Molecular characterization of ‘Candidatus Borrelia tachyglossi’ (family Spirochaetaceae) in echidna ticks, Bothriocroton concolor

Siew-May Loh,1 Amber Gillett,2 Una Ryan,1 Peter Irwin1 and Charlotte Oskam1,*

Abstract
Recently, a novel species of the genus Borrelia was identified in Bothriocroton concolor and Ixodes holocyclus ticks from echidnas. Analyses of 16S rRNA and flaB genes identified three closely related genotypes of this bacterium (Borrelia sp. Aus A-C) that were unique and distinct from previously described borreliae. Phylogenetic analyses of flaB (763 bp), groEL (1537 bp), gyrB (1702 bp) and glpQ (874 bp) gene sequences and concatenated sequences (3585 bp) of three gene loci (16S rRNA, flaB and gyrB) were consistent with previous findings and confirm that this novel species of the genus Borrelia is more closely related to, yet distinct from, the Reptile-associated (REP) and Relapsing Fever (RF) groups. At the flaB locus, genotypes A, B and C shared the highest percentage sequence similarities (87.9, 88 and 87.9 %, respectively) with Borrelia turcica (REP), whereas at the groEL and gyrB loci, these genotypes were most similar (88.2–89.4 %) to Borrelia hermsii (RF). At the glpQ locus, genotypes A and B were most similar (85.7 and 85.4 % respectively) to Borrelia sp. Tortoise14H1 (REP). The presence of the glpQ gene, which is absent in the Lyme Borreliosis group spirochaetes, further emphasises that the novel species of the genus Borrelia characterized in the present study does not belong to this group. Phylogenetic analyses at multiple loci produced consistent topographies revealing the monophyletic grouping of this bacterium, therefore providing strong support for its species status. We propose the name ‘Candidatus Borrelia tachyglossi’, and hypothesize that this species of the genus Borrelia may be endemic to Australia. The pathogenic potential of this bacterium is not yet known.

The family Spirochaetaceae is classified under the order Spirochaetales, belonging to the phylum Spirochaetes. This family consists of genera that are of concern to human health, such as Borrelia and Treponema [1], with common pathogenic species including ‘Treponema pallidum subsp. pallidum’, the causative agent of syphilis worldwide [2], and ‘Treponema pallidum subsp. pertenue’, the bacterium responsible for yaws [3]. The genus Borrelia is a member of the family Spirochaetaceae and through convention is divided into three major clades: Lyme disease/Borreliosis (LB) caused by members of the Borrelia burgdorferi sensu lato complex, the Relapsing Fever (RF) borreliae and the Reptile-associated (REP) borreliae [4]. The LB borreliae currently comprise over 20 recognized species including the primary Northern hemisphere Lyme-disease-causing agents Borrelia afzelii, Borrelia bavariensis, Borrelia burgdorferi sensu stricto, and Borrelia garinii, along with a newly described genospecies ‘Candidatus Borrelia mayonii’ that causes LB in the upper Midwestern USA [5–8]. Members of the LB group are vectored by hard ticks (family Ixodidae), with the pathogenic species commonly transmitted to humans and other animals by ticks within the Ixodes ricinus complex: Ixodes ricinus in Europe, Ixodes persulcatus in Europe and Asia, and Ixodes pacificus and Ixodes scapularis in USA [7]. These pathogenic spirochaetes are dependent on wildlife, particularly rodents and birds, which act as asymptomatic reservoir hosts that maintain their life cycles and transmission [9].

Spirochaetes within the RF group have been reported throughout a number of continents, including Africa [10], Eurasia [11] and North America [12]. These borreliae, in contrast to the LB group, are generally transmitted by soft ticks (family Argasidae), with the exceptions of Borrelia miyamotoi identified in I. persulcatus [13], I. ricinus [14], I. pacificus [15] and I. scapularis [16]; ‘Borrelia lonestari’ in Amblyomma americanum [17]; Borrelia theileri in Rhipicephalus microplus [18]; and ‘Candidatus Borrelia texanensis’ in Dermacentor variabilis [19].

The third major clade of this genus, the REP borreliae, was identified after the discovery of Borrelia turcica in Hyalomma aegyptium ticks collected from tortoises in Turkey.

