Current Research in Parasitology & Vector-Borne Diseases 6 (2024) 100229

Contents lists available at ScienceDirect

Current Research in Parasitology & Vector-Borne Diseases

journal homepage: www.sciencedirect.com/journal/current-research-in-parasitology-and-vector-borne-diseases

Similarities between *Ixodes ricinus* and *Ixodes inopinatus* genomes and horizontal gene transfer from their endosymbionts

Valérie O. Baede [a,*], Oumayma Jlassi [b,c,d], Paulina M. Lesiczka [e], Hend Younsi [f,g], Hans J. Jansen [h], Khalil Dachraoui [b,c], Jane Segobola [h], Mourad Ben Said [i,j], Wouter J. Veneman [h], Ron P. Dirks [h], Hein Sprong [a], Elyes Zhioua [b,c,**]

[a] Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
[b] Unit of Vector Ecology, Pasteur Institute of Tunis, Tunis, Tunisia
[c] Laboratory of Epidemiology and Veterinary Microbiology, Pasteur Institute of Tunis, Tunis, Tunisia
[d] Faculty of Sciences of Bizerte (FSB), University of Carthage, Tunis, Tunisia
[e] Centre for Monitoring of Vectors (CMV), Netherlands Institute for Vectors, Invasive Plants and Plant Health (NIVIP), Netherlands Food and Consumer Product Safety Authority (NVWA), Wageningen, the Netherlands
[f] Higher Institute of Applied Biological Sciences of Tunis, University Tunis El Manar, Tunis, Tunisia
[g] Laboratory of Biodiversity, Parasitology, and Ecology of Aquatic Ecosystems, University Tunis El Manar, Tunis, Tunisia
[h] Future Genomics Technologies BV, Leiden, the Netherlands
[i] Laboratory of Microbiology, National School of Veterinary Medicine of Sidi Thabet, University of Manouba, Manouba, Tunisia
[j] Department of Basic Sciences, Higher Institute of Biotechnology of Sidi Thabet, University of Manouba, Manouba, Tunisia

ARTICLE INFO

Keywords:
Ixodes ricinus complex
Pangenome comparative analyses
Whole genome sequencing
Lyme borreliosis
Spotted fever rickettsiosis
Next-generation sequencing

ABSTRACT

The taxa *Ixodes ricinus* and *Ixodes inopinatus* are sympatric in Tunisia. The genetics underlying their morphological differences are unresolved. In this study, ticks collected in Jouza-Amdoun, Tunisia, were morphologically identified and sequenced using Oxford Nanopore Technologies. Three complete genome assemblies of *I. inopinatus* and three of *I. ricinus* with BUSCO scores of ~98% were generated, including the reconstruction of mitochondrial genomes and separation of both alleles of the TRPA1, TROSPA and calreticulin genes. Deep sequencing allowed the first descriptions of complete bacterial genomes for "*Candidatus* Midichloria mitochondrii", *Rickettsia helvetica* and *R. monacensis* from North Africa, and the discovery of extensive integration of parts of the *Spiroplasma ixodetis* and "*Ca.* M. mitochondrii" into the nuclear genome of these ticks. Phylogenetic analyses of the mitochondrial genome, the nuclear genes, and symbionts showed differentiation between Tunisian and Dutch ticks, but high genetic similarities between Tunisian *I. ricinus* and *I. inopinatus*. Subtraction of the genome assemblies identified the presence of some unique sequences, which could not be confirmed when screening a larger batch of *I. ricinus* and *I. inopinatus* ticks using PCR. Our findings yield compelling evidence that *I. inopinatus* is genetically highly similar, if not identical, to sympatric *I. ricinus*. Defined morphological differences might be caused by extrinsic factors such as micro-climatic conditions or bloodmeal composition. Our findings support the existence of different lineages of *I. ricinus* as well of its symbionts/pathogens from geographically dispersed locations.

1. Background

The causative agents of the two major tick-borne diseases in Europe, Lyme borreliosis and tick-borne encephalitis, are transmitted by *Ixodes ricinus* (Linnaeus, 1758) (Sprong et al., 2018; Köhler et al., 2023). Additionally, *I. ricinus* is an important vector of emerging pathogens such as *Rickettsia helvetica*, *R. monacensis* and *Spiroplasma ixodetis* (Coipan et al., 2013; Krawczyk et al., 2022). As *I. ricinus* spends almost its entire life in the vegetation, its survival and activity strongly depend on (micro)climatic conditions, such as temperature and relative humidity (Lindgren et al., 2000; Medlock et al., 2013). Currently, *I. ricinus* is found all over Europe, including Scandinavia, central Europe and the

* Corresponding author.
** Corresponding author. Unit of Vector Ecology, Pasteur Institute of Tunis, Tunis, Tunisia.
E-mail addresses: valerie.baede@rivm.nl (V.O. Baede), elyes.zhioua@gmail.com (E. Zhioua).

https://doi.org/10.1016/j.crpvbd.2024.100229
Received 29 July 2024; Received in revised form 30 October 2024; Accepted 4 November 2024
Available online 6 November 2024

Mediterranean, with stable populations in various locations across northern Africa (Zhioua et al., 1999; Estrada-Peña et al., 2014; Wint et al., 2023). Its widespread geographical distribution makes *I. ricinus* an environmental generalist, yet it is unclear how it can adapt to such diverse climates.

One possibility is that *I. ricinus* exists as distinct lineages, each genetically adapted to regional conditions (Estrada-Peña et al., 2014). Well-defined phenotypic groups of *I. ricinus* have been described within each of the main European climate zones (Estrada-Peña et al., 1998; Estrada-Peña, 2001; Gilbert et al., 2014). Genetic investigations revealed limited genetic structuring within European *I. ricinus* populations, although these populations exhibit genetic distinctiveness from *I. ricinus* populations in northern Africa (Noureddine et al., 2011; Poli et al., 2020). Some North-African populations may correspond to the recently described *Ixodes inopinatus* (Estrada-Peña et al., 2014), which seems to be adapted to dry areas in the Mediterranean region (Estrada-Peña et al., 2014; Chitimia-Dobler et al., 2018).

Ixodes inopinatus presents clear morphological differences with *I. ricinus* (Estrada-Peña et al., 2014). Several attempts have been made to identify genetic markers to discriminate between *I. ricinus* and *I. inopinatus* (Estrada-Peña et al., 2014; Hekimoğlu, 2022; Hrazdilova et al., 2023; Rollins et al., 2023). While some studies delineated distinct lineages or clades for *I. ricinus* and *I. inopinatus* (Estrada-Peña et al., 2014; Chitimia-Dobler et al., 2018), others suggest that these distinctions arise from geographical clustering rather than genetic differentiation (Hekimoğlu, 2022; Hrazdilova et al., 2023; Rollins et al., 2023). To complicate the matter, these inconsistent findings may result from potential hybridization between these two taxa (Hekimoğlu, 2022; Hrazdilova et al., 2023). A second, but unexplored, possibility is the existence of a (relatively) subtle difference between *I. ricinus* and *I. inopinatus,* for example a mutation in a gene affecting morphogenesis (Alekseev and Dubinina, 2008; Chitimia-Dobler et al., 2017). Thirdly, morphological differences could potentially be explained by endosymbionts. In other organisms, endosymbionts have been shown to alter host morphology and reduce host vulnerability to elevated environmental temperatures (Gehman and Harley, 2019). To the best of our knowledge, while such effects have been demonstrated in several arthropods, evidence of similar phenomena in ticks remains lacking (Stouthamer et al., 1999; Tsuchida et al., 2014; Hirota et al., 2017). Nevertheless, endosymbionts can affect *I. ricinus* physiology and behaviour (Bakker et al., 2023).

In this study, we explore these three possibilities. So far, most phylogenetic studies have been based on specimens originating from different geographical areas, comparing specimens from mainland Europe with those from northern Africa. However, *I. ricinus* and *I. inopinatus* coexist in Tunisia (Younsi et al., 2020). Here, we aim to associate the morphological with genetic differences to better understand the differences between *I. ricinus* and *I. inopinatus* originating from the same biogeographical region with the same climatic conditions.

2. Materials and methods

2.1. Tick specimen collection and morphological identification

Adult ticks were collected by flagging the vegetation during March 2023 from the forest of Jouza-Amdoun (36°52′9.88"N, 9°0′8.42"E) located in the humid biogeographical area of northern Tunisia (Supplementary Fig. S1). Ticks were identified as *I. ricinus* or *I. inopinatus* based on morphology, using the available adult identification keys (Pérez-Eid, 2007; Estrada-Peña et al., 2014; Estrada-Peña, 2017; Chitimia-Dobler et al., 2018). Ticks were transported on dry ice and stored at −80 °C until DNA extraction.

2.2. DNA extraction and genome sequencing

Six whole individual ticks, three *I. ricinus* and three *I. inopinatus* were ground to a fine powder using a pestle and mortar with liquid nitrogen. The genomic-tip 20/G kit (Qiagen Benelux BV, Venlo, The Netherlands) was used to extract high molecular weight DNA. The quality of the DNA was analysed using Genomic DNA ScreenTape on an Agilent 4200 TapeStation System (Agilent Technologies Netherlands BV, Amstelveen, The Netherlands) and the quantity was measured using a Qubit 3.0 Fluorometer (Life Technologies Europe BV, Bleiswijk, The Netherlands). Nanopore sequencing libraries were prepared using the Ligation Sequencing Kit V14 (SQK-LSK114) according to the manufacturer's instructions (Oxford Nanopore Technologies, Oxford, UK). Each library was run on an R10.4.1 PromethION flowcell (FLO-PRO114M; Oxford Nanopore Technologies (ONT), Oxford, UK) and reloaded on a daily basis after a nuclease flush with Flow Cell wash kit (EXP-WSH004). Super-accuracy live basecalling was done using Guppy 6.1.5.

