



HAL
open science

Shotgun assembly of the assassin bug *Brontostoma colossus* mitochondrial genome (Heteroptera, Reduviidae)

Arthur Kocher, Maria Kamilari, Emeline Lhuillier, Eric Coissac, Julie Péneau,
Jérôme Chave, Jérôme Murienne

► **To cite this version:**

Arthur Kocher, Maria Kamilari, Emeline Lhuillier, Eric Coissac, Julie Péneau, et al.. Shotgun assembly of the assassin bug *Brontostoma colossus* mitochondrial genome (Heteroptera, Reduviidae). *Gene*, 2014, 552 (1), pp.184-194. <10.1016/j.gene.2014.09.033>. <hal-04917775>

HAL Id: hal-04917775

<https://hal.science/hal-04917775v1>

Submitted on 28 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

1 **Shotgun assembly of the assassin bug *Brontostoma colossus***
2 **mitochondrial genome (Heteroptera, Reduviidae)**

3 Arthur Kocher^a, Maria Kamilari^{ab}, Emeline Lhuillier^c, Eric Coissac^d, Julie Péneau^e, Jérôme Chave^a,
4 Jérôme Murienne^a.

5 a. Laboratoire EDB (UMR 5174, CNRS/UPS/ENFA) ; Université Paul Sabatier, Bât. 4R1, 118 route de
6 Narbonne, 31062 Toulouse cedex 9, France ; arthur.kocher@gmail.com; jerome.murienne@univ-tlse3.fr;
7 jerome.chave@univ-tlse3.fr; mariakamilari@gmail.com;

8 b. Division of Animal Biology, Department of Biology, School of Natural Sciences, University of Patras, GR-
9 26500 Patras, Greece.

10 c. INRA, GeT-PlaGe, UAR 1209 Département de génétique animale, INRA Auzeville, 31326 Castanet-
11 Tolosan, France ; emeline.lhuillier@udear.cnrs.fr

12 d. Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, 38041 Grenoble, France ;
13 eric.coissac@inrialpes.fr

14 e. Équipe « Écosystèmes Amazoniens et Pathologie Tropicale » (EA 3593), Faculté de Médecine
15 Hyacinthe Bastaraud de l'Université des Antilles et de la Guyane, Campus Saint Denis, av. d'Estrées, F-
16 97306 Cayenne, French Guiana ; peneau.julie@wanadoo.fr

17

18 Corresponding author : Arthur Kocher, Laboratoire EDB (UMR 5174, CNRS/UPS/ENFA) ; address: Université
19 Paul Sabatier, Bât. 4R1, 118 route de Narbonne, 31062 Toulouse cedex 9, France ; email :
20 arthur.kocher@gmail.com; phone : +33 6.14.30.79.46

21

22 **Abstract**

23 The complete mitochondrial genome of the assassin bug *Brontostoma colossus* (Distant, 1902)
24 (Heteroptera: Reduviidae) has been sequenced using a genome-skimming approach on an Illumina
25 HiSeq 2000 platform. Fifty-four additional heteropteran mitogenomes, including five assassin bug
26 species, were retrieved to allow for comparisons and phylogenetic analyses. The mitochondrial
27 genome of *B. colossus* was determined to be 16,625 bp long, and consists of 13 protein-coding genes
28 (PCGs), 23 transfer-RNA genes (tRNAs), two ribosomal-RNA genes (rRNAs), and one control region.
29 The nucleotide composition is biased toward adenine and thymine (A+T=73.4%). Overall,
30 architecture, nucleotide composition and genome asymmetry are similar among all available assassin
31 bugs mitogenomes. All PCGs have usual start-codons (Met and Ile). Three T and two TA incomplete
32 termination codons were identified adjacent to tRNAs, which was consistent with the punctuation
33 model for primary transcripts processing followed by 3' polyadenylation of mature mRNA. All tRNAs
34 exhibit the classic clover-leaf secondary structure except for tRNA_{Ser(AGN)} in which the DHU arm
35 forms a simple loop. Two notable features are present in the *B. colossus* mitogenome: (i) a 131 bp
36 duplicated unit including the complete tRNA_{Arg} gene, resulting in 23 potentially functional tRNAs in
37 total, and (ii) a 857 bp duplicated region comprising 277 bp of the srRNA gene and 580 bp of the
38 control region. A phylogenetic analysis based on 55 true bugs mitogenomes confirmed that *B.*
39 *colossus* belongs to Reduviidae, but contradicted widely accepted hypothesis. This highlights the
40 limits of phylogenetic analyses based on mitochondrial data only.

41

42

43

44

45

46

47

48

49

50

51

52 **Keywords:** genome skimming; next-generation sequencing; phylogenomic.

53

55 Mitochondrial DNA has various interesting properties such as abundance in animal tissues, small size,
56 relatively simple genomic structure, fast rate of evolution, and a straightforward mode of
57 transmission with a low level of recombination (due to its maternal inheritance). This makes it a
58 valuable tool for comparative genomics, population genetics and phylogenetics at various taxonomic
59 resolutions (Avice et al., 1987; Moritz et al., 1987). An increasing number of complete mitochondrial
60 genomes has been made available in the past decade, relying on long range Polymerase Chain
61 Reaction (PCR), but this approach is difficult to perform and time-consuming. The immense yield now
62 provided by Next Generation Sequencing (NGS) helps resolve these issues. The sequencing of the
63 complete nuclear genome remains expensive because it requires a deep sequencing, but a relatively
64 shallow sequencing can be used to recover the high copy fraction of mitochondrial DNA. This
65 «genome skimming» approach, originally developed for plant organelles assemblage (Besnard et al.,
66 2013; Malé et al., 2014; Straub et al., 2012), has been successfully used to assemble a wide variety of
67 animal mitochondrial genomes (Besnard et al., 2014; Doyle et al., 2014; Thompson et al., 2014;
68 Veale et al., 2014). However, mitogenome-based studies have been unbalanced among taxa, and the
69 amount of available data for insects remains limited in comparison with that of vertebrates (Gissi et
70 al., 2008; Salvato et al., 2008).

71 As for most metazoans, the mitogenome of insects is a circular double-stranded molecule of 14-20kb
72 in size and exhibits a typical set of 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs)
73 and 2 ribosomal RNA genes (rRNAs) (Boore, 1999; Wolstenholme, 1992), even though variations in
74 gene content exist (Gissi et al., 2008; Junqueira et al., 2004; Shao and Barker, 2003; Thao et al.,
75 2004). In addition, it contains a large non-coding region, the control region (CR), which is implicated
76 in the initiation of transcription and replication processes (Bernt et al., 2013a; Clayton, 1992; Saito,
77 2005; Zhang and Hewitt, 1997).

78 Heteroptera (true bugs) contains over 40,000 described species to date, constituting one of the most
79 diverse group of non-holometabolous insects (Weirauch and Schuh, 2011). Assassin bugs
80 (Reduviidae) are a large family of mostly predatory land bugs belonging to the infraorder
81 Cimicomorpha. It currently comprises close to 7,000 species worldwide, that exhibit a remarkable
82 diversity in morphological traits and life habits (Weirauch and Munro, 2009; Wheeler, 1997). Some of
83 them are of agricultural or medical importance, the most notorious being part of the hematophagous
84 Triatominae subfamily, known as vectors of Chagas disease in Central and South America (Lent and
85 Wygodzinsky, 1979). Five assassin bugs mitogenomes have been sequenced so far: *Agriosphodrus*
86 *dohrni* (Li et al., 2011), *Oncocephalus breviscutum* (Li et al., 2013), *Sirthena flavipes* (Gao et. al,
87 2013), *Valentia hoffmanni* (Hua et. al, 2009) and *Triatoma dimidiata* (Dotson and Beard, 2001).
88 Among Reduviidae, the neotropical genus *Brontostoma* currently includes around 20 species
89 (Dougherty, 1995; Gil-Santana and Baena, 2009; Maldonado Capriles, 1990). It is characterized by a
90 bright coloration with red, yellow, black and brown. Like all members of the Ectrichiinae subfamily,
91 they are predators specialized on millipedes.

92 In this paper, we describe a genome-skimming approach using Illumina technology to assemble the
93 complete mitochondrial genome of *Brontostoma colossus* (Distant, 1902). Its organization and
94 features are compared to five other mitogenomes of Reduviidae. Fifty-four additional heteropteran
95 mitochondrial genomes are used to perform a phylogenetic analysis.

96 **2. Material and method**

97 *2.1 Specimen, DNA extraction and sequencing*

98 One specimen of *B. colossus* was collected in French Guiana (RN2 Roura-Saint Georges) on April
99 20th 2010. Total genomic DNA was extracted from leg muscle tissue using the DNeasy Blood and
100 Tissue kit (Qiagen, Valencia, CA, USA), following a protocol adapted from the manufacturer's
101 instructions (Supplementary Data 1). The quality and quantity of extracted genomic DNA was
102 evaluated using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham,
103 MA, USA) and a PicoGreen double-stranded DNA quantitation assay kit (Life Technologies,
104 Carlsbad, CA, USA).

105 The genomic DNA was sent for library construction and sequencing to the GeT-PlaGe core facilities of
106 Genotoul (Toulouse, France). 288 ng of DNA were used for library construction using the Illumina
107 TruSeq Nano DNA Sample Prep Kit following the instructions of the supplier (Illumina Inc., San Diego,
108 CA, USA). After shearing by ultrasonication with a Covaris M220 (Covaris Inc., Woburn, MA, USA),
109 purified fragments were A-tailed and ligated to sequencing indexed adapters. Fragments with an
110 insert size around 450 bp were selected with Agencourt Ampure XP beads (Beckman Coulter, Inc.),
111 and enriched with 8 cycles of PCR before library quantification and validation. The library was
112 multiplexed with 23 other libraries (generated in other projects). The pool of libraries was then
113 hybridized on one lane of HiSeq 2000 flow cell using the Illumina TruSeq PE Cluster Kit v.3, and
114 paired-end reads of 100 nucleotides were collected on the HiSeq 2000 sequencer using the Illumina
115 TruSeq SBS Kit v.3 (200 cycles). Quality filtering was performed by the Consensus Assessment of
116 Sequence and Variation (CASAVA) pipeline. Sequence data were stored on the NG6 platform
117 (Mariette et al., 2012) and all the computations were performed on the computer cluster of the
118 Genotoul bioinformatic platform (Toulouse, France).

119 *2.2 Sequence assembly*

120 Mitochondrial genome and nuclear ribosomal clusters were assembled using a previously described
121 strategy (Besnard et al 2013; Besnard et al 2014; Malé et al. 2014). It is in essence similar to that
122 proposed by Hahn et al. (2013). Reads aligning with the mitochondrial protein sequences of the
123 closely related species *Triatoma dimidiata* (NC 002609, Dotson and Beard, 2001) were identified
124 using the PLAST program (Nguyen and Lavenier, 2009). For the nuclear ribosomal cluster, we started
125 from reads aligning with the 28S and 18S rRNA genes of *Eurydema maracandica* (Yu et al., 2013).

