



Mitochondrial genomes from RNA-Seq reveal phylogeny and selection in *Mepraia* (Hemiptera: Reduviidae)

Matheus Cardoso de Siqueira e Silva¹ · José Paulo Leite Guadanucci² · Tiago Belintani²

Received: 7 December 2025 / Accepted: 13 April 2026
© The Author(s) 2026

Abstract

Despite the epidemiological importance of triatomine bugs as vectors of *Trypanosoma cruzi*, mitogenomic resources remain limited for many lineages, restricting evolutionary and comparative analyses. The endemic Chilean genus *Mepraia* is a notable example, with scarce mitochondrial genomic information despite its relevance to Chagas disease ecology and vector evolution. Here, we reconstructed partial mitochondrial genomes, including all 13 protein-coding genes, from RNA-Seq data for three *Mepraia* species. Phylogenetic analyses confirmed the monophyly of *Mepraia* and resolved interspecific relationships, whereas comparative analyses revealed heterogeneous patterns of intraspecific diversity, with elevated variation in *M. parapatrica*. Codon-based tests detected pervasive purifying selection together with localized signals of positive selection in oxidative phosphorylation genes, particularly NAD5, NAD6, NAD1, and COIII. These results show that transcriptomic data provide a reliable source for mitochondrial genome reconstruction in non-model insects and that adaptive signals can coexist with strong functional constraint in triatomine mitogenomes. Overall, this study expands mitogenomic resources for Triatominae, improves evolutionary inference in *Mepraia*, and provides a foundation for future integrative studies on diversification, mitochondrial evolution, and vector biology.

Keywords Mitochondrial genome · Phylogenomics · RNA-seq · Chagas disease vector · Evolutionary adaptation

Introduction

Blood-feeding triatomine bugs are of major biological and public health interest because they include vectors of *Trypanosoma cruzi* Chagas, 1909 (Kinetoplastida: Trypanosomatidae), the causative agent of Chagas disease. Understanding their evolutionary history is important not only for taxonomy and comparative biology, but also for interpreting patterns of diversification, geographic structure, and vector-associated traits. Among the molecular resources available for such questions, mitochondrial genomes are especially useful because they combine broad

phylogenetic utility with sensitivity to population-level and adaptive processes (Eisen and Fraser 2003; Duchêne et al. 2011). However, mitogenomic data remain scarce for several triatomine lineages, including the endemic Chilean genus *Mepraia*, limiting evolutionary inference in a medically relevant group.

High-throughput sequencing (HTS) technologies, such as whole-genome shotgun sequencing (WGS), anchored hybrid enrichment (AHE), transcriptome (RNA-seq), and ultraconserved elements (UCE) capture, have greatly expanded the methods available for recovering mitogenomes (Raposo do Amaral et al. 2015; Moreira et al. 2015; Baeza et al. 2025). RNA-Seq allows for efficient *in silico* assembly of mitochondrial genomes in non-model organisms, overcoming challenges related to poor DNA quality and complex nuclear genomes (Tian and Smith 2016; Forni et al. 2019; Yang et al. 2024). Despite these advances, most available mitogenomic data are limited to a small number of model or economically significant insect groups, leaving substantial gaps for many other taxa (Cameron 2014; Ramírez et al. 2026). Such taxonomic bias limits comparative and evolutionary analyses across insect lineages. The

Communicated by Willis X Li.

✉ Tiago Belintani
tiagobellintani@gmail.com

¹ Universidade de Araraquara (Uniar), Araraquara, SP, Brazil

² Rio Claro Arachnology Laboratory, Department of Biodiversity, Institute of Biosciences, São Paulo State University (Unesp), Rio Claro, Brazil

subfamily Triatominae Jeannel, 1919 consists of blood-feeding insects known for their role as vectors of trypanosomatid parasites, including *T. cruzi* and *Trypanosoma rangeli* Tejera, 1920 (Lent and Wygodzinsky 1979). These insects are predominantly found throughout the Americas, ranging from the southern United States to Central and South America, with some species also present in Asia and Oceania (Monteiro et al. 2018). Triatominae currently includes over 158 species classified into eight tribes and 19 genera (Paiva et al. 2022; Masonick et al. 2025; Oliveira-Correia et al. 2024). Of these, the tribe Triatomini Jeannel, 1909 is the most species-rich and includes several genera of epidemiological significance, such as *Triatoma* Laporte, 1832, *Panstrongylus* Berg, 1879, and *Rhodnius* Stål, 1859.

The genus *Mepraia* Mazza, Gajardo & Jörg, 1940 consists of three species endemics to Chile: *Mepraia spinolai* (Porter, 1933), *M. gajardoi* Frías, Henry & González, 1998, and *M. parapatrica* Frías-Lasserre 2010. These species exhibit distinct ecological traits and geographic distributions and have therefore become the focus of increasing research interest due to their evolutionary and epidemiological relevance (Frías-Lasserre et al. 2018; 2019). *Mepraia* has been investigated using morphological, morphometric, cytogenetic, and molecular approaches, providing strong evidence for the distinctiveness of the three recognized species (Campos-Soto et al. 2011, 2013; Frías-Lasserre 2010; Frías-Lasserre et al. 2019). Although the monophyly of *Mepraia* is well supported (Kieran et al., 2021; Campos-Soto et al., 2022), recent phylogenomic evidence points to rapid diversification and historical introgression among lineages, suggesting that ancient hybridization played a role in shaping present-day genetic structure (Belintani et al. 2023).

Beyond their evolutionary relevance, *Mepraia* species are epidemiologically important in Chile as components of the sylvatic transmission cycle of Chagas disease and represent the only endemic triatomine genus involved in *T. cruzi* circulation in wild habitats (Tapia-Garay et al. 2018; Lambarri et al. 2018). *Mepraia spinolai* is the most widespread and frequently infected species, with high natural infection rates and documented feeding on mammals and humans, highlighting its major epidemiological role (Coronado et al. 2009; Espinoza et al. 2013). *Mepraia gajardoi* has also been reported as naturally infected in northern coastal and peridomestic environments, indicating potential involvement in local transmission dynamics (González et al. 2015; Tapia-Garay et al. 2018). In contrast, *M. parapatrica* occupies more restricted coastal and insular habitats; although less frequently documented, infected individuals recorded on offshore islands confirm its participation in the sylvatic cycle of *T. cruzi* (Rives-Blanchard et al. 2017).

Despite the medical importance of Triatominae and the increasing availability of genomic data for related

taxa, mitochondrial genomic information remains limited (Ramírez et al. 2026). At the same time, large amounts of publicly available transcriptomic data offer an underexplored opportunity to recover mitochondrial sequences without the need for additional sequencing efforts. In this study, we reconstruct partial mitochondrial genomes of *M. gajardoi*, *M. parapatrica*, and *M. spinolai* using RNA-Seq datasets retrieved from public repositories. By integrating these newly recovered mitochondrial sequences into comparative phylogenetic and evolutionary analyses, we aim to clarify interspecific relationships and help fill important gaps in the evolutionary understanding of *Mepraia*. The genomic resources generated here demonstrate the potential of RNA-Seq data as a valuable source for mitogenome reconstruction and provide a foundation for future studies of mitochondrial diversity, lineage diversification, and evolutionary dynamics in this medically relevant genus.

Materials and methods

Taxon sampling and mitogenome assembly from RNA-Seq data

RNA-Seq data used in this study were previously generated and published by Belintani et al. (2023) and are publicly available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject accession PRJNA916468. Briefly, head and salivary gland tissues were carefully dissected from fifth-instar male nymphs of *Mepraia gajardoi* (n=6), *M. spinolai* (n=6), and *M. parapatrica* (n=6), ensuring preservation of salivary gland integrity (Table S1). Specimens were collected from natural populations within the known geographic range of each species, representing their typical ecological contexts, as detailed in Belintani et al. (2023).

In that study, total RNA was extracted individually from each specimen using a Trizol/chloroform protocol (Chomczynski and Mackey 1995). RNA quality was initially assessed by agarose gel electrophoresis and quantified using a Qubit fluorometer, while NanoDrop spectrophotometry was used to select samples with 260/230 and 260/280 absorbance ratios close to 2.0. RNA integrity was further evaluated by capillary electrophoresis using an Agilent 2100 Bioanalyzer, yielding RNA integrity numbers (RIN) of 7.1 ± 0.3 across the 18 samples. Libraries were then prepared using the TruSeq® RNA Sample Prep Kit v2 (Illumina) and sequenced on an Illumina HiSeq 2500 platform with 2×100 bp paired-end reads at the Functional Genomics Laboratory of ESALQ-USP (Brazil). No additional RNA extraction or sequencing was performed for the present study.

In the present study, raw sequencing reads were quality-assessed using FastQC v0.11.9 (Andrews 2010) and trimmed with Trimmomatic v0.39 (Bolger et al., 2014) using the following parameters: LEADING:5, TRAILING:5, SLIDING-WINDOW:5:20, and MINLEN:50.

Partial mitochondrial genomes were reconstructed using an iterative baiting and mapping strategy implemented in MITObim v1.9.1 (Hahn et al. 2013). The mitochondrial genome of *Triatoma infestans* (NC_035547) was used as the initial reference due to its close phylogenetic relationship to *Mepraia*, as well as its completeness and high-quality annotation. During the iterative assembly process, the reference sequence is progressively updated, reducing potential bias from the initial reference.