2.3. Bioinformatic analysis

Draft genome sequences were *de novo* assembled and corrected using Flye vs. 2.9.1-b1780 (Kolmogorov et al., 2019). Benchmarking Universal Single-Copy Orthologs (BUSCO) scores of the draft assemblies were calculated with Compleasm v.0.2.2 (Huang and Li, 2023), using default settings and an Arachnida-specific database of 2934 single-copy orthologs (eudicots_odb10). Custom data analysis was done using CLC Main Workbench v.23.0.4 or CLC Genomics Workbench v.23.0.4 (Qiagen, Aarhus, Denmark) and NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Contigs of mitochondrial DNA and nuclear genes TRPA1, TROSPA and calreticulin were identified in each *de novo* assembled tick genome sequence by blastn and tblastn queries using the following settings: Program = blastn; Match/Mismatch and Gap Costs = Match 2 Mismatch 3 Existence 5 Extension 2; Expectation value = 10.0; Word size = 11; Mask lower case = No; Mask low complexity regions = Yes; Filter out redundant results = No. This resulted in two alleles of a single-copy TRPA1 gene. For TROSPA and calreticulin, a more stringent e-value was chosen i.e. 1.0E-20 or 1.0E-25. Also, for these genes, only single copies were found. The following sequences were used as a reference: *I. ricinus* mitochondrial DNA sequence (GenBank: NC_018369.2), cDNA sequence (partial) of the *I. scapularis* TRPA1 gene (GenBank: XM_040502442) and amino acid sequence of the *Rhipicephalus microplus* TRPA1 protein (GenBank: XP_037273035.1), *I. ricinus* DNA sequence of TROSPA (GenBank: KF041821) and *I. ricinus* DNA sequence of calreticulin (GenBank: AY395272). TROSPA and calreticulin were annotated using CLC. All TRPA1-coding exons were manually annotated. In addition, TRPA1 genes of 4 Dutch *I. ricinus* ticks were annotated (ENA PRJEB70514 (Lesiczka et al., 2024); Supplementary Table S2).

2.4. Data analyses of Ixodes ticks

The genetic diversity among mitochondrial genomes was estimated using DnaSP 5.0 software (Rozas et al., 2017), as indexed by number of variable sites (S), parsimony informative sites, number of haplotypes, haplotype diversity (Hd), nucleotide diversity (π), and average number of nucleotide differences (k).

2.5. Genomic subtractions and confirmation of species-specific sequences

To identify candidate *I. ricinus*- and/or *I. inopinatus*-specific sequences, whole genomes of 3 *I. ricinus* and 3 *I. inopinatus* from this study were compared using improved genome assemblies (Flye v.2.9.3-b1797; Kolmogorov et al., 2019) with two different approaches. First, repeats in all 6 genomes were masked using RepeatMasker (http://www.repeatmasker.org). All *I. inopinatus* reads were aligned against the combined repeat-masked assemblies of *I. ricinus* and *vice versa*. Unaligned reads were mapped against the repeat library and unmapped reads were used for *de novo* contig assembly and then mapped to the combined, repeat-masked assembly from the other species. The remaining unmapped contigs were aligned against the combined, repeat-masked

assembly from the same species, creating an overview of species-specific contigs. Finally, all reads were aligned to the remaining candidate species-specific contigs and checked for species-specificity and presence of repeats. Resulting species-specific contigs were used in blastn searches against the whole genome assemblies to identify corresponding larger contigs. In the second approach, genome assemblies were annotated using HELIXER v.0.3.2 (Stiehler et al., 2021; Holst et al., 2023). Subsequently, protein and coding sequences (CDS) were extracted. All *I. ricinus* CDSs were compared against all three *I. inopinatus* CDS datasets using blastn (CLC Main Workbench v.24.0.1) and visualized in Venn diagrams. To identify *I. ricinus*-specific CDSs, potential candidates were compared with the two other *I. ricinus* CDS datasets and subsequently checked against *I. inopinatus* genome assemblies. The opposite approach was used to identify *I. inopinatus*-specific CDSs. Corresponding contigs of species-specific CDS were identified using multiple alignment. Contig-specific primers were designed and validated to confirm species-specific contigs by conventional PCR, with the TROSPA gene as a positive control (Supplementary Table S1).

2.6. Genomic analyses of tick pathogens and endosymbionts

De novo assembled tick genome sequences were checked *via* blastn search for the presence of pathogens and endosymbionts. These included *Anaplasma phagocytophilum* (GenBank: CP015376.1), *Borrelia lusitaniae* (GenBank: CP124050), "*Candidatus* Midichloria mitochondrii" (GenBank: CP002130.1; further referred to as *M. mitochondrii*), *Neoehrlichia mikurensis* (GenBank: CP089285.1), *Rickettsia helvetica* (GenBank: AP019563.1), *Rickettsia monacensis* (GenBank: NZ_LN794217), *Rickettsiella viridis* (GenBank: AP018005) and *Spiroplasma ixodetis* (GenBank: AP026933.1). Bacterial genomes were annotated using Bakta v.1.7 (Schwengers et al., 2021). Additionally, Oxford Nanopore Technologies (ONT) reads were aligned against the assembled tick genome sequences using Minimap2 v.2.24 (Li, 2018) and unmapped reads were analysed using Kraken2 v.2.1.3 and database k2_pluspf_20231009, which contains bacterial and viral sequences plus protist and fungal sequences (Wood and Salzberg, 2014; Wood et al., 2019). Aligned reads were filtered for species-level rank code to determine the presence of additional miscellaneous endosymbionts.

2.7. Pangenome analysis

Pangenome analyses were conducted separately for *M. mitochondrii*, *R. helvetica*, and *R. monacensis*. The Anvi'o pangenomic workflow (Delmont and Eren, 2018), with an mcl inflation set to 2, was employed using Anvi'o v7 (Eren et al., 2020). Regarding *M. mitochondrii*, two genomes assembled in this study were compared with 11 genomes sourced from Dutch *I. ricinus* (ENA PRJEB70514; Lesiczka et al., 2024; Supplementary Table S2). For *R. helvetica*, two genomes from this study were contrasted with eight genomes of *R. helvetica* assembled from Dutch *I. ricinus* (ENA PRJEB70514; Lesiczka et al., 2024; Supplementary Table S2). Pangenome comparative analyses for *R. monacensis* involved three newly assembled genomes from this study, alongside a reference genome (strain IrR/Munich) (GenBank: NZ_LN794217). The Anvi'o genome databases were annotated using the NCBI COG function (Galperin et al., 2015).

2.8. Phylogenetic analyses

For the phylogenetic analyses, we utilized complete mitochondrial genome sequences from six Tunisian ticks and 18 Dutch *I. ricinus* genomes (ENA PRJEB70514; Lesiczka et al., 2024; Supplementary Table S2). The phylogeny of the complete nuclear TRPA1 gene, TRPA1 cDNA, and the resultant protein sequences was analysed in 6 Tunisian and 4 Dutch *Ixodes* ticks. To align mitochondrial data and TRPA1 cDNA sequences, ClustalW alignments were computed using Geneious 11.1.4 software (Kearse et al., 2012). For the TRPA1 protein sequences, Geneious alignment was performed using Geneious 11.1.4 software (Kearse et al., 2012). The alignment for the TRPA1 gene was calculated using the MAFFT algorithm through the online service at https://mafft.cbrc.jp/alignment/server/. The alignments of TROSPA and calreticulin genes were computed in CLC Genomics Workbench v.23.0.4 using standard alignment with default settings.

Further phylogenetic analyses were conducted on complete genomes of tick endosymbionts, specifically *M. mitochondrii*, *R. helvetica* and *R. monacensis*, each analysed separately. The dataset included two *R. helvetica* and two *M. mitochondrii* from this study, together with 10 *R. helvetica* and 11 *M. mitochondrii* from Dutch *I. ricinus* ticks (ENA PRJEB70514; Lesiczka et al., 2024; Supplementary Table S2). The single-copy core of 12 *R. helvetica* and 13 *M. mitochondrii* genomes were identified in Anvi'o v.7 (Eren et al., 2020) and made up of 1424 and 1022 single-copy gene (SCG) clusters, respectively. Gene alignments from SCG were extracted and concatenated using the program anvi-get-sequences-for-gene-clusters, with parameters: concatenate-gene-cluster; –report-DNA-sequence. The details of all the phylogenetic analyses are listed in Supplementary Table S3. All phylogenies were inferred by IQTREE 2.1.3 (Nguyen et al., 2015), and the best-fit evolution model was selected based on the Bayesian information criterion (BIC) computed by implemented ModelFinder (Kalyaanamoorthy et al., 2017). Branch supports were assessed by the ultrafast bootstrap (UFBoot) approximation (Minh et al., 2013) and the Shimodaira-Hasegawa like approximate likelihood ratio test (SH-aLRT) (Guindon et al., 2010) and rooted at the midpoint. Trees were visualized and edited in FigTree v.1.4.1 and Adobe Illustrator v.28.0.

3. Results

3.1. Tick collection and morphological identification

In total, 41 adult ticks were collected at the study location in Jouza-Amdoun, Tunisia (Supplementary Fig. S1). After morphological examination, 27 specimens were identified as *I. ricinus* and 11 specimens as *I. inopinatus*. Three individuals showed intermediate characters. Collected specimens of *I. ricinus* and *I. inopinatus* were morphologically different based on the following characteristics (Fig. 1 and Supplementary Figs. S2–S4). Males can be differentiated by the number of rows of lateral conscutal setae between the lateral margin of the idiosoma and the marginal groove. In male *I. ricinus*, there are many rows of lateral conscutal setae present (Fig. 1B), while there is only a single row in male *I. inopinatus* (Fig. 1D). The anal groove has with parallel sides in *I. ricinus* (Fig. 1F-H) and divergent sides in *I. inopinatus* (Fig. 1L-P). The internal spur of coxa 1 is longer and curved touching the coxa 2 in the female *I. ricinus* (Fig. 1J), while it is pointed and slender, shorter and straight reaching coxa 2 in the female *I. inopinatus* (Fig. 1N). Finally, there are up to 6 concentric rows of goblets in the spiracular plate of the female *I. ricinus*, while there are 4 rows in the female *I. inopinatus*.

