

SCIENTIFIC REPORTS



OPEN

Characterization of the complete mitochondrial genomes of *Maiestas dorsalis* and *Japananus hyalinus* (Hemiptera: Cicadellidae) and comparison with other Membracoidea

Yimin Du¹, Chunni Zhang¹, Christopher H. Dietrich², Yalin Zhang¹ & Wu Dai¹

Only six mitochondrial genomes (mitogenomes) have been previously published for Cicadellidae, the largest family of Hemiptera. This study provides complete, annotated mitogenomes of two additional cicadellid, species *Maiestas dorsalis* and *Japananus hyalinus*, and the first comparative mitogenome analysis across the superfamily Membracoidea. The mitogenomes of both sequenced species are similar to those of other studied hemipteran mitogenomes in organization and the lengths are 15,352 and 15,364 bp with an A + T content of 78.7% and 76.6%, respectively. In *M. dorsalis*, all sequenced genes are arranged in the putative ancestral insect gene arrangement, while the tRNA cluster *trnW-trnC-trnY* is rearranged to *trnY-trnW-trnC* in *J. hyalinus*, the first reported gene rearrangement in Membracoidea. Phylogenetic analyses of the 11 available membracoid mitogenomes and outgroups representing the other two cicadomorphan superfamilies supported the monophyly of Membracoidea, and indicated that treehoppers are a derived lineage of leafhoppers. ML and BI analyses yielded topologies that were congruent except for relationships among included representatives of subfamily Deltocephalinae. Exclusion of third codon positions of PCGs improved some node support values in ML analyses.

Insect mitochondrial genomes (mitogenomes) are typically double-stranded circular molecules with 37 genes, including 13 protein-coding genes (PCGs), two ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs)¹. Complete mitochondrial genome sequences are not only more phylogenetically informative than shorter sequences of individual genes, but also provide sets of genome-level characters, such as the relative positions of different genes, RNA secondary structures and modes of control of replication and transcription. Because of the abundance of mitochondria in cells, maternal inheritance, absence of introns, and high evolutionary rates, insect mitochondrial genome sequences are the most extensively used genomic marker(s) in insects and are becoming increasingly important for studies of insect molecular evolution, phylogeny and phylogeography^{2–4}. Following the development of next generation sequencing technology, large numbers of mitochondrial genomes are becoming available^{2,5}. Mitochondrial genomes representing each of the major subordinal lineages from each of the 28 recognized insect orders are now available and representation at the family level is steadily improving.

The hemipterous superfamily Membracoidea (leafhoppers and treehoppers) is of interest because it is the most diverse and successful lineage of sap-sucking phytophagous insects, and because of the great variety of behavioral and life-history strategies found within this group.

The phylogenetic relationships of the leafhopper family Cicadellidae to the treehopper families Membracidae, Melizoderidae, and Aetalionidae, all comprising superfamily Membracoidea, have been controversial for decades

¹Key Laboratory of Plant Protection Resources and Pest Management of Ministry of Education, College of Plant Protection, Northwest A&F University, Yangling, Shaanxi, China. ²Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, Illinois, United States of America. Correspondence and requests for materials should be addressed to W.D. (email: daiwu@nwsuaf.edu.cn)

and conflicts among the few morphology-based phylogenetic hypotheses proposed by different authors for Membracoidea have not yet been resolved. Although, traditionally, Cicadellidae has been regarded as the sister group of a lineage comprising the three treehopper families mainly supported by morphological characters of the adults^{6–8}, studies based on morphological^{9,10}, paleontological¹¹ and behavioral¹² evidence suggest that the Cicadellidae are paraphyletic with respect to Membracidae. Molecular phylogenetic analysis of deep-level membracoid relationships based on 28S rRNA sequence data also indicated that Cicadellidae is paraphyletic with respect to Membracidae¹³. Within the Cicadellidae, the relationships among family-group taxa based on recent phylogenetic analyses of DNA sequence and morphological data remain poorly resolved and no family-group classification has yet gained universal acceptance^{9,13–16}.

Until now, most studies of complete mitochondrial genomes in Hemiptera have focused on Heteroptera. To date, only nine complete mitochondrial genomes of Membracoidea have been deposited in GenBank, including the aetalionid *Darthula hardwickii* (KP316404)¹⁷, the membracids *Leptobelus gazella* (JF801955)¹⁸ and *Entylia carinata* (KX495488)¹⁹, and the cicadellids *Drabescoides nuchalis* (KR349344)²⁰, *Nephotettix cincticeps* (KP749836), *Tambocerus sp.* (KT827824)²¹, *Empoasca vitis* (KJ815009)²² (Qin *et al.*²³ recently showed that Chinese specimens previously identified as *E. vitis* are actually *E. onukii* so the specimens used by Zhou *et al.*²² were probably misidentified because true *E. vitis* is a European species not known to occur in China), *Homalodisca vitripennis* (AY875213; listed as “*H. coagulata*”) and *Idioscopus nitidulus* (KR024406). So far, the secondary structures of tRNAs and rRNAs of these species have not been predicted and the nucleotide sequence data have not been used to estimate phylogenetic relationships. The number of sequenced cicadellid mitochondrial genomes also remains very limited relative to the species-richness of Cicadellidae. Because mitochondrial genomes are available for only three species of Deltocephalinae, the largest cicadellid subfamily, representing only three of the 39 recognized tribes, more taxa and more data are needed for future phylogenetic studies to obtain a more resolved and supported phylogeny of the subfamily.

Here we present and analyze the complete mitochondrial genomes of two additional deltocephaline species *Maiestas dorsalis* (Motschulsky) (tribe Deltocephalini) and *Japananus hyalinus* (Osborn) (tribe Opsini), including the gene order, nucleotide composition, codon usage, tRNA secondary structure, rRNA secondary structure, gene overlaps and non-coding regions. Using these new sequences, along with those of previously published mitochondrial genomes of Membracoidea, we reconstructed the phylogenetic relationships among them based on the concatenated nucleotide sequences of 13 protein-coding genes (PCGs) and two ribosomal RNA genes.

