Research article

Reconstruction of mitochondrial genomes from raw sequencing data provides insights on the phylogeny of *Ixodes* ticks but suggests the caution for species misidentification

Wessam Mohamed Ahmed Mohamed[1,2], Mohamed Abdallah Mohamed Moustafa[1,3*], Samuel Kelava[4], Dayana Barker[5], Keita Matsuno[6,7,8], Nariaki Nonaka[1], Renfu Shao[9], Ben J. Mans[10,11,12], Stephen C. Barker[4], Ryo Nakao[1*]

1. Department of Disease Control, Graduate School of Veterinary Medicine, Hokkaido University, Hokkaido 060-0818, Japan.
2. Division of Bioinformatics, Research Center for Zoonosis Control, Hokkaido University, Hokkaido 001-0020, Japan.
3. Department of Animal Medicine, Faculty of Veterinary Medicine, South Valley University, Qena, 83523, Egypt.
4. Department of Parasitology, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia.
5. School of Veterinary Science, University of Queensland, Gatton Qld, 4343, Australia
6. Division of Risk Analysis and Management, International Institute for Zoonosis Control, Hokkaido University, Sapporo 001-0020, Japan.
7. International Collaboration Unit, International Institute for Zoonosis Control, Hokkaido University, Sapporo 001-0020, Japan.
8. One Health Research Center, Hokkaido University, Sapporo 060-0818, Japan.
9. School of Science, Technology and Engineering, GeneCology Research Centre, University of the Sunshine Coast, Sippy Downs, Queensland 4558, Australia
10. Epidemiology, Parasites and Vectors, Agricultural Research Council-Onderstepoort Veterinary Research, Onderstepoort 0110, South Africa.
11. Department of Life and Consumer Sciences, University of South Africa, South Africa.
12. Department of Veterinary Tropical Diseases, University of Pretoria, Pretoria 0110, South Africa.

*Correspondence: ryo.nakao@vetmed.hokudai.ac.jp (R.N.), m.abdallah@vetmed.hokudai.ac.jp (M.A.M.M.)

Abstract

High-throughput sequencing (HTS) technology has profoundly been involved in sequencing whole genomes of several organisms in a fast and cost-effective manner. Although HTS provides an alternative biomonitoring method to the time-consuming and taxonomy-expertise dependent morphological approach,

still we cannot rule out the possibility of the impediment and misidentification biases. In this article we aim to retrieve whole mitochondrial genome (mitogenome) sequences from publicly available raw sequencing data for phylogenetic comparison of *Ixodes persulcatus*. For this comparison, we sequenced whole mitogenomes of four *I. persulcatus* ticks from Japan and constructed mitogenomes from raw sequencing data of 74 *I. persulcatus* ticks from China. Bayesian phylogenetic trees were inferred by the concatenated fifteen mitochondrial genes. We further tested our results by the phylogenetic analysis of cytochrome c oxidase subunit 1 (*cox1*) gene and internal transcribed spacer 2 (ITS2) sequences. Our findings showed that 70 constructed mitogenomes from China were clustered with the sequenced four mitogenomes of *I. persulcatus* from Japan. We also revealed that mitogenome sequences retrieved from two data sets CRR142297 and CRR142298 were clustered with *Ixodes nipponensis*. Moreover, other two mitogenome sequences from CRR142310 and CRR142311 formed a clade with *Ixodes pavlovskyi.* The phylogenetic analysis of *cox1* gene and ITS2 sequences confirmed the identification errors of these four samples. The overall phylogenetics in our study concluded that accurate morphological identification is necessary before implementing HTS to avoid any misidentification biases.

Keywords: High-throughput sequencing; *Ixodes persulcatus*; mitogenome; phylogenetic analysis.

1. Introduction

Ticks are the second most important vectors of human and animal pathogens among arthropods after mosquitoes. For instance, *Ixodes persulcatus* in Asia and Russia is a vector for tick-borne encephalitis virus (TBEV) (Mansfield et al., 2009) and Lyme borreliae (Murase et al., 2013). Climate change and the expansion of the human-animal interface have contributed to the increase of emerging tick-borne diseases all over the globe (Cavicchioli et al., 2019). Therefore, correct identification of several tick species is crucial for studying ticks and tick-borne pathogens (TBPs).

To date, the rapid evolution of high-throughput sequencing (HTS) platforms have provideed opportunities to have insights in many fields related to genomics (Lee et al., 2013). For example, it has allowed the description of microbial communities of ticks from different ecosystems (Estrada-Peña et al., 2018; Tokarz et al., 2019), the characterization of TBEV (Paulsen et al., 2021) and the detection of several viral populations in ticks (Qiu et al., 2019). In addition, these platforms have contributed to sequencing mitochondrial

genomes (mitogenomes) from various species of invertebrates (McKenna et al., 2010) including ticks (Burnard and Shao, 2019). However, phylogenetic relationships of ticks based on mitochondrial and nuclear genomes remain uncertain due to the limited number of genomes available (Liu et al., 2018). Nevertheless, it is challenging to sequence and assemble the nuclear genomes of ticks if compared to the mitochondrial genomes (Nene, 2009). Thus, mitogenomes are used to infer phylogenetic relationships of ticks as well as insects at different taxonomic levels (Burger et al., 2013; Burger et al., 2014; Carpi et al., 2016; Kelava et al., 2021; Mans et al., 2019; Mans et al., 2015; Mans et al., 2021; McCooke et al., 2015; Nakao et al., 2021). Recently, a hybrid assembly approach that utilizes Illumina short-read and Pacific Biosciences long-read data was applied for the genome studies of several tick species including *I. persulcatus*, *Haemaphysalis longicornis*, *Dermacentor silvarum*, *Hyalomma asiaticum*, *Rhipicephalus sanguineus,* and *Rhipicephalus annulatus* (Guerrero et al., 2019, 2021; Jia et al., 2020). Although the output of these studies can provide key features of the tick metabolism, population structure and genetic diversity of ticks and the associated pathogen composition and ecology, bias is possible in sequence data due to the under-sampled regions that may lead to loss of important regions during assembly (Ross et al., 2013). Proper data processing is also prerequisite as exemplified by the misidentification of TBPs due to poor knowledge on tick microbiome (Buysse and Duron, 2021). Nevertheless, making HTS data available for the public can significantly improve the scientific progress (Resnik, 2010). That is, if the data and supporting information are accurate and available to the scientific community. A good example for utilizing publicly available data resources is the recent phylogenomic analysis of the inward rectifier potassium (Kir) channels in ticks (Saelao et al., 2021). The aim of our study was to reconstruct mitogenome sequences of 74 *I. persulcatus* genomic sequences publicly available at the Genomic Sequence Archive (GSA) and to make a comparison with those from four *I. persulcatus* ticks sampled from different regions of Japan. The results provided an insight of the phylogenetic relationships of *I. persulcatus* with other *Ixodes* species but suggests the caution for the biases possibly caused by species misidentification in some samples.