3.2. Genome assemblies of Tunisian ticks

Three *I. ricinus* and three *I. inopinatus* were sequenced individually using Oxford Nanopore Technologies. The nanopore reads were used for *de novo* assembly of six draft genome sequences (Table 1). Contigs of mitochondrial DNA and nuclear genes TRPA1, TROSPA and calreticulin were identified in each *de novo* assembled tick genome sequence. For all ticks, a circular mitochondrial genome was obtained, coding for 13 protein-encoding genes, 22 transfer RNA genes (tRNAs), and 2 ribosomal RNA genes (rRNAs). The mitochondrial genome sizes ranged from 14,564 to 14,585 bp. The guanine-cytosine base content (GC) of the mitochondrial genomes varied from 20.8% to 20.9%. A multiple alignment of the six mitochondrial genomes indicated 64 variable positions (S) of which 22 were informative parsimony sites (excluding indels). These variable sites defined a total of 6 haplotypes (mean haplotype diversity, Hd = 1), with mean nucleotide diversity $\pi = 0.00177$, and average number of nucleotide differences $k = 25.733$.

Fig. 1. Morphology of Tunisian *I. ricinus* (**A**, **B**, **E**, **F**, **I-L**)/*I. inopinatus* (**C**, **D**, **G**, **H**, **M-P**) males and females. **A**, **B** Dorsal side of male *I. ricinus* (32 ×; Estrada-Peña et al., 2014; Chitimia-Dobler et al., 2017; Bursalı et al., 2020), showing many rows of lateral conscutal setae between the lateral margin of the idiosoma and the marginal groove (**B**). **C**, **D** Dorsal side of male *I. inopinatus* (32×; Estrada-Peña et al., 2014; Chitimia-Dobler et al., 2017; Bursalı et al., 2020) with 1 row of lateral conscutal setae between the lateral margin of the idiosoma and the marginal groove (**D**). **E**, **F** Ventral side of male *I. ricinus* (32×; Chitimia-Dobler et al., 2017), showing an anal groove with parallel sides (**F**). **G**, **H** Ventral side of male *I. inopinatus* (32×; Chitimia-Dobler et al., 2017) showing an anal groove with divergent sides (**H**). **I-L** Ventral side of female *I. ricinus* (32×; Estrada-Peña et al., 2014; Chitimia-Dobler et al., 2017; Bursalı et al., 2020), showing that the internal spur of coxa 1 is longer and curved touching the coxa 2 (**J**), the number of concentric rows of goblets in the spiracular plate > 4 (**K**) and an anal groove with parallel sides (**L**). **M-P** Ventral side of female *I. inopinatus* (32×; Estrada-Peña et al., 2014; Chitimia-Dobler et al., 2017; Bursalı et al., 2020) showing that the internal spur of coxa 1 is pointed and slender, shorter and straight reaching coxa 2 (**N**), the number of concentric rows of goblets in the spiracular plate = 4 (**O**), and an anal groove with divergent sides (**P**). Photomicrographs are marked in blue for *I. ricinus* and in orange for *I. inopinatus*.

High read quality enabled complete separation of both alleles of nuclear markers TRPA1, TROSPA and calreticulin. For all assemblies, complete copies of TRPA1, TROSPA and calreticulin alleles were located on single contigs. The TRPA1 gene codes for a candidate infrared sensing protein. All 12 obtained alleles from the 6 Tunisian ticks contained 26 coding exons, divided over 197.0 and > 243 kb and corresponding to a 3555-bp open reading frame (ORF), and showed a high variation in intron sizes. TROSPA encodes for a tick receptor which interacts with outer surface protein A (OspA) of spirochetes. It encompasses 2 coding exons in a 498-bp ORF. The calreticulin gene also consists of 2 coding exons, in a 1242-bp ORF.

3.3. Phylogenetic analyses of mitochondrial genomes and nuclear genes TRPA1, TROSPA and calreticulin

For phylogenetic analysis of the mitochondrial genomes, 18 Dutch *I. ricinus* sequences were included. A phylogenetic tree shows two strongly supported clades, distinguished on the basis of geographical structure (Fig. 2A). The first well supported clade, contains 18 sequences from Dutch *I. ricinus*, while in the second clade, sequences derived from Tunisian ticks are grouped together. However, in this clade, no species- or lineage-based structure was detected. For TRPA1, phylogenetic analyses were computed for the complete TRPA1 gene (Supplementary Fig. S5) and TRPA1 cDNA (Fig. 2B). Sequences from 4 Dutch *I. ricinus* were

Table 1
Specifications of *de novo* assembled draft genome sequences of six Tunisian *I. ricinus/I. inopinatus* ticks and their endosymbionts using the Oxford Nanopore Technologies platform.

Specimen code	Tick1 (Tick1-IR)	Tick2 (Tick2-IR)	Tick3 (Tick3-IO)	Tick4 (Tick4-IO)	Tick5 (Tick5-IR)	Tick6 (Tick6-IO)
Morphological identification/sex	*I. ricinus* ♀	*I. ricinus* ♀	*I. inopinatus* ♀	*I. inopinatus* ♀	*I. ricinus* ♂	*I. inopinatus* ♂
Assembly length (bp)	3,804,189,329	3,807,135,979	3,847,505,909	3,848,037,673	3,880,493,362	3,742,000,746
Number of contigs	11,379	11,323	12,256	11,736	9335	14,088
Largest contig (bp)	8,856,599	13,164,091	14,238,608	8,104,775	53,183,116	6,653,259
N50 contig length (bp)	830,639	788,297	766,319	798,251	1,180,761	528,215
Mean coverage (fold)	20	21	21	21	32	21
Complete BUSCO	98.71%	98.06%	98.13%	98.77%	98.40%	98.09%
"*Candidatus* Midichloria mitochondrii" (CP002130)	Undetectable	1.19 Mb, 1 contig (complete)	0.98 Mb, 2 contigs (partial)	1.19 Mb, 1 contig (complete)	Undetectable	Undetectable
Rickettsia helvetica (FG22047_03)	Undetectable	Undetectable	1.36 Mb, 1 contig (complete)	1.36 Mb, 1 contig (complete)	Undetectable	Undetectable
Rickettsia monacensis (NZ_LN794217)	Undetectable	1.39 Mb, 1 contig (complete)	Undetectable	Undetectable	1.39 Mb, 1 contig (complete)	1.38 Mb, 1 contig (complete)
Borrelia lusitaniae (CP124050)	Undetectable	0.80 Mb, 2 contigs (partial)	0.63 Mb, 3 contigs (partial)	Undetectable	Undetectable	Undetectable

Note: Contigs corresponding to *Anaplasma phagocytophilum*, *Rickettsiella* sp., *Spiroplasma* sp., or *Neoehrlichia mikurensis* were not detected.

included. Both trees presented similar topology, showing two well-supported clades consisting of Tunisian *I. ricinus* and *I. inopinatus* as well as Dutch *I. ricinus*. In contrast, the phylogenetic tree based on the protein sequence revealed a different pattern, with all Tunisian ticks grouped together, while the sequences from the Dutch ticks formed a separate clade (Supplementary Fig. S6).

Multiple alignments were computed for TROSPA and calreticulin, both supplemented with sequences obtained from Dutch *I. ricinus* ticks or GenBank. A deletion of 12 nucleotides, corresponding to an alanine-methionine-valine-alanine sequence, was observed in the N-terminus of the TROSPA gene in all Tunisian *I. ricinus* and *I. inopinatus* ticks (Supplementary Fig. S7). This deletion was not apparent in the Dutch ticks or sequences obtained from GenBank (Supplementary Figs. S7–S8, see Supplementary Figs. S9–S10 for complimentary trees). The alignment of calreticulin sequences from Tunisian ticks was supplemented with a reference sequence (GenBank: AY395272) (Supplementary Fig. S11). Sequences of the Tunisian ticks contained 36 nucleotide substitutions compared to the consensus sequence, and were observed in both Tunisian *I. ricinus* and *I. inopinatus*.

3.4. Subtractive genomics

To identify candidate *I. ricinus*- and/or *I. inopinatus*-specific sequences, whole genomes of *I. ricinus* and *I. inopinatus* were compared *via* two different approaches. In our first approach, repeats were masked in six genome assemblies, after which the nanopore reads of *I. inopinatus* were aligned against the combined repeat-masked assemblies of Tunisian *I. ricinus* (and *vice versa*). For *I. inopinatus*, only contig_4938 was specific and not repetitive, whereas for *I. ricinus*, contigs_1198, _22410, and _24147 were specific and not repetitive. PCR using specific primer pairs to confirm the presence/absence of these contigs in the original DNA lysates detected one discrepancy with the *in-silico* analysis in contig_4938 (present in Tick2-IR) and one in contig_1198 (present in Tick3-IO). PCR confirmation in 15 other Tunisian ticks detected discrepancies in all of the suspected *I. inopinatus/I. ricinus*-specific contigs (Table 2).

In the second approach, genome assemblies were annotated, and coding sequences (CDS) were compared. This yielded five *I. ricinus*-specific CDSs which correspond to various exons of the venom metalloproteinase antarease-like TtrivMP_A gene (GenBank: XP_040068574.1). *Ixodes ricinus*-specific contig_22410 (above) overlaps with exon 2 of this gene. Four *I. inopinatus*-specific CDSs were identified, next to three CDSs that also correspond to various exons of the TtrivMP_A gene. Subsequent analyses showed that presence/absence of members of this gene family in Tunisian ticks is highly variable between individual ticks, but not specific to *I. inopinatus*. *Ixodes inopinatus*-specific CDS 410_40 corresponds to a papilin-like protein encoding gene (*papln*), of which 17 variants are present on contig_410. The papilin-like gene family is a family of protease inhibitors with a putative function in insect embryonic development. Multiple exons of all gene variants were checked for species-specificity. While PCR showed *I. inopinatus*-specific presence of contig_410 in the original DNA lysates, discrepancies were detected in 15 other Tunisian ticks (Table 2). Further results are shown in Supplementary Tables S4–S5.