Paired-end reads were merged with FLASH (Magoč and Salzberg 2011) to recover extended fragments when possible. Assemblies were generated using the integrated MIRA assembler v3.4.1.1, with a maximum of 30 iterative cycles; in most cases, convergence was achieved within five iterations. For protein-coding genes that were fragmented or poorly recovered, MitoGeneExtractor v1.9.5 (Brasseur et al. 2023) was additionally employed to improve the recovery of mitochondrial protein-coding genes (PCGs). Final contigs were annotated using MITOS (Bernt et al., 2013) under the invertebrate mitochondrial genetic code. The resulting mitochondrial maps represent partial reconstructions derived solely from expressed PCGs. As a result, mitochondrial genome structure and gene order were not inferred, and structural validation by PCR was not conducted.

All computational analyses were performed using high-performance computing resources provided by the Center for Scientific Computing (NCC/GridUNESP) at São Paulo State University.

Phylogenetic analyses

Phylogenetic analyses were conducted to investigate evolutionary relationships within the genus *Mepraia* and to determine its placement within the subfamily Triatominae using 13 mitochondrial protein-coding genes (PCGs). Two datasets were assembled: (i) a focal dataset including PCGs recovered in this study from all individuals of *M. gajardoi*, *M. parapatrica*, and *M. spinolai*; and (ii) an expanded dataset that combined the focal dataset with 51 additional Triatominae mitogenomes retrieved from GenBank (Table S2).

Two analytical strategies were applied. First, a concatenated analysis of the 13 PCGs from the focal dataset was conducted to resolve intrageneric relationships within *Mepraia*. Second, the expanded dataset was used to infer the phylogenetic placement of *Mepraia* within a broader Triatominae context.

Multiple sequence alignments for each PCG were generated using MAFFT v7.475 (Katoh and Standley 2013) under the E-INS-i strategy. This approach applies global pairwise alignment with iterative refinement and is suitable for datasets containing conserved and variable regions with indels. Alignments were manually inspected in MEGA11 (Tamura et al. 2021) to confirm reading-frame integrity, detect internal stop codons, and verify codon consistency. Ambiguously aligned positions were filtered using TrimAl v1.2 (Capella-Gutiérrez et al. 2009) with a gap threshold of 0.8 (-gt 0.8), removing alignment columns containing more than 80% gaps. The mitochondrial PCG alignments were concatenated into a partitioned supermatrix using AMAS v1.0 (Borowiec 2016). Gene boundaries were retained during concatenation, and a partition file specifying the exact nucleotide coordinates of each locus was generated simultaneously.

Phylogenetic inference under the maximum likelihood (ML) criterion was conducted using IQ-TREE 3 (Wong et al. 2025) on the concatenated mitochondrial dataset. The best-fitting substitution model and partitioning scheme were jointly estimated using ModelFinder Plus (-m MFP+MERGE; Kalyaanamoorthy et al. 2017), allowing automatic merging of partitions with similar evolutionary patterns under a relaxed clustering algorithm (-recluster 10). Branch support was evaluated using 1000 replicates of the SH-like approximate likelihood ratio test (SH-aLRT; Guindon et al. 2010) and 1000 ultrafast bootstrap replicates (UFBoot; Hoang et al. 2018). Numerical stability was ensured using the safe likelihood optimization mode (-safe), and ultrafast bootstrap branch length optimization was refined using the BNNI correction (-bnni) to reduce potential support overestimation.

Bayesian inference (BI) was conducted in MrBayes v3.2.7 (Ronquist et al. 2012) on the same concatenated dataset. The ML partitioning scheme was retained for Bayesian inference, with all partitions analyzed under GTR+I+ Γ (nst=6; rates=invgamma). Model parameters were unlinked across partitions, and rate heterogeneity among partitions was accommodated using ratepr=variable. Two independent runs were performed, each with four Markov chains, for 20 million generations, sampling every 1,000 generations. Convergence was assessed by monitoring the average standard deviation of split frequencies (ASDSF<0.01) and by examining effective sample sizes (ESS>200) and parameter trace plots in Tracer v1.7.2 (Rambaut et al. 2018). Log-likelihood traces indicated that stationarity was reached after approximately 5 million generations. A conservative burn-in of 25% was applied to exclude pre-convergence samples. Posterior probabilities (PP) were calculated from the remaining trees to estimate nodal support.

Topological congruence between maximum likelihood (ML) and Bayesian inference (BI) trees was assessed with particular emphasis on relationships within *Mepraia* and among major Triatominae lineages. Nodes were considered strongly supported when they received both ultrafast bootstrap values (UFBoot) $\geq$ 95% and posterior probabilities (PP) $\geq$ 0.95. Final trees were visualized and annotated in FigTree v1.4.4 (Rambaut 2012).

Genetic divergence and selection analyses of mitochondrial protein-coding genes

Pairwise genetic distances were estimated for each of the 13 mtDNA PCGs of *Mepraia* using the Kimura 2-Parameter (K2P) model (Kimura 1980) with a gamma distribution (shape parameter=1) to account for rate heterogeneity among sites. The K2P model was selected because it is widely adopted in Triatominae studies and provides an appropriate balance between simplicity and accuracy for mitochondrial (Monteiro et al. 2013; Belintani et al. 2020; Silva et al. 2023). Nucleotide sequences (all codon positions) were aligned prior to distance estimation, and standard errors were calculated via 1000 bootstrap replicates. Analyses were performed in MEGA11 (Tamura et al. 2021).

Genetic structure was explored through a principal component analysis (PCA) based on the K2P distance matrix. PCA was conducted in PAST v4.03 (Hammer et al. 2001) to visualize patterns of variation among individuals in multivariate space and to assist in assessing intra- and interspecific relationships.

Selection patterns and variation in evolutionary rates across the 13 mtDNA PCGs were assessed using codon-based models implemented in HyPhy v2.5 (Kosakovsky Pond et al. 2005, 2020). Three complementary approaches were applied to detect both site-specific and lineage-specific selection.

The Mixed Effects Model of Evolution (MEME) was used to identify sites subject to episodic diversifying selection across the phylogeny (Murrell et al. 2012), employing four ω rate classes and a significance threshold of $p < 0.05$. The Fast Unconstrained Bayesian Approximation (FUBAR) method was then used to detect sites under pervasive positive or purifying selection (Murrell et al. 2013), using 40 grids points and 2 million MCMC iterations, with the first 20% discarded as burn-in; two independent runs were conducted to verify convergence. Finally, the Branch-site Unrestricted Statistical Test for Episodic Diversification (BUSTED) was applied to detect branches experiencing episodic positive selection (Murrell et al. 2015). The model incorporated synonymous rate variation (SRV) with four classes plus an additional class to account for alignment noise, and significance was assessed using a likelihood ratio test ($p < 0.05$).

Data availability

The mitochondrial sequences generated in this study are publicly available in GenBank (NCBI) under accession numbers PV521367–PV521600 (see Table S4). Additional data and materials related to this study are available from the authors upon reasonable request.

Results

The mitochondrial genomes of *M. gajardoi*, *M. parapatrica*, and *M. spinolai* were assembled from previously generated RNA-Seq datasets using an iterative baiting and mapping strategy, which reduces potential bias associated with the initial reference sequence. A total of 18 partial mitochondrial assemblies were recovered (six per species) using a reference-guided strategy. Although the assemblies are partial due to gaps in transcript coverage, all 13 mitochondrial protein-coding genes (PCGs) were successfully reconstructed in every sample. The three highest-quality assemblies are shown in Fig. 1A–C.

The assembled mitochondrial sequences ranged in length from approximately 15,173 bp to 17,762 bp (Table 1), values consistent with mitochondrial genome sizes reported for Triatominae (Pita et al. 2017; Cigarroa-Toledo et al. 2024; Almeida et al. 2025). As expected from transcriptome-based assemblies, recovery of non-protein-coding regions was variable. While all 13 PCGs were consistently recovered, the detection of tRNA and rRNA genes varied among assemblies and is most likely attributable to uneven transcript representation rather than true biological differences.

In *M. gajardoi* (17,138–17,762 bp), five tRNA loci were absent or incomplete in the best-quality assembly (trnE, trnH, trnI, trnL2, and trnV). In *M. parapatrica* (15,813–17,492 bp), eight tRNA loci were missing or incomplete (trnD, trnF, trnH, trnI, trnL1, trnS2, trnT, and trnV). Assemblies of *M. spinolai* (15,975–17,529 bp) showed comparatively higher tRNA recovery, with some samples presenting all 22 tRNAs; however, several loci (*trnF*, *trnP*, *trnQ*, *trnT*, and *trnV*) were absent or incomplete in other assemblies. The order and strand orientation of the 13 PCGs were conserved across all samples and were identical to the reference genome, indicating consistent recovery of these genes.

Mitochondrial AT content ranged from 57.41% to 66.76% (Table 1), consistent with the AT-rich composition typical of hemipteran mitogenomes (Wang et al. 2015). Sequencing depth across recovered mitochondrial regions was high, with average coverage values ranging from 5,470 $\times$ to over 13,800 $\times$. This high coverage supports the robustness of the reconstructed PCGs for downstream phylogenetic and molecular evolutionary analyses.