Materials and Methods

Sample collection and DNA extraction. Adults of *M. dorsalis* used in this study were collected in Guilin, Guangxi province, China (25°63'N, 109°91'E, July 2015), while *Japananus hyalinus* specimens were collected in Yangling, Shaanxi province, China (34°27'N, 108°09'E, September 2014). Fresh specimens were initially preserved in 100% ethanol, and then stored at –20 °C in the laboratory. After morphological identification, voucher specimens with male genitalia prepared were deposited in the Entomological Museum of Northwest A&F University, and total genomic DNA was extracted from muscle tissues of the thorax using the DNeasy DNA Extraction kit (Qiagen).

Next generation sequence assembly. Most of the mitochondrial genome sequences of the two species were generated using Illumina HiSeq™ 2500 with paired reads of 2 × 150 bp. A total of 26,435,488 and 13,989,788 raw paired reads were retrieved and quality-trimmed using CLC Genomics Workbench v7.0.4 (CLC Bio, Aarhus, Denmark) with default parameters for *M. dorsalis* and *J. hyalinus* respectively. Subsequently, with the mitochondrial genome of *D. nuchalis* (KR349344)²⁰ employed as reference, the resultant 26,435,023 and 13,989,519 clean paired reads were used for mitochondrial genome reconstruction using MITObim v1.7 software²⁴ with default parameters. A total of 18,137 individual mitochondrial reads yielded an average coverage of 158.6 × for *M. dorsalis*, while for *J. hyalinus* a total of 17,924 individual mitochondrial reads yielded an average coverage of 70.3 ×.

Gap closing-PCR amplification and sequencing. According to the flanking sequences assembled from the NGS data, we designed two pairs of primers to amplify the control region (CR) (*M. dorsalis*: Forward 5'-3': TGTATAACCGGAATGCTGGCACAA and Reverse 5'-3': TTAGGGTATGAACCTAATAGCT; *J. hyalinus*: Forward 5'-3': ATAGCCAGAATCAAACCT and Reverse 5'-3': AAGTGTCACAGGCTTAGGT). PCR reactions were performed with TaKaRa LA-Taq Kits (TaKaRa Co., Dalian, China) under the following cycling conditions: 5 min at 94 °C, 38 cycles of 30 s at 94 °C, 1 min at 45–50 °C, 3 min at 68 °C, and a final elongation step at 68 °C for 15 min. PCR products were electrophoresed on 1% agarose gels, purified and then sequenced in both directions on an ABI 3730 XL automated sequencer (Applied Biosystems).

Genome annotation and bioinformatic analyses. The two mitochondrial genomes were annotated with GENEIOUS R8 (Biomatters Ltd., Auckland, New Zealand). All 13 protein-coding genes and 2 rRNA genes were determined by comparison with the homologous sequences of other leafhoppers from GenBank. The 22 tRNA genes were identified using both of the tRNAScan-SE server v 1.21²⁵ and MITOS WebSever²⁶, the secondary structure was also predicted by the MITOS WebSever.

The base composition and relative synonymous codon usage (RSCU) values of each protein coding gene (PCG) were calculated with MEGA 6.06²⁷. Strand asymmetry was calculated using the formulas AT skew = [A – T]/[A + T] and GC skew = [G – C]/[G + C]²⁸. The number of nonsynonymous substitutions per nonsynonymous site (Ka) for the two species were calculated with DnaSP 5.0²⁹, using *Magiccada tredecim* (Riley) from Cicadoidea and *Callitettix braconoides* (Walker)³⁰ from Cercopoidea as references. The tandem repeats of the A + T-rich region were identified by the tandem repeats finder online server (<http://tandem.bu.edu/trf/trf.html>)³¹.

Superfamily	Family	Subfamily	Tribe	Species	Accession number	Reference	
Cicadoidea	Cicadidae	Cicadettinae	Taphurini	<i>Magiccada tredecim</i>	KM000130	Direct Submission	
Cercopoidea	Cercopidae	Callitettixinae	Callitettixini	<i>Callitettix braconoides</i>	JX844628	Liu <i>et al.</i> , 2014 ³⁰	
Membracoidea	Membracidae	Centrotinae	Leptobelini	<i>Leptobelus gazella</i>	JF801955	Zhao and Liang, 2016 ¹⁸	
		Smiliinae	Polyglyptini	<i>Entylia carinata</i>	KX495488	Mao <i>et al.</i> , 2016 ¹⁹	
		Aetalionidae	Darthulinae	Darthulini	<i>Darthula hardwickii</i>	KP316404	Liang <i>et al.</i> , 2016 ¹⁷
		Cicadellidae	Cicadellinae	Proconiini	<i>Homalodisca vitripennis</i>	AY875213	Direct Submission
		Typhlocybinae	Empoascini	<i>Empoasca vitis</i>	KJ815009	Zhou <i>et al.</i> , 2016 ²²	
		Eurymelinae	Idiocerini	<i>Idioscopus nitidulus</i>	KR024406	Direct Submission	
		Deltocephalinae	Chiasmini	<i>Nephotettix cincticeps</i>	KP749836	Direct Submission	
		Drabescini	<i>Drabescoides nuchalis</i>	KR349344	Wu <i>et al.</i> , 2016 ²⁰		
		Athysanini	<i>Tambocerus</i> sp.	KT827824	Yu <i>et al.</i> , 2015 ²¹		
		Deltocephalini	<i>Maiestas dorsalis</i>	KX786285	This study		
Opsiini	<i>Japananus hyalinus</i>	KY129954	This study				

Table 1. Taxonomic information and GenBank accession numbers for the species used in this study.