3.5. Endosymbionts in Tunisian ticks

All six *de novo* assembled genome sequences were checked for the presence of known tick endosymbionts by blastn search, leading to two categories of hits. The first category consisted of near-perfect hits with more than 99% identity to the query sequence. These hits led to the identification of one partial genome (Tick3-IO) and two complete genomes (Tick2-IR and Tick4-IO) of *M. mitochondrii*, two complete genomes of *R. helvetica* (Tick3-IO and Tick4-IO), three complete genomes of *R. monacensis* (Tick2-IR, Tick5-IR, and Tick6-IO), and two partial genomes of *B. lusitaniae* (Tick2-IR and Tick3-IO) (Table 1). Except for *R. helvetica*, these endosymbiont genomes were found in both *I. ricinus* and *I. inopinatus* ticks. The *R. monacensis* genome size ranged from 1,406,970 to 1,412,380 bp and coded for 1684 chromosomal genes, 32 transfer RNA genes (tRNAs), and 3 ribosomal RNA genes (rRNAs). The GC base content was 32.4% for all three genomes (Supplementary Table S6). Contigs corresponding to *Anaplasma phagocytophilum*, *Neoehrlichia mikurensis*, *Rickettsiella viridis*, or *Spiroplasma ixodetis* could not be found. The second category of blastn hits revealed tick genome contigs that contain stretches with 75–90% sequence identity to *S. ixodetis*, which are interrupted by stretches with high sequence identity to the *I. ricinus* genome. In addition, we found tick genome contigs that contain sequences with more than 90% identity with parts of the *M. mitochondrii* genome (see below).

After aligning ONT reads against the assembled genomes sequences, 1–2% reads was unmapped. Of these unmapped reads, 4–10% could be classified by Kraken2 analysis. Most prominent species are shown in Supplementary Table S7. This analysis confirmed the presence of the *B. lusitaniae* genome and plasmids in tick specimens Tick2-IR and Tick3-IO, next to presence of the *R. monacensis* genome in Tick6-IO. The presence of endosymbionts detected by Kraken2 was comparable between Tunisian *I. ricinus* and *I. inopinatus*.

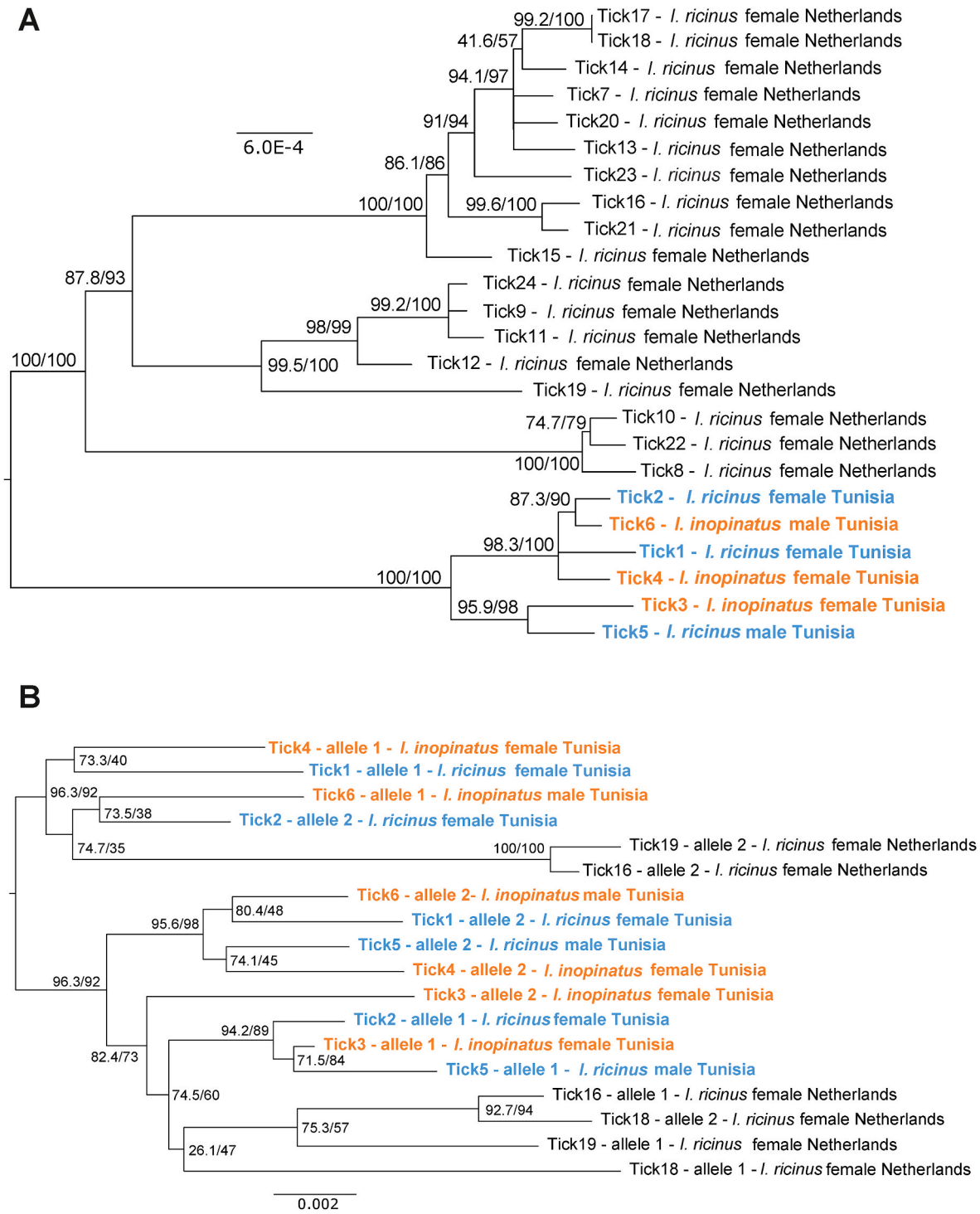

Fig. 2. Maximum likelihood phylogenetic tree of Tunisian *I. ricinus/I. inopinatus* and Dutch *I. ricinus* based on the complete mitochondrial genome (**A**) and the nuclear gene TRPA1 cDNA (**B**). Both trees are midpoint-rooted with bootstrap values (SH-aLRT/UFB) above the 75/75 threshold displayed. All tick sequences are described as *I. ricinus* or *I. inopinatus*, by country of origin and tick sex. All Tunisian *I. ricinus* sequences are marked in blue, while Tunisian *I. inopinatus* sequences are marked in orange.

3.6. Comparative and phylogenetic analyses of tick endosymbionts

A comparative genome analysis (Supplementary Fig. S12) was performed to determine whether the functional genome contents of the *R. monacensis* assemblies differed among themselves or from the reference genome of strain IrR/Munich (GenBank: NZ_LN794217). The Anvi'o pangenome analyses revealed 1236–1257 gene clusters (1623–1661 genes) in analysed genomes. The number of singular gene clusters was equal in assemblies from this study, whereas 20 singular gene clusters were detected in the *R. monacensis* reference genome (GenBank: NZ_LN794217). Complete results of Anvi'o comparative analyses for *R. monacensis* are available in Supplementary Table S8.

A phylogenetic tree for *M. mitochondrii* was constructed using concatenated sequences from Tunisian *I. ricinus* and *I. inopinatus* and 11 sequences extracted from complete genomes of Dutch *I. ricinus* (Fig. 3A). The length of concatenated sequences ranged from 796,266 to 797,913

Table 2
PCR confirmation of *I. inopinatus/I. ricinus*-specific contigs. PCR targeted the four contigs, which were specific for either *I. inopinatus* (IO) or *I. ricinus* (IR) in the *in-silico* genomic subtractions.

DNA sample ID	Tick ID	Sex	*I. inopinatus/I. ricinus*-specific contigs					
			4938 (IO)	410 (IO)	22410 (IR)	24147 (IR)	1198 (IR)	TROSPA (IO/IR)
Tick1-IR	IR	♀	neg	neg	pos	pos	pos	pos
Tick2-IR	IR	♀	**pos**	neg	pos	pos	pos	pos
Tick5-IR	IR	♂	neg	neg	na	na	pos	pos
23014_01	IR	♀	neg	**pos**	pos	pos	**neg**	pos
23014_04	IR	♀	neg	neg	pos	pos	pos	pos
23014_18	IR	♀	**pos**	neg	pos	pos	pos	pos
23014_19	IR	♀	**pos**	neg	pos	pos	**neg**	pos
23014_20	IR	♀	neg	**pos**	pos	pos	pos	pos
23014_21	IR	♀	neg	neg	**neg**	pos	**neg**	pos
23014_22	IR	♀	neg	neg	pos	pos	**neg**	pos
23014_08	IR	♂	**pos**	neg	pos	**neg**	**neg**	pos
Tick3-IO	IO	♀	pos	pos	neg	neg	**pos**	pos
Tick4-IO	IO	♀	pos	pos	neg	neg	neg	pos
Tick6-IO	IO	♂	pos	pos	neg	neg	neg	pos
23014_10	IO	♂	**neg**	pos	**pos**	neg	neg	pos
23014_11	IO	♂	pos	**neg**	neg	neg	neg	pos
23014_13	IO	♂	**neg**	**neg**	neg	**pos**	neg	pos
23014_14	IO	♂	pos	**neg**	**pos**	**pos**	neg	pos
23014_15	IO	♂	pos	**neg**	**pos**	neg	neg	pos
23014_16	IO	♂	**neg**	pos	neg	**pos**	neg	pos
23014_17	IO	♂	**neg**	**neg**	**pos**	neg	**pos**	pos

Notes: DNA samples in bold were also used for ONT sequencing (see Table 1). PCR results in bold highlight the discrepancies with the *in-silico* genomic subtractions. A PCR targeting the TROSPA gene was used as a positive control. Primer pairs and protocols are provided in Supplementary Tables S4–S5.
Abbreviation: na, insufficient DNA of Tick5-IR was left to be included in all analyses.

bp. In this tree, *M. mitochondrii* associated with Tunisian ticks formed a separate clade, which contained sequences from both Tunisian *I. ricinus* and *I. inopinatus*. The phylogenetic analysis for *R. helvetica* was based on concatenated sequences from eight Dutch and two Tunisian ticks (Fig. 3B). The length of concatenated sequences ranged from 1,001,226 to 1,012,326 bp. *Rickettsia helvetica* associated with Tunisian ticks clustered together, whereas the sequences from Dutch ticks formed a separate cluster. For *R. monacensis*, the phylogenetic analysis was based on 3 sequences derived from this study and the reference genome of strain IrR/Munich. *Rickettsia monacensis* associated with Tunisian *I. ricinus* and *I. inopinatus* clustered together, separately from the reference sequence (Fig. 3C).