126 Reads with a match of at least 90% were assembled into contigs using the Velvet assembler (Zerbino
127 and Birney, 2008) with a k-mer length of 81 and all the remaining parameters left at their default
128 values. The resulting contigs were used as seeds to initiate the genome walking strategy (iterative
129 mapping) using the extractreads2 program (included in the Obitools package;
130 <http://metabarcoding.org/obitools>). Reads sharing at least 80 consecutive bp with the seeds were
131 selected and subsequently used as seed to repeat the operation until no new read was identified.
132 The newly selected reads were assembled with the Velvet assembler. The few resulting contigs were
133 assembled using Geneious 6.0.6 Pro (Biomatters, Auckland, New Zealand). Two regions of the
134 genome were not assembled using this procedure, due to ambiguities in the assembling process. At
135 these locations, repeated regions were revealed by the coexistence of two assembling paths with
136 significant coverage support (i.e. of the same magnitude than the average coverage), one path being

137 at the junction between two copies of the repeated element, the other being at the end of the last
138 copy (see Figure 1). We inferred the repeated copy number by comparing the number of reads
139 mapping on a DNA fragment present only once in the mitogenome to that of DNA fragment
140 belonging to the repeated element, assuming that the read coverage of a particular genomic region
141 is proportional to its copy fraction in the sample. This coverage analysis is detailed in the
142 Supplementary Material.

143 Coverage statistics were computed on the assembled genome with Geneious 6.0.6, by mapping
144 the reads using the following mapping parameter: a minimum overlap of 100 bp, a minimum
145 overlap identity of 95%, a word length of 50 and a maximum mismatch per read of 5%.

146 2.3 Genome annotation

147 The mitochondrial genome was first annotated using the MITOS web server (Bernt et al., 2013b)
148 applying the invertebrate mitochondrial genetic code (NCBI code Table 5). The annotations
149 of tRNA genes were kept unchanged. The annotations of PCGs were refined by checking
150 manually for consistent start/stop codons and reading frames. The annotations of rRNA genes
151 were extended until adjacent tRNAs (tRNA_{Ser}(TCN), tRNA_{Leu}(CUN) and tRNA_{Val}), following the
152 punctuation model of mtDNA transcription (Ojala et al., 1981; Stewart and Beckenbach, 2009).
153 The 5' end of srRNA is not flanked by a tRNA. We adjusted it by mapping the srRNA of the
154 *Agriosphodrus dohrni* mitogenome for which the secondary structure was predicted, and
155 presented all expected domains (Li et al. 2011). We used Geneious 6.0.6 Pro with the following
156 parameters: a word length of seven; maximum gap size of 15 and maximum mismatches of
157 40%. This approach conducted to extend lrRNA's annotation by 33 bp until tRNA_{Leu} at its 3' end
158 and by 578 bp until tRNA-Val at its 5' end. srRNA's annotation was extended by 2 bp until tRNA_{Val}
159 and by 47 bp at its 5' end. We further verify the consistency of these new annotations by mapping
160 the rRNAs of *Agriosphodrus dohrni* as described above. The remaining large non-annotated
161 sequence was annotated as the control region in homology with other insect mitogenomes.

162 The 18S and 28S rRNAs were annotated in comparison with that of *Eurydema maracandica* (Yu et al.,
163 2013). The 5.8S rRNA was annotated in comparison with that of *Triatoma dimidiata* (accession
164 number: KF142517).

165 2.4 Sequence analyses and phylogenetics

166 Base composition and codon usage were computed with MEGA6 (Tamura et al., 2013). AT-skew
167 $[(A-T)/(A+T)]$ and GC-skew $[(G-C)/(G+C)]$ were used to measure nucleotide compositional
168 differences between genes (Perna and Kocher, 1995) Relative synonymous codon usage (RSCU)
169 were used to described bias in synonymous codon composition.

170 Tandem repeats were identified using Tandem Repeat Finder webserver (Benson, 1999). The
171 secondary structure of tRNA's was inferred via the MITOS web server pipeline. Putative stem-
172 loop structures were inferred using the RNAstructure web server (Bellaousov et al., 2013). We
173 looked for structures conserved among the six assassin bugs within the 100 bp of the control
174 region flanking tRNA_{Ile} where stem loops structures have already been reported in Reduviidae
175 (Dostson and Beard, 2001; Gao et al. 2013). We used the Turbofold algorithm (Harmanici et al.

176 2011), which infers secondary structure from high base pairing probabilities using the
177 information derived from the sequence itself via the nearest neighbour thermodynamic model
178 and also the information computed by using pairwise-sequence-alignment-based probabilities.

179 Fifty-four additional heteropteran mitogenomes, including five assassin bugs, were
180 downloaded from Genbank (Table 1). Two mitogenomes of Auchenorrhyncha were used as
181 outgroups. The 13 PCGs were used for the analysis to allow for comparison with previous
182 studies (Li et al., 2011; Yang et al., 2013). They were first aligned separately based on amino-
183 acids translation with translatorX (Abascal et al., 2010). Divergent regions were removed with
184 Gblocks.0.91b before back-translation in order to conserve reading frames. All resulting
185 alignments were then concatenated using FASconCAT (Kück and Meusemann, 2010). The best
186 partitioning scheme and substitution model were inferred with PartitionFinder.1.1.1 (Lanfear
187 et al., 2012), using the greedy algorithm for scheme search and the Bayesian information
188 criterion for scheme selection. A maximum-likelihood (ML) analysis was performed with
189 RAxML 8.0 (Stamatakis, 2014), using the rapid bootstrap analysis option with the majority-rule
190 tree based bootstopping criteria. Bootstrap support values were printed on the best ML tree.
191 A Bayesian analysis was conducted using Mr.Bayes 3.2 (Ronquist and Huelsenbeck, 2003),
192 starting from four random trees with 10 Markov chains (nine heated chain and 1 cold chain),
193 2,000,000 generations and all other parameters set to default. Each set was sampled every
194 200 generations with a burn-in of 25% of the sampled trees. At the end of the analysis, the
195 average standard deviation of split frequencies was below the recommended 0.01.

196 **3 Results and discussion**

197 *3.1 Genome sequencing, assembly and annotation*

198 After filtering 4.74% of the initial reads, raw sequence data represented a total of 7,831,929 paired-
199 end reads (15,663,858 reads in total). Among the remaining reads, 34,224 were assembled into a
200 16,625 bp circular sequence, representing the complete mitochondrial genome with an average
201 sequencing depth of 209.6. A circular map of the mitogenome and the assembly coverage are
202 presented on Figure 2. The sequence was deposited in Genbank under the accession number
203 KM044501. A total of 38 genes (13 PCGs, 23 tRNAs, two rRNAs) and one control region were
204 identified. Twenty-four genes are encoded on the majority strand and the others mapped to the
205 minority strand. Seven gene overlaps were observed, the longest being a 8-bp region between
206 tRNA_{Cys} and tRNA_{Trp} (Table 2), which is a peculiar feature in Arthropoda (Bernt et al., 2013a). Apart
207 from the control region, 13 non-coding regions ranging from 1 bp to 46 bp were identified.

208 Three repeated regions were identified: a 131-bp element containing tRNA_{Arg}; an 857-bp element
209 consisting of 277 bp of srRNA and 580 bp of the control region and a 74-bp repeated element
210 followed by a 39 partial copy located in the control region were tandem repeats are also found in
211 other assassin bugs (Li et al., 2011). The existence of these duplications was supported by assembling
212 paths with high coverage (over 100 reads). Details on the results of the coverage analysis are given in
213 the Supplementary material. However, the sequencing technology used in this study does not allow
214 inferring repeats copy numbers with high certainty, especially if a polymorphism exists in the sample.
215 Indeed, heteroplasmy is often associated with tandem repetition (Zhang, 1997).

216 The complete nuclear ribosomal gene cluster was recovered. A total of 21,514 reads were assembled
217 into a 8,287 bp sequence comprising of 18S rRNA (1,893 bp), ITS1 (1,141 bp), 5.8S rRNA (155 bp),
218 ITS2 (941 bp) and 28S rRNA (4,077 bp). The sequence was deposited on Genbank under the accession
219 number KM278219.

220 The mitogenome of *B. colossus* shares the same architecture and orientation as the other six
221 mitogenomes of assassin bugs, except for the presence of an additional tRNA_{Arg} gene, as will be
222 described below. This gene arrangement (without the additional tRNA_{Arg}) is also found in *Drosophila*
223 *melanogaster* and was the first to be determined, differing by a single tRNA translocation, from that
224 of the chelicerate *Limulus polyphemus*, which is considered ancestral for Arthropoda (Boore et al.,
225 1995; Lavrov et al., 2000). This mitogenome organisation is also found in crustaceans, and is thought
226 to be ancestral for the insect-crustacean clade (Boore, 1999; Boore et al., 1998). Among the available
227 data for heteropterans, seven out of 54 species have been found to present a different gene
228 arrangement: *Nabicalis apicalis* mitogenome miss the cluster containing tRNA_{Ile}, tRNA_{Gln} and
229 tRNA_{Met} (Li et al., 2012a). tRNA_{Ile} and tRNA_{Gln} are also missing in *Urochela quadrinotata* (Yuting et
230 al., 2012). The positions of tRNA_{Thr} and tRNA_{Pro} are inverted in *Physopelta gutta* (Hua et al., 2008).
231 The gene order of the *Stenopirates* sp. mitogenome differs largely with the inversion of two tRNA
232 genes (tRNA_{Thr} and tRNA_{Pro}) and translocations of five gene clusters (tRNA_{Thr}-tRNA_{Pro} -ND6, CYTB-
233 tRNA_{Ser}(TCN), ND1- tRNA_{Leu}(CUN), l-rRNA- tRNA_{Val} -s-rRNA and the control region) between ND4L and
234 tRNA_{Ile} (Li et al., 2012b). The positions of tRNA_{Cys} and tRNA_{Trp} are exchanged in *Aradacanthia heissi*
235 (Shi et al., 2012). In the latter, as well as in the two other Aradoidea represented (*Brachyrhynchus*
236 *hsiao* and *Neuroctenus parus*); the positions of tRNA_{Ile} and tRNA_{Gln} are exchanged (Li et al., 2014;
237 Hua et al., 2008).

238 The size of the six assassin bug mitogenomes ranges from 15,625 bp in *V. hoffmannii* to 17,019 bp in
239 *T. dimidiata*. These differences are mostly due to variations in the size of the control region, which is
240 generally observed for all insects. Previous studies have reported control region size ranging from 70
241 bp in *Ruspolia dubia* (Orthoptera) to 4,599bp in *Drosophila melanogaster* (Diptera) (Garesse, 1988;
242 Zhou et al., 2007).