Phylogenetic analyses. *Taxa selection.* Besides the two mitochondrial genomes obtained in this study, another nine available mitochondrial genomes for Membracoidea were downloaded from NCBI for phylogenetic analyses. A cicada, *Magiccada tredecim*, and a cercopid, *Callitettix braconoides*, were selected as outgroups. The ingroup species represented three families, Membracidae, Aetalionidae and Cicadellidae (Table 1).

Sequence alignment and substitution saturation test. Sequences of all 13 PCGs and two rRNA genes were used in our analyses. After excluding stop codons, each PCG was aligned individually with codon-based multiple alignments using the MAFFT algorithm in the TranslatorX online server (<http://translatorx.co.uk/>)³², with gaps and ambiguous sites removed from the protein alignment before back-translating to nucleotides using GBLOCKS under default settings. The sequences of two rRNA genes were aligned separately using the MUSCLE algorithm implemented in MEGA 6.06. We found that MUSCLE and MAFFT have been shown to perform similarly on our rDNA sequence data but because MUSCLE is implemented in MEGA 6.06, which facilitates manual removal of poorly aligned positions, we used MUSCLE for these genes.

Saturation tests for different codon positions of PCGs and the two rRNA genes were performed, with the uncorrected p-distances plotted against the GTR distances. All distances were generated using PAUP/4.0 b10³³. The slope, correlation coefficient and average GTR distance were used as measures of substitution saturation; the lower the slope, the greater the level of saturation^{34–36}.

Dataset concatenation, partitioning and substitution model selection. Alignments of all genes were concatenated using SequenceMatrix 1.7.8³⁷. Five datasets were generated: 1) P123: 13 PCGs with 9846 nucleotides; 2) P12: first and second codon positions of 13 PCGs with 6564 nucleotides; 3) P123R: 13 PCGs and two rRNAs with 11905 nucleotides; 4) P12R: first and second codon positions of 13 PCGs and two rRNAs with 8623 nucleotides; and 5) AA: amino acid sequences of 13 PCGs with 3282 amino acids.

Optimal nucleotide substitution models and partition strategies were chosen by PartitionFinder v1.1.1³⁸. Under the “greedy” search algorithm, we chose “unlinked” to estimate branch lengths and used the Bayesian information criterion (BIC) as the metric for the partitioning scheme. Details of the best-fit schemes for ML and BI analysis are shown in Supplementary Table S3.

Phylogenetic inference. ML analyses were conducted using raxmlGUI 1.5³⁹ under the GTRGAMMAI model, and the node reliability was assessed by performing 1000 rapid bootstrap replicates (BS). Bayesian analysis was performed using MrBayes 3.2.6⁴⁰. Two simultaneous runs with eight independent chains were run for five million generations and trees were sampled every 1000 generations. After the average standard deviation of split frequencies fell below 0.01, the first 25% of the total samples were discarded as burn-in and the remaining trees were used to generate a consensus tree and calculate the posterior probabilities (PP).

Results and Discussion

General features of the two newly sequenced mitochondrial genomes. The complete mitochondrial genomes of *M. dorsalis* (GenBank: KX786285) and *J. hyalinus* (GenBank: KY129954) are double-stranded circular molecules of length of 15352 bp and 15364 bp, respectively (Fig. 1). Both sizes were comparable to other sequenced Membracoidea mitochondrial genomes ranging from 14805 bp in *Nephotettix cincticeps* to 16007 bp in *Leptobelus gazella* (Table 2). Each mitochondrial genome included the 37 typical mitochondrial genes (13 PCGs, 22 tRNAs and two rRNAs) and a control region (A + T-rich region) (Supplementary Tables S1, S2).

Base composition. All the newly obtained mitochondrial genomes exhibited heavy AT nucleotide bias, with A + T% of the whole sequences 78.8% in *M. dorsalis* and 76.6% in *J. hyalinus*, similar to those found in other sequenced Membracoidea (75.6–78.8%)(Table 2). CR sequences in most species showed the strongest A + T% biases, except for *Entylia carinata*, *Drabescoides nuchalis* and *J. hyalinus*. All species had higher A + T% in *rrnL*

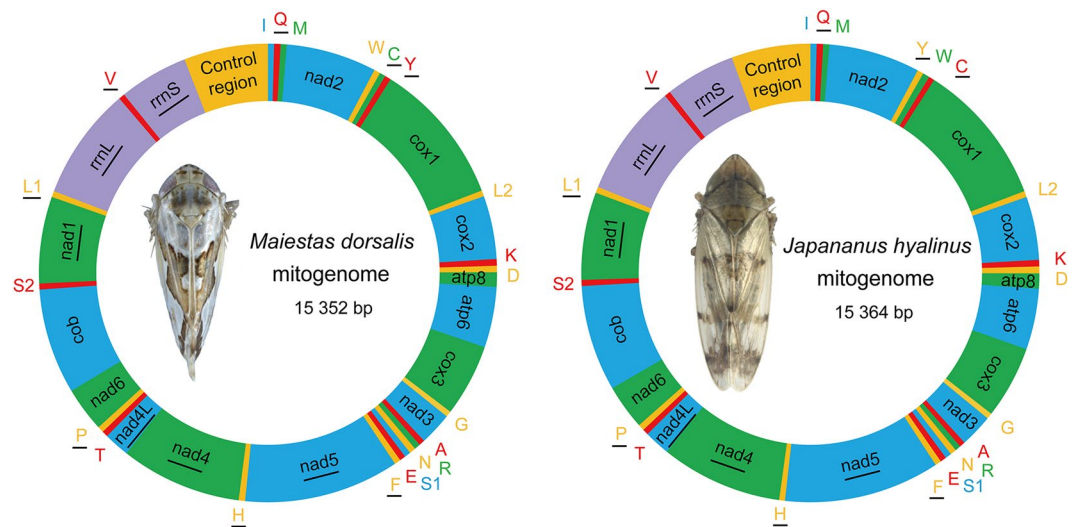


Figure 1. Circular map of the mitochondrial genome of *Maiestas dorsalis* and *Japananus hyalinus*. Protein coding and ribosomal genes are shown with standard abbreviations. Transfer RNA (tRNA) genes are indicated using the IUPAC-IUB single letter amino acid codes. Gene names without underlining indicate the direction of transcription in the major (J) strand, while names with underlining indicate transcription in the minor (N) strand.