3.7. Horizontal gene transfer of tick endosymbionts

As mentioned above, a blastn search with *S. ixodetis sHm* (GenBank: AP026933) revealed four contigs that contain stretches with 70–90% sequence identity to *Spiroplasma* in the draft genome sequence of Tick1-IR (Fig. 4A–C, Supplementary Figs. S13–S15, Supplementary Tables S9A–B). The *Spiroplasma*-like sequences in the tick genome assembly showed homology with chromosomes and plasmids of three Ixodetis clade *Spiroplasma* strains: *S. ixodetis sHm* (GenBank: AP026933, AP026934, AP026935; Arai et al., 2022), *S. ixodetis sAtri* (GenBank: CP117528, CP117532, CP117533; Moore and Ballinger, 2023), and *S. ixodetis Y32* (GenBank: CP127039, CP127040, CP127041, CP127042, CP127043, CP127044). Some homology (67–71% identity) was found with the chromosome of *Pachydiplax longipennis* associated *S. platyhelix PALS-1* (GenBank: CP051215), but not with any of the other *Spiroplasma* genome sequences in the NCBI database (Supplementary Table S10). Contig_6893 of Tick1-IR contains two *Spiroplasma* cassettes. Cassette A-long contains ten *Spiroplasma* sequences, interrupted by nine *Ixodes* sequences (Fig. 4A–Supplementary Table S11A). The *Spiroplasma* sequences have a total size of about 19 kb and can be further subdivided into 26 pieces, each of which has a characteristic pattern of homologies with the reference *Spiroplasma* chromosomes and plasmids (Supplementary Table S9A). Cassette B contains two *Spiroplasma* sequences separated by one *Ixodes* sequence (Supplementary Table S9B). Contig_7467 of Tick1-IR contains cassette C-long, in which two *Spiroplasma* stretches are interrupted by one *Ixodes* sequence (Fig. 4C–Supplementary Table S11C).

We then used cassettes A, B, and C of Tick1-IR and 5 kb of their immediate upstream and downstream flanking regions as queries to map the corresponding regions in the draft genome assemblies of all six Tunisian ticks, and also in the draft genome assemblies of four individual Dutch female *I. ricinus* (Fig. 4A-C). Complete copies of cassette A-long were found in 3 out of 6 Tunisian ticks, whereas 2 Tunisian tick assemblies only contain a truncated version of cassette A, lacking sequences A20-A26 and part of A19. The genome assembly of Tick6-IO contains an even shorter version of cassette A-long, lacking sequences A1-12, A20-26, and part of A19. Cassette A-long was not found in any of the four Dutch tick genome assemblies. Cassette B was only found in Tick1-IR and Tick2-IR. All 6 Tunisian and all 4 Dutch ticks contained at least one complete copy, and one or more partial copies of cassette C-long (Fig. 4C). In addition, Tunisian Tick5-IR and all 4 Dutch ticks contained partial or complete versions of cassette A-C-short, which includes *Spiroplasma* sequences A18-20 and C4b-6. In cassette C-long, extra *Ixodes* sequence insertions were found in e.g. sequences C2 (0.2-kb), C3 (1.6-kb), and C4 (6.0 or 0.4 kb) (Fig. 4C).

Blastn searches using the *M. mitochondrii* chromosome (GenBank: CP002130) as a query against the genome assemblies of the six Tunisian and the four Dutch ticks also resulted in the identification of multiple tick genome contigs that contain short stretches with high homology to *M. mitochondrii* (92–98% sequence identity). About 40 different *Midichloria*-related sequences could be identified, which were named cassette A to cassette AL (Supplementary Table S12, Supplementary Fig. S16). The number of different *Midichloria*-related sequences per genome assembly varied from four in Tunisian Tick1-IR to eleven in Dutch Tick1-IR6. Most *Midichloria*-related sequences had a size between less than 0.1 kb and 0.4 kb. Only few cassettes were longer, namely cassette A (1.0 kb), B (0.6 kb), C (0.5 kb), J_long (16.9 kb), R_long (6.5 kb), S (0.6 kb), AH (3.8 kb), and AK (0.5 kb). Cassette A was found in all ten ticks and cassette C was found in all ticks, except for Dutch Tick1-IR6. All other cassettes were only found in a smaller subset of tick assemblies, some of them both in Tunisian and Dutch ticks and others only in Tunisian or only in Dutch ticks.

To examine if the *Spiroplasma*-derived insertions in the *I. ricinus*

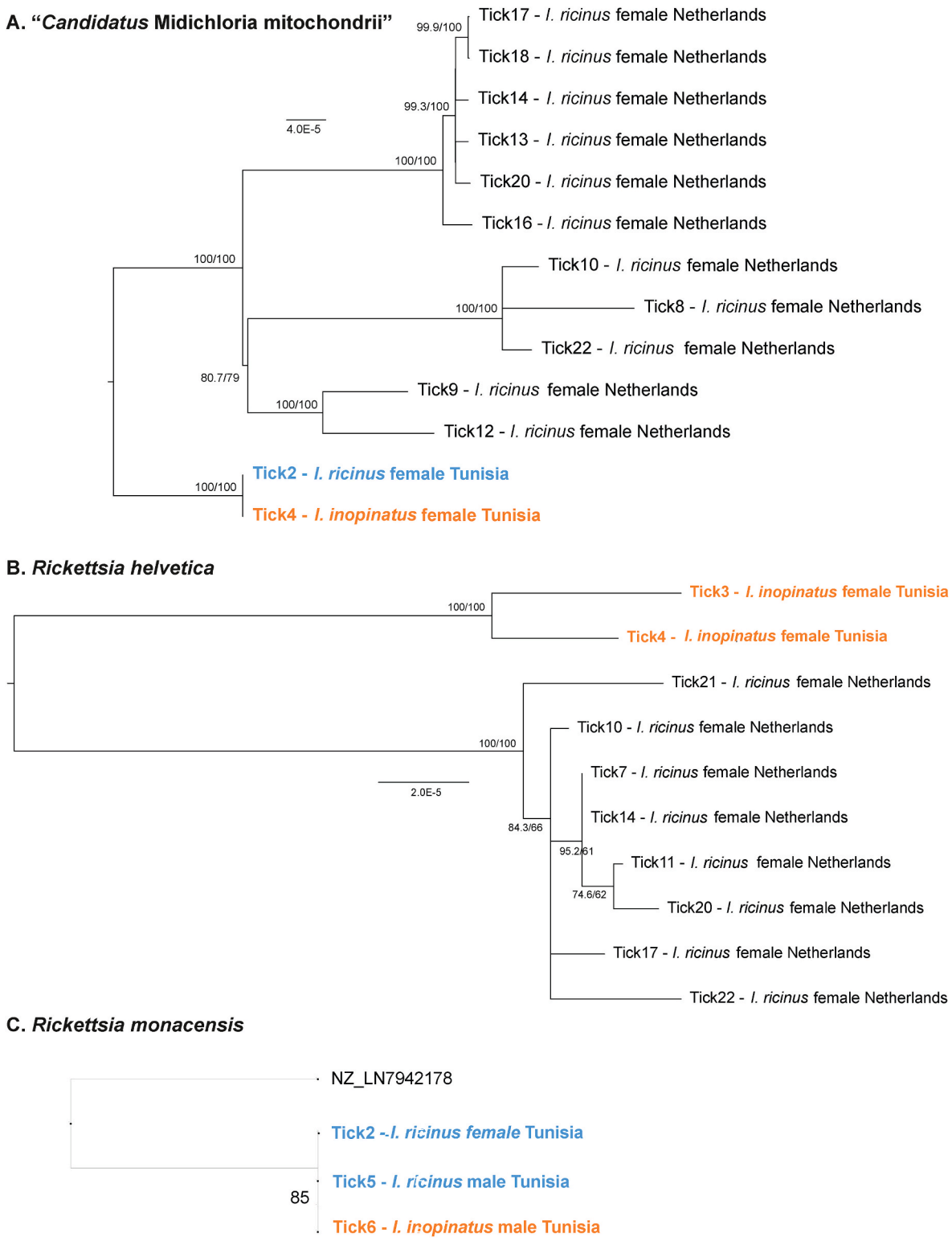

Fig. 3. Maximum likelihood phylogenetic tree based on endosymbiont sequences obtained from Tunisian *I. ricinus/I. inopinatus* and Dutch *I. ricinus* for "*Ca.* M. mitochondrii" (**A**), *R. helvetica* (**B**) and *R. monacensis* (**C**). Trees are midpoint-rooted with bootstrap values (SH-aLRT/UFB) above the 75/75 threshold displayed. All tick sequences are described as *I. ricinus* or *I. inopinatus*, by country of origin and tick sex. For *R. monacensis*, reference genome NZ_LN794217 is included. All Tunisian *I. ricinus* sequences are marked in blue, while Tunisian *I. inopinatus* sequences are marked in orange.