243 3.2 Protein-coding genes

244 The total length of the 13 PCGs was 11,041 bp. Their nucleotide composition is strongly biased
245 toward AT with an overall AT content of 72.8% (Supplementary Table 1). All PCGs have an ATN start
246 codon (Table 2). Six PCGs initiated with ATT (ND2, COX2, ATP8, ND5, ND4L and ND1), five initiated
247 with ATG (COX1, ATP6, COX3, ND4 and CYTB) and two initiated with ATA (ND3 and ND6). Four genes
248 share the same ATG start codon in the six assassin bugs mitogenomes (COXI, ATP6, COXIII and ND4).
249 No GTG start codon was found in *B. colossus* in contrast with other assassin bugs for ND5, ND4L and
250 ND1 genes (Supplementary Table 2). Other unconventional start codons were described in insects
251 such as TTG in heteropterans (Yang et al., 2013), CGA and TTAG in lepidopterans (Lee et al., 2006;
252 Yukuhiro et al., 2002), or ATAA, GTAA and TTAA in dipterans (Ballard, 2000; Clary and Wolstenholme,
253 1985) but none of them were found in assassin bugs.

254 The majority of PCGs have a usual TAA stop codon, but three T and two TA stop codons were
255 identified (ND2, COX1, COX2 and ND5, CYTB respectively). These incomplete stop codons are
256 immediately adjacent to tRNA genes encoded on the same strand, consistent with the punctuation
257 model for primary transcripts processing followed by 3' polyadenylation of mature mRNA that will
258 allow the completion of termination codons (Nagaike, 2005; Ojala et al., 1981; Stewart and
259 Beckenbach, 2009). Incomplete stop codons can be found in all six assassin bugs mitogenomes and
260 are shared with many arthropods (Boore, 2000).

261 3.3 Ribosomal and transfer RNA genes

262 rRNA genes locations and lengths are similar to those of other insects mitogenomes. IrRNA is located
263 between tRNA_{Leu(CUN)} and tRNA_{Val} and is 1,251 bp-long. srRNA is located between tRNA_{Val} and the
264 control region and is 793 bp-long. Their AT content is respectively 79.0% and 74.0%.

265 The classical set of 22 tRNAs found in arthropods is present in *B. colossus*, but an additional copy of
266 the tRNA_{Arg} gene was identified (see below). Their lengths vary between 61 bp (tRNA_{Ala}) and 72 bp
267 (tRNA_{Lys}). Secondary structures of tRNAs are schematized in Supplementary Figure 1. The classical
268 clover leaf structure was observed for each of them, except for tRNA_{Ser(AGN)}, in which the D arm is
269 reduced to a simple loop, as in many insects, and more generally, in most bilaterians (Bernt et al.,
270 2013a; Wolstenholme, 1992).

271 Flanking tRNA_{Ala}, we identified a duplicated element consisting of two identical 131-bp copies
272 separated by eight non-coding base pairs. It includes the entire tRNA_{Arg} and short-flanking regions
273 corresponding to 28 bp of tRNA_{Ala} and 38 bp of tRNA_{Asn} (Figure 3). This unusual feature results in
274 two copies of tRNA_{Arg} for a total of 23 tRNA genes which is relatively rare in insect mitogenomes,
275 even though more than 22 tRNAs were already observed in other species such as *Coreana raphaelis*
276 (Lepidoptera; Kim et al., 2006), *Thrips imaginis* (Thysanoptera; Shao and Barker, 2003), *Chrysomya*
277 *chloropyga* (Diptera; Junqueira et al., 2004) and *Trialeurodes vaporariorum* (Hemiptera; Thao et al.,
278 2004). To our knowledge, the presence of an additional tRNA_{Arg} has only been described in Porifera
279 and Placozoa (Lavrov and Lang, 2005; Signorovitch et al., 2007). The two copies of this duplicated
280 element are strictly identical, which suggests a recent origin of the duplication event. Interestingly, it
281 could lead to gene rearrangement through a duplication/deletion mechanism involving the random
282 deletion of the original copy of the gene (Boore, 2000; Moritz and Brown, 1986; Shao et al., 2006).
283 Most repeated elements are located close to the replication origin, supporting the idea that
284 mitogenomic duplication events are mainly due to replication slippage mechanisms (Macey et al.,
285 1998; Zhang and Hewitt, 1997). However, to our knowledge, no putative replication origin located
286 close to this duplicated region has been mentioned so far.

287 3.4 Non-coding regions

288 Thirteen short intergenic spacers (IGS) and a long control region were identified, matching the usual
289 organisation of insect mitogenomes. Most IGS are very short, with less than four base pairs, and
290 seven overlapping sequences are found (Table 2). The longest IGS are the 41 and 46 bp-flanking
291 regions of the first copy of tRNA_{Arg} and the 38-bp IGS found between ND1 and tRNA_{Ser2}. The latter is

292 remarkably long in every assassin bug mitogenome, ranging from 23 bp in *V. hoffmanni* to 309 bp in
293 *T. dimidiata*. It exhibits tandem repeats in *A. dohrni* and *T. dimidiata* and has been suspected to be
294 one of the replication origins (Dotson and Beard, 2001; Li et al., 2011) comparably to the 193 bp
295 region located between the tRNA_{Ileu}(UUR) and COXII genes in the honeybee *A. mellifera* (Crozier and
296 Crozier, 1993).

297 The control region is 1,896-bp long and is located between the srRNA and tRNA_{Ile} genes. In *B.*
298 *colossus*, as well as in the five other Reduviidae, it exhibits a higher G+C content than that of the
299 whole mitogenome, in contrast with other insect species in which the control region was found to be
300 remarkably A+T rich (Zhang and Hewitt, 1997).

301 The alignment of the six Reduviidae control regions reveals a conserved sequence block (CSB) of 40
302 bp, including a string of 13 Gs (Figure 4a; Li et al., 2011). CSBs have been identified in the control
303 region of various metazoans and are generally thought to play a role in the replication mechanism
304 (Lee et al., 1995; Walberg and Clayton, 1981; Zhang and Hewitt, 1997). However, we did not find
305 similarity between the CSB described here and those reported for other taxa, even though G islands
306 have already been described in other insects (Oliveira et al., 2008). More studies are needed to
307 identify precisely replication origins and to speculate on the role of adjacent sequences, as was
308 proposed by (Saito, 2005) who strongly suspected a “T-strech” sequence conserved in *Drosophila* and
309 other insect species to be involved in the replication process.

310 The control regions of the six assassin bugs present a similar organization (Figure 4b). In all of them,
311 except for *S. flavipes*, tandem-repeats were identified between the CSB and tRNA_{Ile}. In *B. colossus*,
312 they consist of four 74 bp units and one 39 bp unit, the latter corresponding to a partial copy of the
313 74 bp unit. The 100 bp preceding the CSB are remarkably G+C rich (42%) in *B. colossus*, but the
314 nucleotide composition of the whole control region does not show any clear pattern, in contrast with
315 other assassin bugs in which long G+C and A+T rich regions were found preceding and following the
316 CSB respectively (Dotson and Beard, 2001; Gao et al., 2013; Li et al., 2011). In *B. colossus* control
317 region, a notable feature is a large duplicated region consisting of two identical copies of 857 bp
318 separated by seven non-coding nucleotides. It includes 277 bp of srRNA and 580 bp of the control
319 region, comprising the CSB. It is remarkably long and results in two copies of the CSB, which could
320 therefore have implications on the mitogenome replication process.

321 Within the 60 bp of the control region flanking the tRNA_{Ile}, DNA segments have the potential to form
322 stem-loop structures involving at least 11 base-pairings in the six assassin bugs mitogenomes except
323 for *V. hoffmanni* (Figure 5). Such features may be involved in the replication mechanism (Song and
324 Liang, 2009; Zhang et al., 1995). However, this is only speculative. More comprehensive studies
325 would be required to assess the significance of these inferences and the potential role of these
326 structures.

327 3.5 Nucleotide content and codon usage

328 The nucleotide composition is strongly biased toward adenine and thymine in the mitogenome of *B.*
329 *colossus*, with A+T representing 73.5% of the whole sequence and ranging from 70.2% in the control
330 region, 70.8% in protein-coding genes, 76.8% in tRNA genes to 77.1% in rRNA genes. AT-rich codons
331 are predominant, with the most prevalent being in order ATT (Ile), TTA (Leu), TTT (Phe) and ATA

332 (Met). The relative synonymous codon usage (RSCU) clearly indicates that AT rich codons are
333 favoured among synonymous codons (Supplementary Table 3). At the third codon position, AT
334 content is particularly high (82.8%), and G nucleotides are under-represented (GC skew= -0.20). AT
335 content, as well as A-T and G-C skews patterns, are similar among the six assassin bugs mitochondrial
336 genomes (Supplementary Table 1).

337 3.6 Phylogenetic analysis

338 Bayesian inference and Maximum Likelihood analysis (ML) generated phylogenetic trees with very
339 similar topologies. The tree inferred by the Bayesian method is presented in Figure 6 with nodes
340 posterior probabilities and ML bootstrap support values. The topology of the best ML tree for
341 Reduviidae is also presented. The relationships among Reduviidae are conserved in both analyses
342 except for the position of *B. colossus*, which is placed as a sister group of *V. hoffmanni* under
343 Bayesian inference whereas it is the early lineage of Reduviidae in the ML tree. Our results are not
344 well supported and are hardly comparable with those of recent studies that have addressed the
345 relationships among assassin bugs based on nuclear and mitochondrial DNA as well as morphological
346 data for a large number of taxa (Weirauch and Munro, 2009; Weirauch, 2008). However, the higher-
347 level relationships of Reduvida remain poorly resolved and the addition of mitogenomic data for
348 more taxa will surely provide useful phylogenetic information in the future.

349 The 20 superfamilies represented in our dataset are monophyletic except for Miroidea. At the infra-
350 order level, Pentatomorpha is monophyletic with the following relationships between the five
351 superfamilies: Aradoidea + (Pentatomoidea + (Lygaeoidea + (Pyrrhocoroidea + Coreoidea))). These
352 relationships are strongly supported in our analyses and do not confirm the results of previous
353 studies based on a subset of the present mitogenomic data that placed Coreoidea and Lygaeoidea as
354 sister groups (Hua et al., 2008; Yang et al., 2013). Based on the tree generated by Bayesian inference,
355 Nepomorpha is monophyletic with the following relationships between the six superfamilies:
356 (Pleioidea + Corixoidea) + ((Notonectoidea + Naucoroidea) + (Nepoidea + Ochteroidea)). ML analysis
357 differs in positioning Pleoidea as a sister group of the remaining Nepomorpha. These results are
358 inconsistent with those of a recent study based on molecular and morphological data (Hebsgaard et
359 al., 2004). Interestingly, they also contradict a previous analysis of a subset of the present
360 mitogenomic dataset by confirming the monophyly of Nepomorpha including Pleoidea, for which an
361 infraordinal status was proposed (Hua et al., 2009). Gerromorpha and Leptopodomorpha are also
362 monophyletic, but only two species of each were included in the study. Enicocephalomorpha was
363 only represented by *Stenopirates* sp.