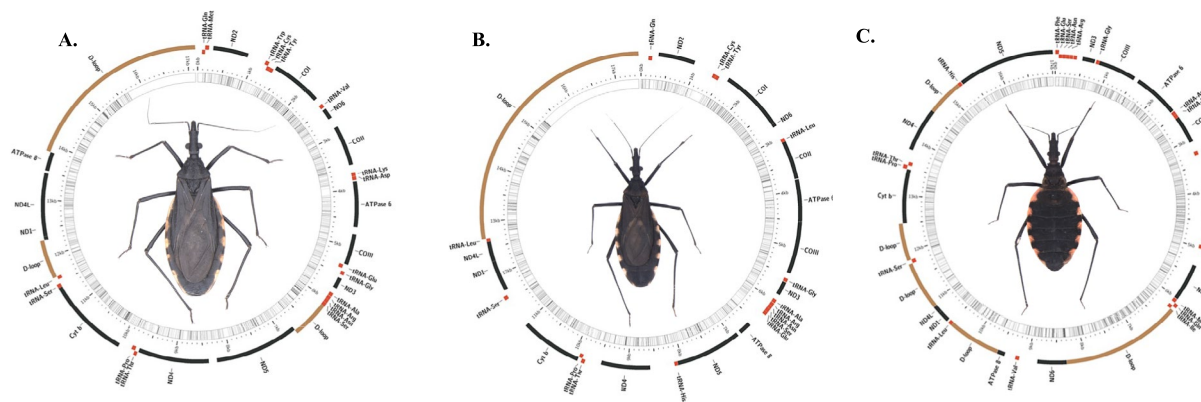


Fig. 1 Partial mitochondrial genome maps reconstructed from RNA-Seq data for the three species of *Mepraia*. **AM.** *gajardoi*, **BM.** *parapatrica*, and **CM.** *spinolai*. Circular diagrams show the arrangement and orientation of mitochondrial genes recovered in the assemblies, including the 13 protein-coding genes (PCGs), transfer RNAs (tRNAs), and ribosomal RNAs (rRNAs) when present. Genes on the majority and

minority strands are shown on the outer and inner tracks, respectively. Black segments indicate protein-coding genes, red segments indicate tRNAs, and brown segments indicate rRNAs. These maps represent partial mitochondrial assemblies; therefore, tRNA and rRNA genes are not fully recovered, and gene order does not necessarily reflect the complete mitochondrial genome

Table 1 Descriptive statistics of mitochondrial assemblies for *Mepraia gajardoi*, *M. parapatrica*, and *M. spinolai*

Species	Samples	Length (bp)	Reads (n)	Max coverage (×)	Mean coverage (×)	GC content (%)	T content (%)
<i>M. spinolai</i>	<i>M. spinolai1</i>	17,529	1,564,229	298,069	11,000	29.6	66.2
<i>M. spinolai</i>	<i>M. spinolai2</i>	15,975	1,096,992	452,552	9,157	29.9	64.1
<i>M. spinolai</i>	<i>M. spinolai3</i>	16,941	1,716,780	577,372	13,562	30.3	62.5
<i>M. spinolai</i>	<i>M. spinolai4</i>	17,195	1,721,226	450,774	13,887	29.9	62.1
<i>M. spinolai</i>	<i>M. spinolai5</i>	17,198	1,628,346	340,765	11,461	29.0	62.1
<i>M. spinolai</i>	<i>M. spinolai6</i>	17,165	1,621,226	350,774	12,886	29.7	62.0
<i>M. gajardoi</i>	<i>M. gajardoi7</i>	17,762	858,105	242,684	5,925	29.3	62.1
<i>M. gajardoi</i>	<i>M. gajardoi8</i>	15,898	940,275	301,135	7,237	29.3	66.8
<i>M. gajardoi</i>	<i>M. gajardoi9</i>	16,398	1,146,516	264,059	9,346	29.2	62.5
<i>M. gajardoi</i>	<i>M. gajardoi10</i>	17,173	1,099,046	270,813	7,744	29.3	63.0
<i>M. gajardoi</i>	<i>M. gajardoi11</i>	17,398	1,244,312	264,054	8,246	29.2	62.5
<i>M. parapatrica</i>	<i>M. gajardoi12</i>	15,173	1,198,077	260,813	7,744	29.3	63.0
<i>M. parapatrica</i>	<i>M. parapatrica13</i>	16,492	1,201,222	312,731	7,500	29.3	59.9
<i>M. parapatrica</i>	<i>M. parapatrica14</i>	16,813	811,235	252,299	6,224	29.1	60.7
<i>M. parapatrica</i>	<i>M. parapatrica15</i>	17,492	901,222	328,73	6,500	29.3	58.9
<i>M. parapatrica</i>	<i>M. parapatrica16</i>	15,813	711,225	202,299	5,724	29.7	61.9
<i>M. parapatrica</i>	<i>M. parapatrica17</i>	16,27	1,240,510	378,017	10,097	29.2	57.4
<i>M. parapatrica</i>	<i>M. parapatrica18</i>	17,094	742,086	257,044	5,470	29.0	57.4

Assembly length corresponds to the total length of recovered mitochondrial contigs (bp). Coverage values (×) represent sequencing depth per nucleotide position, estimated from mapped RNA-Seq reads. High coverage values reflect the elevated expression of mitochondrial transcripts in RNA-Seq datasets. GC and AT contents were calculated from consensus sequences; due to the presence of undetermined bases (Ns), their sum may not equal 100%

Phylogenetic analyses

Phylogenetic analyses based on the concatenated dataset of 13 mitochondrial PCGs yielded well-resolved topologies with high statistical support for most internal nodes (see Supplementary Fig_S1_Phylogenies_ML_BI_Mepraia.pdf). Both focal (Fig. 2) and comprehensive datasets (Fig. 3) showed strong topological congruence between ML and BI.

In the focal dataset, *Mepraia* exhibited clear and well-supported species-level differentiation (SH-aLRT=100%, UFBot=100%, PP=1.0, Fig. 2). *Mepraia spinolai*, *M. parapatric*, and *M. gajardo* were consistently recovered as distinct genetic clusters, each supported by maximal node support values, indicating strong resolution of species boundaries.

Mepraia spinolai exhibited uniformly high internal support, with most nodes receiving maximal values (SH-aLRT=100%, UFBot=100%, PP=1.0). A single

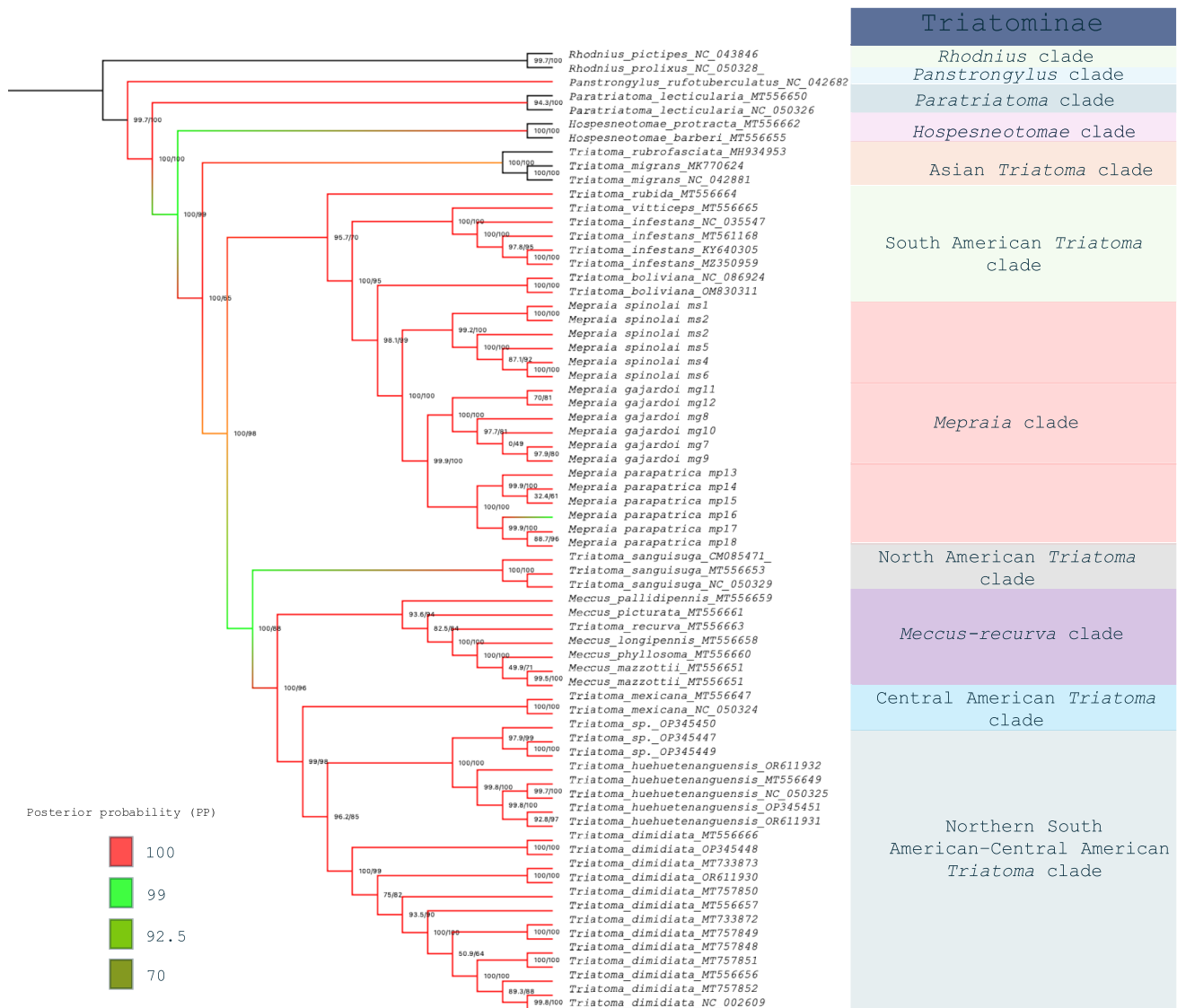


Fig. 3 Maximum likelihood phylogeny of Triatominae inferred from a concatenated alignment of the 13 mitochondrial protein-coding genes, including the *Mepraia* sequences generated in this study and additional mitogenomes retrieved from GenBank. Branch support values

correspond to SH-aLRT and ultrafast bootstrap replicates, while posterior probabilities from Bayesian inference are represented by branch coloration

multivariate structure was highly concordant across genes and consistently reflected species-level differentiation.

