

RESEARCH

Open Access



The repetitive genome of the *Ixodes ricinus* tick reveals transposable elements have driven genome evolution in ticks

Isobel Ronai^{1†} , Rodrigo de Paula Baptista^{2,3,4,5†} , Nicole S. Paulat⁶ , Julia C. Frederick⁷ , Tal Azagi⁸ , Julian W. Bakker⁹ , Katie C. Dillon² , Hein Sprong^{8*†} , David A. Ray^{6†} and Travis C. Glenn^{2,7†}

[†]Isobel Ronai and Rodrigo de Paula Baptista contributed equally to this work.

[†]Hein Sprong, David A. Ray and Travis C. Glenn contributed equally to this work.

*Correspondence: hein.sprong@rivm.nl

¹ University of Sydney, Sydney, NSW, Australia

² Institute of Bioinformatics, University of Georgia, Athens, GA, USA

³ Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA, USA

⁴ Department of Infectious Diseases, Houston Methodist Research Institute, Houston, TX, USA

⁵ Department of Medicine, Weill Cornell Medical College, New York, NY, USA

⁶ Department of Biological Sciences, Texas Tech University, Lubbock, TX, USA

⁷ Department of Environmental Health Science, University of Georgia, Athens, GA, USA

⁸ Centre for Infectious Diseases, National Institute of Public Health and Environment, Bilthoven, The Netherlands

⁹ Laboratory of Entomology, Wageningen University & Research, Wageningen, The Netherlands

Abstract

Background: Ticks are obligate blood-feeding parasites associated with a huge diversity of diseases globally. The hard tick *Ixodes ricinus* is the key vector of Lyme borreliosis and tick-borne encephalitis in Western Eurasia. *Ixodes* ticks have large and repetitive genomes that are not yet well characterized.

Results: Here we generate two high-quality *I. ricinus* genome assemblies, with haploid genome assembly sizes of approximately 2.15 Gbp. We find transposable elements comprise at least 69% of the two *I. ricinus* genome assemblies, amongst the highest proportions found in animals. The transposable elements in ticks are highly diverse and novel, so we constructed a repeat library for ticks using our *I. ricinus* genome assemblies and the high-quality genome assembly of *I. scapularis*, another major tick vector of Lyme borreliosis. To understand the impact of transposable elements on tick genomes we compared their accumulation in the two *Ixodes* sister species. We find transposable elements in these two species to have distinctive post-speciation patterns, suggesting transposable elements are drivers of genome evolution in ticks.

Conclusions: The *I. ricinus* genome assemblies and our tick repeat library will be valuable resources for biological insights into these important ectoparasites. Our findings highlight that further research into the impact of transposable elements on the genomes of blood-feeding parasites is required.

Keywords: Castor bean tick, Ixodidae, Acari, Lyme disease, Genomics, Illumina sequencing, Oxford Nanopore sequencing

Background

Ticks in the genus *Ixodes* are important global parasites. In Western Eurasia, the hard tick *Ixodes ricinus* is associated with the major human diseases Lyme borreliosis and tick-borne encephalitis [1, 2]. This tick species is also associated with a diversity of emerging diseases, including those caused by the pathogens *Anaplasma phagocytophilum*,



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Borrelia miyamotoi, *Neoehrlichia mikurensis*, *Spiroplasma ixodetis*, *Rickettsia helvetica*, *Rickettsia monacensis* and *Babesia* species [3]. These diseases are often also of veterinary relevance, for both livestock and pets [4–6]. Furthermore, the bite of *I. ricinus* can cause non-infectious diseases, such as alpha-gal syndrome (red meat allergy) [7, 8].

The threat of tick-borne diseases associated with *I. ricinus* is escalating [9, 10]. For example, the incidence of Lyme borreliosis and tick-borne encephalitis has increased in several European countries and is appearing in new countries [11–15]. This increase in tick-borne diseases in Europe is likely caused by higher numbers of *I. ricinus* and the spread of *I. ricinus* populations [10]. The habitat expansion of *I. ricinus* is due to reforestation and increased host availability [16]. In addition, climate change is expanding the range of *I. ricinus* at its extremes of altitude in central Europe, and extremes of latitude in Scandinavia [15, 17]. New tick control strategies and vaccines are therefore urgently needed, but have been impeded by our lack of knowledge of the fundamental and evolutionary biology of *I. ricinus* [18].

While *I. ricinus* is of medical and veterinary importance, we have only a preliminary understanding of its genome. The first *I. ricinus* reference genome assembly [19] is partial (Table 1), covering only a quarter of the expected size relative to the genome assemblies for

Table 1 Genome assemblies of *Ixodes ricinus* Maya and Murphy. Comparison with the *I. ricinus* Charles River partial genome assembly [19] and *I. scapularis* current reference genome assembly [20]

	<i>I. ricinus</i> Maya	<i>I. ricinus</i> Murphy	<i>I. ricinus</i> Charles River	<i>I. scapularis</i>
(A) Genome assembly summary statistics				
NCBI accession number	JBDHMV000000000	JBDHMW000000000	GCA_000973045.2	GCF_016920785.2
Number of contigs (scaffolds)	26,099	25,161	205,422 (204,904)	3,313 (648)
Total length (bp)	2,146,299,968	2,140,780,036	514,506,549	2,226,883,318
Largest contig/scaffold (bp)	2,184,647	2,598,665	32,538	252,992,795
GC content (%)	45.86	45.86	44.98	46.16
Contig N50 (bp)	244,759	274,727	3,100	1,420,000
Contig L50	2,661	2,260	51,597	434
Number of uncalled bases (N's) per 100,000 bp	4.9	0.93	7.67	11.97
(B) Genome annotation numbers				
Genes	27,060	26,934	20,075	38,656
mRNA	30,149	30,359	24,654	34,235
Exons (includes all transcripts)	203,949	202,898	45,506	300,723
lncRNA	2,421	2,352	1,415	2,885
rRNA	136	104	4	442
snoRNA	15	13	3	19
snRNA	109	110	22	120
tRNA	1,405	1,430	152	2,747
(C) BUSCO assessment, scores are a proportion of the Arthropoda ortholog database (1,013 genes)				
Complete (%)	91.7	90.7	19.2	95.3
Complete: Single-copy (%)	88.9	86.2	18.7	89.1
Complete: Duplicated (%)	2.8	4.5	0.5	6.2
Fragmented (%)	2.7	3.1	24.6	1.3
Missing (%)	5.6	6.2	56.2	3.4

the other major Lyme borreliosis tick vector and sister species, *I. scapularis* [20, 21]. This lack of tick genome assemblies impedes comparative analyses of major tick vectors.

The genomes of ticks have been challenging to assemble due in part to their large genome sizes, abundant biological genome contamination and genome complexity [20–26]. In particular, a large proportion of tick genomes are repetitive, with repeats comprising more than two-thirds of the genome assembly in *Ixodes* species [20, 24, 27]. However, the major component of these repetitive regions, the transposable elements (TEs), remain largely undescribed in ticks [22]. For example, we previously identified 46% of the TEs in the *I. scapularis* genome assembly as unknown [20]. In addition, the current genome assemblies for ticks have used pools of ticks [23] or focused on producing a consensus sequence from a single individual [20, 24], independent genome assemblies from individual ticks have not yet been generated. The generation of genome assemblies from multiple individual ticks would enable intraspecies comparisons and improve our understanding of tick biology and genomics. Here we generate genome assemblies for two individual field-caught *I. ricinus* females (Maya and Murphy) using a combination of long-read Oxford Nanopore sequencing and Illumina short-reads. We annotate these assemblies and characterize the novel TEs present in ticks. These two new *I. ricinus* genome assemblies enable comparisons at both the intraspecies and interspecies levels.

Results

Two high-quality *I. ricinus* genome assemblies

To obtain *I. ricinus* genome assemblies with minimal biological genome contamination we conducted stringent filtering of the sequencing reads from *I. ricinus* Maya and *I. ricinus* Murphy. From the total number of reads, 78.6% of reads for *I. ricinus* Maya and 80.2% for Murphy mapped to *Ixodes* genome assemblies (Additional file 1: Fig. S1). In addition, 10.4% of the total reads for Maya and 17.6% for Murphy were taxonomically unclassified (Additional file 1: Fig. S1) and were therefore retained as they could potentially be from *I. ricinus*. Both *I. ricinus* Maya and Murphy included biological contamination from the host blood they had been fed on as nymphs (0.8% and 1.3%, respectively; Additional file 1: Fig. S1) and microbiota (10.2% and 0.8%, respectively; Additional file 1: Fig. S1). The host DNA contamination was filtered out from the reads and the microbiota DNA contamination filtered out from the genome assemblies.