2. Materials and methods

2.1 Specimen collection and DNA extraction

A total of four adult ticks were collected from Hokkaido prefecture (n = 2), Nagano prefecture (n = 1), Fukushima prefecture (n = 1), Japan between 2013 and 2014. The collected ticks were morphologically identified as two adult males and two adult females of *I. persulcatus* based on a standard key under a

97 stereomicroscope (Yamaguti et al., 1971). More specifically, in order to distinguish *I. persulcatus* from
98 *Ixodes pavlovskyi* and *Ixodes nipponensis*, the following morphological characteristics were investigated:
99 the length of the internal spur of coxa I, the shape of the spiracular plate, the length and color of the legs,
100 and the apex of the hypostome as reported elsewhere (Nakao et al., 1992). Individual tick specimens were
101 cut into half with a sterile blade. A half was crushed with stainless beads using a Micro Smash MS-100R
102 (TOMY, Tokyo, Japan) at 2,500 rpm for 30 s. The DNA was extracted using a blackPREP Tick DNA/RNA
103 Kit (Analytikjena, Germany) according to the manufacturer's instructions.
104 *2.2 Construction of NGS libraries and whole mitogenome sequencing*
105 The entire mitogenome sequence of *I. persulactus* was amplified in two overlapping PCRs (long-range and
106 short). Long-range PCR primers: mtG_K23: 5′-TCCTACATGATCTGAGTTYAGACCG-3′ and mtG_K26:
107 (5′- ACGGGCGATATGTRCATATTTTAGAGC-3′) and short PCR primers (I_gap_F3: 5′-
108 TTTYWAATTAAGATAGAAACCAACCTG-3′ and I_gap_R3 5′-AAATGTAAGGAGCATCACTCADA-
109 3′) were designed by aligning complete mitogenomes of genus *Ixodes* deposited in the database. The long-
110 range and short PCRs were performed as previously described (Kelava et al., 2021) with modifications.
111 Briefly, a 50 µl-reaction mixture of long-range PCR was performed containing 10 µl of 5 × PrimeSTAR
112 GXL Buffer (Mg^{2+} Plus) (TaKaRa Bio Inc., Shiga, Japan), 4.0 µl of dNTP Mixture (2.5 mM each), 200 nM
113 of each primer, 1.0 µl of PrimeSTAR® GXL DNA Polymerase (TaKaRa Bio Inc.), and 2.0 µl of template
114 DNA. The reaction conditions were 45 cycles of 98 °C for 10 s, 60 °C for 15 s, and 68 °C for 10 min. Short
115 PCR was performed in a 25 µl-reaction mixture containing 12.5 µl of 2 × Gflex PCR Buffer ($Mg^{2+,}$ dNTP
116 plus) (TaKaRa Bio Inc.), 0.5 µl of Tks Gflex DNA Polymerase (1.25 units/µl) (TaKaRa Bio Inc.), 200 nM
117 of each primer, and 1.0 µl of template DNA. The reaction conditions were 94 °C for 60 s, 45 cycles of 98 °C
118 for 10 s, 55 °C for 15 s, 68 °C for 60s, and a final extension of 68 °C for 5 min. Product of the PCR were
119 analyzed by electrophoresis in a 1.5% agarose gel stained with Gel-Red™ (Biotium, Hayward, CA). PCR
120 products were purified with a NucleoSpin Gel and PCR Clean-Up Kit (TaKaRa Bio Inc.).
121 Illumina sequencing libraries were constructed from the purified amplicons of two universal tick
122 mitogenome PCRs (Kelava et al., 2021) using the Nextera DNA Library Prep Kit (Illumina, Hayward, CA)
123 and were sequenced with the MiSeq reagent kit v3 for 600 cycles on an Illumina MiSeq platform. Geneious
124 v10.2.6 (Biomatters Ltd., Auckland, New Zealand) was used to map the reads against a reference
125 mitochondrial genome (accession number: NC_004370). The complete mitogenome sequences obtained

were submitted to the DNA Data Bank of Japan (http://www.ddbj.nig.ac.jp) with the accession numbers LC595234-LC595237.

2.3 Retrieving mitogenome sequences from the database

We downloaded the forward and reverse raw sequence reads of the 74 *I. persulcatus* ticks sampled from China available at GSA (https://bigd.big.ac.cn/gsa/browse/CRA002715). Pair-end raw sequence reads were paired, merged and mapped to *I. persulcatus* reference mitogenome sequence (accession number: NC_004370) using Geneious v10.2.6.

2.4 Phylogenetic analysis of the mitogenome sequences

Consensus sequences were aligned with the four assembled *I. persulcatus* mitogenomes from Japan, and reference mitochondrial genomes of *I. persulcatus* (accession numbers: NC_004370 and KU935457), *I. nipponensis* (accession number: MT371808), *I. pavlovskyi* (accession numbers: NC_023831 and LC595233), *Ixodes simplex* (accession number: KY457531), *Ixodes hexagonus* (accession number: AF081828), *Ixodes rubicundus* (accession number: KY457530), *Ixodes holocyclus* (accession number: MH043266), *Ixodes uriae* (accession number: AB087746), *Ixodes ricinus* (accession number: KF197132) and *Ixodes tasmani* (accession number: MH043271). Fifteen mitochondrial genes were used to infer phylogenetic trees. We concatenated the sequences of 13 protein-coding (*cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *cytb*, *atp6*, and *atp8*) and two ribosomal RNA (12S rRNA and 16S rRNA) genes and used the concatenated sequence to construct phylogenetic trees. We used BEAST version 1.4, a program for Bayesian analysis of molecular sequences using Markov Chain Monte Carlo (MCMC) to create the Bayesian phylogeographic trees (Drummond and Rambaut, 2007).