Across all loci, samples were partitioned into three well-defined groups corresponding to *M. gajardoi*, *M. parapatrica*, and *M. spinolai* (Supplementary_PCA.pdf, Figures S1–S13). For most genes (COI, COII, COIII, CytB, NAD1, NAD2, NAD4, NAD4L, NAD5, and NAD6), species separation occurred primarily along the first principal component (PC1), with minimal overlap among clusters. These loci exhibited compact intraspecific groupings and clear interspecific differentiation (Table 2).

Mepraia spinolai consistently formed the most compact cluster across loci, reflecting low intraspecific genetic

distances (generally 0.000–0.010 substitutions per site; Supplementary_Distance_genetics.xlsx, Table S1). Slightly higher values were observed in faster-evolving loci such as ATP8, where intraspecific distances ranged from 0.000 to 0.010.

Mepraia gajardoi showed moderately broader dispersion in PCA space for some loci, with intraspecific distances reaching 0.000–0.040 in ATP8 and 0.000–0.020 in NAD3, but remaining 0.000–0.010 for most genes.

Mepraia parapatrica exhibited comparatively greater intraspecific dispersion for specific loci, particularly ATP8 and NAD3. In ATP8, intraspecific distances ranged from 0.000 to 0.110, whereas in NAD3 values ranged from 0.000

Table 2 Summary of pairwise Kimura 2-parameter (K2P) genetic distances (substitutions per site) among mitochondrial protein-coding genes in *Mepraia*

Gene	Length (bp)	<i>M. spinolai</i>	<i>M. gajardoi</i>	<i>M. parapatraca</i>	<i>spinolai-gajardoi</i>	<i>spinolai-parapatraca</i>	<i>gajardoi-parapatraca</i>
COI	1,542	0.000–0.000	0.000–0.010	0.000–0.080	0.050–0.060	0.060–0.080	0.060–0.080
COII	687	0.000–0.000	0.000–0.000	0.000–0.090	0.050–0.060	0.080–0.120	0.080–0.100
COIII	784	0.000–0.010	0.000–0.010	0.000–0.090	0.040–0.050	0.090–0.090	0.090–0.090
CytB	1,14	0.000–0.000	0.000–0.020	0.000–0.080	0.050–0.060	0.080–0.090	0.080–0.090
ATP6	684	0.000–0.010	0.000–0.010	0.000–0.080	0.060–0.060	0.080–0.110	0.080–0.100
ATP8	165	0.000–0.010	0.000–0.040	0.000–0.110	0.000–0.160	0.060–0.160	0.040–0.160
NAD1	918	0.000–0.000	0.000–0.010	0.000–0.060	0.050–0.080	0.060–0.080	0.070–0.080
NAD2	1,029	0.000–0.000	0.000–0.010	0.000–0.060	0.050–0.080	0.060–0.080	0.070–0.080
NAD3	357	0.000–0.010	0.000–0.020	0.000–0.130	0.060–0.070	0.070–0.130	0.070–0.100
NAD4	1,379	0.000–0.010	0.000–0.010	0.000–0.080	0.040–0.050	0.080–0.090	0.080–0.090
NAD4L	297	0.000–0.010	0.000–0.010	0.000–0.080	0.040–0.050	0.080–0.090	0.080–0.090
NAD5	1,716	0.000–0.010	0.000–0.010	0.000–0.080	0.050–0.060	0.070–0.080	0.070–0.080
NAD6	531	0.000–0.007	0.000–0.011	0.000–0.045	0.032–0.037	0.039–0.053	0.030–0.045

Intraspecific distances are reported separately for each species, and interspecific distances are shown for all species pairs. Values represent minimum–maximum ranges across individuals

Table 3 Summary of positive selection analysis on *Mepraia* mitochondrial protein-coding genes

Gene	MEME (sites)	FUBAR (sites)	BUSTED (p-value)
NAD5	17	37	0.5000
NAD6	5	10	0.5000
NAD1	5	10	0.5000
COIII	2	9	0.5000
COII	0	1	0.5000
NAD4L	1	0	0.5000
ATP6	1	0	0.5000
NAD4	1	0	0.5000
CytB	0	0	0.5000
ATP8	1	0	0.5000
NAD2	1	0	0.5000
NAD3	1	0	0.5000
COI	1	0	0.5000

to 0.130, indicating higher variability relative to more conserved loci such as COI, NAD1, and NAD2. Nevertheless, intraspecific divergence remained within the range typically observed for mitochondrial protein-coding genes in closely related taxa.

Patterns of interspecific genetic distances were concordant with PCA structure. Pairwise distances between *M. spinolai* and the other species ranged from 0.032 to 0.037 (NAD6) to 0.060–0.160 (ATP8) depending on the gene, with most loci falling between 0.050 and 0.090. Distances between *M. gajardoi* and *M. parapatraca* ranged from 0.030 to 0.045 (NAD6) to 0.040–0.160 (ATP8) across genes, with the majority of loci clustering between 0.070 and 0.100. The highest values (up to 0.160) were observed in more variable genes, particularly ATP8.

Overall, both PCA and K2P distance analyses consistently support clear species-level differentiation among the three *Mepraia* taxa, while also revealing heterogeneous

patterns of intraspecific variability across loci, particularly in shorter and more rapidly evolving mitochondrial genes.

Analysis of selective pressure on mitochondrial protein-coding genes revealed heterogeneous selection profiles across loci, indicating that selective signals are unevenly distributed among genes (Table 3). The strongest evidence of positive selection was detected in NAD5, NAD6, NAD1, and COIII, which together concentrated the majority of candidate sites identified across all analyses.

NAD5 exhibited the strongest signal, with 17 sites detected by MEME and 37 by FUBAR. NAD5 showed the highest number of candidate sites, with 17 detected by MEME and 37 by FUBAR. Similarly, NAD6 and NAD1 each showed five sites under episodic selection and ten under pervasive selection. The COIII gene displayed fewer candidate sites overall, with two sites identified by MEME and nine by FUBAR.

Despite the presence of multiple site-level signals detected by MEME and FUBAR, BUSTED analyses did not provide support for gene-wide positive selection in any of these loci ($p=0.5000$), indicating that positive selection is restricted to specific codons rather than acting across entire genes.

The remaining mitochondrial genes (COII, NAD4L, ATP6, NAD4, ATP8, NAD2, NAD3, and COI) showed limited evidence of positive selection. For each of these genes, MEME identified only a single candidate site under episodic selection, while FUBAR detected no sites under pervasive selection, with the exception of COII, which exhibited one site exceeding the posterior probability threshold. Consistently, BUSTED analyses for all these loci were non-significant ($p=0.5000$), supporting the absence of gene-wide selection signals. No evidence of positive selection was detected for CytB by any method (MEME, FUBAR, or

BUSTED), indicating strong evolutionary constraint on this gene in the analyzed dataset.

Discussion

Phylogenetic and evolutionary analyses based on transcriptome-derived mitochondrial genomes provide a robust framework for interpreting genetic diversity and lineage relationships in this endemic Chilean lineage of Triatominae. As key vectors of *T. cruzi*, the etiological agent of Chagas disease, these insects remain comparatively understudied despite their epidemiological importance (Ramírez et al. 2026). The genus *Mepraia* comprises three species with distinct ecological and geographic distributions, and reconstructing their evolutionary history is essential for refining taxonomy and informing vector biology (Chacón et al. 2016; Frías-Lasserre et al. 2018; Garrido et al. 2019).

Partial mitochondrial genomes of *M. gajardoi*, *M. parapatrica*, and *M. spinolai* enabled the recovery of a complete and colinear set of the 13 mtDNA PCGs. The conserved gene order, consistent with that of *T. infestans* (Pita et al. 2017), supports the structural stability of triatomine mitogenomes. Variation in tRNA recovery is most parsimoniously explained by differences in transcript representation rather than genomic rearrangement. Additionally, AT content values were consistent with previously reported patterns in Triatominae and other insects (Cigarroa-Toledo et al. 2024; Almeida et al. 2025; Behura 2015; Desalle 2017), supporting the reliability of the reconstructed sequences for downstream evolutionary analyses.

The mtDNA PCGs are particularly relevant due to their extensive use in phylogenetic inference (Justi et al. 2014; Belintani et al. 2020; Sun et al. 2022; Silva et al. 2023; Groppo et al. 2025) and their widespread application in molecular identification of insect vectors through DNA barcoding approaches (Hebert et al. 2003; Cywinska et al. 2006; Monteiro et al. 2004). The strong phylogenetic resolution obtained in this study empirically supports the robustness of transcriptome-derived mitochondrial assemblies for evolutionary analyses in Triatominae. Although transcriptome-based approaches may underrepresent non-coding regions, coding sequences typically show high concordance with DNA-derived mitogenomes (Nabholz et al. 2010; Forni et al. 2019).