Author affiliations: 1Vector and Water-Borne Pathogen Research Group, School of Veterinary and Life Sciences, Murdoch University, Perth, Western Australia, Australia; 2Australia Zoo Wildlife Hospital, Beerwah, Queensland, Australia.
*Correspondence: Charlotte Oskam, c.oskam@murdoch.edu.au
Keywords: Australian ticks; Candidatus Borrelia tachyglossi; Spirochaetaceae; Bothriocroton concolor; Echidna.
Abbreviations: LB, Lyme disease/Borreliosis; REP, Reptile-associated; RF, Relapsing Fever.
All sequences generated in this study were deposited in GenBank/EMBL/DDJB under accession numbers KY586964–KY586974.
One supplementary figure and seven supplementary tables are available with the online Supplementary Material.
Although LB borreliae have never been identified in Australian ticks, wildlife or people [23], two RF borreliae are recognized: *B. heilneri* and *Borrelia anserina* that are transmitted by Rhipicephalus (Boophilus) australis [24, 25] and Argas persicus [26, 27], respectively. In addition, *Borrelia queenslandica* was the first species of the genus *Borrelia* to be reported from native long-haired rats, *Rattus villosissimus*, in north-west Queensland, Australia. While the soft tick, *Ornithodorus gurneyi*, was considered to be the vector of this species due to its presence in the region, transmission experiments were not successful, and molecular characterization was never conducted to reliably identify the species of the genus *Borrelia* [28, 29].

Sequence analysis of multiple loci offers the advantage over morphological characterization of being highly discriminatory, therefore serving as a reliable method for accurate identification, characterization, and population, and epidemiological analyses in numerous bacterial studies [30, 31]. This technique was first used on *Borrelia burgdorferi* in 2008 [32] and has become an increasingly common technique for taxonomic and epidemiological studies of this genus [33–36].

Recently, a novel species of the genus *Borrelia* was detected in a number of echidna ticks, *Bothriocroton concolor*, collected [37]. An additional representative was detected in an *Ixodes holocyclus* tick [38]. Based on the 16S rRNA (1097 bp) and flagellin (*flaB*, 400 bp) gene phylogenetic analyses, this species of the genus *Borrelia* formed a distinct clade from other well-described borreliae, indicating this organism, designated *Borrelia sp. Aus*, to be unique [37]. In the present study, we conducted sequence analyses of the *flaB*, *groEL*, *glpQ* and *gyrB* genes, in addition to the 16S rRNA and short *flaB* loci reported previously [37], to study the relationship of this novel species to other of the genus *Borrelia*. Phylogenetic analysis confirmed its species status, and we hereby propose to designate this species as *Candidatus Borrelia tachyglossi*.

This study was conducted under the compliance of the Australian Code for the Responsible Conduct of Research, 2007 and Australian Code for the Care and Use of Animals for Scientific Purposes, 2013. Tick collection was carried out opportunistically with the approval from the Murdoch University Animal Ethics Committee.

Genomic DNA was extracted previously from 97 *Bothriocroton concolor* ticks, with 38 (39 %) ticks testing positive at the *Borrelia*-specific *flaB* locus [37]. These 38 *Borrelia*-positive ticks were included in this study. Nested- and semi-nested PCRs were conducted using primers targeting the housekeeping genes *flaB* (763 bp), *groEL* (1537 bp), *glpQ* (874 bp) and *gyrB* (1702 bp) (Table S1, available in the online Supplementary Material). Primers were designed to amplify short fragments of each gene with overlapping regions in order to obtain maximum coverage of the genes analysed for accurate characterization. PCR cycling conditions were as described by Loh *et al.* [37] (see Table S1 for respective annealing temperatures), with the exception of *groEL* primers: initial denaturation at 95 °C for 5 mins, 35 cycles of denaturation at 95 °C for 30 s, annealing at 48 °C for 40 s, extension at 72 °C for 2 mins, and a final extension at 72 °C for 7 mins. Amplification products of the targeted DNA products were electrophoresed in 1–2 % agarose gel with blue light using safe (Invitrogen), and positive samples were purified and sent for Sanger sequencing.