2.5 Analysis of cox1 gene and internal transcribed spacer 2 (ITS2) sequences

We inferred a phylogeny from the *cox1* gene, which is considered a reliable marker for identifying tick species using the maximum clade credibility (MCC) phylogenetic analysis. A total of 74 *cox1* gene sequences were obtained from the raw reads of the *I. persulcatus* ticks reported in Jia et al (2020), after being merged and mapped to *I. persulcatus* reference *cox1* gene sequence (accession number: NC_004370) using Geneious v10.2.6. The *cox1* gene sequences were aligned and the MCC phylogenetic trees were inferred using a generalized skyline plot model, that was embedded in a Bayesian MCMC analysis in BEAST version 1.4.

To exclude the possibility that mitochondrial introgression has recently occurred in these ticks, we also

studied the internal transcribed spacer 2 (ITS2) sequences of CRR142297, CRR142298, CRR142310, CRR142311, CRR142260, and CRR142300. Briefly, the merged raw reads were mapped to *I. persulcatus* reference ITS2 sequence (accession number: JQ737128) using Geneious v10.2.6. Consensus sequences were aligned with the reference ITS2 sequences of *I. persulcatus* (accession numbers: JQ625713, JF703107, JQ737128, AB032834, D88868, and D88874)*, I. nipponensis* (accession numbers: D88851, D88846, and D88850*)*, *I. pavlovskyi* (accession numbers: D88859, and D88860), *Ixodes laguri* (accession number: MF979542). The 18 ITS2 sequences were aligned with the MAFFT analytical tool (Katoh and Standley, 2013). Neighbor-joining consensus phylogenetic tree was constructed by the Tamura-Nei 93 as a substitution model, with 1,000 bootstrap replications.

3. Results and discussion

The mean number of total pair-end reads obtained from MiSeq was 282,721 per sample. After filtration, an average of 65.4% of reads were successfully mapped against the reference mitogenome. The length of complete mitogenomes of this study ranged between 14,542 bp and 14,549 bp with a mean sequencing depth of × 2,709 (min = × 947, max = × 4,407). Each mitogenome encoded 13 protein-coding, two ribosomal RNA (rRNA) (12S and 16S), and 22 transfer RNA genes with one non-coding control region in the same arrangement with that of the *I. persulcatus* reference mitogenome sequence (accession number: NC_004370). Mapping of 74 published sample sequences from China was successful in detecting the mitogenomes with the length between 14,543 bp and 14,571 bp (Table 1). In average, 0.001095% (min = 0.000039%, max = 0.004567%) of the reads were mapped on each mitogenome and the sequencing depth ranged between × 11 and × 2345 (mean = × 469). The detailed results of the mapping are provided in Supplementary Table S1. All genes were encoded in accordance with *I. persulcatus* reference mitogenome sequence.

MCC phylogenetic tree of the consensus mitogenomes revealed that sequences CRR142297 and CRR142298 (identified as *I. persulcatus* (Jia et al., 2020)) clustered with *I. nipponensis*. Moreover, two other sequences, CRR142310 and CRR142311 (also identified as *I. persulcatus* (Jia et al., 2020)) clustered with *I. pavlovskyi* (Figure 1). In addition, MCC phylogenetic analysis of 90 *cox1* gene sequences (1,534 nt) confirmed the incorrect identification of the above-mentioned four samples (Figure 2). Moreover, neighbor-joining consensus phylogenetic analysis revealed that ITS2 sequences from CRR142297 and CRR142298 clustered with *I. nipponensis*. Similarly, ITS2 sequences from CRR142310 and CRR142311 clustered with *I. pavlovskyi* reference sequences separately from *I. persulcatus* reference sequences (Figure 3).

Although *I. nipponensis* is more common in Korea and Japan, it has been detected from Hunan province in China (Cheng et al., 2018). In addition, *I. pavlovskyi* was recorded among the *Ixodes* ticks of China for the first time in 2016 from samples deposited in the medical entomology gallery of China (Guo et al., 2016). Unfortunately, the authors of Jia et al (2020) have not provided clear supplementary data that will accurately identify the geographic localities of the specimens used for genome sequencing. However, we noticed that the sample IDs of each misidentified *I. nipponensis* (JLip#) and *I. pavlovskyi* (NXip#) were similar (Figure 1), which may suggest that these samples were probably collected from the same geographical locations. However, natural hybridization between *I. persulcatus* and *I. pavlovskyi* has been reported from Siberia (Rar et al., 2019) where putative hybrids (i.e. identified as one species based on the morphological appearance and to another species based on *cox1* gene and ITS2) were detected. This phenomenon could explain the misidentification of CRR142310 and CRR142311, but to the best of our knowledge, no natural hybridization between *I. persulcatus* and *I. nipponensis* has ever been reported. Thus, the misidentification of CRR142297 and CRR142298 is still difficult to explain, especially since mitochondrial introgression could be excluded based on co-segregation of the *cox1* and ITS2 genes.

Although *I. persulcatus, I. pavlovskyi*, and *I. nipponensis* can carry similar communities of TBPs (Masuzawa et al., 1999; Seo et al., 2021; St. John et al., 2021), the prevalence of these pathogens are highly variable between these tick species (Rar et al., 2017) as well between the vertebrate hosts of ticks (Swei and Kwan, 2017). Hence, accurate identification of ticks is essential for providing correct epidemiology of the associated TBPs.