Species	whole		PCGs		16S		12S		CR	
	length	AT%	length	AT%	length	AT%	length	AT%	length	AT%
<i>Leptobelus gazella</i>	16007	78.8	10920	77.0	1188	81.8	736	79.0	1750	88.4
<i>Entylia carinata</i>	15662	78.1	10924	77.3	1171	82.0	722	78.3	1474	79.4
<i>Darthula hardwickii</i>	15355	78.0	10916	76.8	1198	82.3	737	78.0	1077	83.6
<i>Homalodisca vitripennis</i>	15304	78.4	10965	77.1	1201	80.9	728	77.9	1033	88.1
<i>Empoasca vitis</i>	15154	78.3	10947	76.7	1149	81.9	725	81.3	977	89.0
<i>Idioscopus nitidulus</i>	15287	78.7	10943	77.3	1196	80.3	734	78.6	991	89.1
<i>Nephotettix cincticeps</i>	14805	77.7	10901	76.9	1201	80.8	741	78.6	399	83.9
<i>Drabesoides nuchalis</i>	15309	75.6	10933	74.6	1197	79.1	740	78.9	956	77.3
<i>Tambocerus</i> sp.	15955	76.4	10963	74.3	1217	80.5	732	78.0	1581	86.0
<i>Maiestas dorsalis</i>	15352	78.8	10961	78.1	1217	81.4	745	79.7	908	81.5
<i>Japananus hyalinus</i>	15364	76.6	10954	75.8	1208	79.8	752	79.5	848	76.3

Table 2. Nucleotide compositions in regions of the Membracoidea mitochondrial genomes.

than *rrnS*, and the PCGs showed the lowest A + T% among whole genes (Table 2). Except for *Empoasca vitis* and *J. hyalinus*, all species showed positive AT-skews (0.007–0.250), for both whole sequences and individual gene sequences. Except for the CR sequences in *L. gazella*, *E. carinata*, *Idioscopus nitidulus* and *J. hyalinus*, all other regions showed negative GC-skews (−0.073 to −0.345) (Supplementary Fig. S1).

Gene rearrangement in Membracoidea. Among the 11 sequenced mitochondrial genomes, the gene order of most species was highly conserved and identical to the putative ancestral insect (*Drosophila yakuba*) mitochondrial genome arrangement¹. The only exception was *J. hyalinus*, the first known member of Membracoidea with a tRNA gene rearrangement, in which the tRNA cluster *trnW-trnC-trnY* was rearranged to *trnY-trnW-trnC* (Fig. 2). Although such gene rearrangements were previously unknown in the superfamily Membracoidea, they have often been found in other insect groups^{41,42}.

Protein-coding genes, codon usage and substitution rates. Of the 13 PCGs in *M. dorsalis* and *J. hyalinus*, nine were located on the majority strand (J-strand) while the other four PCGs were located on the minority strand (N-strand), as observed in other Membracoidea species (Supplementary Tables S1; S2). Among the concatenated 13 PCGs of each species of Membracoidea, the third codon position had an A + T content (84.3–91.7%) much higher than that of the first (70.1–74.7%) and second (67.8–70.0%) positions (Supplementary Fig. S2).

Most of the PCGs started with the standard ATN codon, except for *nad5*, which began with TTG, a pattern also observed in *Tambocerus* sp., *I. nitidulus* and *E. carinata*. Also, *atp8* in *H. vitripennis*, *D. hardwickii* and *I. nitidulus* started with TTG, and *cox2* in *Empoasca vitis* began with GTG. In *M. dorsalis*, *nad5* terminated with TAG, *cox2* and *nad1* ended with an incomplete T codon, with all other 10 PCGs using TAA as the T codon. In *J.*

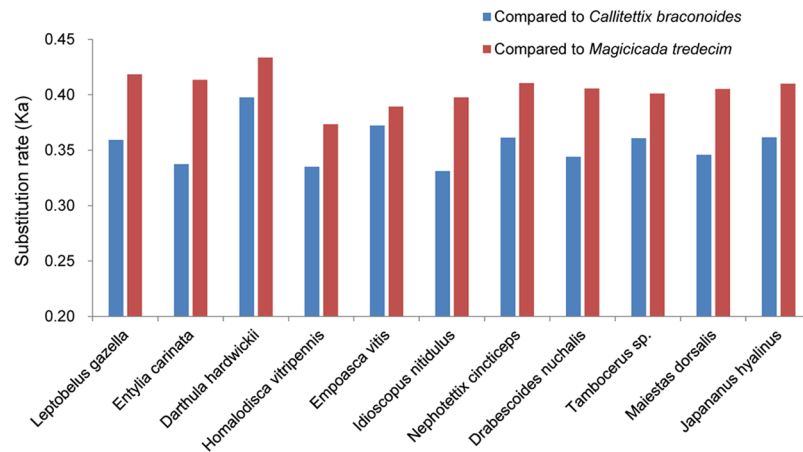


Figure 4. Comparison of substitution rates among Membracoidea mitochondrial genomes. The nonsynonymous substitution rate (Ka) was calculated in a pairwise fashion, using *Callitettix braconoides* and *Magicicada tredecim* as references.



Figure 5. Organization of the control region in Membracoidea mitochondrial genomes. The colored ovals with Arabic numerals indicate the tandem repeats, the remaining regions are shown with colored boxes.