genome are transcribed, 88 *I. ricinus* RNAseq datasets were downloaded from NCBI's SRA database. The datasets contained in total about 500 Gb of Illumina RNAseq data derived from 88 RNA samples in 17 different Bioprojects (Supplementary Table S10). Twelve RNAseq samples contained reads that aligned against cassette C. One whole nymph RNA sample contained reads aligning against cassette A. We also examined if the *Midichloria*-derived insertions are transcribed. In total 53 out of the 88 RNAseq datasets contained reads aligning to one or more of the *Midichloria*-derived cassettes. In total 23 cassettes had corresponding RNAseq reads in 1–5 RNAseq datasets, whereas 7 cassettes had RNAseq

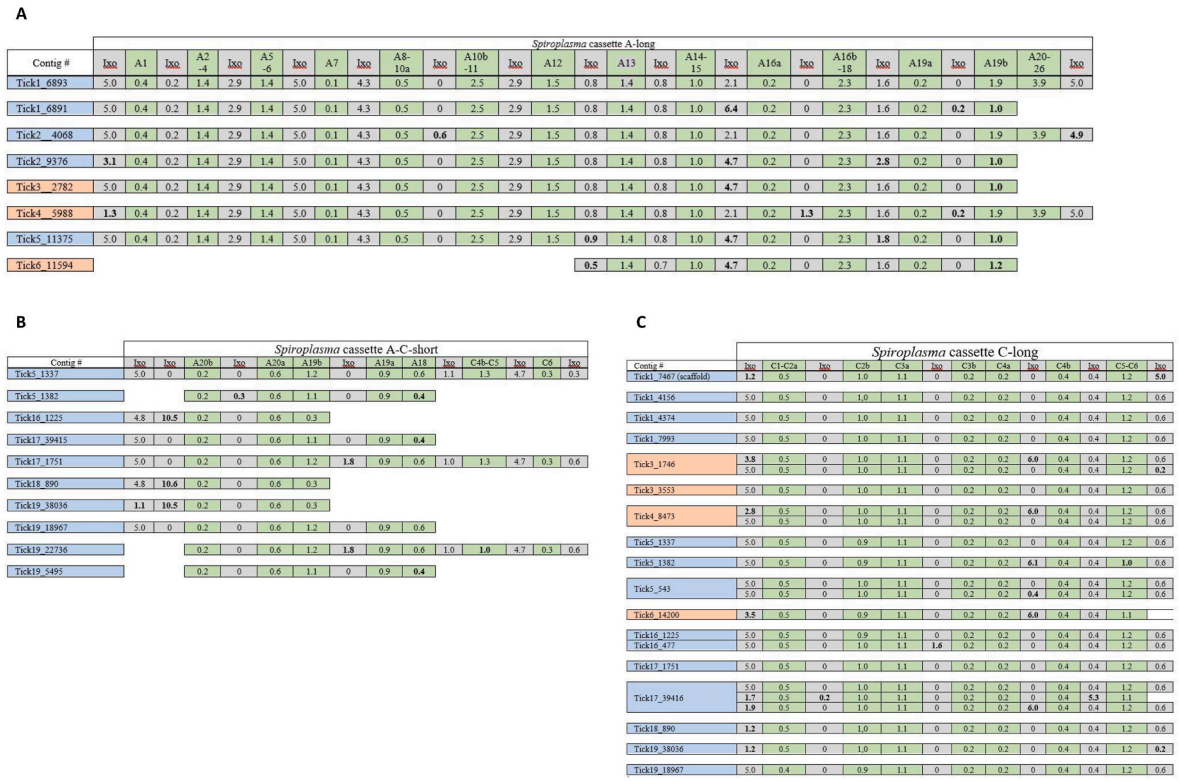

Fig. 4. Cassettes of *Spiroplasma*-like sequences in tick genomes. **A** *Spiroplasma* cassette A-long. Numbers indicate the sizes (kb) of the alternating *Ixodes* (Ixo) and *Spiroplasma* (A1-A26) sequences. The query sequence that was used for mapping the A-long cassette in all tick specimens is ric_fem_32_contig_6893_Spiroplasma_cassette_A47.70. **B** *Spiroplasma* cassette A-C-short. Numbers indicate the sizes (kb) of the alternating *Ixodes* (Ixo) and *Spiroplasma* (A18-A20 and C4b-C6) sequences. The query sequences that were used for mapping the A-C-short cassette in all tick specimens are ric_male_1_contig_1337-RC_Spiroplasma_cassette_A13.38 and ric_fem_32_scaffold_7467_Spiroplasma_cassette_C14.61. **C** *Spiroplasma* cassette C-long. Numbers indicate the sizes (kb) of the alternating *Ixodes* (Ixo) and *Spiroplasma* (C1-C6) sequences. The query sequence that was used for mapping the C-long cassette in all tick specimens is ric_fem_32_scaffold_7467_Spiroplasma_cassette_C14.61.

reads in 10 or more RNAseq datasets. Reads aligning to cassette J_long were found in 28 datasets and reads aligning to cassette A were found in 34 datasets. In 12 datasets, only reads mapping to cassette A were detected.

4. Discussion

In this study, we performed genomic analyses to explain the morphological differences between *I. ricinus* and *I. inopinatus*. An earlier study showed the sympatric occurrence of *I. ricinus* and *I. inopinatus* in Jouza-Amdoun, Tunisia (Younsi et al., 2020). To eliminate the risk of spatial and temporal variation, we analysed *I. ricinus* and *I. inopinatus* ticks collected simultaneously from this location. Therefore, if morphological differences between these ticks stem from genetic variations, they should be evident in this dataset. Our genomic analyses included the complete mitochondrial genome and nuclear markers TROSPA and calreticulin, as these markers were used in comparable studies (Hrazdilova et al., 2023; Rollins et al., 2023). We introduced additional complexity by incorporating the nuclear gene TRPA1 (~220 kb), comprising 26 coding exons corresponding to a 3555-bp ORF. We also looked for *I. ricinus*- and/or *I. inopinatus*-specific sequences within the complete tick genome assemblies by genomic subtractions. Finally, we investigated the influence of tick endosymbionts on tick morphology through qualitative and phylogenetic analyses. However, neither of our genomic analyses was able to distinguish between sympatric *I. ricinus* and *I. inopinatus* ticks from Tunisia.

Phylogenetic analysis of the complete mitochondrial genome indicated that the Tunisian ticks were more closely related to each other than to Dutch *I. ricinus*. This confirmed earlier observations showing a genetic distance between *I. ricinus* from Europe and Northern Africa (de Meeûs, 2002; Noureddine et al., 2011; Poli et al., 2020). Based on microsatellite markers, preliminary analysis of both datasets of *I. ricinus* and *I. inopinatus* revealed two isolated populations, which can correspond to two different species (Velez et al., 2023). Furthermore, analysis of the complete mitochondrial genome failed to distinguish Tunisian *I. ricinus* and *I. inopinatus*, confirming the findings of Rollins et al. (2023). For TRPA1, phylogenetic analyses performed on both the complete gene and cDNA, did not reveal any potential differences between *I. ricinus* and *I. inopinatus* from Tunisia or the Netherlands. However, a clearer distinction based on geographical origin was observed in the protein sequence analysis, resulting in two separate clades. Similar to Hrazdilova et al. (2023), our analysis of calreticulin demonstrated its incapability for distinguishing between *I. ricinus* and *I. inopinatus*. Hrazdilova et al. (2023) suggested differentiation based on the nuclear marker TROSPA. Nevertheless, a multiple sequence alignment of TROSPA showed no apparent differences between the Tunisian *I. ricinus* and *I. inopinatus* ticks examined in this study. Also, two separate approaches to compare whole genomes, including repeat-masking and CDS comparison, did not result in the identification of *I. ricinus*- and/or *I. inopinatus*-specific sequences. Based on these findings, we must reject the hypothesis that sympatric *I. ricinus* and *I. inopinatus* are genetically distinct. Thus, it is improbable that the morphological disparities between these ticks are determined by genetic differentiation into separate species or lineages.

Our alternative hypothesis posited that these morphological differences are influenced by tick endosymbionts. Consequently, we investigated the presence of well-known symbionts of *Ixodes* ticks, including *A. phagocytophilum*, *B. lusitaniae*, *M. mitochondrii*, *N. mikurensis*, *R. helvetica*, *R. monacensis*, *R. viridis*, and *S. ixodetis*. While we successfully identified partial or complete genomes of *B. lusitaniae*,

M. mitochondrii, *R. helvetica*, and *R. monacensis*, presence nor absence of any of these bacteria was exclusive to either *I. ricinus* or *I. inopinatus*, suggesting these endosymbionts do not explain the *I. inopinatus* phenotype. Subsequently, phylogenetic analyses were performed separately on obtained genomes of *M. mitochondrii*, *R. helvetica* and *R. monacensis*. For all three bacteria, bacterial genomes obtained from Tunisian ticks clustered separately from those obtained from Dutch ticks or the reference genome. However, bacterial genomes from Tunisian *I. ricinus* and *I. inopinatus* formed a single clade. This implies that these endosymbionts evolved within either the Tunisian or Dutch population of their invertebrate host, but their genetic background does not explain morphological differences between *I. ricinus* and *I. inopinatus*.