We have not scrutinised the identification of other species studied before, but we urge caution when using the publicly available data from this study. It may also be noted that while the quality of the genomes published in in Jia et al. (2020) are reported as above 90% complete using BUSCO analysis as criteria, the deposited protein coding datasets show completeness from 60-80% as determined with BUSCO, which put these genomes on par with that of other sequenced tick genomes (Mans, 2020). This would imply that while the genes are present in the assemblies, the identification and extraction of the protein coding sequences, and the annotation is currently incomplete.

Funding

This research was supported by KAKENHI (16H06431, 19H03118, 19F19097, 20K21358 and 20KK0151) and the Japan Program for Infectious Diseases Research and Infrastructure (20wm0225016j0001) from the

Japan Agency for Medical Research and Development (AMED).

References

Burger, T.D., Shao, R.F., Barker, S.C., 2013. Phylogenetic analysis of the mitochondrial genomes and nuclear rRNA genes of ticks reveals a deep phylogenetic structure within the genus *Haemaphysalis* and further elucidates the polyphyly of the genus *Amblyomma* with respect to *Amblyomma sphenodonti* and *Amblyomma elaphense*. Ticks Tick Borne Dis. 4, 265-274.

Burger, T.D., Shao, R.F., Labruna, M.B., Barker, S.C., 2014. Molecular phylogeny of soft ticks (Ixodida: Argasidae) inferred from mitochondrial genome and nuclear rRNA sequences. Ticks Tick Borne Dis. 5, 195-207.

Burnard, D., Shao, R., 2019. Mitochondrial genome analysis reveals intraspecific variation within Australian hard tick species. Ticks Tick Borne Dis. 10, 677-681.

Buysse, M., Duron, O., 2021. Evidence that microbes identified as tick-borne pathogens are nutritional endosymbionts. Cell 184, 2259-2260.

Carpi, G., Kitchen, A., Kim, H.L., Ratan, A., Drautz-Moses, D.I., McGraw, J.J., Kazimirova, M., Rizzoli, A., Schuster, S.C., 2016. Mitogenomes reveal diversity of the European Lyme borreliosis vector *Ixodes ricinus* in Italy. Mol. Phylogenet. Evol. 101, 194-202.

Cavicchioli, R., Ripple, W.J., Timmis, K.N., Azam, F., Bakken, L.R., Baylis, M., Behrenfeld, M.J., Boetius, A., Boyd, P.W., Classen, A.T., Crowther, T.W., Danovaro, R., Foreman, C.M., Huisman, J., Hutchins, D.A., Jansson, J.K., Karl, D.M., Koskella, B., Mark Welch, D.B., Martiny, J.B.H., Moran, M.A., Orphan, V.J., Reay, D.S., Remais, J.V., Rich, V.I., Singh, B.K., Stein, L.Y., Stewart, F.J., Sullivan, M.B., van Oppen, M.J.H., Weaver, S.C., Webb, E.A., Webster, N.S., 2019. Scientists' warning to humanity: microorganisms and climate change. Nat. Rev. Microbiol. 17, 569-586.

Cheng, T.Y., Chen, Z., Li, Z.B., Liu, G.H., 2018. First report of *Ixodes nipponensis* infection in goats in China. Vector Borne Zoonotic Dis. 18, 575-578.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

241 Estrada-Peña, A., Cabezas-Cruz, A., Pollet, T., Vayssier-Taussat, M., Cosson, J.-F., 2018. High throughput
242 sequencing and network analysis disentangle the microbial communities of ticks and hosts within and
243 between ecosystems. Front. Cell Infect. Microbiol. 8, 236.
244 Guerrero, F.D., Bendele, K.G., Ghaffari, N., Guhlin, J., Gedye, K.R., Lawrence, K.E., Dearden, P.K., Harrop,
245 T.W.R., Heath, A.C.G., Lun, Y., Metz, R.P., Teel, P., Perez de Leon, A., Biggs, P.J., Pomroy, W.E., Johnson,
246 C.D., Blood, P.D., Bellgard, S.E., Tompkins, D.M., 2019. The Pacific Biosciences *de novo* assembled
247 genome dataset from a parthenogenetic New Zealand wild population of the longhorned tick, *Haemaphysalis*
248 *longicornis* Neumann, 1901. Data Brief 27, 104602.
249 Guerrero, F.D., Ghaffari, N., Bendele, K.G., Metz, R.P., Dickens, C.M., Blood, P.D., Tidwell, J., Miller, R.J.,
250 de León, A.A.P., Teel, P.D., Johnson, C.D., 2021. Raw pacific biosciences and illumina sequencing reads
251 and assembled genome data for the cattle ticks *Rhipicephalus microplus* and *Rhipicephalus annulatus*. Data
252 Brief 35, 106852.
253 Guo, Y., Sun, Y., Xu, R., 2016. The genus *Ixodes* (Acari: Ixodidae) in China with three new record species.
254 Acta Parasitol. 61, 729-742.
255 Jia, N., Wang, J., Shi, W., Du, L., Sun, Y., Zhan, W., Jiang, J.F., Wang, Q., Zhang, B., Ji, P., Bell-Sakyi, L.,
256 Cui, X.M., Yuan, T.T., Jiang, B.G., Yang, W.F., Lam, T.T., Chang, Q.C., Ding, S.J., Wang, X.J., Zhu, J.G.,
257 Ruan, X.D., Zhao, L., Wei, J.T., Ye, R.Z., Que, T.C., Du, C.H., Zhou, Y.H., Cheng, J.X., Dai, P.F., Guo, W.B.,
258 Han, X.H., Huang, E.J., Li, L.F., Wei, W., Gao, Y.C., Liu, J.Z., Shao, H.Z., Wang, X., Wang, C.C., Yang,
259 T.C., Huo, Q.B., Li, W., Chen, H.Y., Chen, S.E., Zhou, L.G., Ni, X.B., Tian, J.H., Sheng, Y., Liu, T., Pan,
260 Y.S., Xia, L.Y., Li, J., Tick, G., Microbiome, C., Zhao, F., Cao, W.C., 2020. Large-scale comparative analyses
261 of tick genomes elucidate their genetic diversity and vector capacities. Cell 182, 1328-1340.e13.
262 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements
263 in performance and usability. Mol. Biol. Evol. 30, 772-780.
264 Kelava, S., Mans, B.J., Shao, R., Moustafa, M.A.M., Matsuno, K., Takano, A., Kawabata, H., Sato, K., Fujita,
265 H., Ze, C., Plantard, O., Hornok, S., Gao, S., Barker, D., Barker, S.C., Nakao, R., 2021. Phylogenies from
266 mitochondrial genomes of 120 species of ticks: Insights into the evolution of the families of ticks and of the
267 genus *Amblyomma*. Ticks Tick Borne Dis. 12, 101577.
268 Lee, C.-Y., Chiu, Y.-C., Wang, L.-B., Kuo, Y.-L., Chuang, E.Y., Lai, L.-C., Tsai, M.-H., 2013. Common
269 applications of next-generation sequencing technologies in genomic research. Transl. Cancer Res. 2, 33-45.

