McFadzean H, Barlow AM. 2020. An outbreak of bovine babesiosis in February, 2019, triggered by above average winter temperatures in southern England and co-infection with *Babesia divergens* and *Anaplasma phagocytophilum*. *Parasit Vectors* 13:305. <https://doi.org/10.1186/s13071-020-04174-3>
3. Jackson AP, Otto TD, Darby A, Ramaprasad A, Xia D, Echaide IE, Farber M, Gahlot S, Gamble J, Gupta D, et al. 2014. 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. <https://doi.org/10.1093/nar/gku322>
4. González LM, Estrada K, Grande R, Jiménez-Jacinto V, Vega-Alvarado L, Sevilla E, Barrera J de la, Cuesta I, Zaballos Á, Bautista JM, Lobo CA, Sánchez-Flores A, Montero E. 2019. Comparative and functional genomics of the protozoan parasite *Babesia divergens* highlighting the invasion and egress processes. *PLoS Negl Trop Dis* 13:e0007680. <https://doi.org/10.1371/journal.pntd.0007680>
5. McFadzean H, Johnson N, Phipps LP, Hobbs RL. 2021. High morbidity associated with an outbreak of tick - borne disease in a dairy herd, Cornwall. *Vet Record Case Rep* 9:e171. <https://doi.org/10.1002/vrc2.171>
6. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
7. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol* 20:257. <https://doi.org/10.1186/s13059-019-1891-0>
8. Meleshko D, Hajirasouliha I, Korobeynikov A. 2021. coronaSPAdes: from biosynthetic gene clusters to RNA viral assemblies. *Bioinformatics* 38:1–8. <https://doi.org/10.1093/bioinformatics/btab597>
9. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>
10. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. <https://doi.org/10.1093/gigascience/giab008>
11. Garber AL, Armbruster CR, Lee SE, Cooper VS, Bomberger JM, McAllister SM. 2022. SprayNPray: user-friendly taxonomic profiling of genome and metagenome contigs. *BMC Genomics* 23:202. <https://doi.org/10.1186/s12864-022-08382-2>
12. Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11:e0163962. <https://doi.org/10.1371/journal.pone.0163962>
13. Mikheenko A, Pribelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34:i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>
14. Cuesta I, González LM, Estrada K, Grande R, Zaballos Á, Lobo CA, Barrera J, Sanchez-Flores A, Montero E. 2014. High-Quality Draft Genome Sequence of *Babesia divergens*, the Etiological Agent of Cattle and Human Babesiosis. *Genome Announc* 2. <https://doi.org/10.1128/genomeA.01194-14>