Moreover, the topological congruence observed between our results and previous mitogenomic studies (Dotson and Beard 2001; Pita et al. 2017), as well as with whole-genome data from *Triatoma mexicana* (Herrich-Schaeffer, 1848) (Aguilera-Urbe et al. 2020) and *Panstrongylus geniculatus* (Latreille 1811) (Canizales-Silva et al. 2025), further supports the biological accuracy of the reconstructed sequences

and supports the reliability of transcriptome-derived mitochondrial data for recovering evolutionary patterns in Triatominae. Within this framework, the sampling design is consistent with the objectives of reconstructing mitochondrial genes from RNA-Seq data. It is also appropriate for evaluating species-level evolutionary relationships (Belintani et al. 2023). Because mitochondrial genomes are haploid and maternally inherited, a moderate number of individuals is sufficient to recover species-level lineages and conserved gene organization patterns, as reflected in the well-resolved phylogenetic structure obtained (Ballard and Whitlock 2004). Although this design was not intended to exhaustively characterize intraspecific diversity, the dataset provides a robust structure for phylogenetic inference and comparative evolutionary analyses.

Transcriptome-derived mitochondrial datasets provide a reliable and efficient resource for evolutionary analyses in non-model insects. Because mitochondrial transcripts are often abundant and expressed as polycistronic units, RNA-Seq data can yield high coverage across coding regions, enabling robust recovery of mitochondrial protein-coding genes without additional targeted sequencing (Smith 2013; Tian and Smith 2016; Hahn et al. 2013). In our study, all 13 mitochondrial protein-coding genes were consistently recovered across samples, and the resulting phylogenetic patterns were highly congruent with previous mitochondrial and nuclear-based hypotheses for Triatominae. Together, these results support the reliability of transcriptome-derived mitochondrial coding sequences for phylogenetic inference in *Mepraia*, consistent with previous studies demonstrating strong concordance between RNA-derived and DNA-derived mitogenomes (Nabholz et al. 2010; Forni et al. 2019).

Nevertheless, transcriptome-based assemblies have important limitations. Lowly expressed regions, particularly some tRNAs and the control region, may be incompletely recovered, which restricts the reconstruction of complete mitochondrial genomes (Tian and Smith 2016; Forni et al. 2019). For this reason, the variation observed here in tRNA recovery is more plausibly explained by differences in transcript representation rather than by true structural differences among mitogenomes. In addition, transcriptomic data alone do not allow definitive inference of complete genome architecture without complementary DNA-based validation (Hahn et al. 2013). Accordingly, the present dataset should be interpreted primarily as a robust mitochondrial coding resource for phylogenetic and molecular evolutionary analyses rather than as a complete structural characterization of the mitogenome.

Phylogenetic analyses based on the concatenation of the 13 mtDNA PCGs recovered the three *Mepraia* species as reciprocally monophyletic and strongly supported clades.

M. parapatrica and *M. gajardoi* were consistently inferred as sister taxa, whereas *M. spinolai* was positioned as the earliest diverging lineage within the genus. This topology corroborates previous hypotheses based on both mitochondrial and nuclear markers (Campos-Soto et al. 2015, 2020) and agrees with recent transcriptome-based inferences (Belintani et al. 2023), reinforcing the evolutionary distinctiveness of *M. spinolai*. The consistency of these relationships across inference methods, together with their agreement with previous multilocus studies (Justi et al. 2014; Kieran et al., 2021; Masonick et al. 2025) and the conserved gene order and composition observed here, supports the robustness of the mitochondrial coding dataset and the reliability of transcriptome-derived assemblies for resolving recent divergences within Triatominae.

Thus, the pattern observed in the comprehensive phylogeny with 69 taxa of Triatominae is consistent and demonstrates that mitogenomes provide sufficient resolution for recovering shallow and moderately deep divergences within Triatominae. The resolution obtained here contrasts with studies based on a limited number of isolated markers, which often produce reduced node support or incongruent topologies in Triatominae (Justi et al. 2014) or other groups (Kwon et al. 2012; Zarrei et al. 2015). Complete mitogenomes and extended multilocus data generally provide stronger phylogenetic signal and more reliable inference of species relationships than single gene fragments (Fraumene et al. 2006; Finnegan et al. 2025).

Furthermore, the topology recovered for *Mepraia* in the comprehensive analysis is consistent with phylogenomic inferences based on ultraconserved elements (UCEs) (Kieran et al. 2021) and with multilocus nuclear datasets (Belintani et al. 2023). Although UCE-based studies emphasize broader sampling across Triatominae and include lower intrageneric representation within *Mepraia*, the relative position of the Chilean lineages remains congruent with the general structure observed in nuclear analyses. This concordance among mitochondrial, transcriptomic, and UCE-derived datasets indicates that the phylogenetic signal in *Mepraia* is not solely attributable to mitochondrial inheritance.

At a broader scale within Triatominae, the topology inferred from the concatenated 13 PCGs shows strong agreement with multilocus and phylogenomic hypotheses (Justi 2014; Justi et al. 2016; Kieran et al. 2021; Masonick et al. 2025). In the comprehensive analysis, *Mepraia* was recovered within the South American *Triatoma* radiation, forming a sister relationship with the clade comprising Central and North American species. This placement mirrors the geographic structuring recovered using thousands of UCE loci (Kieran et al. 2021), in which major regional clades within Triatomini are consistently resolved. Similarly, the

recovery of monophyletic clades corresponding to *Rhodnius*, *Panstrongylus*, *Paratriatoma*, and *Hospesneotomae* is consistent with previous phylogenetic studies based on mitochondrial and nuclear markers (Justi et al. 2014; Kieran et al. 2021; Paiva et al. 2025), reinforcing the stability of these generic lineages.

Some incongruences were restricted to deeper nodes, particularly in the relationships among major geographic lineages. These patterns are consistent with known limitations of mitochondrial datasets, in which substitutional saturation and compositional heterogeneity may compromise phylogenetic resolution at older evolutionary scales (Cameron 2014; Justi et al. 2014). In contrast, multilocus nuclear datasets, especially those comprising hundreds to thousands of loci, have demonstrated greater stability at basal nodes in Triatominae (Kieran et al. 2021; Filée et al. 2022) and across diverse animal groups (Lemmon et al. 2012; Cannon and Kocot 2016; Foley et al. 2019).

Despite these limitations, the strong topological congruence between the topology recovered here and previously published nuclear phylogenomic frameworks supports the reliability of complete mitochondrial coding datasets as an efficient tool for phylogenetic reconstruction in Triatominae, particularly at species-level and intrageneric scales.

Based on the well-established phylogenetic structure recovered here, genetic distance and PCA analyses further reinforce the distinction among the three *Mepraia* species, while revealing contrasting patterns of intraspecific variability. Previous phylogeographic studies demonstrated strong population structuring within the genus, identifying three lineages congruent with the currently recognized species (Campos-Soto et al. 2013). More recent analyses using single nucleotide polymorphisms detected subtle structuring within *M. spinolai*, including differences in genetic diversity between central and peripheral populations (San Juan et al. 2021).

Our estimates indicate that *M. spinolai* exhibits consistently low intraspecific genetic distances and forms a compact cluster in PCA space. This pattern aligns with earlier population genetic studies suggesting relatively high connectivity and limited genetic subdivision across its distribution range. The low mitochondrial divergence observed in *M. spinolai* contrasts with the structuring detected using nuclear SNPs (San Juan et al. 2021), indicating potential discordance between mitochondrial and nuclear histories. Such discordance may reflect demographic processes, such as recent population expansions, that homogenized the maternal lineage (Avice 2000).

In contrast, *M. parapatrica* and, to a lesser extent, *M. gajardoi* exhibit higher intraspecific variability, particularly in rapidly evolving mitochondrial genes such as ATP8. *M. gajardoi* shows moderate and locus-specific variability, with

most genes displaying low divergence and higher values restricted to faster-evolving loci such as ATP8 and NAD3. *M. parapatrica* exhibits comparatively greater intraspecific dispersion for certain loci, particularly ATP8 and NAD3, while most other genes remain comparatively conserved. Their broader dispersion in PCA space and the gene-specific distance patterns detected in ATP8, NAD3, and ATP6 are consistent with patterns of strong population subdivision previously reported for *Mepraia* populations (Campos-Soto et al. 2015, 2020, 2022). This pattern is also consistent with the K2P distance ranges, which show elevated maximum values in these loci compared to more conserved genes. These studies identified pronounced geographic structuring across the distribution range of the genus, suggesting that long-term demographic subdivision may contribute to the accumulation of mitochondrial variation within species. Phylogenomic studies based on transcriptomes have further shown that *Mepraia* underwent rapid diversification and ancient introgression events, resulting in reticulated phylogenetic patterns and overall low genetic diversity (Belintani et al. 2023). Within this framework, the elevated variability observed in *M. parapatrica* likely reflects a more structured demographic history and the accumulation of divergence in fast-evolving loci, consistent with the retention of locus-specific evolutionary signals potentially associated with introgression.