We obtained high-quality genome assemblies for *I. ricinus* Maya and Murphy. The starting genome assembly sizes were 3.3 Gb and after purging the haplotigs, observed a haploid genome assembly size of 2.146 Gbp for Maya and 2.140 Gbp for Murphy (Table 1A). Both Maya and Murphy have similar GC content distribution and no sequence gaps. These genome assemblies are some of the first released tick assemblies that use Oxford Nanopore sequencing. While our Oxford Nanopore based genome assemblies generated more contigs than the PacBio based *I. scapularis* reference genome assembly (Table 1A), the highest quality *Ixodes* genome assembly currently available, these genome assemblies have comparable BUSCO completeness (Table 1C). Overall, the *I. ricinus* Maya genome assembly was of slightly higher completeness than Murphy (Table 1), so we designate Maya as our *I. ricinus* reference genome assembly. Both assemblies are of higher completeness than the first *I. ricinus* reference genome assembly [19].

The genome assemblies of *I. ricinus* Maya and Murphy had a similar number of gene predictions, with approximately 27,000 genes and 204,000 exons (Table 1B). While the number of genes for *I. ricinus* Maya and Murphy is less than the *I. scapularis* genome assembly (Table 1B), most likely this is due to the *I. scapularis* genome assembly being annotated with NCBI's RefSeq pipeline rather than a biological difference. The gene density per Gbp in *I. ricinus* is approximately 12,600 and there is an average of 7.5 exons per gene. The average intron length for *I. ricinus* Maya was 3,308 bp and for *I. ricinus* Murphy 3,296 bp. The total coding sequence length for *I. ricinus* Maya is 37,691,999 bp and *I. ricinus* Murphy 37,663,835 bp, which is only 1.76% of both genome assemblies. In addition, the introns are 24.63% of *I. ricinus* Maya and 24.53% *I. ricinus* Murphy.

Comparative genome analysis of key Lyme borreliosis tick vectors (*Ixodes* species)

We compared the whole genome assemblies of the tick vector for Lyme borreliosis in Western Eurasia (*I. ricinus*) and the high-quality genome assembly of the tick vector in North America (*I. scapularis*). Direct mapping of the assembled *I. ricinus* Maya and Murphy against *I. scapularis* gave an average identity of 84%. As a chromosome-anchored *Ixodes* genome assembly is not yet available, we mapped the *I. ricinus* contigs against the 14 chromosome-scale scaffolds of *I. scapularis* and each *I. ricinus* scaffold was labelled with a unique letter (Fig. 1 and Additional file 1: Figs. S2–S4). A total of 18.6% (53% of contigs) and 17.3% (54% of contigs) of the total genome assembly length were not included in the chromosome-scale scaffolds from *I. ricinus* Maya and Murphy, respectively. The top 14 scaffolds represent the majority of the assembly, missing less than 6% of the genome completeness (Additional file 2: Table S1A). In addition, these top 14 scaffolds complete BUSCO scores are over 94% and 96%, from *I. ricinus* Maya and Murphy, respectively (Additional file 2: Table S1B).

We had the opportunity to compare the genome assemblies of two individual ticks from the same species. Consistent with the strong chromosome-scale synteny (Additional file 1: Fig. S4), the two *I. ricinus* assemblies were highly concordant: the 14 Maya pseudochromosomes align one-to-one with their Murphy counterparts that were scaffolded using the *I. scapularis* genome assembly. Most discordances were concentrated in smaller, unscaffolded (unplaced) contigs, whereas the major pseudochromosome scaffolds were largely collinear. In line with this, a Mash analysis using all contigs from both assemblies (including those not shown in Additional file 1: Fig. S4) yielded an estimated ANI of 96.96% between Maya and Murphy, indicating limited sequence divergence.

The two *I. ricinus* genome assemblies and the *I. scapularis* genome assembly share 13,963 orthogroups (Additional file 1: Fig. S5). These shared orthogroups are determined by 18,468 transcripts in *I. ricinus* Maya, 18,487 transcripts in *I. ricinus* Murphy and 22,206 transcripts in *I. scapularis*. There are 14,942 unique gene IDs in *I. ricinus* Maya for the shared orthogroups, 14,918 in *I. ricinus* Murphy and 17,807 in *I. scapularis*.

(See figure on next page.)

Fig. 1 Chromosome-scale scaffold synteny of *Ixodes scapularis* and *I. ricinus*. **A** The 14 chromosome-scale scaffolds of *I. scapularis* (gray) with the 14 corresponding scaffolds of *I. ricinus* Maya (colored); **B** the largest scaffold of *I. scapularis* with scaffold “a” of *I. ricinus* Maya and *I. ricinus* Murphy. Colors denote shared sequences between the chromosome-scale scaffolds of the assemblies, white gaps denote a lack of shared sequences and gray lines denote the repetitive hits

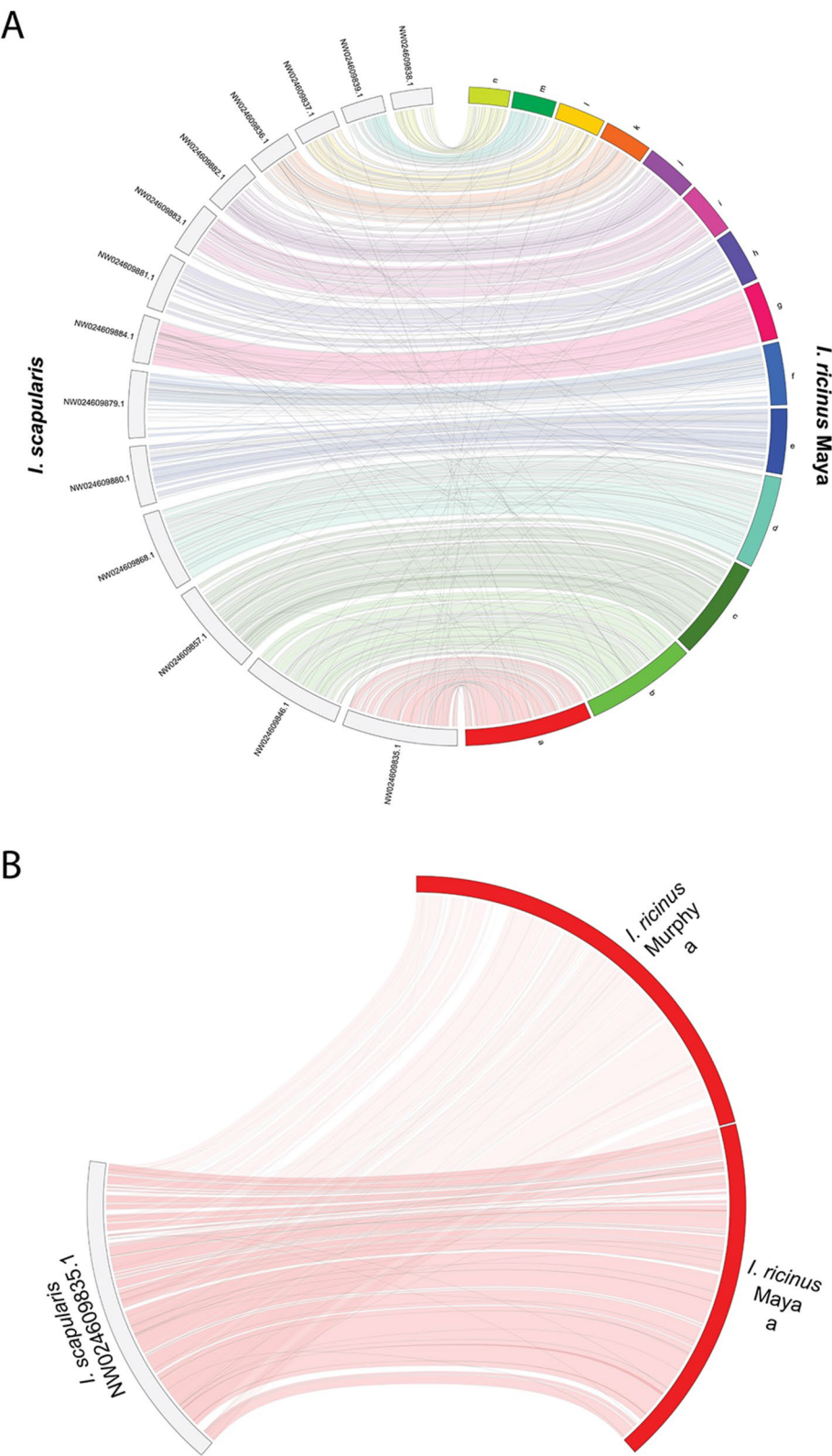


Fig. 1 (See legend on previous page.)