Liu, Z.Q., Liu, Y.F., Kuermanali, N., Wang, D.F., Chen, S.J., Guo, H.L., Zhao, L., Wang, J.W., Han, T., Wang, Y.Z., Wang, J., Shen, C.F., Zhang, Z.Z., Chen, C.F., 2018. Sequencing of complete mitochondrial genomes confirms synonymization of *Hyalomma asiaticum asiaticum* and *kozlovi*, and advances phylogenetic hypotheses for the *Ixodidae*. PLOS One 13, e0197524.

Mans, B., Featherston, J., Kvas, M., Pillay, K.A., de Klerk, D.G., Pienaar, R., de Castro, M.H., Schwan, T.G., Lopez, J.E., Teel, P., de Leon, A.A.P., Sonenshine, D.E., Egekwu, N.I., Bakkes, D.K., Heyne, H., Kanduma, E.G., Nyangiwe, N., Bouattour, A., Latif, A.A., 2019. Argasid and ixodid systematics: Implications for soft tick evolution and systematics, with a new argasid species list. Ticks Tick Borne Dis. 10, 219-240.

Mans, B.J., 2020. Quantitative visions of reality at the tick-host interface: biochemistry, genomics, proteomics, and transcriptomics as measures of complete inventories of the tick sialoverse. Front. Cell Infect. Microbiol. 10, 574405-574405.

Mans, B.J., de Klerk, D., Pienaar, R., de Castro, M.H., Latif, A.A., 2015. Next-generation sequencing as means to retrieve tick systematic markers, with the focus on *Nuttalliella namaqua* (Ixodoidea: Nuttalliellidae). Ticks Tick Borne Dis. 6, 450-462.

Mans, B.J., Kelava, S., Pienaar, R., Featherston, J., de Castro, M.H., Quetglas, J., Reeves, W.K., Durden, L.A., Miller, M.M., Laverty, T.M., Shao, R., Takano, A., Kawabata, H., Moustafa, M.A.M., Nakao, R., Matsuno, K., Greay, T.L., Evasco, K.L., Barker, D., Barker, S.C., 2021. Nuclear (18S-28S rRNA) and mitochondrial genome markers of *Carios* (*Carios*) *vespertilionis* (Argasidae) support *Carios Latreille*, 1796 as a lineage embedded in the Ornithodorinae: re-classification of the *Carios sensu* Klompen and Oliver (1993) clade into its respective subgenera. Ticks Tick Borne Dis. 12, 101688.

Mansfield, K.L., Johnson, N., Phipps, L.P., Stephenson, J.R., Fooks, A.R., Solomon, T., 2009. Tick-borne encephalitis virus - a review of an emerging zoonosis. J. Gen. Virol. 90(Pt 8), 1781-1794.

Masuzawa, T., Fukui, T., Miyake, M., Oh, H.B., Cho, M.K., Chang, W.H., Imai, Y., Yanagihara, Y., 1999. Determination of members of a *Borrelia afzelii*-related group isolated from *Ixodes nipponensis* in Korea as *Borrelia valaisiana.* Int. J. Syst. Bacteriol. 49, 1409-1415.

McCooke, J.K., Guerrero, F.D., Barrero, R.A., Black, M., Hunter, A., Bell, C., Schilkey, F., Miller, R.J., Bellgard, M.I., 2015. The mitochondrial genome of a Texas outbreak strain of the cattle tick, *Rhipicephalus* (*Boophilus*) *microplus*, derived from whole genome sequencing Pacific Biosciences and Illumina reads. Gene 571, 135-141.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297-1303.

Murase, Y., Konnai, S., Githaka, N., Hidano, A., Taylor, K., Ito, T., Takano, A., Ando, S., Kawabata, H., Tsubota, T., Murata, S., Ohashi, K., 2013. Prevalence of Lyme borrelia in *Ixodes persulcatus* ticks from an area with a confirmed case of Lyme disease. J. Vet. Med. Sci. 75, 215-218.

Nakao, M., Miyamoto, K., Kitaoka, S., 1992. A new record of *Ixodes pavlovskyi* Pomerantzev from Hokkaido, Japan (Acari: Ixodidae). Med. Entomol. Zool. 43, 229-234.

Nakao, R., Shinjo, K., Sakiyama, T., Ogata, S., Kusakisako, K., Kinoshita, G., Naguib, D., Chatanga, E., Mohamed, W.M.A., Moustafa, M.A.M., Matsuno, K., Ito, T., Nonaka, N., Sashika, M., Tsubota, T., Shimozuru, M., 2021. *Amblyomma testudinarium* infestation on a brown bear (*Ursus arctos yesoensis*) captured in Hokkaido, a northern island of Japan. Parasitol. Int. 80, 102209.

Nene, V., 2009. Tick genomics--coming of age. Front Biosci (Landmark edition) 14, 2666-2673.

Paulsen, K.M., Lamsal, A., Bastakoti, S., Pettersson, J.H., Pedersen, B.N., Stiasny, K., Haglund, M., Smura, T., Vapalahti, O., Vikse, R., Alfsnes, K., Andreassen Å, K., 2021. High-throughput sequencing of two European strains of tick-borne encephalitis virus (TBEV), Hochosterwitz and 1993/783. Ticks Tick Borne Dis. 12, 101557.