364 Our analysis supports the paraphyly of Cimicomorpha that consisted of four different clades:
365 (Cimicoidea + Naboidea), Reduvidae, Miridae and Tingidae. Tingidae is placed as a sister group to
366 all remaining Heteroptera. However, the monophyly of Cimicomorpha has been widely accepted and
367 is supported by various analyses that have been using mitochondrial and nuclear data for a larger
368 number of taxa (Li et al., 2012; Schuh et al., (2009); Tian et al., 2008). Infraordinal relationships are
369 conserved in both ML and Bayesian analyses: Tingidae + (Miridae + (Pentatomomorpha +
370 ((Cimicoidea + Naboidea) + (Leptopodomorpha + ((Enicocephalomorpha + Gerromorpha) +
371 Nepomorpha)))). These results are questioning the general consensus that considers
372 Enicocephalomorpha as the early infraorder of Heteroptera (Weirauch et al., 2011). However, the
373 relationships between the infraorders of Heteroptera remain controversial. Only few phylogenetic
374 studies have addressed the question and most of these have only included a small number of taxa
375 (Mahner, 1993; Wheeler et al., 1993; Xie et al., 2008). Mitogenomic data provide a new insight in this
376 regard, but more taxa should be added to the current database, especially in the poorly represented
377 infra-orders Enicocephalomorpha, Gerromorpha, Leptopodomorpha and Dipsocoromorpha.

378 Our results are incongruent with current phylogenetic hypothesis of Heteroptera (e.g. Schuh et al.
379 2009). On the other hand, they are in accordance with previous studies based on a subset of the
380 present mitogenomic data (Li et al., 2011; Yang et al., 2013). Analyses performed on individual genes
381 by Tian et al (2008) and Schuh et al. (2009) indicate that the monophyly of Cimicomorpha is only
382 supported by nuclear DNA. The incongruence of phylogenetic analyses among different genomic
383 regions is a well-known issue that can have various biological causes such as incomplete lineage
384 sorting or rate variation among partitions (Som, 2014). Lin and Danforth (2004) studied the
385 differences in the pattern of nucleotide substitution among nuclear and mitochondrial genes and
386 concluded that insect phylogenetic studies should increasingly focus on nuclear data. This raises the
387 limitations of phylogenetic inferences from complete mitochondrial genomes only. While reducing
388 stochastic errors by providing large molecular datasets, this approach is susceptible to potential site-
389 specific bias and would benefit from a conjoint analysis with other genes. One advantage of the
390 genome skimming approach used in this paper is that it allows the recovery of nuclear genes of
391 phylogenetic interest in addition to the full mitogenome sequence. However, 18S and 28S data
392 available to date on Genbank are too scarce (14 species out of 55 represented in our mitochondrial
393 dataset) to perform a combined analysis.

394 **4** *Concluding remarks*

395 This study provides further evidence that NGS can be used efficiently to generate mitogenomic data
396 with a low amount of DNA. We successfully recovered the full mitogenome sequence of *Brontostoma*
397 *colossus*, which included two unusual duplicated regions. The Illumina technology used in this study
398 identified repeated elements without ambiguities, but their copy number can only be estimated
399 using the sequence coverage information. In a near future, the rapid evolution of sequencing
400 technologies (especially read length) and bioinformatics tools will probably bring improvements in
401 this regard. The increasing number of full mitochondrial genome sequences brings precious
402 phylogenetic information. However, our study highlights the limits of analyses based on
403 mitochondrial DNA only. Currently, there is a lack of correspondence between publicly available
404 mitogenomic and nuclear data. The genome skimming-approach could provide an interesting
405 improvement in this regard. In a single experimentation, it allows to recover the full mitogenome
406 sequence as well as nuclear genes that are classically used for phylogenetic inference. We argue that
407 future studies reporting full organelles sequencing with genome skimming approaches should
408 systematically report the assembly of the nuclear genes of phylogenetic interest.
409

411 Table 1: Complete or near-complete mitochondrial genomes used in this study

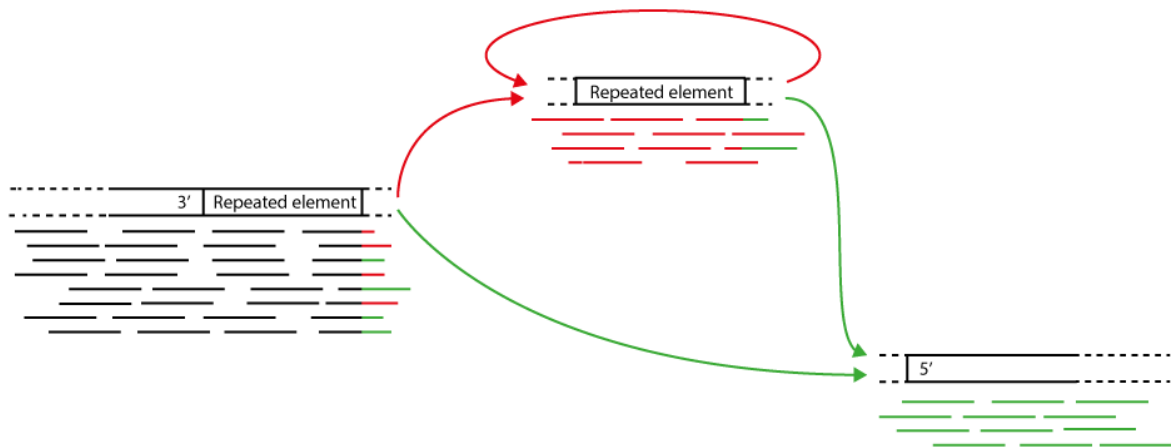
Suborder	Infra-order/ Superfamily	Family	Species	Accession num.	Reference	
Auchenorrhyncha	Fulgoroidea	Fulgoridae	<i>Lycorma delicatula</i>	EU909203	(Song et al., 2012)	
		Flatidae	<i>Geisha distinctissima</i>	NC_012617	(Song and Liang, 2009)	
Heteroptera	Cimicomorpha	Cimicoidea	<i>Orius niger</i>	EU427341	(Hua et al., 2008)	
			Miroidea	<i>Adelphocoris fasciaticollis</i>	NC_023796	(Wang et al., 2014)
	Naboidea	Anthocoridae	<i>Apolygus lucorum</i>	NC_023083	(Wang et al., 2013)	
			<i>Lygus lineolaris</i>	NC_021975	Unpublished	
			<i>Nesidiocoris tenuis</i>	NC_022677	(Dai et al., 2012)	
		Tingidae	<i>Corythucha ciliata</i>	NC_022922	(Yang et al., 2013)	
			Nabidae	<i>Alloeorhynchus bakeri</i>	HM235722	(Li et al., 2012a)
				<i>Gorpis annulatus</i>	JF907591	(Li et al., 2012a)
			<i>Gorpis humeralis</i>	JF927830	(Li et al., 2012a)	
			<i>Himacerus apterus</i>	JF927831	(Li et al., 2012a)	
			<i>Himacerus nodipes</i>	JF927832	(Li et al., 2012a)	
			<i>Nabis apicalis</i>	JF907590	(Li et al., 2012a)	
		Reduivoidea	Reduviidae	<i>Agriosphodrus dohrni</i>	NC_015842	(Li et al., 2011)
				<i>Brontostoma colossus</i>	KM044501	This study
				<i>Oncocephalus breviscutum</i>	NC_022816	(Li et al., 2013)
	<i>Sirthena flavipes</i>			HQ645959	(Gao et al., 2013)	
	<i>Triatoma dimidiata</i>			NC_002609	(Dotson and Beard, 2001)	
			<i>Valentia hoffmanni</i>	NC_012823	(Hua et al., 2009)	
		Enicocephalomorpha				
		Enicocephaloidea	Enicocephalidae	<i>Stenopirates</i> sp.	NC_016017	(Li et al., 2012b)
		Gerromorpha				
		Gerroidea	Gerridae	<i>Aquarius paludum</i>	NC_012841	(Hua et al., 2009)
		Hydrometroidea	Hydrometridae	<i>Hydrometra</i> sp.	NC_012842	(Hua et al., 2009)
		Leptopodomorpha				
		Leptopodoidea	Leptopodidae	<i>Leptopus</i> sp.	FJ456946	(Hua et al., 2009)
		Saldoidea	Saldidae	<i>Saldula arsenjevi</i>	EU427345	(Hua et al., 2008)
		Nepomorpha				
		Corixoidea	Corixidae	<i>Sigara septemlineata</i>	FJ456941	(Hua et al., 2009)
		Naucoroidea	Aphelocheiridae	<i>Aphelocheirus ellipsoideus</i>	FJ456939	(Hua et al., 2009)
	Naucoridae		<i>Ilyocoris cimicoideus</i>	NC_012845	(Hua et al., 2009)	
		Nepoidea	Belostomatidae	<i>Diplonychus rusticus</i>	FJ456940	(Hua et al., 2009)
	Nepidae		<i>Laccotrephes robustus</i>	NC_012817	(Hua et al., 2009)	
		Notonectoidea	Notonectidae	<i>Enithares tibialis</i>	NC_012819	(Hua et al., 2009)
		Ochtheroidea	Gelastocoridae	<i>Nerthra</i> sp.	NC_012838	(Hua et al., 2009)
	Ochtheridae		<i>Ochterus marginatus</i>	NC_012820	(Hua et al., 2009)	
		Pleioidea	Helotrephidae	<i>Helotrephes</i> sp.	FJ456951	(Hua et al., 2009)
		Pentatomomorpha				
		Aradoidea	Aradidae	<i>Aradacanthia heissi</i>	HQ441233	(Shi et al., 2012)
	<i>Brachyrhynchus hsiaoi</i>			NC_022670	(Li et al., 2014)	
	<i>Neuroctenus parus</i>			EU427340	(Hua et al., 2008)	
		Coreoidea	Alydidae	<i>Riptortus pedestris</i>	EU427344	(Hua et al., 2008)
	Coreidae			<i>Hydaropsis longirostris</i>	EU427337	(Hua et al., 2008)
	Rhopalidae		<i>Aeschyntelus notatus</i>	EU427333	(Hua et al., 2008)	
			<i>Stictopleurus subviridis</i>	NC_012888	(Hua et al., 2009)	
	Lygaeoidea	Berytidae	<i>Yemmalysus parallelus</i>	EU427346	(Hua et al., 2008)	
			Colobathristidae	<i>Phaenacantha marcida</i>	EU427342	(Hua et al., 2008)
		Geocoridae	<i>Geocoris pallidipennis</i>	EU427336	(Hua et al., 2008)	
Malcidae			<i>Chauliops fallax</i>	NC_020772	(Hua et al., 2008)	
<i>Malcus inconspicuus</i>			EU427339	(Hua et al., 2008)		
Pentatomoidea	Cydniidae	<i>Macroscyrtus gibbulus</i>	EU427338	(Hua et al., 2008)		
		Dinidoridae	<i>Coridius chinensis</i>	JQ739179	(Liu et al., 2012)	
	Pentatomidae	<i>Dolycoris baccarum</i>	NC_020373	(Zhang et al., 2013)		
		<i>Halyomorpha halys</i>	NC_013272	(Lee et al., 2009)		
		<i>Nezara viridula</i>	NC_011755	(Hua et al., 2008)		
		Plataspidae	<i>Coptosoma bifaria</i>	EU427334	(Hua et al., 2008)	
		<i>Megacopta cribraria</i>	NC_015842	(Hua et al., 2008)		
	Tessaratomidae	<i>Eusthenes cupreus</i>	NC_022449	(Song et al., 2013)		
	Urostylididae	<i>Urochela quadrinotata</i>	NC_020144	(Yuting et al., 2012)		
Pyrrhocoroidea	Largidae	<i>Physopelta gutta</i>	EU427343	(Hua et al., 2008)		
		Pyrrhocoridae	<i>Dysdercus cingulatus</i>	EU427335	(Hua et al., 2008)	