Within the shared orthogroups there are 9,703 that have a single transcript for all three genome assemblies, all other shared orthogroups have varying transcript amounts among the species. Looking at orthogroups unique to each *Ixodes* species, there are 4,408 *I. ricinus*-specific orthogroups which are made up of 4,619 transcripts with 4,119 unique gene IDs for Maya, whereas Murphy consists of 4,588 transcripts with 4,084 unique gene IDs. There are 856 *I. scapularis*-specific orthogroups which are made up of 3,340 transcripts, and all orthogroups were determined by multiple transcripts with 2,460 unique gene IDs. In comparison, 1,008 unique orthogroups for the *I. ricinus* genome assemblies were determined by a single transcript each.

Annotation of repetitive DNA in *Ixodes* ticks and creation of a tick repeat library

Currently an “*Ixodes*” query in RepBase, the key repository of TE consensus sequences, retrieves only 65 entries and 77% of these entries belong to a single Long Terminal Repeat (LTR) superfamily, Gypsy. Using RepeatModeler we recovered over 4,500 putative TE consensus sequences across the two *I. ricinus* assemblies and the *I. scapularis* assembly. Our final tick repeat library contains 3,684 consensus sequences that met the 80–80–80 rule [28] and were therefore classified as belonging to families assignable to known TE orders and families (Additional file 2: Table S2).

The *Ixodes* repeat library has 30 TE superfamilies (Additional file 2: Table S2). The library is dominated by Terminal Inverted Repeat (TIR) transposons (43% of the library). The most numerous superfamily in the library was TIR-like, which have consensus sequences that harbor TIRs but do not have any obvious target site duplications (29% of the entire library and 67% of TIRs in the library). Gypsy LTR elements were the second most prevalent superfamily in the library (15% of the entire library) and the third was R1 Long Interspersed Elements (LINEs) (8% of the entire library).

Characterization of the repetitive *Ixodes* genome

We find the *I. ricinus* genome assemblies are at least 73% repetitive DNA, of which 69% is TEs (Fig. 2 and Additional file 1: Fig. S6A, Additional file 2: Table S2). The *I. scapularis* genome assembly is at least 74% repetitive DNA, of which 68% is TEs (Additional file 1: Fig. S6B, Additional file 2: Table S2). Our genome assembly proportions for repetitive DNA in *Ixodes* are slightly higher than identified in previous studies [20, 23–25, 27], likely due to their use of traditional annotation pipelines, which use homology-based approaches or apply the raw output of TE-finders to the genome assembly. These traditional pipelines underestimate TE content, especially for non-model species [29, 30].

The TEs in the *Ixodes* genome assemblies (Additional file 2: Table S2) are dominated by TIR transposons (34–35% of each genome assembly), predominantly the aforementioned TIR-like elements (21–22%), but also the TcMariner and piggyBac superfamilies. The next most common TEs are LINE retrotransposons (15% of each genome assembly), mostly CR1, R1, I/Jockey, L2, and L1/Tx1 superfamilies. The third most common TEs are LTR retrotransposons (13–14% of each genome assembly), which are predominantly solo LTRs and members of the Gypsy superfamilies. Fourth are Penelope-like Element (PLE) retrotransposons (3% of each genome assembly). The fifth most common TEs are the rolling-circle (RC) transposons (2% of each genome assembly), mainly Helentrons.

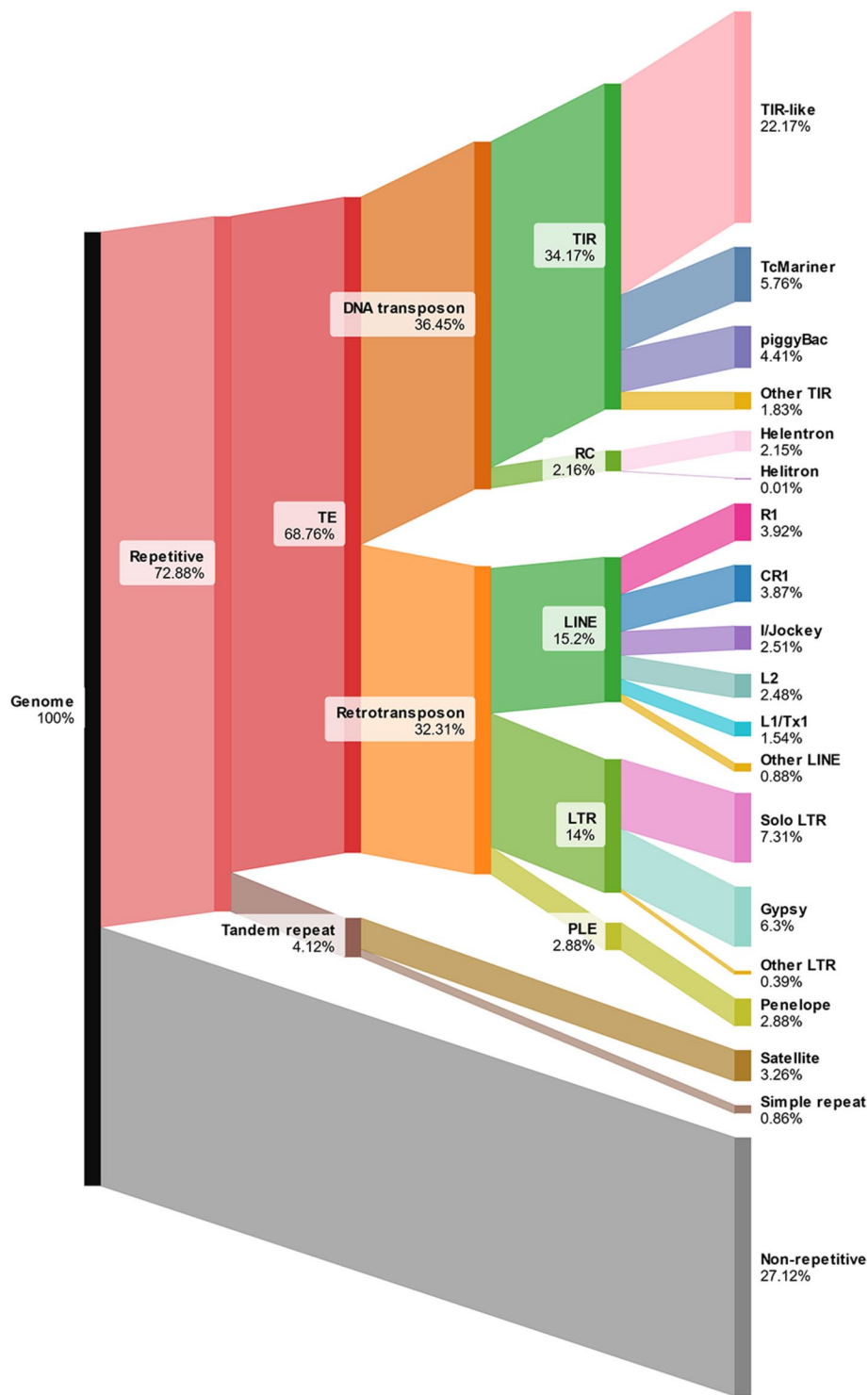


Fig. 2 Genome assembly proportions of repetitive and non-repetitive DNA in *Ixodes ricinus* Maya. The majority of the *I. ricinus* genome assembly is repetitive DNA, dominated by transposable elements (TEs): Terminal Inverted Repeat (TIR) transposons; Long Interspersed Element (LINE) retrotransposons; Long Terminal Repeat (LTR) retrotransposons; Penelope-like (PLE) retrotransposons; and rolling-circle (RC) transposons. Genome assembly proportions of repetitive and non-repetitive DNA for *I. ricinus* Murphy and *I. scapularis* see Supplemental Material (Additional file 1: Fig. S6)

The tandem repeats in the *Ixodes* genome assemblies (Additional file 2: Table S2) are dominated by satellite DNA (3–4% of each genome assembly). This satellite DNA is often associated with non-LTR LINEs, comprising approximately one quarter of the total satellite content in *I. ricinus* and one third in *I. scapularis*.

Interspecies accumulation patterns of TEs in *Ixodes*

Recent TE expansions in the *Ixodes* genomes have occurred, indicated by the left-skew of the TE landscapes (Fig. 3). Post-divergence of *I. ricinus* and *I. scapularis* each species has a distinct TE accumulation landscape. For example, most recently in *I. ricinus* there has been an increase in Helentron elements, whereas this superfamily has lower rates of accumulation in *I. scapularis* (Fig. 3). Also, most recently in *I. ricinus* there has been a decrease in TcMariner accumulation, which is not seen in *I. scapularis* (Fig. 3). In *I. ricinus* the Sola superfamily has been largely dormant throughout its evolutionary history and in *I. scapularis* this superfamily recently expanded (Fig. 3).