Qiu, Y., Abe, T., Nakao, R., Satoh, K., Sugimoto, C., 2019. Viral population analysis of the taiga tick, *Ixodes persulcatus*, by using Batch Learning Self-Organizing Maps and BLAST search. J. Vet. Med. Sci. 81, 401-410.

Rar, V., Livanova, N., Sabitova, Y., Igolkina, Y., Tkachev, S., Tikunov, A., Babkin, I., Golovljova, I., Panov, V., Tikunova, N., 2019. *Ixodes persulcatus*/*pavlovskyi* natural hybrids in Siberia: Occurrence in sympatric areas and infection by a wide range of tick-transmitted agents. Ticks Tick Borne Dis. 10, 101254.

Rar, V., Livanova, N., Tkachev, S., Kaverina, G., Tikunov, A., Sabitova, Y., Igolkina, Y., Panov, V., Livanov, S., Fomenko, N., Babkin, I., Tikunova, N., 2017. Detection and genetic characterization of a wide range of infectious agents in *Ixodes pavlovskyi* ticks in Western Siberia, Russia. Parasit. Vectors 10, 258.

Resnik, D.B., 2010. Genomic research data: open vs. restricted access. IRB 32, 1-6.

Ross, M.G., Russ, C., Costello, M., Hollinger, A., Lennon, N.J., Hegarty, R., Nusbaum, C., Jaffe, D.B., 2013. Characterizing and measuring bias in sequence data. Genome Biol. 14, R51.

Saelao, P., Hickner, P.V., Bendele, K.G., Pérez de León, A.A., 2021. Phylogenomics of Tick Inward Rectifier Potassium Channels and Their Potential as Targets to Innovate Control Technologies. Front. Cell Infection Microbiol. 11, 647020-647020.

Seo, M.-G., Kwon, O.-D., Kwak, D., 2021. Molecular detection of *Rickettsia raoultii*, *Rickettsia tamurae*, and associated pathogens from ticks parasitizing water deer (*Hydropotes inermis argyropus*) in South Korea. Ticks Tick Borne Dis. 12, 101712.

St. John, H.K., Masuoka, P., Jiang, J., Takhampunya, R., Klein, T.A., Kim, H.-C., Chong, S.-T., Song, J.-W., Kim, Y.-J., Farris, C.M., Richards, A.L., 2021. Geographic distribution and modeling of ticks in the Republic of Korea and the application of tick models towards understanding the distribution of associated pathogenic agents. Ticks Tick Borne Dis. 12, 101686.

Swei, A., Kwan, J.Y., 2017. Tick microbiome and pathogen acquisition altered by host blood meal. ISME J. 11, 813-816.

Tokarz, R., Tagliafierro, T., Sameroff, S., Cucura, D.M., Oleynik, A., Che, X., Jain, K., Lipkin, W.I., 2019. Microbiome analysis of *Ixodes scapularis* ticks from New York and Connecticut. Ticks Tick Borne Dis. 10, 894-900.

Yamaguti, N., Tipton, V.J., Keegan, H.L., Toshioka, S., 1971. Ticks of Japan, Korea, and the Ryukyu Islands. Brigham Young University Science Bulletin, Biological Series 15 : 1. 1.

Zhao, G.-P., Wang, Y.-X., Fan, Z.-W., Ji, Y., Liu, M.-j., Zhang, W.-H., Li, X.-L., Zhou, S.-X., Li, H., Liang, S., Liu, W., Yang, Y., Fang, L.-Q., 2021. Mapping ticks and tick-borne pathogens in China. Nat. Commun. 12, 1075.

Figure Legends

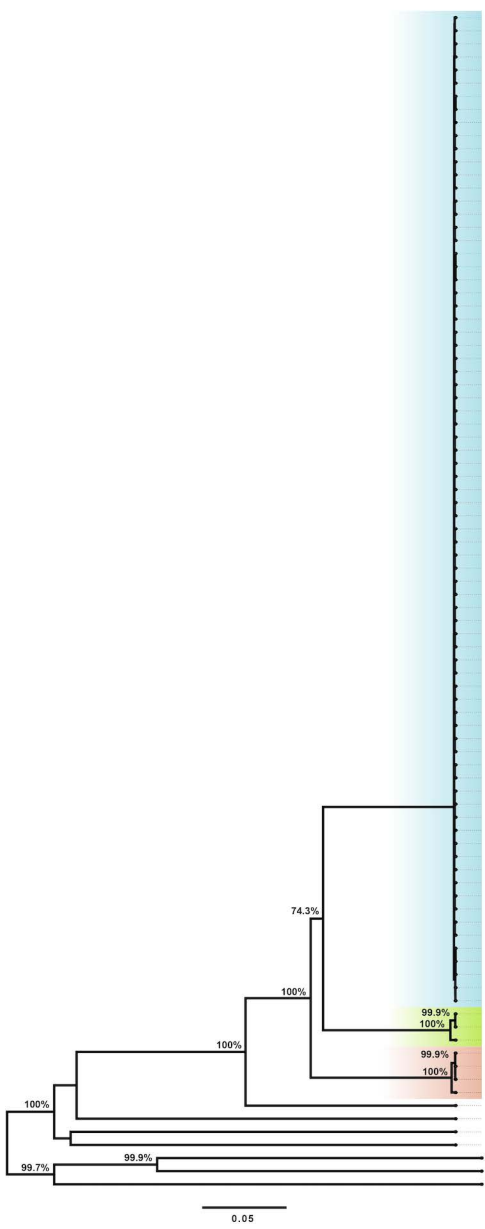
Figure 1. Bayesian Maximum Credibility (MCC) tree of 15 concatenated mitochondrial gene sequences of 10 *Ixodes* tick species. Cyan, light red and green colors highlight the *Ixodes persulcatus*, *I. pavlovskyi,* and *I. nipponensis* clades, respectively. The other 7 *Ixodes* species; *I. simplex*, *I. hexagonus*, *I. rubicundus, I. holocyclus, I. uriae, I. ricinus* and *I. tasmani* are not highlighted. Brown and cyan colored text (sequence names) indicates sequences obtained from China and Japan, respectively.