The markedly elevated intraspecific divergence observed for ATP8 in *M. parapatrica* merits specific consideration. Technical artefacts are unlikely, as sequencing coverage was high, open reading frames were intact, and the signal was restricted to ATP8 rather than affecting multiple loci. A more parsimonious explanation relates to the intrinsic properties of ATP8, one of the shortest and fastest-evolving mitochondrial protein-coding genes, characterized by relaxed functional constraints and elevated substitution rates in insects (Boore 1999; Dong et al. 2021). In lineages experiencing prolonged population structure and isolation, these characteristics may promote locus-specific divergence (Avisé et al. 1987). This interpretation is further supported by previous evidence of long-term population structuring within *Mepraia* (Campos-Soto et al. 2015, 2020) and by documented ancient introgression events (Belintani et al. 2023). Accordingly, the high intraspecific variation detected here most likely represents genuine biological signal rather than methodological bias.

Differences relative to earlier population-level studies are primarily attributable to variation in marker choice and genomic resolution. Whereas previous analyses focused on demographic patterns using a limited number of mitochondrial fragments or nuclear markers, the present study examines variation at the mitogenomic scale across all 13 mitochondrial PCGs. This broader genomic perspective

allows the detection of heterogeneity among individual mitochondrial loci, revealing that intraspecific variability in *M. parapatrica* is unevenly distributed across the mitogenome. Consequently, loci with faster evolutionary rates, such as ATP8 and NAD3, contribute disproportionately to the overall signal of intraspecific divergence detected here, whereas more conserved genes such as COI, NAD1, and NAD2 retain comparatively low levels of variation.

Selection analyses revealed heterogeneous evolutionary pressures across mitochondrial PCGs. Site-level approaches (FUBAR and MEME) identified candidate codons under positive selection in NAD5, NAD6, NAD1, and COIII, whereas most other genes showed little evidence of diversifying selection. In contrast, BUSTED did not detect gene-wide episodic positive selection in any locus, indicating that adaptive signals are localized to specific codons rather than acting across entire genes. This pattern is consistent with expectations for mitochondrial genomes, which are typically subject to strong purifying selection due to their central role in oxidative phosphorylation (Ballard and Whitlock 2004; Rand et al. 2004). Empirical studies have shown that adaptive evolution in mitochondrial genes is usually confined to discrete functional sites rather than occurring gene-wide (Meiklejohn et al. 2007; Fonseca et al. 2008).

At the pathway level, genes showing the strongest selection signals encode subunits of Complex I (NAD1, NAD5, NAD6) and Complex IV (COIII), suggesting that selective pressures may preferentially affect components of the oxidative phosphorylation machinery. Complex I, in particular, has repeatedly been implicated in adaptive mitochondrial evolution across animal taxa (Mishmar et al. 2003; Murray and Horscroft 2016). However, in the absence of structural modeling or experimental validation, functional consequences cannot be inferred, and selection signals are interpreted conservatively as evidence of localized adaptive divergence amid pervasive purifying selection.

In the absence of structural modeling or experimental validation, functional effects cannot be confidently attributed to individual amino acid substitutions. Accordingly, signals of selection are interpreted conservatively, supported by concordant patterns across independent analyses, including genetic distances and PCA. Overall, these results are consistent with localized adaptive divergence occurring within a background of strong purifying selection, with fine-scale adaptive processes restricted to specific codons.

These molecular patterns provide additional context for existing evolutionary hypotheses proposed for the genus. The coexistence of localized positive selection signals and heterogeneous intraspecific diversity aligns with scenarios of historical population subdivision and lineage diversification, including the rapid diversification and ancient hybridization framework suggested by Belintani et al. (2023).

Specifically, the elevated variability in *M. parapatrica* is consistent with biogeographic models invoking isolation and dispersal between insular and continental populations (Campos-Soto et al. 2022). Furthermore, the combination of strong purifying selection with codon-specific adaptive signals fits models of divergence under partial reproductive isolation, supported by cytogenetic differentiation (Frías-Lasserre 2010) and experimental evidence of both pre- and post-zygotic barriers within the genus (Campos-Soto et al. 2016).

By integrating phylogenetic inference, genetic distance analyses, and codon-level selection tests, this study provides a comprehensive view of lineage divergence and evolutionary dynamics in *Mepraia*. The combined evidence supports a refined understanding of species boundaries and is consistent with previously proposed scenarios of population structuring, introgression, and diversification within the genus (Campos-Soto et al. 2016, 2022; Belintani et al. 2023). Patterns of intraspecific variability across mitochondrial loci further highlight the importance of genome-wide approaches for capturing heterogeneous evolutionary signals. In this context, the interplay between purifying selection and localized adaptive processes is consistent with models of divergence under partial reproductive isolation, as suggested by cytogenetic and experimental studies (Frías-Lasserre 2010; Campos-Soto et al. 2016). Collectively, these findings provide a framework for future integrative studies combining genomic, ecological, and functional approaches, and establish a mitogenomic reference for population genetics, phylogeography, and comparative analyses of Triatominae vectors.

In conclusion, the transcriptome-derived mitochondrial genomes analyzed here provide robust evidence for the monophyly and internal relationships of *Mepraia*, while also revealing heterogeneous intraspecific variation and localized signals of adaptive evolution within a background of strong purifying selection. Beyond clarifying evolutionary patterns in an endemic and epidemiologically relevant triatomine lineage, our results demonstrate that publicly available RNA-Seq data can be effectively used to recover mitochondrial coding datasets in non-model insects. These findings expand the genomic resources available for Triatominae and provide a comparative framework for future studies integrating mitochondrial, nuclear, ecological, and functional data to better understand diversification and vector biology in this group.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00438-026-02434-y>.

Acknowledgements We thank the Center for Scientific Computing (NCC/GridUNESP) at São Paulo State University for providing the high-performance computing infrastructure used in this study. We are

also grateful to colleagues of the Rio Claro Arachnology Laboratory (UNESP) for helpful discussions and technical support during the development of this work.

Author contributions M.C.S.S. performed the bioinformatic analyses and interpreted the data. T.B. performed the bioinformatic analyses and interpreted the data, conceived and supervised the study, and wrote the manuscript. J.P.L.G. contributed to the conceptual development of the study and critically revised the manuscript. All authors discussed the results and approved the final version of the manuscript.

Funding The Article Processing Charge (APC) for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614). This research was supported by the São Paulo Research Foundation (FAPESP) under grant numbers 2024/20480-1, 2023/15240-9, and 22/12596-4, and by the Coordination for the Improvement of Higher Education Personnel (CAPES) – Finance Code 001.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

References

- Aguilera-Urbe M, Meza-Lázaro RN, Kieran TJ et al (2020) Phylogeny of the North-Central American clade of blood-sucking reduviid bugs of the tribe Triatomini (Hemiptera: Triatominae) based on the mitochondrial genome. *Infect Genet Evol* 84:104373. <https://doi.org/10.1016/j.meegid.2020.104373>
- Almeida CE, Du L, Wang J et al (2025) The mitogenome of *Triatoma brasiliensis brasiliensis* (Hemiptera: Reduviidae), the main Chagas disease vector in the semi-arid region of northeastern Brazil. *Parasit Vectors* 18:131. <https://doi.org/10.1186/s13071-025-06769-0>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Avise JC (2000) *Phylogeography*. Harvard University Press, Cambridge
- Avise JC, Arnold J, Ball RM et al (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst* 18:489–522. <https://doi.org/10.1146/annurev.es.18.110187.002421>
- Baeza JA, Minish JJ, Michael TP (2025) Assembly of mitochondrial genomes using nanopore long-read technology in three sea chubs (Teleostei: Kyphosidae). *Mol Ecol Resour* 25:e14034. <https://doi.org/10.1111/1755-0998.14034>

- Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Mol Ecol* 13:729–744. <https://doi.org/10.1046/j.1365-294X.2003.02063.x>
- Behura SK (2015) Insect phylogenomics. *Insect Mol Biol* 24:403–441. <https://doi.org/10.1111/imb.12152>
- Belintani T, Congrains C, Pinotti H et al (2023) Transcriptome-based phylogenomic analysis reveals evidence of ancient hybridization events in the *Mepraia* genus (Hemiptera: Reduviidae: Triatominae). *Front Ecol Evol* 11:1215319. <https://doi.org/10.3389/fevo.2023.1215319>
- Belintani T, Oliveira J, Pinotti H et al (2020) Phylogenetic and phenotypic relationships of the *Triatoma sordida* subcomplex (Hemiptera: Reduviidae: Triatominae). *Acta Trop* 212:105679. <https://doi.org/10.1016/j.actatropica.2020.105679>
- Bernt M, Donath A, Jühling F, et al (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* 69:313–319. <https://doi.org/10.1016/j.ympev.2012.08.023>
- Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Res* 27:1767–1780. <https://doi.org/10.1093/nar/27.8.1767>
- Borowiec ML (2016) AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4:e1660. <https://doi.org/10.7717/peerj.1660>
- Brasseur MV, Astrin JJ, Geiger MF, Mayer C (2023) MitoGeneExtractor: efficient extraction of mitochondrial genes from next-generation sequencing libraries. *Methods Ecol Evol* 14:1017–1024. <https://doi.org/10.1111/2041-210X.14075>
- Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu Rev Entomol* 59:95–117. <https://doi.org/10.1146/annurev-ento-011613-162007>
- Campos R, Botto-Mahan C, Coronado X, et al (2011) Wing shape differentiation of *Mepraia* species (Hemiptera: Reduviidae). *Infection, Genetics and Evolution* 11:329–333. <https://doi.org/10.1016/J.MEEGID.2010.11.002>
- Campos R, Botto-Mahan C, Coronado X et al (2013) Phylogenetic relationships of the *spinolai* complex and other Triatomini based on mitochondrial DNA sequences (Hemiptera: Reduviidae). *Vector Borne Zoonotic Dis* 13:73–76. <https://doi.org/10.1089/vbz.2011.0954>
- Campos-Soto R, Torres-Pérez F, Solari A (2015) Phylogenetic incongruence inferred with two mitochondrial genes in *Mepraia* spp. and *Triatoma eratyrusiformis* (Hemiptera, Reduviidae). *Genet Mol Biol* 38:390–395. <https://doi.org/10.1590/S1415-475738320140301>
- Campos-Soto R, Panzera F, Pita S et al (2016) Experimental crosses between *Mepraia gajardoi* and *M. spinolai* and hybrid chromosome analyses reveal the occurrence of several isolation mechanisms. *Infect Genet Evol* 45:205–212. <https://doi.org/10.1016/j.meegid.2016.09.001>
- Campos-Soto R, Díaz-Campusano G, Rives-Blanchard N et al (2020) Biogeographic origin and phylogenetic relationships of *Mepraia* (Hemiptera, Reduviidae) on islands of northern Chile. *PLoS ONE* 15:e0234056. <https://doi.org/10.1371/journal.pone.0234056>
- Campos-Soto R, Rodríguez-Valenzuela E, Díaz-Campusano G et al (2022) Testing phylogeographic hypotheses in *Mepraia* (Hemiptera: Reduviidae) suggests a complex spatio-temporal colonization in the coastal Atacama Desert. *Insects* 13:419. <https://doi.org/10.3390/insects13050419>
- Canizales-Silva C, Alvarado-Lopez MA, Hernández C et al (2025) The genome sequence of the kissing bug, *Panstrongylus geniculatus* (Latreille, 1811). *Wellcome Open Res* 10:43. <https://doi.org/10.12688/wellcomeopenres.23631.1>
- Cannon JT, Kocot KM (2016) Phylogenomics using transcriptome data. In: Shain DH (ed) *Methods in Molecular Biology*, vol 1452. Humana Press, New York, pp 65–80
- Chacón F, Bacigalupo A, Quiroga JF et al (2016) Feeding profile of *Mepraia spinolai*, a sylvatic vector of Chagas disease in Chile. *Acta Trop* 162:171–173. <https://doi.org/10.1016/j.actatropica.2016.06.027>
- Chomczynski P, Mackey K (1995) Substitution of chloroform by bromochloropropane in the single-step method of RNA isolation. *Anal Biochem* 225:163–164. <https://doi.org/10.1006/abio.1995.1126>
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/BIOINFORMATICS/BTP348>
- Cigarroa-Toledo N, Baak-Baak CM, Chan-Pérez JI et al (2024) Dataset of assembly and annotation of the mitogenomes of *Triatoma dimidiata* and *Triatoma huehuetenanguensis* captured from Yucatán. *México Data Brief* 52:109866. <https://doi.org/10.1016/j.dib.2023.109866>
- Coronado X, Rozas M, Botto-Mahan C et al (2009) Molecular epidemiology of Chagas disease in the wild transmission cycle: the evaluation in the sylvatic vector *Mepraia spinolai* from an endemic area of Chile. *Am J Trop Med Hyg* 81:656–659. <https://doi.org/10.4269/ajtmh.2009.09-0053>
- Cywinska A, Hunter FF, Hebert PDN (2006) Identifying Canadian mosquito species through DNA barcodes. *Med Vet Entomol* 20:413–424. <https://doi.org/10.1111/j.1365-2915.2006.00653.x>
- DeSalle R (2017) mtDNA: the small workhorse of evolutionary studies. *Front Biosci* 22:4522. <https://doi.org/10.2741/4522>
- Dong Z, Wang Y, Li C et al (2021) Mitochondrial DNA as a molecular marker in insect ecology: current status and future prospects. *Ann Entomol Soc Am* 114:470–476. <https://doi.org/10.1093/aesa/saa020>
- Dotson EM, Beard CB (2001) Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. *Insect Mol Biol* 10:205–215. <https://doi.org/10.1046/j.1365-2583.2001.00258.x>
- Duchêne S, Archer FI, Vilstrup J et al (2011) Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. *PLoS ONE* 6:e27138. <https://doi.org/10.1371/journal.pone.0027138>
- Eisen JA, Fraser CM (2003) Phylogenomics: intersection of evolution and genomics. *Science* 300:1706–1707. <https://doi.org/10.1126/science.1084684>
- Espinoza B, Martínez-Ibarra JA, Villalobos G et al (2013) Genetic variation of North American triatomines (Insecta: Hemiptera: Reduviidae): initial divergence between species and populations of Chagas disease vector. *Am J Trop Med Hyg* 88:275–284. <https://doi.org/10.4269/ajtmh.2012.12-0105>
- Filée J, Merle M, Bastide H et al (2022) Phylogenomics for Chagas disease vectors of the *Rhodnius* genus (Hemiptera, Triatominae): what we learn from mito-nuclear conflicts and recommendations. *Front Ecol Evol* 9:938. <https://doi.org/10.3389/fevo.2021.750317>
- Finnegan N, Lima MGM, Lynch JW (2025) Mitochondrial DNA for phylogeny building: assessing individual and grouped mtGenes as proxies for the mtGenome in platyrrhines. *Am J Primatol* 87:e70017. <https://doi.org/10.1002/ajp.70017>
- Foley S, Lüddecke T, Cheng D-Q et al (2019) Tarantula phylogenomics: a robust phylogeny of deep theraphosid clades inferred from transcriptome data sheds light on the prickly issue of urticating setae evolution. *Mol Phylogenet Evol* 140:106573. <https://doi.org/10.1016/j.ympev.2019.106573>
- Fonseca RR, Johnson WE, O'Brien SJ et al (2008) The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics* 9:119. <https://doi.org/10.1186/1471-2164-9-119>
- Forni G, Puccio G, Bourguignon T et al (2019) Complete mitochondrial genomes from transcriptomes: assessing pros and cons of data mining for assembling new mitogenomes. *Sci Rep* 9:1–9. <https://doi.org/10.1038/s41598-019-51313-7>