Intraspecies accumulation patterns of TEs in *I. ricinus*

The generation of two individual *I. ricinus* genome assemblies provides the opportunity to compare the TE accumulation patterns in ticks within a comparably short timescale. Overall, the TE accumulations in the two *I. ricinus* genome assemblies are an almost exact match, with only minor differences in the genome assembly proportions (Fig. 3; Additional file 3; Additional file 4; Additional file 5). The two *I. ricinus* genome assemblies have small differences in the total content of some TE orders, for example, Helentron differs by approximately 61 Kbp and 747 insertions (Additional file 3; Additional file 4). While some of this difference in TE content between individuals may be due to scaffold assembly errors, these differences between the two *I. ricinus* assemblies are comparable with the extent of TE polymorphisms found in highly contiguous genome assemblies of other organisms [32–34].

Discussion

We have generated two high-quality genome assemblies for the major disease vector *I. ricinus*. These assemblies have comparable quality to the genome assemblies of other tick species such as *I. scapularis* [20, 21] and *I. persulcatus* [24]. The genome assembly size of *I. ricinus* is approximately 2.15 Gbp. This assembly size is many fold larger than the genome assemblies of the most well-studied insects *Drosophila melanogaster* 0.14 Gbp [35] and *Anopheles gambiae* 0.26 Gbp [36], plus larger than closer related arthropods such as *Varroa destructor* 0.37 Gbp [37] and *Parasteatoda tepidariorum* 1.4 Gbp [38]. The large genome size of *Ixodes* ticks is mostly due to the high proportion of repetitive DNA their genomes contain, which is commonly found in arthropods with gigantic genome sizes [22, 39–41].

Our repetitive DNA curation finds nearly three quarters of the *I. ricinus* genome assemblies are repeat-derived, with almost all the repetitive DNA being derived from TEs. *Ixodes* ticks have one of the highest proportions of TEs currently reported for an animal genome assembly [22, 29, 41]. Other tick species also have high proportions of

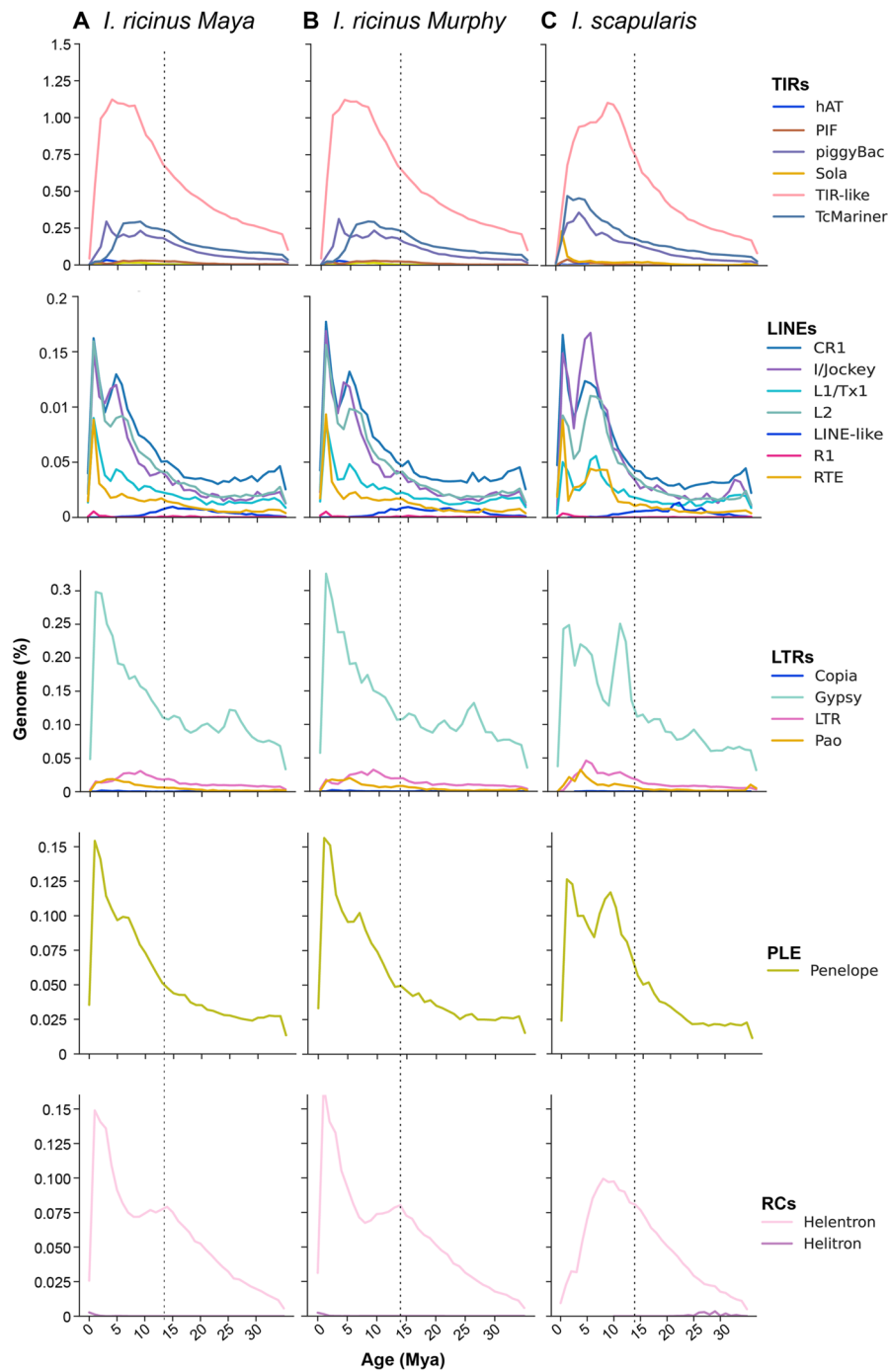


Fig. 3 Transposable element timeline for *Ixodes* species. **A** *I. ricinus* Maya, **B** *I. ricinus* Murphy, and **C** *I. scapularis* reference genome assembly. The genome assembly percentage of each TE superfamily was calculated using the total TE superfamily content in bp for each age bin and dividing by the total genome assembly size. The estimated time of accumulation is based on a neutral mutation rate of 7.68×10^{-9} substitutions/site/million years and divergences between individual elements and their respective consensus sequences. Each row designates a TE order (TIR DNA transposons (TIRs); non-Long Terminal Repeats retrotransposons (non-LTRs); Long Terminal Repeats retrotransposons (LTRs); Penelope-like elements retrotransposons (PLEs); and rolling circle transposons (RCs). The vertical dashed line indicates the approximate divergence time between *I. ricinus* and *I. scapularis* of 13 million years ago [31]

TEs, ranging from over 25% to 64% of the genome assembly [24, 42]. In contrast, the proportion of TEs in mosquitoes can range from almost zero to 50% of the genome assembly [43].

TE-driven genome evolution has occurred in *Ixodes* ticks relatively recently, within the last approximately 13 million years [31]. In two sister lineages, *I. ricinus* and *I. scapularis*, we find the TE proportions have distinctive post-speciation trajectories. The TEs in these *Ixodes* tick species are often highly active and rapidly turned over. This TE pattern suggests that TEs accumulate and then are quickly removed from the genomes of *Ixodes* ticks via nonhomologous recombination, which is known to occur in other animals [40, 44–47]. Therefore, rapid structural change in the genomes of the two *Ixodes* sister lineages likely occurred once the two *Ixodes* species separated. TEs are known drivers of species diversity [48–55] and we propose this might be occurring for *Ixodes* ticks.

There is typically a positive relationship between TE content and genome size [56]. Given this relationship, one would expect the other tick genera that have larger genomes than *Ixodes* [24] to exhibit higher proportions of TEs. As our analysis of *Ixodes* TEs is the most extensive and detailed examination of tick TEs to date, previous analyses may have underestimated TE content in other tick species. Therefore, future work should focus on characterizing the TEs in tick species with the largest genomes, such as *Haemaphysalis longicornis*, which has a genome assembly size of 3.16 Gbp [57].