Other potential explanations for the morphological differences between *I. ricinus* and *I. inopinatus* include hybridization, alternative splicing, or epigenetic mechanisms influenced by external factors, in addition to more subtle genetic variation. Our phylogenetic analysis of TRPA1 cDNA suggests the absence of *I. ricinus/I. inopinatus* hybrids in this study. If hybrids were present, a monophyletic *I. inopinatus* clade would have been expected in the TRPA1 cDNA tree, alongside clades containing both *I. ricinus* and *I. inopinatus* (Fig. 2B). The lack of a monophyletic *I. inopinatus* clade in this tree indicates that hybridization likely did not occur. External factors such as bloodmeal composition or microclimatic conditions could influence gene expression through alternative splicing or epigenetic mechanisms (Alekseev and Dubinina, 2008; Rimphanitchayakit and Tassanakajon, 2010; Shiina and Shimizu, 2020). More extensive transcriptomic and epigenetic research is needed to investigate these hypotheses.

It is possible that the morphological differences between *I. ricinus* and *I. inopinatus* can be explained by minor genetic differences in genes not investigated in this study. Nevertheless, our use of advanced sequencing techniques in this study yielded high data quality. Through NGS data, we successfully assembled and annotated six complete tick genomes, distinguished between alleles of multiple nuclear genes, and generated seven complete genomes of tick endosymbionts and tick-borne pathogens. Notably, this effort includes the first complete draft genomes of *M. mitochondrii*, *R. helvetica* and *R. monacensis* originating from northern Africa. These genomes offer valuable resources for the development of diagnostics and vaccines against emerging tick-borne pathogens such as *R. monacensis*. This pathogen, an etiologic agent of the Mediterranean spotted fever, was detected in six human patients since 2007, most recently in Portugal (de Sousa et al., 2022). *Rickettsia monacensis* is prevalent in *Ixodes* ticks across Europe, Africa and Asia and has recently been detected in *Rhipicephalus sanguineus* (*sensu lato*) tick from the USA (Parola et al., 2013; Lineberry et al., 2022).

In our study, the use of advanced sequencing techniques also yielded some unexpected findings. Specifically, all Tunisian ticks showed a deletion of 4 codons in the N-terminus of the TROSPA gene compared to Dutch *I. ricinus*. This deletion sequence falls within a putative transmembrane domain, which is not directly involved in the binding of OspA (Pal et al., 2004; Figlerowicz et al., 2013). Nevertheless, TROSPA mutants with a partial or complete deletion of this domain had a slightly increased capacity to bind OspA *in vitro*, potentially due to altered charge (Figlerowicz et al., 2013). Regardless, Figlerowicz et al. (2013) suggested TROSPA-OspA interaction involves more than electrostatic interaction. Therefore, further research is needed to determine the effect of this mutation on vector competence of Tunisian ticks. Alternatively, this mutation could serve as a marker for monitoring *I. ricinus* migrating by birds from northern Africa to Europe.

We also detected evidence of horizontal gene transfer. Although occurrences of HGT from bacteria to eukaryotes remain relatively rare, functional events have been described (Iwanaga et al., 2014; Chou et al., 2015; Xiong et al., 2022). One such example is the horizontal transfer of type VI secretion amidase effector (Tae) gene families into eukaryotes, including ticks and mites (Chou et al., 2015). These Tae proteins are potent bactericidal molecules capable of degrading bacterial cell walls. One of these proteins, Dae2, expressed in the salivary glands and midgut of *I. scapularis*, has been shown to cleave peptidoglycan *in vitro*, including the one incorporated into the *Borrelia burgdorferi* sacculus. Consequently, Dae2 contributes to the innate ability of *I. scapularis* to regulate *B. burgdorferi* levels post-acquisition (Chou et al., 2015). Thus, HGT events transferring bacterial genes into ticks can lead to functional genes affecting tick immunity and potentially their vector competence. In the present study, HGT of *Spiroplasma* and *Midichloria* sequences was observed in all Tunisian ticks. For *Spiroplasma,* three different cassettes were identified, next to 40 different *Midichloria*-related cassettes. Alignment against RNAseq reads from public databases showed that some of these bacterial cassettes were expressed in ticks. However, further investigation is needed to determine the functionality of these HGT events.

5. Conclusions

Overall, our findings provide strong evidence indicating that *I. ricinus* and *I. inopinatus* are genetically similar despite morphological differences. We found no discernible correlation between the presence of endosymbionts and morphological disparities. It is therefore likely that ticks which appear *I. inopinatus* based on morphology represent a subpopulation of *I. ricinus* with a distinct phenotype. These morphological differences might have implications for vector competence or tick biology, such as adaptation to warm and dry climates. Our findings highlight the utility of next-generation sequencing as a powerful tool for investigating ticks and tick-borne pathogens.

CRediT authorship contribution statement

Valérie O. Baede: Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Oumayma Jlassi:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing. **Paulina M. Lesiczka:** Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Hend Younsi:** Conceptualization, Methodology, Investigation, Resources, Writing – review & editing. **Hans J. Jansen:** Methodology, Validation, Formal analysis, Data curation, Writing – review & editing. **Khalil Dachraoui:** Methodology, Investigation, Resources, Writing – review & editing. **Jane Segobola:** Formal analysis, Data curation, Writing – review & editing. **Mourad Ben Said:** Investigation, Writing – review & editing. **Wouter J. Veneman:** Formal analysis, Data curation, Writing – review & editing. **Ron P. Dirks:** Methodology, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Hein Sprong:** Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Elyes Zhioua:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing.

Ethical approval

Not applicable.

Funding

This research was financially supported by the Dutch Ministry of Health, Welfare, and Sport (VWS).

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crpvbd.2024.100229.

Data availability

The data supporting the conclusions of this article are included within the article and its supplementary file. The sequence data and assemblies generated in this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB74028 (https://www.ebi.ac.uk/ena/browser/view/PRJEB74028) with sample accessions ERS18497391, ERS18497392, ERS18497393, ERS18497394, ERS18497395, and ERS18497396. The sequencing data are available in ERR12930748, ERR12930749, ERR12930750, ERR12930751, ERR12930752, and ERR12930753. The assemblies are available in ERZ24798805, ERZ24798806, ERZ24798807, ERZ24798808, ERZ24798809, ERZ24798810.

References

Alekseev, A.N., Dubinina, H.V., 2008. Enhancement of risk of tick-borne infection: Environmental and parasitological aspects of the problem. J. Med. Entomol. 45, 812–815.

Arai, H., Inoue, M.N., Kageyama, D., 2022. Male-killing mechanisms vary between *Spiroplasma* species. Front. Microbiol. 13, 1075199. https://doi.org/10.3389/fmicb.2022.1075199.

Bakker, J.W., Begemann, H.L.M., Fonville, M., Esser, H.J., de Boer, W.F., Sprong, H., Koenraadt, C.J.M., 2023. Differential associations of horizontally and vertically transmitted symbionts on *Ixodes ricinus* behaviour and physiology. Parasites Vectors 16, 443. https://doi.org/10.1186/s13071-023-06025-3.

Bursalı, A., Tekin, Ş., Keskin, A., 2020. A contribution to the tick (Acari: Ixodidae) fauna of Turkey: The first record of *Ixodes inopinatus* Estrada-Peña, Nava & Petney. Acarol. Stud. 2, 126–130. https://doi.org/10.47121/acarolstud.706768.

Chitimia-Dobler, L., Bestehorn, M., Bröker, M., Borde, J., Molcanyi, T., Andersen, N.S., et al., 2017. Morphological anomalies in *Ixodes ricinus* and *Ixodes inopinatus* collected from tick-borne encephalitis natural foci in Central Europe. Exp. Appl. Acarol. 72, 379–397.

Chitimia-Dobler, L., Rieß, R., Kahl, O., Wölfel, S., Dobler, G., Nava, S., Estrada-Peña, A., 2018. *Ixodes inopinatus* – Occurring also outside the Mediterranean region. Ticks Tick Borne Dis. 9, 196–200.

Chou, S., Daugherty, M.D., Peterson, S.B., Biboy, J., Yang, Y., Jutras, B.L., et al., 2015. Transferred interbacterial antagonism genes augment eukaryotic innate immune function. Nature 518, 98–101.

Coipan, E.C., Jahfari, S., Fonville, M., Maassen, C.B., Van Der Giessen, J., Takken, W., et al., 2013. Spatiotemporal dynamics of emerging pathogens in questing *Ixodes ricinus*. Front. Cell. Infect. Microbiol. 3, 36. https://doi.org/10.3389/fcimb.2013.00036.

de Meeûs, T., 2002. Sex-biased genetic structure in the vector of Lyme disease, *Ixodes ricinus*. Evolution 56, 1802–1807.

de Sousa, R., dos Santos, M.L., Cruz, C., Almeida, V., Garrote, A.R., Ramirez, F., et al., 2022. Rare case of rickettsiosis caused by *Rickettsia monacensis*, Portugal, 2021. Emerg. Infect. Dis. 28, 1068–1071.

Delmont, T.O., Eren, A.M., 2018. Linking pangenomes and metagenomes: The *Prochlorococcus* metapangenome. Peer J. 6, e4320. https://doi.org/10.7717/peerj.4320.

Eren, A.M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M.S., et al., 2020. Community-led, integrated, reproducible multi-omics with anvi'o. Nat. Microbiol. 6, 3–6.

Estrada-Peña, A., 2001. Distribution, abundance, and habitat preferences of *Ixodes ricinus* (Acari: Ixodidae) in northern Spain. J. Med. Entomol. 38, 361–370.

Estrada-Peña, A., 2017. *Ixodes inopinatus* Estrada-Peña, Nava & Petney, 2014. In: Estrada-Peña, A., Petney, T. (Eds.), Ticks of Europe and North Africa: A Guide to Species Identification. Springer, Switzerland, pp. 203–206.

Estrada-Peña, A., Daniel, M., Frandsen, F., Gern, L., Gettinby, G., Gray, J.S., et al., 1998. *Ixodes ricinus* strains in Europe. Z. Bakteriol. 287, 185–189.

Estrada-Peña, A., Nava, S., Petney, T., 2014. Description of all the stages of *Ixodes inopinatus* n. sp. (Acari: Ixodidae). Ticks Tick Borne Dis. 5, 734–743.