Figure 2. Bayesian Maximum Credibility (MCC) tree of *cox1* gene sequences (1,534 bp) of 10 *Ixodes* species. Cyan, light red and green colors highlight the *Ixodes persulcatus*, *I. pavlovskyi,* and *I. nipponensis*

clades, respectively. The other 7 *Ixodes* species; *I. simplex*, *I. hexagonus*, *I. rubicundus, I. holocyclus, I. uriae, I. ricinus* and *I. tasmani* are not highlighted. Brown and cyan colored text (sequence names) indicates sequences obtained from China and Japan, respectively.

Figure 3. Neighbor-joining consensus phylogenetic tree of ITS2 sequences of four *Ixodes* species. Bootstrap values are shown on each node. Cyan, light red and green colors highlight the *Ixodes persulcatus*, *I. pavlovskyi,* and *I. nipponensis* clades, respectively. Text (sequence names) in brown indicates sequences obtained from China.


CRR142253 HLJIp.830
CRR142288 HLJip.327
CRR142282 HLJip.264
CRR142287 HLJip.326
CRR142270 HLJip.169
CRR142296 JLIp.485
CRR142301 JLIp.623
CRR142272 HLJip.246
CRR142291 HLJip.347
CRR142278 HLJip.252
CRR142305 JLip.210
CRR142255 HLJip.101
CRR142279 HLJip.260
CRR142263 HLJip.160
CRR142242 HLJIp.809
CRR142254 HLJIp.831
CRR142250 HLJIp.827
CRR142266 HLJip.163
CRR142243 HLJIp.819
CRR142284 HLJip.266
CRR142304 JLIp.209
CRR142292 HLJip.77
CRR142293 HLJip.78
CRR142259 HLJip.106
CRR142261 HLJip.13
CRR142299 JLIp.621
CRR142280 HLJip.261
CRR142309 NMGip.245
CRR142273 HLJip.247
CRR142248 HLJIp.825
CRR142249 HLJIp.826
CRR142252 HLJIp.829
CRR142264 HLJip.161
CRR142283 HLJip.265
CRR142306 JIip.253
CRR142307 LNip.256
CRR142267 HLJip.164
CRR142308 NMGip.244
CRR142245 HLJIp.821
CRR142244 HLJIp.820
CRR142251 HLJIp.828
CRR142268 HLJip.165
CRR142286 HLJip.325
CRR142269 HLJip.168
CRR142277 HLJip.251
CRR142246 HLJIp.822
CRR142275 HLJIp.249
CRR142247 HLJIp.824
CRR142262 HLJip.159
CRR142281 HLJip.262
CRR142300 JLIp.622
CRR142302 JLIp.751
CRR142257 HLJip.103
CRR142289 HLJip.328
CRR142265 HLJip.162
CRR142285 HLJip.323
CRR142276 HLJip.250
CRR142290 HLJip.346
CRR142238 HLJIp.805
CRR142294 HLJip.79
CRR142260 HLJip.12
CRR142271 HLJip.213
CRR142239 HLJIp.806
CRR142241 HLJIp.808
CRR142258 HLJip.104
CRR142274 HLJip.248
CRR142295 HLJip.80
LC595237 I. persulcatus
LC595236 I. persulcatus
CRR142256 HLJip.102
LC595234 I. persulcatus
KU935457 I. persulcatus
NC_004370 I. persulcatus
CRR142303 JLIp.832
LC595235 I. persulcatus
CRR142240 HLJIp.807
CRR142310 NXip.288
CRR142311 NXip.341
LC595233 I. pavlovskyi
NC_023831 I. pavlovskyi
CRR142297 JLIp.590
CRR142298 JLIp.591
MT371808 I. nipponensis
KF197132 I. ricinus
AF081828 I. hexagonus
KY457531 I. simplex
KY457530 I. rubicundus
AB087746 I. uriae
MH043266 I. holocyclus
MH043271 I. tasmani
100%
0.05


LC595237 I. persulcatus
CRR142293 HLJip.78
CRR142250 HLJip.827
CRR142252 HLJip.829
CRR142289 HLJip.328
CRR142299 JLIp.621
LC595235 I. persulcatus
CRR142240 HLJIp.807
NC_004370 I. persulcatus
CRR142284 HLJip.266
CRR142304 JLip.209
CRR142259 HLJip.106
CRR142296 JLIp.485
CRR142278 HLJip.252
CRR142246 HLJip.822
CRR142295 HLJip.80
CRR142261 HLJip.13
CRR142277 HLJip.251
LC595234 I. persulcatus
LC595236 I. persulcatus
CRR142256 HLJip.102
CRR142249 HLJIp.826
CRR142264 HLJip.161
CRR142309 NMGip.245
CRR142242 HLJIp.809
CRR142266 HLJip.163
CRR142265 HLJip.162
CRR142270 HLJip.169
CRR142247 HLJIp.824
CRR142290 HLJip.346
CRR142292 HLJip.77
CRR142301 JLIp.623
CRR142243 HLJIp.819
CRR142308 NMGip.244
CRR142287 HLJip.326
CRR142305 JLip.210
CRR142303 JLIp.832
CRR142279 HLJip.260
CRR142302 JLIp.751
CRR142253 HLJIp.830
CRR142294 HLJip.79
CRR142262 HLJip.159
CRR142275 HLJip.249
CRR142269 HLJip.168
CRR142307 LNip.256
CRR142255 HLJip.101
CRR142276 HLJip.250
CRR142257 HLJip.103
CRR142288 HLJip.327
CRR142268 HLJip.165
CRR142283 HLJip.265
CRR142263 HLJip.160
CRR142273 HLJip.247
CRR142286 HLJip.325
CRR142291 HLJip.347
CRR142300 JLIp.622
CRR142306 JIip.253
CRR142238 HLJIp.805
CRR142248 HLJIp.825
CRR142244 HLJIp.820
CRR142282 HLJip.264
CRR142285 HLJip.323
KU935457 I. persulcatus
CRR142254 HLJIp.831
CRR142281 HLJip.262
CRR142267 HLJip.164
CRR142241 HLJIp.808
CRR142251 HLJIp.828
CRR142272 HLJip.246
CRR142245 HLJIp.821
CRR142280 HLJip.261
CRR142239 HLJIp.806
CRR142271 HLJip.213
CRR142260 HLJip.12
CRR142258 HLJip.104
CRR142274 HLJip.248
CRR142297 JLIp.590
CRR142298 JLIp.591
MT371808 I. nipponensis
CRR142310 NXip.288
CRR142311 NXip.341
NC_023831 I. pavlovskyi
LC595233 I. pavlovskyi
KF197132 I. ricinus
KY457531 I. simplex
AF081828 I. hexagonus
KY457530 I. rubicundus
AB087746 I. uriae
MH043266 I. holocyclus
MH043271 I. tasmani
74.3%
100%
100%
99.9%
100%
99.9%
100%
100%
99.9%
99.7%
0.05