The abundance and diversity of TEs in the genomes of ticks could be due to their obligate blood-feeding parasitic lifestyle. Ticks have a close association with their vertebrate hosts (hard ticks take three prolonged blood meals during their lifecycle) and subsequently contain a plethora of microbiota. These close associations are reflected in our *I. ricinus* sequencing reads, as we found DNA contamination from the tick's vertebrate host and microbiota. We suggest tick DNA, with its close association to the DNA of other organisms, may be a conducive environment for horizontal transfer of TEs. There is evidence that horizontal transfer of TEs can occur from eukaryote parasites to their hosts and viruses to eukaryotes [58–62]. For example the LINE BovB was likely horizontally transferred to different vertebrates via a tick [63]. Also, the host preferences of a blood-feeder might be important for TE abundance and diversity, as not all blood-feeders have high proportions of TEs [43]. We suggest future studies should investigate whether tick species and other blood-feeders that are host-generalists have greater TE diversity compared to those that are host-specialists. Other environments that might be associated with TE abundance, include the aquatic environment with the gigantic genomes of krill [41], lungfish [64, 65], caecilians [66] and axolotls [67].

Conclusions

Our two genome assemblies for the tick *I. ricinus* will facilitate a deeper understanding of the fundamental biology of *Ixodes* species in general and *I. ricinus* specifically. Comparative genomics for ticks will be further enhanced with newly available *Ixodes* genome assemblies [68, 69] and population genetic studies^{c.f.} [70]. More broadly our findings indicate further studies of TEs in ticks and other blood-feeding parasites are required for a more comprehensive understanding of the evolutionary history of these medically and veterinary important parasites.

Methods

Tick biological samples

We collected questing *I. ricinus* nymphs in September 2020 by blanket dragging in De Dorschkamp, a deciduous forest between Wageningen and Renkum, the Netherlands (51°58′39.5″N, 5°41′59.3″E). These nymphs are representative of the wild population of *I. ricinus* and were morphologically identified to species level using an identification key [71]. To obtain adult *I. ricinus* we fed the nymphs on an artificial membrane blood feeding system, as described in [72]. The bovine blood was obtained from Carus (Wageningen University, The Netherlands), under animal ethics protocol no. AVD1040020173624. Engorged nymphs were collected from the feeding membrane and incubated at room temperature until they molted to adults. The adults were collected and stored at −80° C. We designated *I. ricinus* adult female #1 as biospecimen Maya1009 and *I. ricinus* adult female #2 as biospecimen Murphy0812.

DNA isolation and sequencing (Oxford Nanopore Technologies and Illumina)

DNA extraction and sequencing was performed at Future Genomics Technologies BV (Leiden, the Netherlands). The two individual, whole *I. ricinus* adults were ground to a fine powder using a pestle and mortar with liquid nitrogen. To extract high molecular weight DNA a Genomic-tip 20/G kit (Qiagen Benelux BV, Venlo, the Netherlands) was used according to the standard protocol. DNA quality was measured via electrophoresis in Genomic DNA ScreenTape on an Agilent 4200 TapeStation System (Agilent Technologies Netherlands BV, Amstelveen, the Netherlands) and total DNA measured using a Qubit 3.0 Fluorometer (Life Technologies Europe BV, Bleiswijk, the Netherlands); we obtained between 3.05 µg to 5.96 µg of total DNA.

The DNA samples were used to prepare a 1D ligation library using the Ligation Sequencing Kit SQK-LSK110 according to the manufacturer's instructions (Oxford Nanopore Technologies, Oxford, United Kingdom). Oxford Nanopore Technologies (ONT) libraries were first tested on a MinION flowcell (FLO-MIN106) and subsequently run on an R9.4.1 PromethION flowcell (FLO-PRO002) using the following settings: basecall model: high-accuracy; basecaller version: 4.0.11 (PromethION).

Parallel aliquots of the DNA samples used for ONT sequencing, were used to prepare Illumina libraries using the Nextera DNA Flex Library Prep Kit according to the manufacturers' instructions (Illumina Inc. San Diego, CA, USA). Library quality was measured via electrophoresis in D1000 ScreenTape on an Agilent 4200 TapeStation System (Agilent Technologies Netherlands BV, Amstelveen, the Netherlands). The genomic paired-end (PE) libraries were sequenced with a read length of 2 × 150 nt using the Illumina NovaSeq 6000 system. The ONT sequences underwent quality assessment using Nanoplot v1.42 [73] and Illumina sequences were assessed using MultiQC v1.14 [74].

Genome assembly and quality assessment

The DNA sequence data of Maya1009 and Murphy0812 were submitted to a custom hybrid genome sequence pipeline (Additional file 1: Fig. S7). First, ONT reads were filtered using yacrd v.0.6.2 [75] to remove potential chimeric reads generated by the sequencing ligation kit. Then bovine blood meal DNA contamination was removed from

the filtered reads using a custom database containing *Bos taurus* (GCA_002263795.3). The ONT reads were then assembled using Flye 2.9.1 [76]. The final scaffolds generated by the pipeline was then polished using Illumina short reads and NextPolish v1.4.0 [77] with three reruns. To predict the putative haploid genome size of the assemblies, we used purge_haplotigs v1.1.2 [78], to remove redundant contigs.

To check for any DNA contamination on the polished deduplicated assemblies we used NCBI's Foreign Contamination Screen tool [79], gx-db version 0.4.0 and taxonid 6944 for *Ixodes* species. This tool checks if contamination is detected as a contig or within an assembled contig. Any contaminant we detected in the assemblies was removed or hardmasked using bedtools v2.30.0 [80]. To quantify the amount of biological contamination we calculated the proportion of blood meal and microbiota reads in the two genome assemblies.

To assess the completeness of the two genome assemblies and the first *I. ricinus* reference genome assembly [19], we submitted them to Benchmarking Universal Single-copy Orthologs (BUSCO) v5 [81] using the Arthropoda database. In parallel, the genome assemblies were annotated using a combination of AUGUSTUS v3.4.0 [82] gene prediction and Liftoff v1.6.3 [83] annotation transfer tool using the recent *I. scapularis* genome assembly annotation (NCBI Assembly GCA_016920785.2 [20]) as a template.

Gene and gene family annotation

The annotated Maya1009 and Murphy0812 gff files, plus the gff file associated with the current *I. scapularis* assembly (NCBI Assembly GCA_016920785.2 [20]) were submitted to agat_sp_extract_sequences.pl from the AGAT v1.0 pipeline [84] to obtain all annotated protein sequences in a FASTA. Those protein FASTA files were used to run an ortholog comparison for the *I. ricinus* samples (Maya1009, Murphy0812) using Orthofinder v2.5.2 [85]. The overlap of the resulting Orthogroups were then plotted using R package VennDiagram v1.7.3 [86]. To examine the gene families that were unique to the *Ixodes* species, we ran Orthofinder v2.5.2 to compare the *I. scapularis* current reference assembly and our *I. ricinus* assembly (Maya1009). Genes that were only found in the *I. ricinus* sample Maya1009 were extracted from the dataset with the corresponding transcript information. These steps were repeated for the orthogroups unique to the *I. scapularis* genome assembly.

Whole genome comparison

We compared the Maya1009 and Murphy0812 genome assemblies to the current reference genome assembly of *I. scapularis* (NCBI Assembly GCA_016920785.2 [20]), which is the highest quality *Ixodes* genome assembly with chromosome-scale scaffolds currently available. The average identity between the *I. ricinus* genome assemblies and *I. scapularis* was calculated using FastANI v1.33 [87]. To obtain an overall comparison at a chromosome-scale level between the two *Ixodes* species, we first hard masked for repeats in both *I. ricinus* genome assemblies using RepeatModeler v2.0.3 [88]. Then the masked *I. ricinus* contigs were "scaffolded" to the 14 chromosome-scale scaffolds (i.e. the 14 longest scaffolds) from the *I. scapularis* reference genome assembly (representing almost 90% of the *I. scapularis* total genome size [20]) using RagTag v2.1.0 [89]. This step was required because the majority of the *I. ricinus* scaffolds are shorter than the *I.*

scapularis reference genome assembly scaffolds. The scaffolded contig IDs for Maya1009 (Additional file 2: Table S3A) and Murphy0812 (Additional file 2: Table S3B). In addition, we assessed assembly quality using QUAST-LG v.5.2.0 [90] and determined their BUSCO score using compleasm v.0.2.5 [91]. We conducted synteny analysis using minimap2 v.2.2.24 [92] and Circos [93] was used to make the synteny plots. We also compared the Maya1009 and Murphy0812 assemblies using Mash [94].