413 **Table 2: Summary of the mitochondrial genome of *Brontostoma colossus***

Locus	Direction	Location (bp)	Size (bp)	Anticodon	Start codon	Stop codon	Interlocus nucleotides
tRNA ^{Ile}	F	1-63	63	GAT			0
tRNA ^{Gln}	R	66-134	69	TTG			2
tRNA ^{Met}	F	133-202	70	CAT			-2
ND2	F	203-1193	991		ATT	T--	0
tRNA ^{Trp}	F	1194-1260	67	TCA			0
tRNA ^{Cys}	R	1253-1317	65	GCA			-8
tRNA ^{Tyr}	R	1319-1382	64	GTA			1
COX1	F	1384-2917	1534		ATG	T--	1
tRNA ^{Leu(UUR)}	F	2918-2982	65	TAA			0
COX2	F	2983-3661	679		ATT	T--	0
tRNA ^{Lys}	F	3662-3733	72	CTT			0
tRNA ^{Asp}	F	3734-3796	63	GTC			0
ATP8	F	3797-3955	159		ATT	TAA	0
ATP6	F	3949-4620	672		ATG	TAA	-7
COX3	F	4620-5408	789		ATG	TAA	-1
tRNA ^{Gly}	F	5413-5477	65	TCC			4
ND3	F	5478-5831	354		ATA	TAA	0
tRNA ^{Arg}	F	5873-5937	65	TCG			41
tRNA ^{Ala}	F	5984-6044	61	TGC			46
tRNA ^{Arg}	F	6049-6113	65	TCG			4
tRNA ^{Asn}	F	6118-6187	70	GTT			4
tRNA ^{Ser(AGN)}	F	6187-6255	69	GCT			-1
tRNA ^{Glu}	F	6258-6319	62	TTC			2
tRNA ^{Phe}	R	6318-6388	71	GAA			2
ND5	R	6389-8094	1706		ATT	TA-	0
tRNA ^{His}	R	8095-8158	64	GTG			0
ND4	R	8159-9493	1335		ATG	TAA	0
ND4L	R	9487-9771	285		ATT	TAA	-7
tRNA ^{Thr}	F	9774-9837	64	TGT			2
tRNA ^{Pro}	R	9838-9903	66	TGG			0
ND6	F	9906-10397	492		ATA	TAA	2
CYTB	F	10397-11529	1133		ATG	TA-	-1
tRNA ^{Ser(TCN)}	F	11530-11598	69	TGA			0
ND1	R	11637-12548	912		ATT	TAA	38
tRNA ^{Leu(CUN)}	R	12549-12613	65	TAG			0
lrrRNA	R	12614-13864	1251				0
tRNA ^{Val}	R	13865-13935	71	TAC			0
srRNA	R	13936-14728	793				0
Control region		14729-16625	1896				0

414

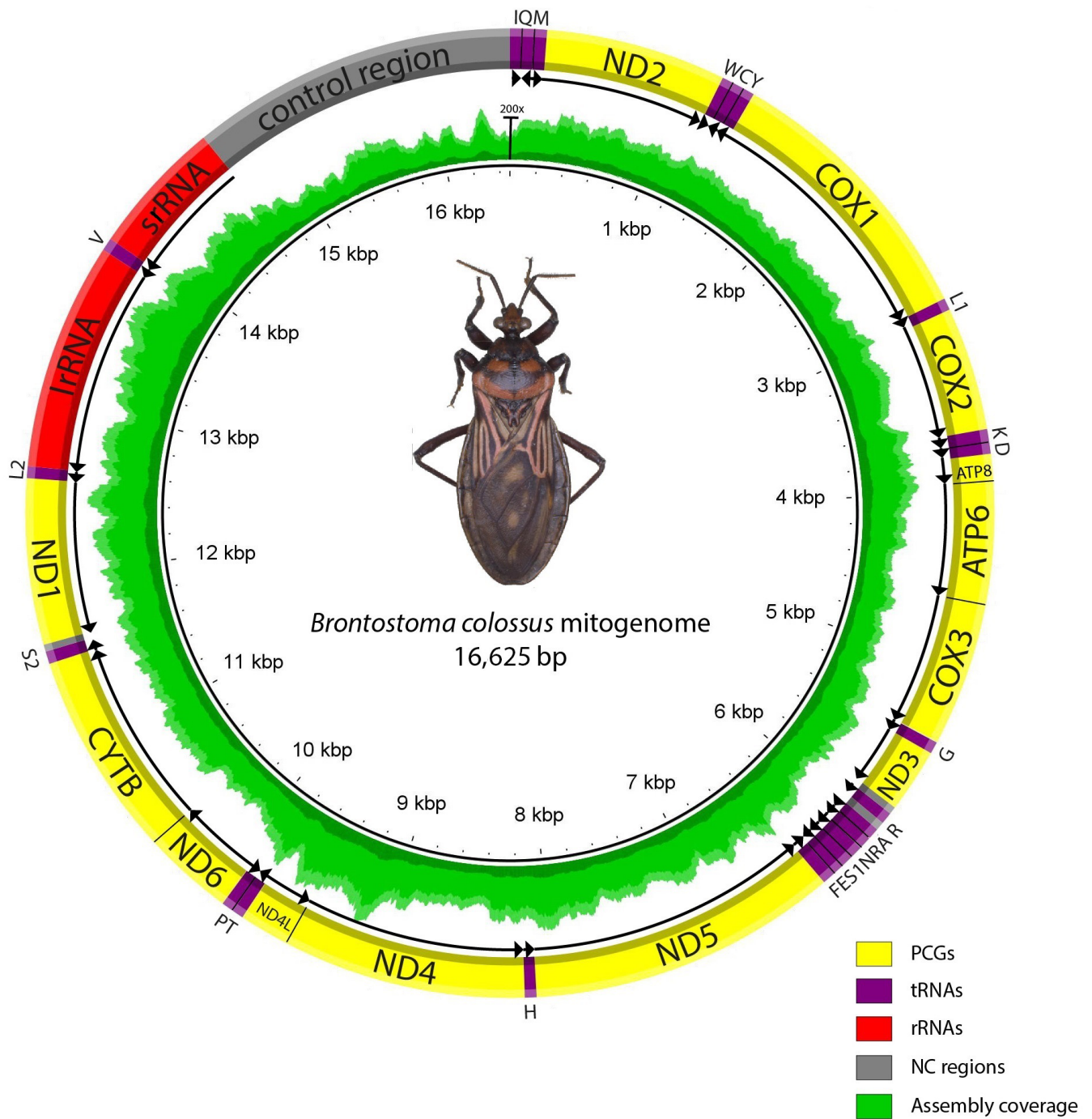
415



417

418 **Figure 1 : Schematization of an assembly ambiguity revealing a repeated element in the sequence. When assembling the**
419 **5' end of the repeated element, two different "assembly paths" are supported by a significant number of reads: (i) a path**
420 **leading to the beginning of a novel repeated element (depicted in red) (ii) a path leading to the region flanking the 5' end**
421 **of the last repetition (depicted in green).**

422

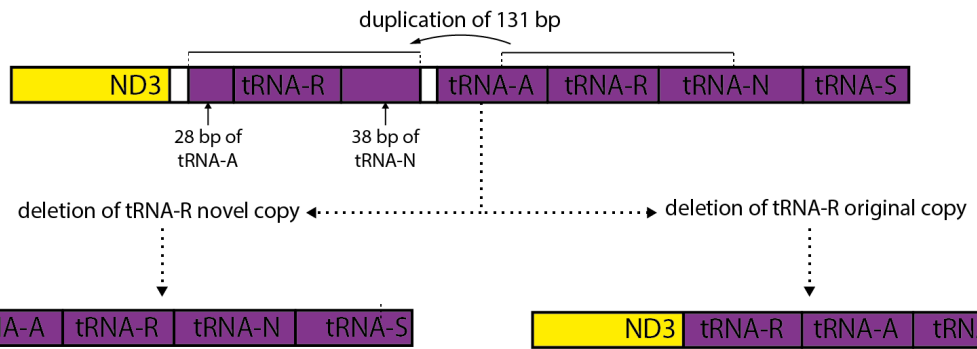


424

425 **Figure 2: Schematic representation of *Brontostoma colossus* mitogenome. tRNAs are labelled according to the IUPAC-IUB**
 426 **single-letter amino acid codes. Arrows indicate directions of genes. The scale on the assembly coverage ring indicates**
 427 **200x coverage.**

428

429



430

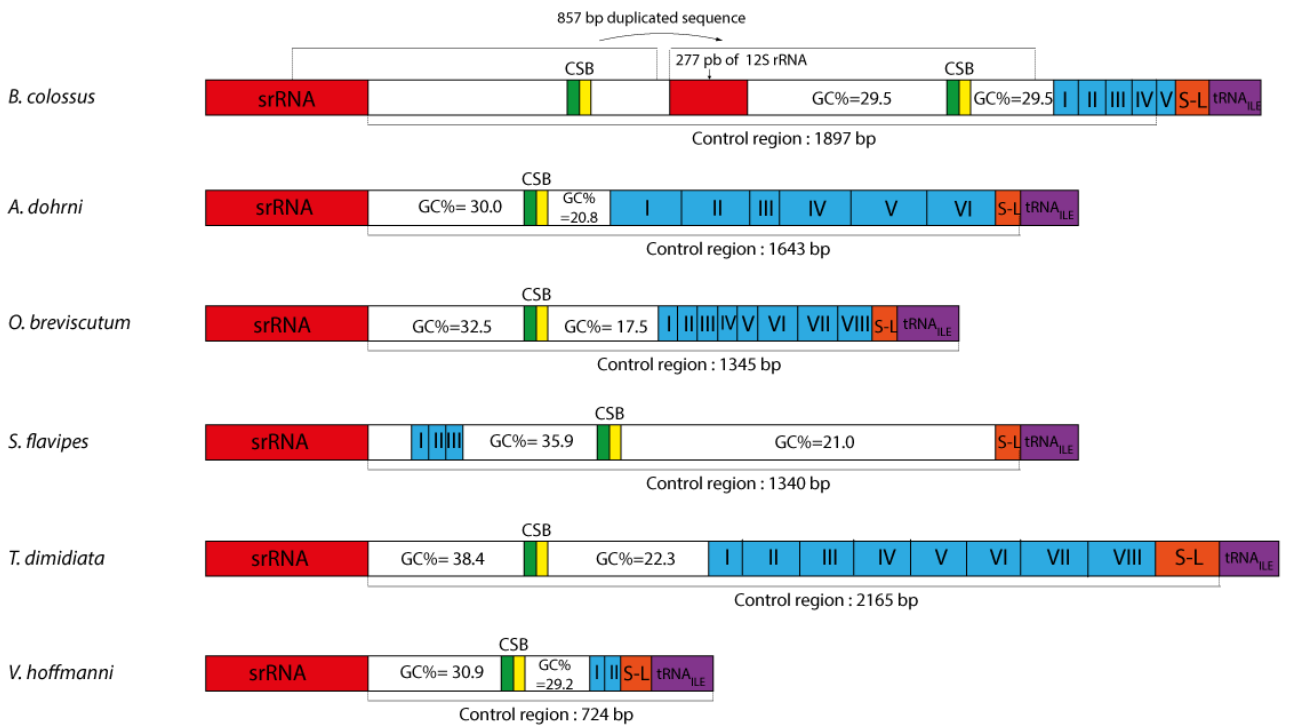
431 **Figure 3: Organization of the 131 bp duplicated sequence comprising tRNA-Arg and potential gene rearrangements after**
 432 **random deletion of one tRNA-arg copy.**