DNA sequences generated at *flaB*, *groEL*, *glpQ* and *gyrB* loci were aligned and analysed together with sequences representing species of the genus *Borrelia* retrieved from GenBank. All sequences were aligned using MAFFT v.7.017 [39] and then refined using MUSCLE [40]. The best-fit model for each locus was assessed using MEGA6 [41] and was selected based on the Bayesian Information Criterion (BIC). Bayesian phylogenetic reconstructions using sequence alignments of all four loci were generated using MrBayes 3.2.6 [42], and concatenated phylogenetic reconstructions using concatenated sequence alignments were produced using the CIPRES Science Gateway V.3.3 [43]. GTR and HKY substitution models were selected, with gamma categories of five, MCMC length of 1 100 000, burn-in length of 10 000 and subsampling frequency of 200.

In the present study, *Borrelia*-specific *flaB* (763 bp), *groEL* (1537 bp), *gyrB* (1702 bp) and *glpQ* (874 bp) DNA sequences were successfully amplified and sequenced from 38 *Borrelia*-positive *Bothriocroton concolor* ticks described by Loh *et al.* [37]. Thirty samples were positive for all *flaB* fragments; 24 samples were positive for all *glpQ* fragments; 13 samples were positive for all *groEL* fragments; and 10 samples were positive for all *gyrB* fragments.

Previously, three closely related genotypes were distinguished in the *Bothriocroton concolor* ticks using the 16S rRNA gene sequences, tentatively given the designations *Borrelia sp. Aus A*, *Borrelia sp. Aus B* and *Borrelia sp. Aus C* [37]. However, in the present study, these genotypes are referred as *Candidatus Borrelia tachyglossi* genotypes A, B and C, respectively. At the *flaB* locus, genotypes A, B and C consisted of nine, three and five identical samples, respectively. The *flaB* gene alignment (787 bp) between *Candidatus Borrelia tachyglossi* genotypes and other described species of the genus *Borrelia* showed that the novel genotypes shared highest percentage sequence identities with *Borrelia turcica* from the REP group (87.9–88 %); similarities to the LB *Borrelia* group ranged from 82.1–83.2 %; with the least similarity to *Borrelia hispanica* from the RF group (79.8–80 %). The percentage identities within the *Candidatus Borrelia tachyglossi* genotypes ranged from 99.6–99.9 %. The percentage nucleotide identities between *Candidatus Borrelia tachyglossi* genotypes and *Borrelia turcica* (87.9–
formed their own monophyletic clade, separate from *Borrelia turcica* and the RF group, with high bootstrap support (98%) (Fig. 1). The percentage nucleotide identities between *Candidatus Borrelia tachyglossi* genotypes and *Borrelia hermsii* (88.1–88.2%) were higher than that between *Borrelia turcica* and *Borrelia hermsii* (86.7%) (Table S5).

A Bayesian phylogenetic tree reconstructed using the concatenated alignment (3585 bp) consisting of three genes, 16S rRNA, *flaB* and *gyrB*, available for each of the main borreliae (Fig. 1), illustrated that the *Candidatus Borrelia tachyglossi* genotypes from Bothriocroton concorlack ticks grouped separately, with *Borrelia turcica* as the closest relative (91.1–91.2% nucleotide identities) (Table S6). Likewise, the concatenated alignment (5154 bp), which excluded *Borrelia turcica* (REP), based on the four loci amplified in the present study (16S rRNA, *flaB*, *groEL* and *gyrB*) also produced a similar tree topology with *Candidatus Borrelia tachyglossi* genotypes forming a monophyletic clade supported by high posterior probabilities (Fig. S1), sharing the highest percentage identities with *Borrelia hermsii* (90.3%) (Table S7).

All phylogenetic trees reconstructed revealed similar topologies with the REP group species of the genus *Borrelia* as the closest sister clade. The recently established REP group has been detected in various reptiles from several countries and in ticks that parasitize them [4, 21]. The recent discovery of a novel species of the genus *Borrelia* in Amblyomma vareness*e* collected from the reticulated python (*Python reticulatus*) showed that the species of the genus *Borrelia* identified clustered together with the REP-associated *Borrelia* group, along with *Borrelia turcica*, based on phylogenetic analyses of 16S rRNA and *flaB* genes [44]. However, the pathogenic potential of the species of the genus *Borrelia* belonging to the REP group is unknown.

A number of members within the genus *Borrelia* are well known to cause diseases in humans outside of Australia; nonetheless, borrelial tick-borne disease in humans still remains highly speculative and controversial in this country. *Candidatus Borrelia tachyglossi* was first reported in 2016 [37], hence the urgency to further characterize this bacterium on the basis of multi-loci gene sequencing in order to confirm the taxonomic position of this new member in the genus *Borrelia*. Our results, based on sequence and phylogenetic characterization of multiple loci, provide conclusive evidence that *Candidatus Borrelia tachyglossi*, identified in *Bothriocroton concorlack* ticks from echidnas, is distinct from other described species of the genus *Borrelia* and constitutes to a new clade in this genus.