Repetitive DNA identification and characterization

The repetitive DNA in *Ixodes* ticks was subjected to detailed manual annotation. We had previously characterized the repetitive DNA of *I. scapularis* using a standard RepeatModeler/RepeatMasker pipeline [20]. Here we generated a de novo repeat library, using a custom pipeline, for the genome assemblies of *I. ricinus* (Maya1009 and Murphy0812) and *I. scapularis* [20]. First we used RepeatModeler v.2.0.3 [88] to produce putative consensus sequences that we extracted (https://github.com/davidaray/bioinfo_tools/blob/master/extract_align.py) and subjected to extension using RAM [95]. The putative consensus sequences were categorized as a LTR retrotransposon, LINE retrotransposon, Short INterspersed Element (SINE) retrotransposon, PLE retrotransposon, TIR transposon, RC transposon, Maverick transposon or unidentified. This categorization was done using a custom bash script (TEcurate.sh) of RepeatClassifier (part of the RepeatModeler package) and BLASTP searches (e-values $> 1e^{-50}$) against a database of known autonomous TEs. We then used a custom curation script (https://github.com/davidaray/bioinfo_tools/blob/master/TEcurate.sh) and the TE-Aid (<https://github.com/clemgoub/TE-Aid>) package to generate genome assembly coverage plots, self-alignment dot-plots, structure and ORF plots, and copy number estimates. For sequences categorized as 'unidentified', we determined likely group membership using the TE-Aid plots to identify structural hallmarks (i.e. TIRs and LTRs) and sequence characteristics (repetitive tails, Helitron-specific CTRR motifs, SINE A-B boxes, piggyBac TTAA target site duplications, etc.), if the TE remained unidentified it was categorized as unknown. The resulting tick repeat library was combined with known arthropod TEs in RepBase and the TEs we had previously identified in *I. scapularis* [20] to create a custom library that was applied to Maya, Murphy, and the *I. scapularis* assembly. This combined library was subjected to analysis using cd-hit-est and a 90% similarity cutoff to eliminate highly similar consensus sequences.

For the TEs we used the identifiers provided by TEcurate.sh to generate a unique identifier that included the species of origin, RepeatModeler ID, and TE Class/Family information. For example, iRic1.5.2409#LINE/I: identified in Maya1009 (Murphy0812 is designated as iRic2), RepeatModeler ID rnd-5_family-2409 (5.2409) and RepeatClassifier/blastp analysis identified it as a LINE element of the I family. For the unidentified TEs we used visual examination and categorization to assign an appropriate identifier.

We determined the TE consensus sequences. The LTR and PLE retrotransposons we identified were subjected to postprocessing to obtain consensus sequences. LTRs were processed by hand to subdivide them into their LTR and internal segments [96]. PLEs often insert as tandemly repeated arrays, so we used the TE-Aid output to split the tandem arrays into a single representative full-length consensus. RAM occasionally falsely identifies segmental duplications as TEs. Therefore, any low-copy (< 50

copies) consensus sequences greater than 10 kb and with no consistent TE structural hallmarks were assumed to be segmental duplications or assembly artifacts and removed from the library. Consensus sequences with fewer than 10 full-length copies were also removed from the library. We submitted all the novel TE consensus sequences we identified to the Dfam TE database.

We identified that the LINE superfamilies CR1 and L2 were commonly associated with satellite repeats at their 5' ends. Ten randomly selected consensus sequences were subjected to further characterization to check whether these repeats were an artifact of genome assembly. The ONT reads were searched using 400 bp queries consisting of 200 bp of the upstream satellite sequence and 200 bp of the downstream LINE sequence. In all ten test cases, the queries were identified in the original reads, confirming that these satellite/LINE chimeras exist in the genome of these ticks. To aid with submission to the Dfam repeat database, we separated the satellite sequences from the LINE consensus sequences and included both in the final tick repeat library. For example, the CR1 consensus iRic2.5.1902 was subdivided into its satellite component (iRic2.5.1902S) and its LINE component (iRic2.5.1902L). These satellite repeats are a significant portion of the consensus sequences. For example, in iRic2.5.1902 the satellite consists of a 316 bp repeated 13.2 times and comprises more than half of the original consensus (4,211 bp versus 3,760 bp derived from the CR1 portion).

Our tick repeat library was used to mask the *I. ricinus* and *I. scapularis* genome assemblies with RepeatMasker. The output was processed to eliminate any overlapping hits using RM2Bed.py, part of the RepeatMasker installation package, generating BED files for downstream analyses. In addition, we used the tick repeat library to calculate the proportion of each TE superfamily in the three genome assemblies and then created a Sankey diagram using SankeyMATIC.

We obtained TE accumulation time estimates by estimating a neutral mutation rate for *Ixodes*. We aligned randomly selected orthologous introns (example, Additional file 2: Table S4) from *I. ricinus* and *I. scapularis* assemblies in MEGA v11 [97] under the maximum composite likelihood model (0.999). Then, using a species divergence time of 13 million years ago (see Supplementary Material of [31]), we calculated a neutral mutation rate of approximately 7.68×10^{-9} substitutions/site/million years for *Ixodes*.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-025-03909-8>.

Additional file 1. Supplementary figures.
Additional file 2. Supplementary tables.
Additional file 3. RepeatMasker output for *Ixodes ricinus* Maya.
Additional file 4. RepeatMasker output for *Ixodes ricinus* Murphy.
Additional file 5. RepeatMasker output for *Ixodes scapularis*.

Acknowledgements

We thank Ron Dirks (Future Genomics Technologies, Leiden, The Netherlands) for invaluable advice on sample preparation and next generation sequencing. The High Performance Computing Center at Texas Tech University contributed valuable bioinformatic resources.

Peer review information

Tim Sands was the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions

Conceptualization: IR, HS, TCG Validation: RPB, NSP, JCF, KCD, DAR Formal analysis: RPB, NSP, JCF, KCD, DAR Investigation: TA, JWB Resources: TA, JWB Data Curation: RPB, NSP, DAR Writing—Original Draft: IR, DAR Writing—Review & Editing: All Visualization: IR, RPB, NSP, JCF, DAR Supervision: HS, DAR, TCG Project administration: IR, RPB, HS, DAR, TCG.

Funding

HS and TA employment was funded by The Netherlands Organization for Health Research and Development (ZonMw, project number 52200-30-07), which peer-reviewed the grant application, and by the Dutch Ministry of Health, Welfare, and Sports (VWS).

Data availability

The datasets generated during the current study are available in NCBI, <https://www.ncbi.nlm.nih.gov/>. Raw reads and the genome assemblies are available under the BioProject PRJNA816462 [98], with Biosamples (Maya1009/SAMN26676983 and Murphy0812/SAMN26677070). The Whole Genome Shotgun projects are available under DDBJ/ENA/GenBank accessions JBDHMHV0000000000 (Maya1009) [99] and JBDHMHV0000000000 (Murphy0812) [100]. The versions described in this published article are version JBDHMHV0000000000 and JBDHMHV0000000000. The transposable element datasets (tick transposable element consensus sequences and supporting information) generated during the current study have been submitted to the Dfam transposable element repository and database, <https://dfam.org> [101]. RepeatMasker output datasets generated by DAR during this study are included in this published article's Additional files 3–5.

Declarations**Ethics approval and consent to participate**

The bovine blood, which was used for an artificial membrane blood feeding system was obtained from Carus (Wageningen University, The Netherlands), and the procedures were approved by the Animal Ethics Committees of Wageningen Research under animal ethics protocol no. AVD1040020173624. The animal experiments were approved by the Dutch Central Authority for Scientific Procedures on Animals.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 27 May 2024 Accepted: 11 December 2025