433

A

B. colossus CCCTTAAATAGTCTCATACCCGGACAAGCATGGGGGGGTGTGGGG
A. dohrni CCTGAATAAGTITATACCCGGACAGACTTGGGGGTG--GGGGT.
O. breviscutum CCCTAAATAGTATCATACCCGGATAGGATGGGGGGG--GGGT
S. flavipes TCCCTAAATAGTCTCATACCCGGAGA---TGGGGGGG--GGGT
T. dimidiata CCTTAGATRGTCTCATACCCGGATAGGATGGGGGGG--GGGG
V. hoffmanni TCCAAAATAAGTITCATACCCGGATACGGTGGGGGGTCCCGG--

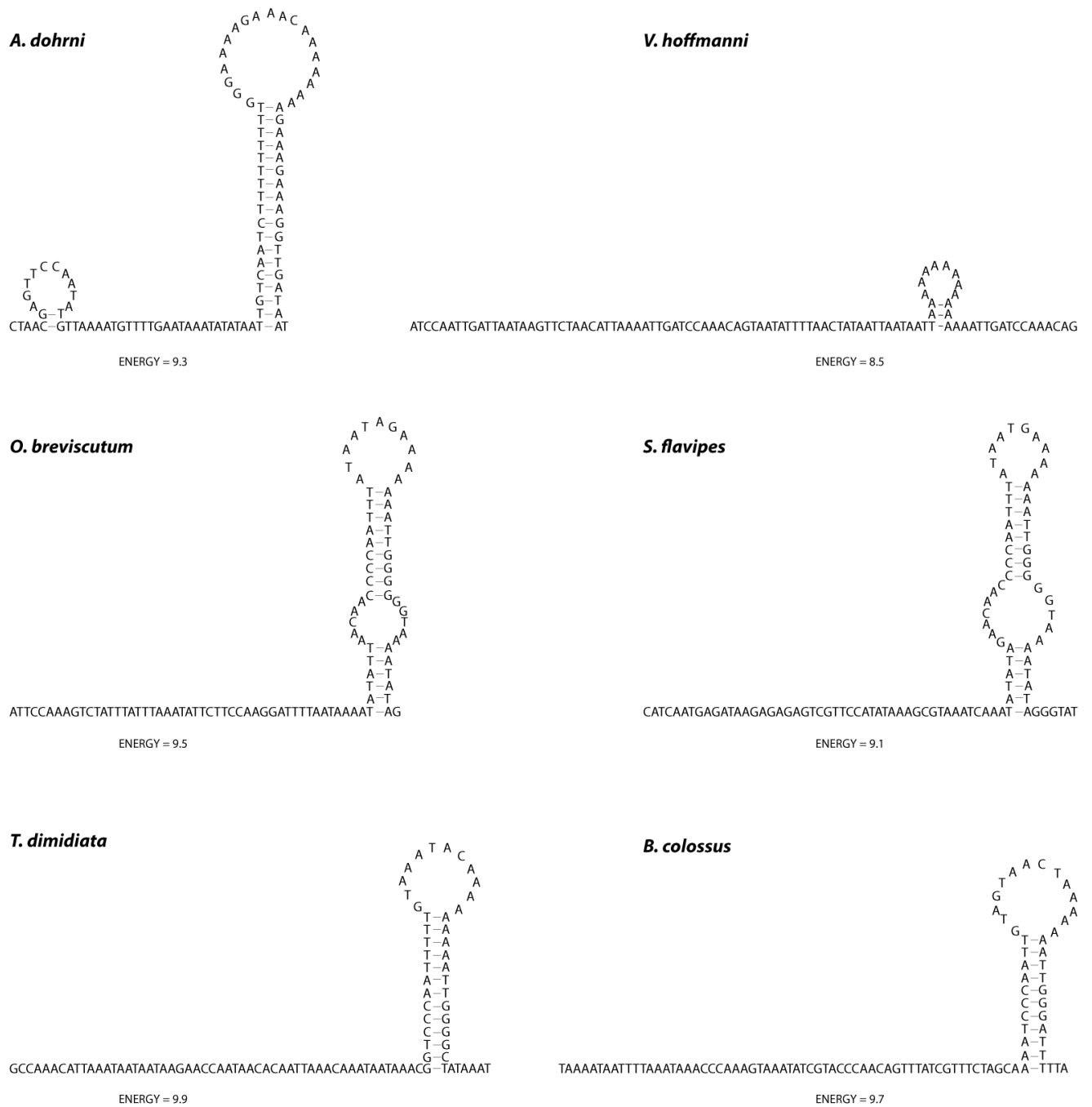
B



434

435 **Figure 4: (A)** Alignment of the conserved sequence blocks identified in the mitochondrial control region of the six assassin
 436 bugs. Nucleotide positions that are conserved among the six assassin bugs are highlighted. **(B)** Structural organization of
 437 the mitochondrial control region of the six assassin bugs. The blue boxes with roman numerals indicate tandem repeat
 438 units. The CSB box indicates the conserved sequence block, and the yellow part indicates the "G element". The orange
 439 box "S-L" indicates the region where potential stem loops are found. GC content is shown in the white boxes previously
 440 described as GC-rich and AT-rich regions.

441



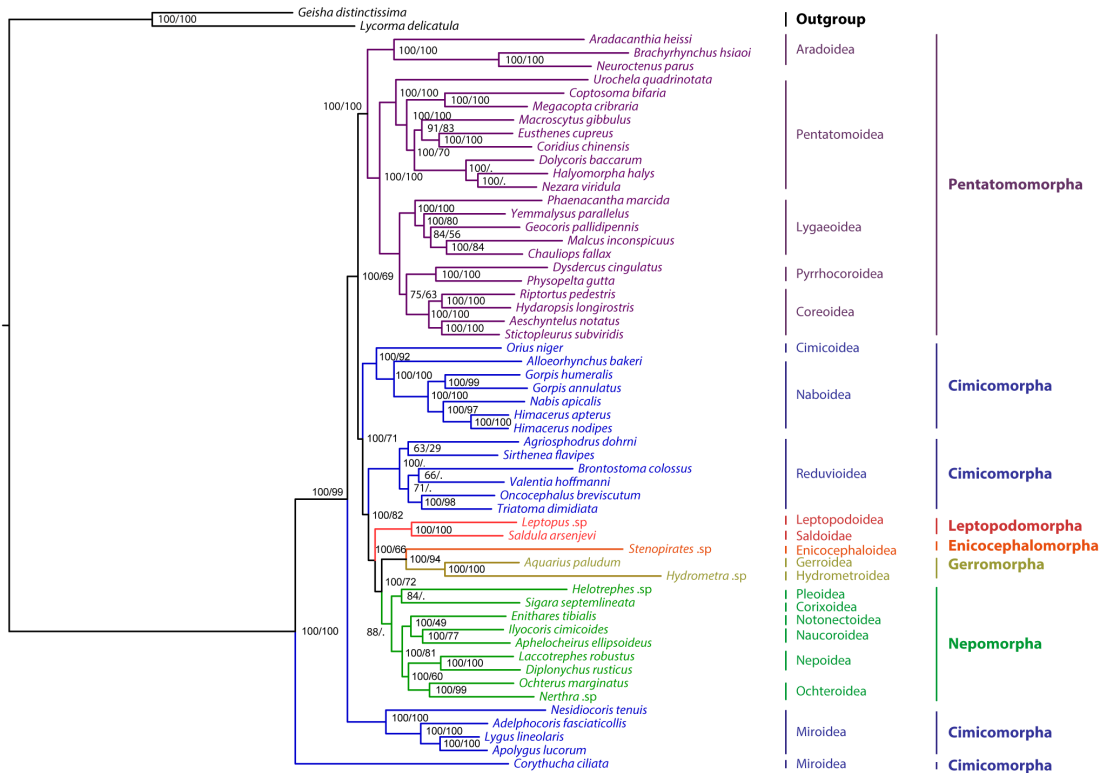
442

443 **Figure 5: Conserved stem loop structures inferred by the TurboFold algorithm on the the 100 bp of the control region**

444 **flanking tRNA_{Ile} of the six assassin bugs. Values of free energy are indicated above each structure.**

445

A



B



446

447

448

449

450

451

452

Figure 6: (A) Phylogenetic tree inferred by Bayesian analysis from 55 heteropteran mitogenomes. Left numbers at the nodes indicate Bayesian posterior probabilities expressed in percentages. When the node was also present on the tree inferred by ML analysis, right numbers indicate bootstrap support values. A dot indicates that the node was absent on the tree inferred by ML analysis (B) Maximum likelihood (ML) sub-tree corresponding to Reduviidae family, with bootstrap values depicted on nodes.

453 **Acknowledgements**

454 This work was supported by PO-FEDER TIMGED N°30195 (Trypanosomes d'Intérêt Médical en
455 Guyane française – Epidémiologie et Diagnostic) and “Investissement d'Avenir” grants managed by
456 Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; TULIP, ANR-10-LABX-41, ANR-11-
457 IDEX-0002-02) as well as project METABAR (ANR-11-BSV7-0020). We would like to thank Christine
458 Aznar, Denis Blanchet, Jean-Pierre Dujardin and Jean-Michel Bérenger for the help with obtaining
459 and identifying the specimen. We are grateful to the Genotoul bioinformatics platform Toulouse
460 Midi-Pyrenees for providing computing and storage resources.