Figs S3, S4). Two extra single A nucleotides, 23 G-U mismatches, six U-U mismatches, two A-A mismatches and one C-U mismatch were found in *M. dorsalis* tRNA genes. While in *J. hyalinus* one extra single A nucleotide, 26 G-U mismatches, four U-U mismatches, one A-A mismatch, one A-C mismatch and one A-G mismatch were found.

Similar to other insect mitochondrial genomes, two *rrn* genes were encoded on the N-strand, and located between *trnL1* and *trnV*, and between *trnV* and the A + T-rich region, respectively. The lengths of the *rrnL* and *rrnS* in *M. dorsalis* were 1217 bp and 745 bp, with the A + T content 81.4% and 79.7%, respectively, while in *J.*

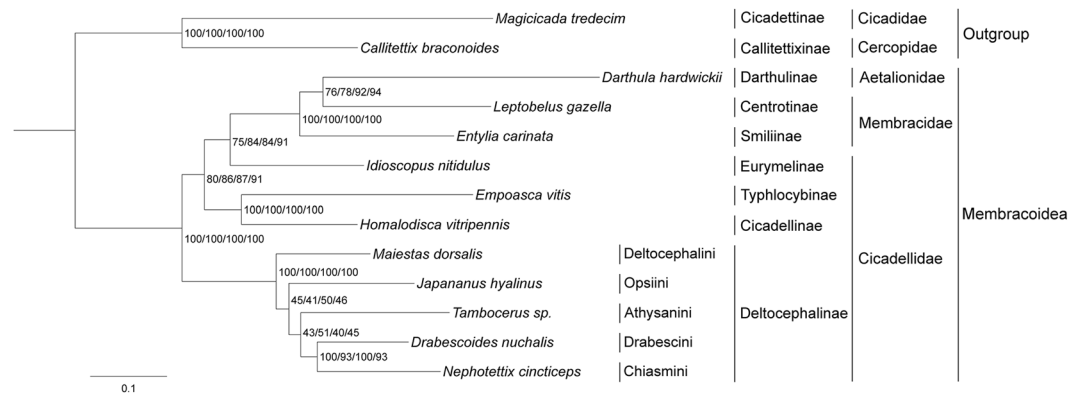


Figure 6. Membracoidea phylogeny based on the P123/P12/P123R/P12R datasets inferred from RAxML. Numbers on branches are bootstrap values (BS).

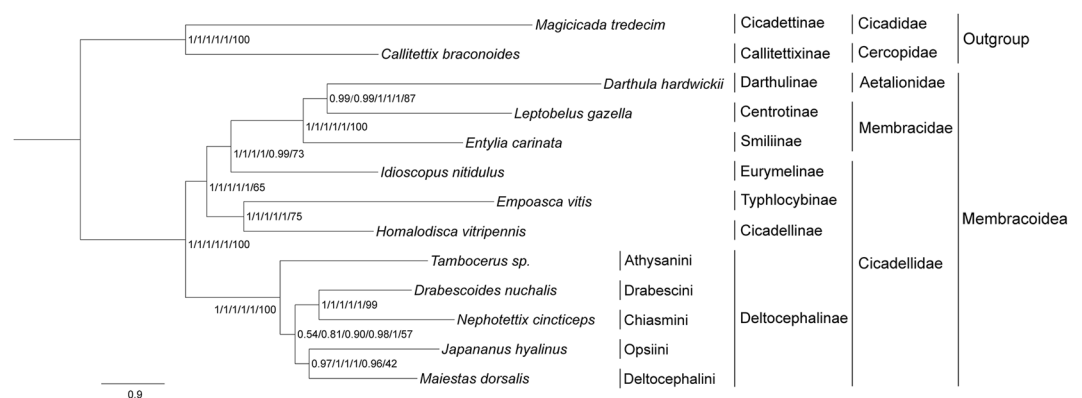


Figure 7. Membracoidea phylogeny based on the P123/P12/P123R/P12R/AA datasets obtained with MrBayes and AA dataset inferred from RAxML. Numbers on branches are Bayesian posterior probabilities (PP) and bootstrap values (BS).

hyalinus are 1208 bp and 752 bp, with the A + T content 79.8% and 79.5%, respectively, which were consistent with the data available for other Membracoidea species (Table 2).

Gene overlaps and non-coding regions. The whole *M. dorsalis* mitochondrial genome had a total of 35 bp in overlaps between 11 gene junctions, while *J. hyalinus* had 27 bp overlaps between eight gene junctions. The longest overlap (8 bp) occurs between *trnW* and *trnC* in both two species. The two common pairs of gene overlaps: *atp8-atp6* (7 bp) and *nad4-nad4l* (7 bp), also were found in two other species (Supplementary Tables S1, S2).

There were 13 intergenic spacers totaling 92 bp non-coding bases in the *M. dorsalis* mitochondrial genome, ranging in size from 1 bp to 16 bp, and the longest two intergenic spacers (16 bp) were between *nad2-trnW* and *cox1-trnL2* respectively. For *J. hyalinus*, 15 intergenic spacers occupied 178 bp non-coding bases, the longest being 73 bp between *trnY* and *trnW*, caused by the tRNA gene rearrangement (Supplementary Tables S1, S2).

The putative control region, located between *rrnS* and *trnI*, was the longest intergenic spacer in the mitochondrial genome. The lengths of this region in *M. dorsalis* and *J. hyalinus* were 908 bp and 848 bp respectively, well within the range of other sequenced Membracoidea (399 bp in *N. cincticeps* to 1750 bp in *L. gazella*) (Table 1).

The control region sequence of *M. dorsalis* included one large tandem repeat, two 239 bp tandem repeat units and a partial third (179 bp) were beginning with the first nucleotide of this region (Fig. 5). *J. hyalinus* also had two different 159 bp tandem repeat units, consisting of the main portion of the control region. Tandem repeats identified in the control region of other sequenced Membracoidea mitochondrial genomes include four in *H. vitripennis*, three in *D. nuchalis*, and two different types of tandem repeats in *L. gazella*, *E. carinata* and *Tambocerus sp.* Characteristics of this region in Membracoidea were taxon-specific, and the different size or copy numbers of repeat units had some influence on the size of the region (Fig. 5).