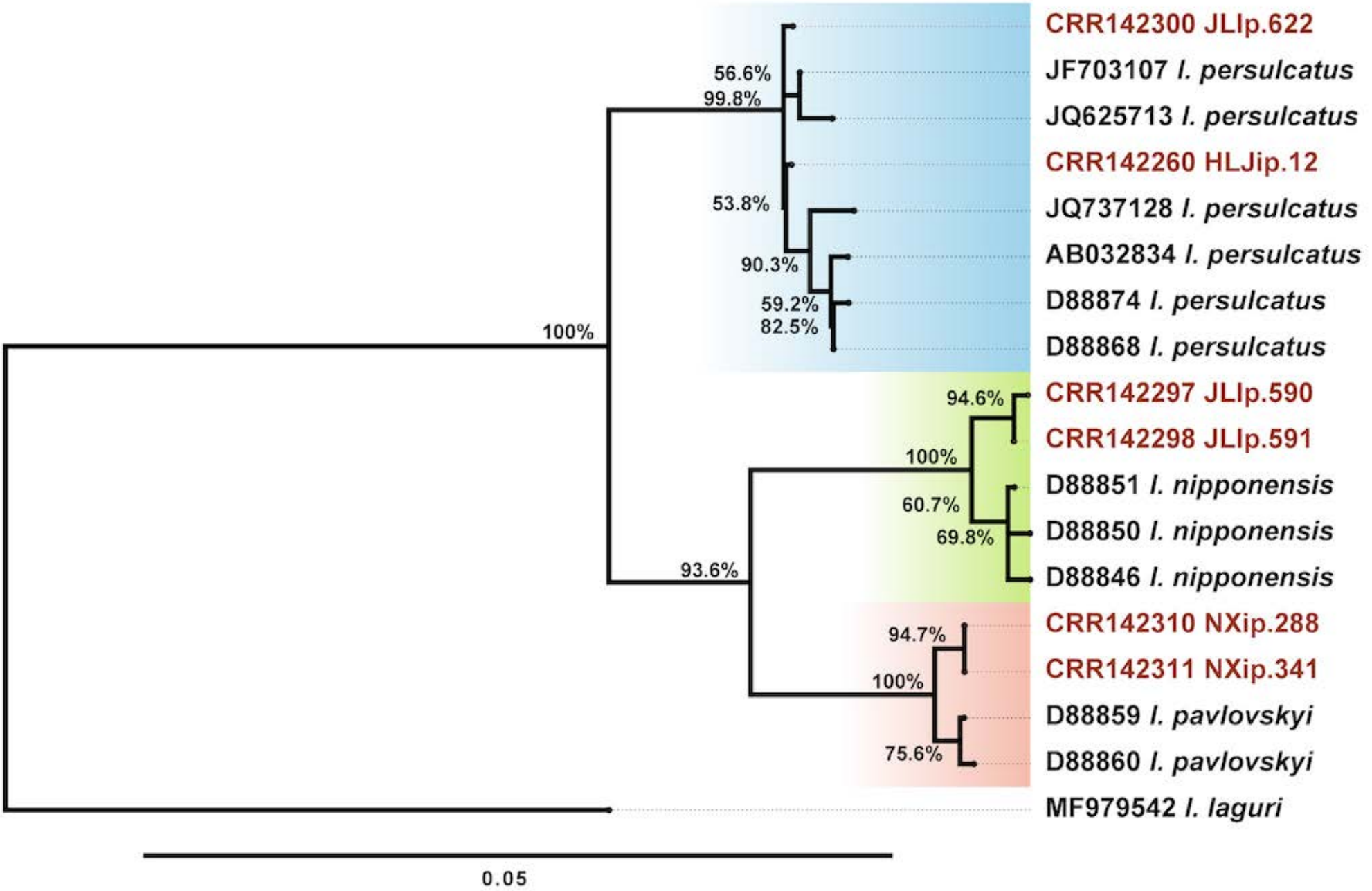

CRR142300 JLIp.622
JF703107 I. persulcatus
JQ625713 I. persulcatus
CRR142260 HLJip.12
JQ737128 I. persulcatus
AB032834 I. persulcatus
D88874 I. persulcatus
D88868 I. persulcatus
CRR142297 JLIp.590
CRR142298 JLIp.591
D88851 I. nipponensis
D88850 I. nipponensis
D88846 I. nipponensis
CRR142310 NXip.288
CRR142311 NXip.341
D88859 I. pavlovskyi
D88860 I. pavlovskyi
MF979542 I. laguri
56.6%
99.8%
53.8%
90.3%
59.2%
82.5%
100%
94.6%
100%
60.7%
69.8%
93.6%
94.7%
100%
75.6%
0.05

Table 1: Summary of mapping of high-throughput sequencing reads to *I. persulcatus* reference mitogenome sequence.

	Mean	Min.	Max.
Mitogenome length (bp)	14,547	14,543	14,571
Total number of reads	28,027,633	860,432	160,707,036
Number of reads mapped to mitogenome	30,153	623	194,603
Percent of mapped reads	0.001095	0.000039	0.004567
Mapped nucleotides	6,820,950	155,999	34,097,385
Sequencing depth	469	11	2,345