Published online: 02 January 2026

References

1. Cook MJ, Puri BK. Estimates for Lyme borreliosis infections based on models using sentinel canine and human seroprevalence data. *Infect Dis Model.* 2020;5:871–88.
2. European Centre for Disease Prevention and Control. Tick-borne encephalitis. In: ECDC. Annual epidemiological report for 2019. Stockholm: ECDC; 2021. <https://www.ecdc.europa.eu/en/publications-data/tick-borne-encephalitis-annual-epidemiological-report-2019>.
3. Sprong H, Azagi T, Hoornstra D, Nijhof AM, Knorr S, Baarsma ME, et al. Control of Lyme borreliosis and other *Ixodes ricinus*-borne diseases. *Parasit Vectors.* 2018;11:145.
4. Gilbert L. Louping ill virus in the UK: a review of the hosts, transmission and ecological consequences of control. *Exp Appl Acarol.* 2016;68:363–74.
5. Leschnik MW, Kirtz GC, Thalhammer JG. Tick-borne encephalitis (TBE) in dogs. *Int J Med Microbiol.* 2002;291(Suppl 33):66–9.
6. Pantchev N, Pluta S, Huisinga E, Nather S, Scheufelen M, Vrhovec MG, et al. Tick-borne diseases (Borreliosis, Anaplasmosis, Babesiosis) in German and Austrian dogs: status quo and review of distribution, transmission, clinical findings, diagnostics and prophylaxis. *Parasitol Res.* 2015;114(Suppl 1):S19–54.
7. Apostolovic D, Tran TAT, Starkhammar M, Sánchez-Vidaurre S, Hamsten C, Van Hage M. The red meat allergy syndrome in Sweden. *Allergo J Int.* 2016;25:49–54.
8. Bircher AJ, Hofmeier KS, Link S, Heijnen I. Food allergy to the carbohydrate galactose- α -1,3-galactose (α -gal): four case reports and a review. *Eur J Dermatol.* 2017;27:3–9.
9. Lindgren E, Andersson Y, Suk JE, Sudre B, Semenza JC. Monitoring EU emerging infectious disease risk due to climate change. *Science.* 2012;336:418–9.
10. Medlock JM, Hansford KM, Bormane A, Derdakova M, Estrada-Peña A, George J-C, et al. Driving forces for changes in geographical distribution of *Ixodes ricinus* ticks in Europe. *Parasit Vectors.* 2013;6:1.
11. Hofhuis A, Bennema S, Harms M, van Vliet AJ, Takken W, van den Wijngaard CC, et al. Decrease in tick bite consultations and stabilization of early Lyme borreliosis in the Netherlands in 2014 after 15 years of continuous increase. *BMC Public Health.* 2016;16:425.
12. Kreusch TM, Holding M, Hewson R, Harder T, Medlock JM, Hansford KM, et al. A probable case of tick-borne encephalitis (TBE) acquired in England, July 2019. *Euro Surveill.* 2019;24:1900679.

13. Kunze U. The international scientific working group on tick-borne encephalitis (ISW TBE): review of 17 years of activity and commitment. *Ticks Tick Borne Dis.* 2016;7:399–404.
14. Jaenson TG, Hjertqvist M, Bergström T, Lundkvist A. Why is tick-borne encephalitis increasing? A review of the key factors causing the increasing incidence of human TBE in Sweden. *Parasit Vectors.* 2012;5:184.
15. Nuttall P. *Climate. Ticks and Disease.* CAB; 2021.
16. Sprong H, Hofhuis A, Gassner F, Takken W, Jacobs F, van Vliet AJ, et al. Circumstantial evidence for an increase in the total number and activity of *Borrelia*-infected *Ixodes ricinus* in the Netherlands. *Parasit Vectors.* 2012;5:294.
17. Garcia-Vozmediano A, Krawczyk AI, Sprong H, Rossi L, Ramassa E, Tomassone L. Ticks climb the mountains: ixodid tick infestation and infection by tick-borne pathogens in the Western Alps. *Ticks Tick Borne Dis.* 2020;11:101489.
18. Rego ROM, Trentelman JJA, Anguita J, Nijhof AM, Sprong H, Klempa B, et al. Counterattacking the tick bite: towards a rational design of anti-tick vaccines targeting pathogen transmission. *Parasit Vectors.* 2019;12:229.
19. Cramaro WJ, Hunewald OE, Bell-Sakyi L, Muller CP. Genome scaffolding and annotation for the pathogen vector *Ixodes ricinus* by ultra-long single molecule sequencing. *Parasit Vectors.* 2017;10:71.
20. De S, Kingan SB, Kitsou C, Portik DM, Foor SD, Frederick JC, et al. A high-quality *Ixodes scapularis* genome advances tick science. *Nat Genet.* 2023;55:301–11.
21. Nuss AB, Lomas JS, Reyes JB, Garcia-Cruz O, Lei W, Sharma A, et al. The highly improved genome of *Ixodes scapularis* with X and Y pseudochromosomes. *Life Sci Alliance.* 2023;6:e202302109.
22. Petersen M, Armisén D, Gibbs RA, Hering L, Khila A, Mayer G, et al. Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Ecol Evol.* 2019;19:11.
23. Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nat Commun.* 2016;7:10507.
24. Jia N, Wang J, Shi W, Du L, Sun Y, Zhan W, et al. Large-scale comparative analyses of tick genomes elucidate their genetic diversity and vector capacities. *Cell.* 2020. <https://doi.org/10.1016/j.cell.2020.07.023>.
25. Geraci NS, Spencer Johnston J, Paul Robinson J, Wikel SK, Hill CA. Variation in genome size of argasid and ixodid ticks. *Insect Biochem Mol Biol.* 2007;37:399–408.
26. Pagel Van Zee J, Geraci NS, Guerrero FD, Wikel SK, Stuart JJ, Nene VM, et al. Tick genomics: the *Ixodes* genome project and beyond. *Int J Parasitol.* 2007;37:1297–305.
27. Ullmann A, Lima C, Guerrero F, Piesman J, Black IWW. Genome size and organization in the blacklegged tick, *Ixodes scapularis* and the Southern cattle tick, *Boophilus microplus*. *Insect Mol Biol.* 2005;14:217–22.
28. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capi P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet.* 2007;8:973–82.
29. Osmanski AB, Paulat NS, Korstian J, Grimshaw JR, Halsey M, Sullivan KAM, et al. Insights into mammalian TE diversity through the curation of 248 genome assemblies. *Science.* 2023;380:eabn1430.
30. Platt RN II, Blanco-Berdugo L, Ray DA. Accurate transposable element annotation is vital when analyzing new genome assemblies. *Genome Biol Evol.* 2016;8:403–10.
31. Mans BJ, Featherston J, Kvas M, Pillay K-A, de Klerk DG, Pienaar R, et al. Argasid and ixodid systematics: implications for soft tick evolution and systematics, with a new argasid species list. *Ticks Tick-Borne Dis.* 2019;10:219–40.
32. Springer NM, Anderson SN, Andorf CM, Ahern KR, Bai F, Barad O, et al. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat Genet.* 2018;50:1282–8.
33. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature.* 2015;526:75–81.
34. Kim S, Mun S, Kim T, Lee K-H, Kang K, Cho J-Y, et al. Transposable element-mediated structural variation analysis in dog breeds using whole-genome sequencing. *Mamm Genome.* 2019;30:289–300.
35. Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, Galle SE, et al. The release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome Res.* 2015;25:445–58.
36. Habtewold T, Wagah M, Tambwe MM, Moore S, Windbichler N, Christophides G, et al. A chromosomal reference genome sequence for the malaria mosquito, *Anopheles gambiae*, Giles, 1902, Ifakara strain. *Wellcome Open Res.* 2023;8:74.
37. Techer MA, Rane RV, Grau ML, Roberts JMK, Sullivan ST, Liachko I, et al. Divergent evolutionary trajectories following speciation in two ectoparasitic honey bee mites. *Commun Biol.* 2019;2:357.
38. Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, et al. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biol.* 2017;15:62.
39. Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, et al. The locust genome provides insight into swarm formation and long-distance flight. *Nat Commun.* 2014;5:2957.
40. Shah A, Hoffman JI, Schielzeth H. Comparative analysis of genomic repeat content in Gomphocerine grasshoppers reveals expansion of satellite DNA and Helitrons in species with unusually large genomes. *Genome Biol Evol.* 2020;12:1180–93.
41. Shao C, Sun S, Liu K, Wang J, Li S, Liu Q, et al. The enormous repetitive Antarctic krill genome reveals environmental adaptations and population insights. *Cell.* 2023;186:1279–1294.e1219.
42. Ribeiro JMC, Bayona-Vásquez NJ, Budachetri K, Kumar D, Frederick JC, Tahir F, et al. A draft of the genome of the Gulf Coast tick, *Amblyomma maculatum*. *Ticks Tick Borne Dis.* 2023;14:102090.
43. Melo ES, Wallau GL. Mosquito genomes are frequently invaded by transposable elements through horizontal transfer. *PLoS Genet.* 2020;16:e1008946.
44. Blass E, Bell M, Boissinot S. Accumulation and rapid decay of non-LTR retrotransposons in the genome of the three-spine stickleback. *Genome Biol Evol.* 2012;4:687–702.
45. Tollis M, Boissinot S. The transposable element profile of the *Anolis* genome. *Mob Genet Elem.* 2011;1:107–11.
46. Lavoie CA, Platt RN, Novick PA, Counterman BA, Ray DA. Transposable element evolution in *Heliconius* suggests genome diversity within Lepidoptera. *Mob DNA.* 2013;4:21.
47. Palacios-Gimenez OM, Koelman J, Palmada-Flores M, Bradford TM, Jones KK, Cooper SJB, et al. Comparative analysis of morabine grasshopper genomes reveals highly abundant transposable elements and rapidly proliferating satellite DNA repeats. *BMC Biol.* 2020;18:199.