461 **References**

- 462 Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences
463 guided by amino acid translations. *Nucleic Acids Res.* 38, W7–13.
- 464 Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C.,
465 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and
466 systematics. *Annu. Rev. Ecol. Syst.* 18, 489–522.
- 467 Ballard, J.W.O., 2000. Comparative genomics of mitochondrial DNA in members of the *Drosophila*
468 *melanogaster* subgroup. *J. Mol. Evol.* 51, 48–63.
- 469 Bellaousov, S., Reuter, J.S., Seetin, M.G., Mathews, D.H., 2013. RNAstructure: web servers for RNA
470 secondary structure prediction and analysis. *Nucleic Acids Res.* 41, W471–W474.
- 471 Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.*
472 27, 573.
- 473 Bernt, M., Braband, A., Schierwater, B., Stadler, P.F., 2013a. Genetic aspects of mitochondrial
474 genome evolution. *Mol. Phylogenet. Evol.* 69, 328–338.
- 475 Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsche, G., Pütz, J., Middendorf, M.,
476 Stadler, P.F., 2013b. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Mol.*
477 *Phylogenet. Evol.* 69, 313–319.
- 478 Besnard, G., Christin, P.-A., Malé, P.-J.G., Coissac, E., Ralimanana, H., Vorontsova, M.S., 2013.
479 Phylogenomics and taxonomy of Lecomtelleae (Poaceae), an isolated panicoid lineage from
480 Madagascar. *Ann. Bot.* 112, 1057–1066.
- 481 Besnard, G., Jühling, F., Chapuis, É., Zedane, L., Lhuillier, É., Mateille, T., Bellafiore, S., 2014. Fast
482 assembly of the mitochondrial genome of a plant parasitic nematode (*Meloidogyne graminicola*)
483 using next generation sequencing. *C. R. Biol.* 337, 295–301.
- 484 Boore, J.L., 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27, 1767–1780.
- 485 Boore, J.L., 2000. The duplication/random loss model for gene rearrangement exemplified by
486 mitochondrial genomes of deuterostome animals. In: Sankoff, D. and Nadeau, J.H. eds. *Comparative*
487 *Genomics*. Springer Netherlands, pp. 133–147.
- 488 Boore, J.L., Collins, T.M., Stanton, D., Daehler, L.L., Brown, W.M., 1995. Deducing the pattern of
489 arthropod phylogeny from mitochondrial-DNA rearrangements. *Nature* 376, 163-165.

490 Boore, J.L., Lavrov, D.V., Brown, W.M., 1998. Gene translocation links insects and crustaceans.
491 Nature 392, 667–668.

492 Clary, D.O., Wolstenholme, D.R., 1985. The mitochondrial DNA molecule of *Drosophila yakuba*:
493 nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22, 252–271.

494 Clayton, D.A., 1992. Transcription and replication of animal mitochondrial DNAs. Int Rev Cytol 141,
495 217–232.

496 Crozier, R.H., Crozier, Y.C., 1993. The mitochondrial genome of the honeybee *Apis mellifera*:
497 complete sequence and genome organization. Genetics 133, 97–117.

498 Dai, X., Xun, H., Chang, J., Zhang, J., Hu, B., Li, H., Yuan, X., Cai, W., 2012. The complete mitochondrial
499 genome of the plant bug *Nesidiocoris tenuis* (Reuter) (Hemiptera: Miridae: Bryocorinae: Dicyphini).
500 Zootaxa (3554), 30–44.

501 De Bruijn, M.H., 1983. *Drosophila melanogaster* mitochondrial DNA, a novel organization and genetic
502 code. Nature 304, 234–241.

503 Dotson, E.M., Beard, C.B., 2001. Sequence and organization of the mitochondrial genome of the
504 Chagas disease vector, *Triatoma dimidiata*. Insect Mol. Biol. 10, 205–215.

505 Dougherty, V., 1995. A review of the new world Ectrichodiinae genera (Hemiptera: Reduviidae).
506 Trans. Am. Entomol. Soc. 121, 173–225.

507 Doyle, S.R., Griffith, I.S., Murphy, N.P., Strugnell, J.M., 2014. Low-coverage MiSeq next generation
508 sequencing reveals the mitochondrial genome of the Eastern Rock Lobster, *Sagmariasus verreauxi*.
509 Mitochondrial DNA 0, 1–2.

510 Gao, J., Li, H., Truong, X.L., Dai, X., Chang, J., Cai, W., 2013. Complete nucleotide sequence and
511 organization of the mitochondrial genome of *Sirthenea flavipes* (Hemiptera: Reduviidae: Peiratinae)
512 and comparison with other assassin bugs. Zootaxa 3669, 1–16.

513 Garesse, R., 1988. *Drosophila melanogaster* mitochondrial DNA: gene organization and evolutionary
514 considerations. Genetics 118, 649–663.

515 Gil-Santana, H., Baena, M., 2009. Two new species of *Brontostoma* Kirkaldy (Hemiptera: Heteroptera:
516 Reduviidae: Ectrichodiinae) from Bolivia, with description of the male genitalia of two other species
517 of the genus, and description of the female of *B. doughertyae* Gil-Santana, Lopes, Marques &
518 Jurberg. Zootaxa 1979, 41–52.

519 Gissi, C., Iannelli, F., Pesole, G., 2008. Evolution of the mitochondrial genome of Metazoa as
520 exemplified by comparison of congeneric species. Heredity 101, 301–320.

521 Hahn C, Bachmann L, Chevreux B, 2013. Reconstructing mitochondrial genomes directly from
522 genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids
523 Research, 41, e129.

524 Harmanci, A.O., Sharma, G., Mathews, D.H., 2011. TurboFold: Iterative probabilistic estimation of
525 secondary structures for multiple RNA sequences. BMC Bioinformatics 12, 108.

526 Hebsgaard, M.B., Andersen, N.M., Damgaard, J., 2004. Phylogeny of the true water bugs
527 (Nepomorpha: Hemiptera–Heteroptera) based on 16S and 28S rDNA and morphology. Syst. Entomol.
528 29, 488–508.

529 Hua, J., Li, M., Dong, P., Cui, Y., Xie, Q., Bu, W., 2008. Comparative and phylogenomic studies on the
530 mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). BMC Genomics 9,
531 610.

532 Hua, J., Li, M., Dong, P., Cui, Y., Xie, Q., Bu, W., 2009. Phylogenetic analysis of the true water bugs
533 (Insecta: Hemiptera: Heteroptera: Nepomorpha): evidence from mitochondrial genomes. BMC Evol.
534 Biol. 9, 134.

535 Junqueira, A.C.M., Lessinger, A.C., Torres, T.T., da Silva, F.R., Vettore, A.L., Arruda, P., Azeredo Espin,
536 A.M.L., 2004. The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera:
537 Calliphoridae). Gene 339, 7–15.

538 Kim, I., Lee, E.M., Seol, K.Y., Yun, E.Y., Lee, Y.B., Hwang, J.S., Jin, B.R., 2006. The mitochondrial
539 genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). Insect Mol. Biol. 15,
540 217–225.

541 Kopp, A., True, J.R., 2002. Phylogeny of the Oriental *Drosophila melanogaster* species group: a
542 multilocus reconstruction. Syst. Biol. 51, 786–805.

543 Kück, P., Meusemann, K., 2010. FASconCAT: Convenient handling of data matrices. Mol. Phylogenet.
544 Evol. 56, 1115–1118.

545 Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. Partitionfinder: combined selection of
546 partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–
547 1701.

548 Lavrov, D.V., Boore, J.L., Brown, W.M., 2000. The complete mitochondrial DNA sequence of the
549 horseshoe crab *Limulus polyphemus*. Mol. Biol. Evol. 17, 813–824.

550 Lavrov, D.V., Lang, B.F., 2005. Transfer RNA gene recruitment in mitochondrial DNA. Trends Genet.
551 21, 129–133.

552 Lee, W.-J., Conroy, J., Howell, W.H., Kocher, T.D., 1995. Structure and evolution of teleost
553 mitochondrial control regions. J. Mol. Evol. 41, 54–66.

554 Lee, E.-S., Shin, K.S., Kim, M.-S., Park, H., Cho, S., Kim, C.-B., 2006. The mitochondrial genome of the
555 smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae). Gene 373, 52–57.

556 Lee, W., Kang, J., Jung, C., Hoelmer, K., Lee, S.H., Lee, S., 2009. Complete mitochondrial genome of
557 brown marmorated stink bug *Halyomorpha halys* (Hemiptera: Pentatomidae), and phylogenetic
558 relationships of hemipteran suborders. Mol. Cells 28, 155–165.

559 Lent, H., Wygodzinsky, P.W., 1979. Revision of the Triatominae (Hemiptera, Reduviidae), and their
560 significance as vectors of Chagas' disease. Bulletin of the AMNH; v. 163, article 3.

561 Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform.
562 Bioinformatics 25, 1754–1760.

563 Li, H., Gao, J., Cai, W., 2013. Complete mitochondrial genome of the assassin bug *Oncocephalus*
564 *breviscutum* (Hemiptera: Reduviidae). Mitochondrial DNA 0,1–2.

565 Li, H., Gao, J., Liu, H., Liu, H., Liang, A., Zhou, X., Cai, W., 2011. The architecture and complete
566 sequence of mitochondrial genome of an assassin bug *Agriosphodrus dohrni* (Hemiptera:
567 Reduviidae). Int. J. Biol. Sci. 7, 792.

- 568 Li, H., Liu, H., Song, F., Shi, A., Zhou, X., Cai, W., 2012a. Comparative mitogenomic analysis of damsel
569 bugs representing three tribes in the family Nabidae (Insecta: Hemiptera). PLoS One 7, e45925.
- 570 Li, H., Liu, H., Shi, A., Štys, P., Zhou, X., Cai, W., 2012b. The complete mitochondrial genome and novel
571 gene arrangement of the unique-headed bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). PLoS
572 One 7, e29419.
- 573 Li, M., Tian, Y., Zhao, Y., Bu, W., 2012c. Higher Level Phylogeny and the First Divergence Time
574 Estimation of Heteroptera (Insecta: Hemiptera) Based on Multiple Genes. PLoS ONE 7, e32152.
- 575 Li, H., Shi, A., Song, F., Cai, W., 2014. Complete mitochondrial genome of the flat bug *Brachyrhynchus*
576 *hsiao* (Hemiptera: Aradidae). Mitochondrial DNA 0,1–2.
- 577 Lin, C.-P., Danforth, B.N., 2004. How do insect nuclear and mitochondrial gene substitution patterns
578 differ? Insights from Bayesian analyses of combined datasets. Mol. Phylogenet. Evol. 30,
- 579 Liu, L., Li, H., Song, F., Song, W., Dai, X., Chang, J., Cai, W., 2012. The mitochondrial genome of
580 *Coridius chinensis* (Hemiptera: Dinidoridae). Zootaxa 3537, 29–40.
- 581 Macey, J.R., Schulte, J.A., Larson, A., Papenfuss, T.J., 1998. Tandem duplication via light-strand
582 synthesis may provide a precursor for mitochondrial genomic rearrangement. Mol. Biol. Evol. 15, 71–
583 75.
- 584 Mahner, M., 1993. Systema Cryptoceratorum Phylogenticum (Insecta, Heteroptera). Stuttgart,
585 Germany: Eds. Schweizerbart.
- 586 Maldonado Capriles, J., 1990. Systematic catalogue of the Reduviidae of the world (Insecta:
587 Heteroptera). Caribb. J. Sci. Spec. Ed.
- 588 Malé, P.-J.G., Bardon, L., Besnard, G., Coissac, E., Delsuc, F., Engel, J., Lhuillier, E., Scotti-Saintagne, C.,
589 Tinaut, A., Chave, J., 2014. Genome skimming by shotgun sequencing helps resolve the phylogeny of
590 a pantropical tree family. Mol. Ecol. Resour.
- 591 Mariette, J., Escudí, F., Allias, N., Salin, G., Noirot, C., Thomas, S., Klopp, C., 2012. NG6: Integrated
592 next generation sequencing storage and processing environment. BMC Genomics 13, 462.
- 593 Moritz, C., Brown, W.M., 1986. Tandem duplication of D-loop and ribosomal RNA sequences in lizard
594 mitochondrial DNA. Science 233, 1425–1427.
- 595 Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA: relevance for
596 population biology and systematics. Annu. Rev. Ecol. Syst. 269–292.
- 597 Nagaike, T., 2005. Human Mitochondrial mRNAs Are Stabilized with Polyadenylation Regulated by
598 Mitochondria-specific Poly(A) Polymerase and Polynucleotide Phosphorylase. J. Biol. Chem. 280,
599 19721–19727.
- 600 Nguyen, V.H., Lavenier, D., 2009. PLAST: parallel local alignment search tool for database
601 comparison. BMC Bioinformatics 10, 329.
- 602 Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human
603 mitochondria. Nature 290, 470-474.
- 604 Oliveira, M.T., Barau, J.G., Junqueira, A.C.M., Feijão, P.C., Rosa, A.C. da, Abreu, C.F., Azeredo-Espin,
605 A.M.L., Lessinger, A.C., 2008. Structure and evolution of the mitochondrial genomes of *Haematobia*