Phylogenetic relationships. We performed 10 independent phylogenetic analyses to evaluate the influence of different datasets and inference methods on tree topology and nodal support. These analyses yielded two different tree topologies, with incongruence restricted to relationships among members of cicadellid subfamily Deltocephalinae (Figs 6,7).

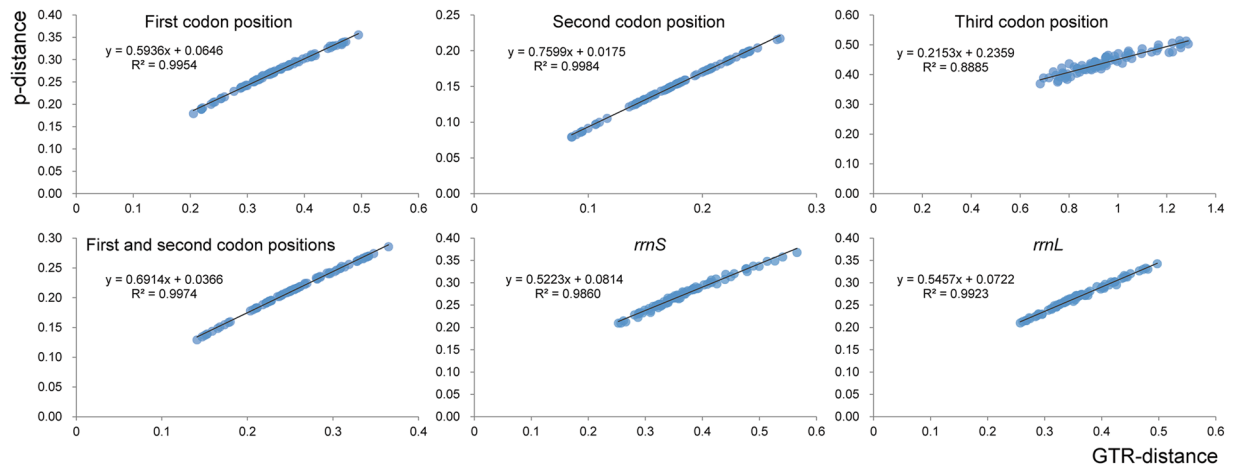


Figure 8. Saturation of the nucleotide substitution tests for 13 protein-coding genes and two rRNA genes (*rrnL* and *rrnS*) of mitochondrial genomes in Membracoidea.

Monophyly at the superfamily level within Membracoidea was strongly supported in both trees (Figs 6,7). Derivation of Membracidae and Aetalionidae from within a paraphyletic Cicadellidae was well supported by all results, as suggested by previous analyses^{9,13}. Although other recent analyses based on mitogenomic data supported treehoppers as a sister group to Cicadellidae^{19–21}, none of these studies included all full available mitochondrial genomes within Membracoidea and, thus, the incongruence may be due to sample bias. For the three treehoppers, the two species of Membracidae *E. carinata* and *L. gazella* formed a paraphyletic grade giving rise to *D. hardwickii* (Aetalionidae), indicating paraphyly of Membracidae (Figs 6,7). This result differs from previous studies, in which Membracidae was usually placed as the sister clade to Aetalionidae^{13,43,44}.

Within Cicadellidae, the eight species sampled in this study represent four subfamilies, Eurymelinae, Cicadellinae, Typhlocybinae and Deltocephalinae. The inferred relationships (Deltocephalinae + (Eurymelinae + (Cicadellinae + Typhlocybinae))) were incongruent with a previous phylogeny based on 28S sequences, in which the relationships (Cicadellinae + (Deltocephalinae + (Typhlocybinae + Eurymelinae))) were found¹³ (Figs 6,7). However, the latter topology received low branch support, which may explain the incongruence.

Within Deltocephalinae, the five species (*M. dorsalis*, *J. hyalinus*, *D. nuchalis*, *N. cincticeps* and *Tambocerus* sp.) representing five tribes (Deltocephalini, Opsiini, Drabescini, Chiasmini and Athysanini, respectively) formed a monophyletic group with high support. For all datasets in BI analyses and the AA dataset in ML analyses, the topology (Athysanini + ((Opsiini + Deltocephalini) + (Drabescini + Chiasmini))) was recovered, while except for the AA dataset, all other datasets yield the topology (Deltocephalini + (Opsiini + (Athysanini + (Drabescini + Chiasmini)))) (Figs 6, 7). Thus, relationships among the included Deltocephalinae were inconsistently resolved in the different analyses, with support for some branches relatively low (BS < 50, PP < 0.9). Thus, as in a previous phylogenetic study based on combined morphological, 28S and Histone H3 data¹⁴ our analyses were unable to resolve some relationships with confidence. Apparent topological discordances between our results and those of Zahniser and Dietrich¹⁴ may be due to the small taxon sample of the current study. Classification of Deltocephalinae has been unstable, with *D. nuchalis* formerly placed in a separate subfamily, Selenocephalinae^{45,46}, but more recently placed into Deltocephalinae based on morphological characters^{47,48}. Recent phylogenetic studies indicate that the subfamily as defined by Oman *et al.*⁴⁹ was not monophyletic and that several other leafhopper subfamilies defined by Oman *et al.*⁴⁹ had their closest relatives within the Deltocephalinae lineage^{13,14,50}.

Previously studies confirmed that inclusion or exclusion of third codon positions had a strong influence on phylogenetic reconstruction^{51,52}. Our saturation tests on different codon positions of PCGs and two rRNA genes (Fig. 8) showed that third codon positions are more saturated than first and second codon positions, with slopes of 0.2135, 0.5936 and 0.7599, respectively. Nevertheless, in our phylogenetic results, tree topologies were consistent regardless of whether third codon positions were excluded, but excluding third positions slightly increased support for some nodes in ML analyses (BS values: 75 to 84, 84 to 91, 80 to 86 and 87 to 91) (Fig. 8).