48. Jurka J, Bao W, Kojima KK, Kohany O, Yurka MG. Distinct groups of repetitive families preserved in mammals correspond to different periods of regulatory innovations in vertebrates. *Biol Direct*. 2012;7:36.
49. Böhne A, Brunet F, Galiana-Arnoux D, Schultheis C, Volff J-N. Transposable elements as drivers of genomic and biological diversity in vertebrates. *Chromosome Res*. 2008;16:203–15.
50. Oliver KR, Greene WK. Transposable elements and viruses as factors in adaptation and evolution: an expansion and strengthening of the TE-thrust hypothesis. *Ecol Evol*. 2012;2:2912–33.
51. Klein SJ, O'Neill RJ. Transposable elements: genome innovation, chromosome diversity, and centromere conflict. *Chromosome Res*. 2018;26:5–23.
52. Serrato-Capuchina A, Matute DR. The role of transposable elements in speciation. *Genes*. 2018. <https://doi.org/10.3390/genes9050254>.
53. Zeh DW, Zeh JA, Ishida Y. Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays*. 2009;31:715–26.
54. Kazazian HH. Mobile elements: drivers of genome evolution. *Science*. 2004;303:1626–32.
55. Talla V, Suh A, Kalsoom F, Dincă V, Vila R, Friberg M, et al. Rapid increase in genome size as a consequence of transposable element hyperactivity in wood-white (*Leptidea*) butterflies. *Genome Biol Evol*. 2017;9:2491–505.
56. Kidwell MG. Transposable elements and the evolution of genome size in eukaryotes. *Genetica*. 2002;115:49–63.
57. Yu Z, He B, Gong Z, Liu Y, Wang Q, Yan X, et al. The new *Haemaphysalis longicornis* genome provides insights into its requisite biological traits. *Genomics*. 2022;114:110317.
58. Suh A, Witt CC, Menger J, Sadanandan KR, Podsiadlowski L, Gerth M, et al. Ancient horizontal transfers of retro-transposons between birds and ancestors of human pathogenic nematodes. *Nat Commun*. 2016;7:11396.
59. Gilbert C, Schaack S, Pace JK II, Brindley PJ, Feschotte C. A role for host–parasite interactions in the horizontal transfer of transposons across phyla. *Nature*. 2010;464:1347–50.
60. Kambayashi C, Kakehashi R, Sato Y, Mizuno H, Tanabe H, Rakotoarison A, et al. Geography-dependent horizontal gene transfer from vertebrate predators to their prey. *Mol Biol Evol*. 2022;39:msac052.
61. Gilbert C, Feschotte C. Horizontal acquisition of transposable elements and viral sequences: patterns and consequences. *Curr Opin Genet Dev*. 2018;49:15–24.
62. Paulat NS, Storer JM, Moreno-Santillán DD, Osmanski AB, Sullivan KAM, Grimshaw JR, et al. Chiropterans are a hotspot for horizontal transfer of DNA transposons in mammalia. *Mol Biol Evol*. 2023;40:msad092.
63. Walsh AM, Kortschak RD, Gardner MG, Bertozzi T, Adelson DL. Widespread horizontal transfer of retrotransposons. *Proc Natl Acad Sci U S A*. 2013;110:1012–6.
64. Wang K, Wang J, Zhu C, Yang L, Ren Y, Ruan J, et al. African lungfish genome sheds light on the vertebrate water-to-land transition. *Cell*. 2021;184:1362–1376.e1318.
65. Meyer A, Schloissnig S, Franchini P, Du K, Wolterting JM, Irisarri I, et al. Giant lungfish genome elucidates the conquest of land by vertebrates. *Nature*. 2021;590:284–9.
66. Wang J, Itgen MW, Wang H, Gong Y, Jiang J, Li J, et al. Gigantic genomes provide empirical tests of transposable element dynamics models. *Genomics, Proteomics & Bioinformatics*. 2021;19:123–39.
67. Nowoshilow S, Schloissnig S, Fei J-F, Dahl A, Pang AWC, Pippel M, et al. The axolotl genome and the evolution of key tissue formation regulators. *Nature*. 2018;554:50–5.
68. de Cerqueira Araujo A, Noel B, Bretaudeau A, Labadie K, Boudet M, Tadrent N, et al. Genome sequences of four *Ixodes* species expands understanding of tick evolution. *BMC Biol*. 2025;23:17.
69. Dillon KC, Frederick JC, Sprong H, Glenn TC, Ronai I. Tick Genome Assemblies: Overcoming biological limitations through advances in sequencing and assembly. *bioRxiv*. 2025:2025.2011.2011.687924.
70. Frederick JC, Thompson AT, Sharma P, Dharmarajan G, Ronai I, Pesapane R, et al. Phylogeography of the black-legged tick (*Ixodes scapularis*) throughout the USA identifies candidate loci for differences in vectorial capacity. *Mol Ecol*. 2023;32:3133–49.
71. Hillyard PD. Ticks of North-West Europe: keys and notes for identification of the species. Shrewsbury, UK: Field Studies Council; 1996.
72. Bakker JW, Esser HJ, Sprong H, Godeke G-J, Hoornweg TE, de Boer WF, et al. Differential susceptibility of geographically distinct *Ixodes ricinus* populations to tick-borne encephalitis virus and louping ill virus. *Emerg Microbes Infect*. 2024;13:2321992.
73. De Coster W, Rademakers R. Nanopack2: population-scale evaluation of long-read sequencing data. *Bioinformatics*. 2023;39:btad311.
74. Ewels P, Magnusson M, Lundin S, Käller M. Multiqc: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32:3047–8.
75. Marijon P, Chikhi R, Varré JS. Yacrd and fpa: upstream tools for long-read genome assembly. *Bioinformatics*. 2020;36:3894–6.
76. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 2019;37:540–6.
77. Hu J, Fan J, Sun Z, Liu S. Nextpolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*. 2020;36:2253–5.
78. Roach MJ, Schmidt SA, Borneman AR. Purge haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*. 2018;19:460.
79. Astashyn A, Tvedte ES, Sweeney D, Sapojnikov V, Bouk N, Joukov V, et al. Rapid and sensitive detection of genome contamination at scale with FCS-GX. *bioRxiv*. 2023. <https://doi.org/10.1101/2023.06.02.543519>.
80. Quinlan AR, Hall IM. BEDtools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–2.
81. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. Busco update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38:4647–54.
82. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*. 2008;24:637–44.

83. Shumate A, Salzberg SL. Liftoff: accurate mapping of gene annotations. *Bioinformatics*. 2020;37:1639–43.
84. Dainat J, Hereñú D, Davis E, Crouch K, Sol L, Agostinho N, pascal-git, & tayyrov. NBISweden/AGAT: AGAT-v1.0.0 (v1.0.0). Zenodo. 2022. <https://doi.org/10.5281/zenodo.7255559>.
85. Emms DM, Kelly S. Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20:238.
86. Chen H, Boutros PC. Venndiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*. 2011;12:35.
87. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9:5114.
88. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al. Repeatmodeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci USA*. 2020;117:9451–7.
89. Alonge M, Lebeigle L, Kirsche M, Aganezov S, Wang X, Lippman ZB, et al. Automated assembly scaffolding elevates a new tomato system for high-throughput genome editing. *bioRxiv*. 2021:2021.2018.469135.
90. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*. 2018;34:i142–50.
91. Huang N, Li H. Compleasm: a faster and more accurate reimplement of BUSCO. *Bioinformatics*. 2023;39:btad595.
92. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34:3094–100.
93. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19:1639–45.
94. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol*. 2016;17:132.
95. Dfam-consortium/RepeatAfterMe: RepeatAfterMe_V0.0.4 [<https://doi.org/10.5281/zenodo.7076442>]
96. Storer JM, Hubley R, Rosen J, Smit AFA. Curation guidelines for *de novo* generated transposable element families. *Curr Protoc*. 2021;1:e154.
97. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 2021;38:3022–7.
98. de Paula Baptista, R. *Ixodes ricinus* (castor bean tick). PRJNA816462. NCBI Sequence Read Archive. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA816462/> (2025).
99. de Paula Baptista, R. *Ixodes ricinus* isolate Maya1009, whole genome shotgun sequencing project. JBD-HMV000000000. GenBank. <https://www.ncbi.nlm.nih.gov/nuccore/JBDHMV000000000> (2025).
100. de Paula Baptista, R. *Ixodes ricinus* isolate Murphy0812, whole genome shotgun sequencing project. JBD-HMW000000000. GenBank. <https://www.ncbi.nlm.nih.gov/nuccore/JBDHMW000000000> (2025).
101. Ray, DA. *Ixodes* TE Library. Dfam. <https://dfam.org> (2025).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.