606 *irritans* and *Stomoxys calcitrans*: the Muscidae (Diptera: Calyptratae) perspective. Mol. Phylogenet.
607 Evol. 48, 850–857.

608 Perna, N.T., Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of
609 animal mitochondrial genomes. J. Mol. Evol. 41, 353–358.

610 Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed
611 models. Bioinformatics 19, 1572–1574.

612 Saito, S., 2005. Replication origin of mitochondrial DNA in insects. Genetics 171, 1695–1705.

613 Salvato, P., Simonato, M., Battisti, A., Negrisolò, E., 2008. The complete mitochondrial genome of the
614 bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). BMC Genomics 9, 331.

615 Schuh, R.T., Weirauch, C., Wheeler, W.C., 2009. Phylogenetic relationships within the Cimicomorpha
616 (Hemiptera: Heteroptera): a total-evidence analysis. Syst. Entomol. 34, 15–48.

617 Shao, R., Barker, S.C., 2003. The highly rearranged mitochondrial genome of the plague thrips, *Thrips*
618 *imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary
619 arrangement of rRNA genes. Mol. Biol. Evol. 20, 362–370.

620 Shao, R., Barker, S.C., Mitani, H., Takahashi, M., Fukunaga, M., 2006. Molecular mechanisms for the
621 variation of mitochondrial gene content and gene arrangement among chigger mites of the genus
622 *Leptotrombidium* (Acari: Acariformes). J. Mol. Evol. 63, 251–261.

623 Shi, A.M., Li, H., Bai, X.S., Dai, X., Chang, J., Guilbert, E., Cai, W., 2012. The complete mitochondrial
624 genome of the flat bug *Aradacanthia heissi* (Hemiptera: Aradidae). Zootaxa 3238, 23–38.

625 Signorovitch, A.Y., Buss, L.W., Dellaporta, S.L., 2007. Comparative Genomics of Large Mitochondria in
626 Placozoans. PLoS Genet. 3, e13.

627 Som, A., 2014. Causes, consequences and solutions of phylogenetic incongruence. Brief. Bioinform.
628 doi:10.1093/bib/bbu015

629 Song, N., Liang, A.-P., 2009. Complete Mitochondrial Genome of the Small Brown Planthopper,
630 *Laodelphax striatellus* (Delphacidae: Hemiptera), with a Novel Gene Order. Zoolog. Sci. 26, 851–860.

631 Song, N., Liang, A.-P., Bu, C.-P., 2012. A molecular phylogeny of Hemiptera inferred from
632 mitochondrial genome sequences. PloS One 7, e48778.

633 Song, W., Li, H., Song, F., Liu, L., Wang, P., Xun, H., Dai, X., Chang, J., Cai, W., 2013. The complete
634 mitochondrial genome of a tessaratomid bug, *Eusthenes cupreus* (Hemiptera: Heteroptera:
635 Pentatomomorpha: Tessaratomidae). Zootaxa 3620, 260–272.

636 Stamatakis, A., 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large
637 Phylogenies. Bioinformatics btu033.

638 Stewart, J.B., Beckenbach, A.T., 2009. Characterization of mature mitochondrial transcripts in
639 *Drosophila*, and the implications for the tRNA punctuation model in arthropods. Gene 445, 49–57.

640 Straub, S.C., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C., Liston, A., 2012. Navigating the tip of
641 the genomic iceberg: Next-generation sequencing for plant systematics. Am. J. Bot. 99, 349–364.

642 Tamura, K., Stecher, G., Peterson, D., Filipowski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary
643 Genetics Analysis Version 6.0. Mol. Biol. Evol. 30, 2725–2729.

- 644 Thao, M.L., Baumann, L., Baumann, P., 2004. Organization of the mitochondrial genomes of
645 whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). *BMC Evol. Biol.* 4, 25.
- 646 Thompson, K.F., Patel, S., Williams, L., Tsai, P., Constantine, R., Baker, C.S., Millar, C.D., 2014. High
647 coverage of the complete mitochondrial genome of the rare Gray's beaked whale (*Mesoplodon grayi*)
648 using Illumina next generation sequencing. *Mitochondrial DNA* 0,1–2.
- 649 Tian, Y., Zhu, W., Li, M., Xie, Q., Bu, W., 2008. Influence of data conflict and molecular phylogeny of
650 major clades in Cimicomorphan true bugs (Insecta: Hemiptera: Heteroptera). *Mol. Phylogenet. Evol.*
651 47, 581–597.
- 652 Veale, A.J., Williams, L., Tsai, P., Thakur, V., Lavery, S., 2014. The complete mitochondrial genomes of
653 two chiton species (*Sypharochiton pelliserpentis* and *Sypharochiton sinclairi*) obtained using Illumina
654 next generation sequencing. *Mitochondrial DNA* 0,1–2.
- 655 Walberg, M.W., Clayton, D.A., 1981. Sequence and properties of the human KB cell and mouse L cell
656 D-loop regions of mitochondrial DNA. *Nucleic Acids Res.* 9, 5411–5421.
- 657 Wang, P., Li, H., Wang, Y., Zhang, J.-H., Dai, X., Chang, J., Hu, B.-W., Cai, W.-Z., 2013. The
658 mitochondrial genome of the plant bug *Apolygus lucorum* (Hemiptera: Miridae): Presently known as
659 the smallest in Heteroptera. *Insect Sci.* 21, 159-173.
- 660 Wang, Y., Li, H., Xun, H., Cai, W., 2014. Complete mitochondrial genome sequence of the plant bug
661 *Adelphocoris fasciaticollis* (Hemiptera: Heteroptera: Miridae). *Mitochondrial DNA* 0,1–2.
- 662 Weirauch, C., 2008. Cladistic analysis of Reduviidae (Heteroptera: Cimicomorpha) based on
663 morphological characters. *Syst. Entomol.* 33, 229–274.
- 664 Weirauch, C., Munro, J.B., 2009. Molecular phylogeny of the assassin bugs (Hemiptera: Reduviidae),
665 based on mitochondrial and nuclear ribosomal genes. *Mol. Phylogenet. Evol.* 53, 287–299.
- 666 Weirauch, C., Schuh, R.T., 2011. Systematics and Evolution of Heteroptera: 25 Years of Progress.
667 *Annu. Rev. Entomol.* 56, 487–510.
- 668 Wheeler, A.G., 1997. Zoological Catalogue of Australia 27.3 A: Hemiptera: Heteroptera
669 (Coleorrhyncha to Cimicomorpha). *Ann. Entomol. Soc. Am.* 90, 703–704.
- 670 Wheeler, W.C., Schuh, R.T., Bang, R., 1993. Cladistic relationships among higher groups of
671 Heteroptera: congruence between morphological and molecular data sets. *Insect Syst. Evol.* 24, 121–
672 137.
- 673 Wolstenholme, D.R., 1992. Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* 141,
674 173–216.
- 675 Xie, Q., Tian, Y., Zheng, L., Bu, W., 2008. 18S rRNA hyper-elongation and the phylogeny of
676 Euhemiptera (Insecta: Hemiptera). *Mol. Phylogenet. Evol.* 47, 463–471.
- 677 Yang, W., Yu, W., Du, Y., 2013. The complete mitochondrial genome of the sycamore lace bug
678 *Corythucha ciliata* (Hemiptera: Tingidae). *Gene* 532, 27–40.
- 679 Yu, S., Wang, Y., Rédei, D., Xie, Q., Bu, W., 2013. Secondary structure models of 18S and 28S rRNAs of
680 the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871
681 (Heteroptera, Pentatomidae). *ZooKeys* 319, 363–377.

682 Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K., Banno, Y., 2002. Significant levels of sequence
683 divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild
684 Mulberry silkworm, *Bombyx mandarina*, and its close relative, the domesticated silkworm, *Bombyx*
685 *mori*. Mol. Biol. Evol. 19, 1385–1389.

686 Yuting, D., Li, Jiang, P., Song, F., Ye, Z., Yuang, X., Dai, X., Chang, J., Cai, W., 2012. Sequence and
687 organization of the mitochondrial genome of an urostylidid bug, *Urochela quadrinotata* Reuter
688 (Hemiptera: Urostylididae). Entomotaxonomia 34, 613–623.

689 Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn
690 graphs. Genome Res. 18, 821–829.

691 Zhang, D.-X., Hewitt, G.M., 1997. Insect mitochondrial control region: a review of its structure,
692 evolution and usefulness in evolutionary studies. Biochem. Syst. Ecol. 25, 99–120.

693 Zhang, D.-X., Szymura, J.M., Hewitt, G.M., 1995. Evolution and structural conservation of the control
694 region of insect mitochondrial DNA. J. Mol. Evol. 40, 382–391.

695 Zhang, Q.-L., Yuan, M.-L., Shen, Y.-Y., 2013. The complete mitochondrial genome of *Dolycoris*
696 *baccarum* (Insecta: Hemiptera: Pentatomidae). Mitochondrial DNA 24, 469–471.

697 Zhou, Z., Huang, Y., Shi, F., 2007. The mitochondrial genome of *Ruspolia dubia* (Orthoptera:
698 Conocephalidae) contains a short A+T-rich region of 70 bp in length. Genome Natl. Res. Counc. Can.
699 50, 855–866.

700