Figlerowicz, M., Urbanowicz, A., Lewandowski, D., Jodynis-Liebert, J., Sadowski, C., 2013. Functional insights into recombinant TROSPA protein from *Ixodes ricinus*. PLoS One 8, e76848. https://doi.org/10.1371/journal.pone.0076848.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., Koonin, E.V., 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res. 43, D261–D269.

Gehman, A.M., Harley, C.D.G., 2019. Symbiotic endolithic microbes alter host morphology and reduce host vulnerability to high environmental temperatures. Ecosphere 10, e02683. https://doi.org/10.1002/ecs2.2683.

Gilbert, L., Aungier, J., Tomkins, J.L., 2014. Climate of origin affects tick (*Ixodes ricinus*) host-seeking behavior in response to temperature: Implications for resilience to climate change? Ecol. Evol. 4, 1186–1198.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Hekimoğlu, O., 2022. Phylogenetic placement of Turkish populations of *Ixodes ricinus* and *Ixodes inopinatus*. Exp. Appl. Acarol. 88, 179–189.

Hirota, B., Okude, G., Anbutsu, H., Futahashi, R., Moriyama, M., Meng, X.-Y., et al., 2017. A novel, extremely elongated, and endocellular bacterial symbiont supports cuticle formation of a grain pest beetle. mBio 8, e01482. https://doi.org/10.1128/mBio.01482-17, 17.

Holst, F., Bolger, A., Günther, C., Maß, J., Triesch, S., Kindel, F., et al., 2023. Helixer - *de novo* prediction of primary eukaryotic gene models combining deep learning and a hidden Markov model. bioRxiv. https://doi.org/10.1101/2023.02.06.527280.

Hrazdilova, K., Danek, O., Hrbatova, A., Cervena, B., Noskova, E., Adamik, P., et al., 2023. Genetic analysis challenges the presence of *Ixodes inopinatus* in Central Europe: Development of a multiplex PCR to distinguish *I. inopinatus* from *I. ricinus*. Parasites Vectors 16, 354. https://doi.org/10.1186/s13071-023-05971-2.

Huang, N., Li, H., 2023. Compleasm: A faster and more accurate reimplementation of BUSCO. Bioinformatics 39, btad595. https://doi.org/10.1093/bioinformatics/btad595.

Iwanaga, S., Isawa, H., Yuda, M., 2014. Horizontal gene transfer of a vertebrate vasodilatory hormone into ticks. Nat. Commun. 5, 3373. https://doi.org/10.1038/ncomms4373.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.

Köhler, C.F., Holding, M.L., Sprong, H., Jansen, P.A., Esser, H.J., 2023. Biodiversity in the lyme-light: Ecological restoration and tick-borne diseases in Europe. Trends Parasitol. 39, 373–385.

Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P.A., 2019. Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 37, 540–546.

Krawczyk, A.I., Röttjers, S., Coimbra-Dores, M.J., Heylen, D., Fonville, M., Takken, W., et al., 2022. Tick microbial associations at the crossroad of horizontal and vertical transmission pathways. Parasites Vectors 15, 380. https://doi.org/10.1186/s13071-022-05519-w.

Lesiczka, P.M., Azagi, T., Krawczyk, A.I., Scott, W.T., Dirks, R.P., Šimo, L., et al., 2024. Deep sequencing of 16 *Ixodes ricinus* ticks unveils insights into their interactions with endosymbionts. bioRxiv. https://doi.org/10.1101/2024.04.22.590557.

Li, H., 2018. Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics 34, 3094–3100.

Lindgren, E., Tälleklint, L., Polfeldt, T., 2000. Impact of climatic change on the northern latitude limit and population density of the disease-transmitting European tick *Ixodes ricinus*. Environ. Health Perspect. 108, 119–123.

Lineberry, M.W., Grant, A.N., Sundstrom, K.D., Little, S.E., Allen, K.E., 2022. Diversity and geographic distribution of rickettsial agents identified in brown dog ticks from across the United States. Ticks Tick Borne Dis. 13, 102050. https://doi.org/10.1016/j.ttbdis.2022.102050.

Medlock, J.M., Hansford, K.M., Bormane, A., Derdakova, M., Estrada-Peña, A., George, J.-C., et al., 2013. Driving forces for changes in geographical distribution of *Ixodes ricinus* ticks in Europe. Parasites Vectors 6, 1. https://doi.org/10.1186/1756-3305-6-1.

Minh, B.Q., Nguyen, M.A.T., Von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.

Moore, L.D., Ballinger, M.J., 2023. The toxins of vertically transmitted *Spiroplasma*. Front. Microbiol. 14, 1148263. https://doi.org/10.3389/fmicb.2023.1148263.

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Noureddine, R., Chauvin, A., Plantard, O., 2011. Lack of genetic structure among Eurasian populations of the tick *Ixodes ricinus* contrasts with marked divergence from north-African populations. Int. J. Parasitol. 41, 183–192.

Pal, U., Li, X., Wang, T., Montgomery, R.R., Ramamoorthi, N., deSilva, A.M., et al., 2004. TROSPA, an *Ixodes scapularis* receptor for *Borrelia burgdorferi*. Cell 119, 457–468.

Parola, P., Paddock, C.D., Socolovschi, C., Labruna, M.B., Mediannikov, O., Kernif, T., et al., 2013. Update on tick-borne rickettsioses around the world: A geographic approach. Clin. Microbiol. Rev. 26, 657–702.

Pérez-Eid, C., 2007. Les tiques: Identification, biologie, importance médicale et vétérinaire. Lavoisier, Paris, France.

Poli, P., Lenoir, J., Plantard, O., Ehrmann, S., Røed, K.H., Leinaas, H.P., et al., 2020. Strong genetic structure among populations of the tick *Ixodes ricinus* across its range. Ticks Tick Borne Dis. 11, 101509. https://doi.org/10.1016/j.ttbdis.2020.101509.

Rimphanitchayakit, V., Tassanakajon, A., 2010. Structure and function of invertebrate Kazal-type serine proteinase inhibitors. Dev. Comp. Immunol. 34, 377–386.

Rollins, R.E., Margos, G., Brachmann, A., Krebs, S., Mouchet, A., Dingemanse, N.J., et al., 2023. German *Ixodes inopinatus* samples may not actually represent this tick species. Int. J. Parasitol. 53, 751–761.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302.

Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., Goesmann, A., 2021. Bakta: Rapid and standardized annotation of bacterial genomes *via* alignment-free sequence identification. Microb. Genom. 7, 000685. https://doi.org/10.1099/mgen.0.000685.

Shiina, T., Shimizu, Y., 2020. Temperature-dependent alternative splicing of precursor mRNAs and its biological significance: A review focused on post-transcriptional regulation of a cold shock protein gene in hibernating mammals. Int. J. Mol. Sci. 21, 7599. https://doi.org/10.3390/ijms21207599.

Sprong, H., Azagi, T., Hoornstra, D., Nijhof, A.M., Knorr, S., Baarsma, M.E., Hovius, J.W., 2018. Control of Lyme borreliosis and other *Ixodes ricinus*-borne diseases. Parasites Vectors 11, 145. https://doi.org/10.1186/s13071-018-2744-5.

Stiehler, F., Steinborn, M., Scholz, S., Dey, D., Weber, A.P.M., Denton, A.K., 2021. Helixer: Cross-species gene annotation of large eukaryotic genomes using deep learning. Bioinformatics 36, 5291–5298.

Stouthamer, R., Breeuwer, J.A.J., Hurst, G.D.D., 1999. *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. Annu. Rev. Microbiol. 53, 71–102.

Tsuchida, T., Koga, R., Fujiwara, A., Fukatsu, T., 2014. Phenotypic effect of "*Candidatus* Rickettsiella viridis", a facultative symbiont of the pea aphid (*Acyrthosiphon pisum*), and its interaction with a coexisting symbiont. Appl. Environ. Microbiol. 80, 525–533.

Velez, R., de Meeûs, T., Beati, L., Younsi, H., Zhioua, E., Antunes, S., et al., 2023. Development and testing of microsatellite loci for the study of population genetics of *Ixodes ricinus* Linnaeus, 1758 and *Ixodes inopinatus* Estrada-Peña, Nava & Petney, 2014 (Acari: ixodidae) in the western Mediterranean region. Acarologia 63, 356–372.

Wint, G.R.W., Balenghien, T., Berriatua, E., Braks, M., Marsboom, C., Medlock, J., et al., 2023. VectorNet: Collaborative mapping of arthropod disease vectors in Europe and surrounding areas since 2010. Euro Surveill. 28, 2200666. https://doi.org/10.2807/1560-7917.ES.2023.28.26.2200666.

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 257. https://doi.org/10.1186/s13059-019-1891-0.

Wood, D.E., Salzberg, S.L., 2014. Kraken: Ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 15, R46. https://doi.org/10.1186/gb-2014-15-3-r46.

Xiong, Q., Wan, A.T.-Y., Liu, X., Fung, C.S.-H., Xiao, X., Malainual, N., et al., 2022. Comparative genomics reveals insights into the divergent evolution of astigmatic mites and household pest adaptations. Mol. Biol. Evol. 39, msac097. https://doi.org/10.1093/molbev/msac097.

Younsi, H., Fares, W., Cherni, S., Dachraoui, K., Barhoumi, W., Najjar, C., Zhioua, E., 2020. *Ixodes inopinatus* and *Ixodes ricinus* (Acari: Ixodidae) are sympatric ticks in North Africa. J. Med. Entomol. 57, 952–956.

Zhioua, E., Bouattour, A., Hu, C.M., Gharbi, M., Aeschliman, A., Ginsberg, H.S., Gern, L., 1999. Infection of *Ixodes ricinus* (Acari: Ixodidae) by *Borrelia burgdorferi sensu lato* in North Africa. J. Med. Entomol. 36, 216–218.