Conclusions

The first comparative analysis of mitochondrial genomes across the superfamily Membracoidea revealed a high level of conservatism in gene order, with all sequenced genes arranged in the putative insect ancestral gene arrangement, but one of the taxa newly sequenced for this study, *J. hyalinus*, had the putative ancestral tRNA cluster *trnW-trnC-trnY* rearranged to *trnY-trnW-trnC*. This is the first reported gene rearrangement in a mitogenome of Membracoidea. Phylogenomic analysis of the concatenated nucleotide sequences of 13 protein-coding genes (PCGs) and two ribosomal RNA genes from all available membracoid mitochondrial genomes supported the monophyly of Membracoidea and paraphyly of Cicadellidae with respect to the treehopper lineage (Aetalionidae + Membracidae). ML and BI analyses yielded topologies that were congruent except for relationships among included representatives of subfamily Deltocephalinae. Exclusion of third codon positions of PCGs improved some node support values in ML analyses. These results suggest that mitochondrial genome sequences are informative of higher level phylogenetic relationships within Membracoidea but may not be sufficient to resolve relationships within some membracoid lineages.

References

- Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767–1780 (1999).
- Cameron, S. L. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu. Rev. Entomol.* **59**, 95–117 (2014).
- Abascal, F., Posada, D., Knight, R. D. & Zardoya, R. Parallel evolution of the genetic code in arthropod mitochondrial genomes. *PLoS Biol.* **4**, e127 (2006).
- Saccone, C., De Giorgi, C., Gissi, C., Pesole, G. & Reyes, A. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene* **238**, 195–209 (1999).
- Jex, A. R., Hu, M., Littlewood, D. T. J., Waeschenbach, A. & Gasser, R. B. Using 454 technology for long-PCR based sequencing of the complete mitochondrial genome from single *Haemonchus contortus* (Nematoda). *BMC Genomics* **9**, 11 (2008).
- Strumpel, H. Beitrag zur phylogenie der Membracidae Rafinesque. *Zoologische Jahrbücher: Abteilung für Systematik, Ökologie und Geographie der Tiere* **99**, 313–407 (1972).
- Evans, J. W. The leafhoppers and froghoppers of Australia and New Zealand (Homoptera: Cicadelloidea and Cercopoidea). *Rec. Aust. Mus.* **31**, 83–129 (1977).
- Dietrich, C. H. & Deitz, L. L. Superfamily Membracoidea (Homoptera: Auchenorrhyncha). II. Cladistic analysis and conclusions. *Syst. Entomol.* **18**, 297–311 (1993).
- Hamilton, K. G. A. Classification, morphology and phylogeny of the family Cicadellidae (Rhynchota: Homoptera). Proceeding of the 1st International Workshop on Biotaxonomy, Classification and Biology of Leafhoppers and Planthoppers of Economic Importance (eds Knight, W. J., Pant, N. C., Robertson, T. S. & Wilson, M. R.) 15–37 (Commonwealth Institute of Entomology, 1983).
- Hamilton, K. G. A. The ground-dwelling leafhoppers Sagmatiini and Myerslopiidae (Rhynchota: Homoptera: Membracoidea). *Invertebr. Taxon.* **13**, 207–235 (1999).
- Shcherbakov, D. E. The earliest leafhoppers (Hemiptera: Karajassidae n. fam.) from the Jurassic of Karatau. *Neues Jahrbuch für Geologie und Paläontologie Monatshefte* **1**, 39–51 (1992).
- Rakitov, R. On differentiation of cicadellid leg chaetotaxy (Homoptera: Auchenorrhyncha: Membracoidea). *Russian Entomol. J.* **6**, 7–27 (1998).
- Dietrich, C. H., Rakitov, R. A., Holmes, J. L. & Black, W. C. Phylogeny of the major lineages of Membracoidea (Insecta: Hemiptera: Cicadomorpha) based on 28S rDNA sequences. *Mol. Phylogenet. Evol.* **18**, 293–305 (2001).
- Zahniser, J. N. & Dietrich, C. A review of the tribes Deltocephalinae (Hemiptera: Auchenorrhyncha: Cicadellidae). *Eur. J. Taxon.* **45**, 1–121 (2013).
- Dietrich, C. H. Keys to the families of Cicadomorpha and subfamilies and tribes of Cicadellidae (Hemiptera: Auchenorrhyncha). *Fla. Entomol.* **88**, 502–517 (2005).
- Dietrich, C. H. The role of grasslands in the diversification of leafhoppers (Homoptera: Cicadellidae): a phylogenetic perspective. *Proceedings of the Fifteenth North American Prairie Conference*, 44–49 (1999).
- Liang, A. P., Gao, J. & Zhao, X. Characterization of the complete mitochondrial genome of the treehopper *Darthula hardwickii* (Hemiptera: Aetalionidae). *Mitochondrial DNA Part A* **27**, 3291–3292 (2016).
- Zhao, X. & Liang, A. P. Complete DNA sequence of the mitochondrial genome of the treehopper *Leptobelus gazella* (Membracoidea: Hemiptera). *Mitochondrial DNA Part A* **27**, 3318–3319 (2016).
- Mao, M., Yang, X. & Bennett, G. The complete mitochondrial genome of *Entylia carinata* (Hemiptera: Membracidae). *Mitochondrial DNA Part B* **1**, 662–663 (2016).
- Wu, Y. F., Dai, R. H., Zhan, H. P. & Qu, L. Complete mitochondrial genome of *Drabescoides nuchalis* (Hemiptera: Cicadellidae). *Mitochondrial DNA Part A* **27**, 3626–3627 (2016).
- Yu, P. F., Wang, M. X., Cui, L., Chen, X. X. & Han, B. Y. The complete mitochondrial genome of *Tambocerus* sp. (Hemiptera: Cicadellidae). *Mitochondrial DNA Part A* **28**, 133–134 (2015).
- Zhou, N. N., Wang, M. X., Cui, L., Chen, X. X. & Han, B. Y. Complete mitochondrial genome of *Empoasca vitis* (Hemiptera: Cicadellidae). *Mitochondrial DNA Part A* **27**, 1052–1053 (2016).
- Qin, D., Zhang, L., Xiao, Q., Dietrich, C. & Matsumura, M. Clarification of the identity of the tea green leafhopper based on morphological comparison between Chinese and Japanese specimens. *PLoS One* **10**, e0139202 (2015).
- Hahn, C., Bachmann, L. & Chevreaux, B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - a baiting and iterative mapping approach. *Nucleic Acids Res.* **41**, e129 (2013).
- Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
- Bernt, M. et al. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **69**, 313–319 (2013).
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
- Perna, N. T. & Kocher, T. D. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* **41**, 353–358 (1995).
- Librado, P. & Rozas, J. DnaSPv5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
- Liu, J., Bu, C., Wipfler, B. & Liang, A. Comparative analysis of the mitochondrial genomes of callitettixini spittlebugs (Hemiptera: Cercopidae) confirms the overall high evolutionary speed of the at-rich region but reveals the presence of short conservative elements at the tribal level. *PLoS One* **9**, e109140 (2014).
- Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
- Abascal, F., Zardoya, R. & Telford, J. M. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* **38**, 7–13 (2010).
- Swofford, D. L. PAUP*: phylogenetic analysis using parsimony, version 4.0 b10. Sinauer Associates, Sunderland, Massachusetts (2003).
- Negrisoló, E., Minelli, A. & Valle, G. The mitochondrial genome of the house centipede *Scutigera* and the monophyly versus paraphyly of myriapods. *Mol. Biol. Evol.* **21**, 770–780 (2004).
- Li, H. et al. Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. *Sci. Rep.* **5**, 8527 (2015).
- Yuan, M. L. et al. High-level phylogeny of the Coleoptera inferred with mitochondrial genome sequences. *Mol. Phylogenet. Evol.* **104**, 99–111 (2016).
- Vaidya, G., Lohman, D. J. & Meier, R. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* **27**, 171–180 (2011).
- Lanfear, R., Calcott, B., Ho, S. Y. & Guindon, S. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **29**, 1695–1701 (2012).
- Silvestro, D. & Michalak, I. RaxmlGUI: a graphical front-end for RAXML. *Org. Divers. Evol.* **12**, 335–337 (2012).
- Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
- Wei, S. J., Li, Q., van Achterberg, K. & Chen, X. X. Two mitochondrial genomes from the families Bethyliidae and Mutillidae: Independent rearrangement of protein-coding genes and higher-level phylogeny of the Hymenoptera. *Mol. Phylogenet. Evol.* **77**, 1–10 (2014).