- Fraumene C, Belle EMS, Castri L et al (2006) High resolution analysis and phylogenetic network construction using complete mtDNA sequences in Sardinian genetic isolates. *Mol Biol Evol* 23:2101–2111. <https://doi.org/10.1093/molbev/msl084>
- Frias-Lasserre D (2010) A new species of *Mepraia* (Hemiptera: Reduviidae: Triatominae) from Chile. *Zootaxa* 2623:57–62. <https://doi.org/10.11646/zootaxa.2623.1.4>
- Frias-Lasserre D, González CR, Reyes C, Lecaros Alvarado A (2018) The interspecific hemelytra differences among *Mepraia* species males (Hemiptera, Reduviidae: Triatominae) in Chile. *J Med Entomol* 55:1478–1485. <https://doi.org/10.1093/jme/tjy119>
- Frias-Lasserre DF, de Oliveira J, Pinotti H, da Rosa JA (2019) Morphological description of *Mepraia* spp. females (Hemiptera: Reduviidae, Triatominae). *Acta Trop* 190:389–394. <https://doi.org/10.1016/j.actatropica.2018.11.028>
- Garrido R, Bacigalupo A, Peña-Gómez F et al (2019) Potential impact of climate change on the geographical distribution of two wild vectors of Chagas disease in Chile: *Mepraia spinolai* and *Mepraia gajardoi*. *Parasit Vectors* 12:1–16. <https://doi.org/10.1186/s13071-019-3744-9>
- González CR, Reyes C, Canals A et al (2015) An entomological and seroepidemiological study of the vectorial-transmission risk of Chagas disease in the coast of northern Chile. *Med Vet Entomol* 29:387–392. <https://doi.org/10.1111/mve.12131>
- Groppi BB, Morales MJA, Belintani T, Guadanucci JPL (2025) High diversity and population structure in a widespread tarantula *Sickius longibulbi* (Mygalomorphae: Theraphosidae). *Zool Anz* 319:31–40. <https://doi.org/10.1016/j.jcz.2025.08.007>
- Guindon S, Dufayard J-F, Lefort V et al (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res* 41:e129. <https://doi.org/10.1093/nar/gkt371>
- Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron* 4:9
- Hebert PDN, Cywinska A, Ball SL, de Waard JR (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270:313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hoang DT, Chernomor O, von Haeseler A et al (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>
- Justi SA (2014) Molecular phylogeny of Triatomini (Hemiptera: Reduviidae: Triatominae). *Parasit Vectors* 7:149. <https://doi.org/10.1186/1756-3305-7-149>
- Justi SA, Dale C, Galvão C (2014) DNA barcoding does not separate South American *Triatoma* (Hemiptera: Reduviidae), Chagas disease vectors. *Parasit Vectors* 7:519. <https://doi.org/10.1186/s13071-014-0519-1>
- Justi SA, Galvão C, Schrago CG (2016) Geological changes of the Americas and their influence on the diversification of the Neotropical kissing bugs (Hemiptera: Reduviidae: Triatominae). *PLoS Negl Trop Dis* 10:e0004527. <https://doi.org/10.1371/journal.pntd.0004527>
- Kalyanamorthy S, Minh BQ, Wong TKF et al (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589. <https://doi.org/10.1038/nmeth.4285>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Kieran TJ (2021) Ultraconserved elements reconstruct the evolution of Chagas disease-vectoring kissing bugs (Reduviidae: Triatominae). *Syst Entomol* 46:46–60. <https://doi.org/10.1111/syen.12485>
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120. <https://doi.org/10.1007/BF01731581>
- Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679. <https://doi.org/10.1093/bioinformatics/bti079>
- Kosakovsky Pond SL, Poon AFY, Velazquez R et al (2020) HyPhy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies. *Mol Biol Evol* 37:295–299. <https://doi.org/10.1093/molbev/msz197>
- Kwon YS, Kim JH, Choe JC, Park YC (2012) Low resolution of mitochondrial COI barcodes for identifying species of the genus *Larus* (Charadriiformes: Laridae). *Mitochondrial DNA* 23:157–166. <https://doi.org/10.3109/19401736.2012.660921>
- Lambarri MC, Cifuentes AC, Ayala S et al (2018) Eco-epidemiology of Chagas disease in Chile. In: Afonso LCC (ed) *Chagas Disease — Basic Investigations and Challenges*. IntechOpen, London
- Lemmon AR, Emme SA, Lemmon EM (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst Biol* 61:727–744. <https://doi.org/10.1093/sysbio/sys049>
- Lent H, Wygodzinsky P (1979) Revision of the Triatominae (Hemiptera: Reduviidae) and their significance as vectors of Chagas disease. *Bull Am Mus Nat Hist* 163:123–520
- Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>
- Masonick PK, Knyshov A, Gordon ERL et al (2025) A revised classification of the assassin bugs (Hemiptera: Heteroptera: Reduviidae) based on combined analysis of phylogenomic and morphological data. *Syst Entomol* 50:102–138. <https://doi.org/10.1111/syen.12646>
- Meiklejohn CD, Montooth KL, Rand DM (2007) Positive and negative selection on the mitochondrial genome. *Trends Genet* 23:259–263. <https://doi.org/10.1016/j.tig.2007.03.008>
- Mishmar D, Ruiz-Pesini E, Golik P et al (2003) Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci U S A* 100:171–176. <https://doi.org/10.1073/pnas.0136972100>
- Monteiro FA, Barrett TV, Fitzpatrick S et al (2004) Molecular phylogeography of the Amazonian Chagas disease vectors *Rhodnius prolixus* and *R. robustus*. *Mol Ecol* 13:997–1007. <https://doi.org/10.1046/j.1365-294X.2004.02143.x>
- Monteiro FA, Peretolchina T, Lazoski C et al (2013) Phylogeographic pattern and extensive mitochondrial DNA divergence disclose a species complex within the Chagas disease vector *Triatoma dimidiata*. *PLoS ONE* 8:e70974. <https://doi.org/10.1371/journal.pone.0070974>
- Monteiro FA, Weirauch C, Felix M et al (2018) Evolution, systematics, and biogeography of the Triatominae, vectors of Chagas disease. *Adv Parasitol* 99:265–344. <https://doi.org/10.1016/bs.apar.2017.12.002>
- Moreira DA, Furtado C, Parente TE (2015) The use of transcriptomic next-generation sequencing data to assemble mitochondrial genomes of *Ancistrus* spp. (Loricariidae). *Gene* 573:171–175. <https://doi.org/10.1016/j.gene.2015.08.059>
- Murray AJ, Horscroft JA (2016) Mitochondrial function at extreme high altitude. *J Physiol* 594:1137–1149. <https://doi.org/10.1113/JP270079>
- Murrell B, Wertheim JO, Moola S et al (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 8:e1002764. <https://doi.org/10.1371/journal.pgen.1002764>
- Murrell B, Moola S, Mabona A et al (2013) FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. *Mol Biol Evol* 30:1196–1205. <https://doi.org/10.1093/molbev/mst030>

- Murrell B, Weaver S, Smith MD et al (2015) Gene-wide identification of episodic selection. *Mol Biol Evol* 32:1365–1371. <https://doi.org/10.1093/molbev/msv035>
- Nabholz B, Jarvis ED, Ellegren H (2010) Obtaining mtDNA genomes from next-generation transcriptome sequencing: a case study on the basal Passerida (Aves: Passeriformes) phylogeny. *Mol Phylogenet Evol* 57:466–470. <https://doi.org/10.1016/j.ympev.2010.06.009>
- Oliveira-Correia JPS, de Oliveira J, Gil-Santana HR et al (2024) Taxonomic reassessment of *Rhodnius zeledoni* Jurberg, Rocha & Galvão: a morphological and morphometric analysis comparing its taxonomic relationship with *Rhodnius domesticus* Neiva & Pinto. *BMC Zool* 9:6. <https://doi.org/10.1186/s40850-024-00197-w>
- Paiva VF, Belintani T, de Oliveira J et al (2022) A review of the taxonomy and biology of Triatominae subspecies (Hemiptera: Reduviidae). *Parasitol Res* 121:1–14. <https://doi.org/10.1007/s00436-021-07414-2>
- Paiva VF, de Oliveira J, Belintani T et al (2025) *Hospesneotomae* n. gen. of the Triatomini tribe presents a turnaround in the taxonomy of the *Triatoma protracta* species complex. *Sci Rep* 15:8143. <https://doi.org/10.1038/s41598-025-91399-w>
- Pita S, Panzera F, Vela J et al (2017) Complete mitochondrial genome of *Triatoma infestans* (Hemiptera, Reduviidae, Triatominae), main vector of Chagas disease. *Infect Genet Evol* 54:158–163. <https://doi.org/10.1016/j.meegid.2017.06.003>
- Rambaut A, Drummond AJ, Xie D et al (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904. <https://doi.org/10.1093/sysbio/syy032>
- Rambaut A (2012) FigTree v1.4.2. Institute of Evolutionary Biology, University of Edinburgh. <http://tree.bio.ed.ac.uk/software/figtree/>
- Ramírez JD, Gutierrez S, Rozo M et al (2026) Advances, pitfalls, and challenges in triatomine genomics. *Trends Parasitol* 42:34–50. <https://doi.org/10.1016/j.pt.2025.11.004>
- Rand DM, Haney RA, Fry AJ (2004) Cytonuclear coevolution: the genomics of cooperation. *Trends Ecol Evol* 19:645–653. <https://doi.org/10.1016/j.tree.2004.10.003>
- Raposo do Amaral F, Neves LG, Resende MFR et al (2015) Ultra-conserved elements sequencing as a low-cost source of complete mitochondrial genomes and microsatellite markers in non-model amniotes. *PLoS ONE* 10:e0138446. <https://doi.org/10.1371/journal.pone.0138446>
- Rives-Blanchard N, Torres-Pérez F, Ortiz S et al (2017) *Trypanosoma cruzi* over the ocean: insular zones of Chile with presence of infected vector *Mepraia* species. *Acta Trop* 172:229–231. <https://doi.org/10.1016/j.actatropica.2017.05.020>
- Ronquist F, Teslenko M, van der Mark P et al (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>
- San Juan E, Araya-Donoso R, Véliz D et al (2021) Genetic diversity in a restricted-dispersal kissing bug: the centre–periphery hypothesis halfway. *Mol Ecol* 30:4660–4672. <https://doi.org/10.1111/mec.16093>
- Silva LA, Belintani T, de Paiva VF et al (2023) Integrative taxonomy and phylogenetic systematics of the *Triatoma rubrovaria* subcomplex (Hemiptera, Triatominae). *Acta Trop* 237:106699. <https://doi.org/10.1016/j.actatropica.2022.106699>
- Smith DR (2013) RNA-Seq data: a goldmine for organelle research. *Brief Funct Genomics* 12:454–456. <https://doi.org/10.1093/bfpg/els066>
- Sun S, Jiang W, Yuan Z, Sha Z (2022) Mitogenomes provide insights into the evolution of Thoracotremata (Brachyura: Eubrachyura). *Front Mar Sci* 9:483. <https://doi.org/10.3389/fmars.2022.848203>
- Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Mol Biol Evol* 38:3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Tapia-Garay V, Figueroa DP, Maldonado A et al (2018) Assessing the risk zones of Chagas' disease in Chile, in a world marked by global climatic change. *Mem Inst Oswaldo Cruz* 113:e170172. <https://doi.org/10.1590/0074-02760170172>
- Tian Y, Smith DR (2016) Recovering complete mitochondrial genome sequences from RNA-Seq: a case study of *Polytomella* non-photosynthetic green algae. *Mol Phylogenet Evol* 98:57–62. <https://doi.org/10.1016/j.ympev.2016.01.017>
- Wang Y, Chen J, Jiang L-Y, Qiao G-X (2015) Hemipteran mitochondrial genomes: features, structures and implications for phylogeny. *Int J Mol Sci* 16:12382–12404. <https://doi.org/10.3390/ijms160612382>
- Wong T, Ly-Trong N, Ren H et al (2025) IQ-TREE 3: phylogenomic inference software using complex evolutionary models. *Mol Biol Evol*. <https://doi.org/10.1093/molbev/msaf015>
- Yang S, Zhu H, Yang Z et al (2024) The completed mitochondrial genomes of *Globodera vulgaris* reveals new insights into the genus *Globodera* phylogeny. *Sci Rep* 14:7253. <https://doi.org/10.1038/s41598-024-57736-1>
- Zarrei M, Talent N, Kuzmina M et al (2015) DNA barcodes from four loci provide poor resolution of taxonomic groups in the genus *Crataegus*. *AoB Plants* 7:plv045. <https://doi.org/10.1093/aobpla/plv045>

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