Borrelial spirochaetes are well known to be associated closely with wildlife and utilize tick vectors to maintain a sylvatic life cycle [45, 46]. Australian wildlife are also known to be involved in spill-over of various zoonotic parasites [47]. Therefore, it is plausible that this bacterium is also likely to persist in the environment through circulation.
Phylogenetic reconstructions based on *Borrelia tachyglossi* (AY241832), *Treponema pallidum* (AY241832), flaB (16S rRNA, Concatenated), flaB, groEL, gplQ and gyrB gene sequences, and concatenated gene sequences of 'Candidatus Borella tachyglossi' genotypes A, B, and C identified in *Bothriocroton concolor* ticks from echidnas, *Brachyspira pilosicoli* (AY241832), *Treponema pallidum* (NZ_CP010566), *Escherichia coli* (X56907) and *Spirrochaeta lutea* (UNUP01000064) were used as
among native ticks and native mammalian (including marsupial) hosts. Unlike the Australian paralysis tick *I. holocyclus*, *Bothriocroton concolor* is a highly specialized tick, with echidnas as their primary host, and with a geographic distribution known only in Australia and Papua New Guinea [48]. *Candidatus* Borrelia tachyglossi* has previously been identified in one human-biting tick (*I. holocyclus*) removed from an echidna [38] and its prevalence in *I. holocyclus* or other human-biting ticks remains to be determined. The morphological characteristics and the pathogenicity of this bacterium are also unknown.

**DESCRIPTION OF ‘CANDIDATUS BORRELLA TACHYGLOSSI’**

*Candidatus* Borrelia tachyglossi (ta.chy.glos’si. N.L. gen. n. tachyglossi of *Tachyglossus aculeatus*, the monotreme host of the ticks in which the bacterium was first identified)).

Species can be differentiated from other borreliae based on sequence and phylogenetic analyses of five genomic loci (16S rRNA, flaB, groEL, gyrB and glpQ). Comparisons of the flaB gene sequences among the *Candidatus* Borrelia tachyglossi* genotypes showed two single nucleotide polymorphisms (SNPs) in genotype A at bases 485 and 519 (GenBank accession no. KY586966); and one SNP in genotype B at base 227 (KY586964). As for the groEL gene, analysis revealed two SNPs in genotype C at bases 656 and 1143 (KY586970). Analysis of the gyrB gene showed two SNPs in genotype B at bases 985 and 1416 (KY586972); and five SNPs in genotype C at bases 405, 757, 1075, 1093 and 1420 (KY586973). At the glpQ locus, genotype A (KY586968) and genotype B (KY586967) showed 12 base differences at bases 3, 43, 103, 112, 264, 515, 520, 530, 532, 702, 819 and 852.

The DNA G+C contents for 16S rRNA, flaB, groEL, gemb, and glpQ genes of *Candidatus* Borrelia tachyglossi* genotype A are 47.3, 40.9, 38.5, 33.8 and 34.1 mol%, respectively. The DNA G+C contents for 16S rRNA, flaB, groEL, gemb, and glpQ genes of *Candidatus* Borrelia tachyglossi* genotype B are 47.2, 40.9, 38.5, 33.9 and 34.7 mol%, respectively. The DNA G+C contents for 16S rRNA, flaB, groEL, and gemb genes of *Candidatus* Borrelia tachyglossi* genotype C are 47.2, 41, 38.5 and 34 mol%, respectively.

**Funding information**
This study was part-funded by the Australian Research Council (LP130100050), Bayer Healthcare, and Bayer Australia Ltd., and the School of Veterinary and Life Sciences Small Grants Scheme. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

**Acknowledgements**
The authors wish to acknowledge Frances Brigg and the Western Australian State Agriculture Biotechnology Centre for Sanger sequencing. The authors also acknowledge the veterinarians and veterinary nurses at the Australian Zoo Wildlife Hospital, Queensland, as well as the volunteers at Waagga Wagga, New South Wales, for the submission of tick samples.

**Conflicts of interest**
The authors declare that there are no conflicts of interest.

**References**