42. Song, S. N., Tang, P., Wei, S. J. & Chen, X. X. Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. *Sci. Rep.* **6**, 20972 (2016).
43. Cryan, J. R., Wiegmann, B. M., Deitz, L. L., Dietrich, C. H. & Whiting, M. F. Treehopper trees: phylogeny of Membracidae (Hemiptera: Cicadomorpha: Membracoidea) based on molecules and morphology. *Syst. Entomol.* **29**, 441–454 (2004).
44. Cryan, J. R. & Urban, J. M. Higher-level phylogeny of the insect order Hemiptera: is Auchenorrhyncha really paraphyletic? *Syst. Entomol.* **37**, 7–21 (2012).
45. Linnavuori, R. Revision of the African Cicadellidae (subfamily Selenocephalinae) (Homoptera, Auchenorrhyncha). *Acta Zool. Fennica* **168**, 1–105 (1983).
46. Zhang, Y. L. & Webb, M. D. A revised classification of the Asian and Pacific selenocephaline leafhoppers (Homoptera: Cicadellidae). *Bull. Nat. Hist. Mus. (London) Entomol.* **65**, 1–103 (1996).
47. Dietrich, C. H. & Rakitov, R. A. Some remarkable new deltocephaline leafhoppers (Hemiptera: Cicadellidae: Deltocephalinae) from the Amazonian rainforest canopy. *J. New York Entomol. S.* **110**, 1–48 (2002).
48. Dmitriev, D. A. Nymphs of some species of the tribes Drabescini and Paraboloponini with a proposed synonymy of Paraboloponini with Drabescini (Hemiptera: Cicadellidae: Deltocephalinae). *Orient. Insects* **38**, 235–244 (2004).
49. Oman, P. W., Knight, W. J. & Nielson, M. W. Leafhoppers (Cicadellidae): A bibliography, generic check-list and index to the world literature 1956–1985. (CAB International Institute of Entomology, 1990).
50. Zahniser, J. N. & Dietrich, C. H. Phylogeny of the leafhopper subfamily Deltocephalinae (Hemiptera: Cicadellidae) based on molecular and morphological data with a revised family-group classification. *Syst. Entomol.* **35**, 489–511 (2010).
51. Wei, S. J., Shi, M., Sharkey, M. J., van Achtenberg, C. & Chen, X. X. Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects. *BMC Genomics* **11**, 371 (2010).
52. Song, N., Liang, A. P. & Bu, C. P. A molecular phylogeny of Hemiptera inferred from mitochondrial genome sequences. *PLoS One* **7**, e48778 (2012).

Acknowledgements

We are grateful to Dr. Delong Guan (Shaanxi Normal University, Xian) for kind help in the sequencing and annotation progress. This project was supported by the National Natural Science Foundation of China (Nos 31272343, 31572306, 31420103911, 30970385) and the “West Light Foundation of the Chinese Academy of Sciences (2012DF06)” and the “Chinese Universities Scientific Fund (2452015135)”.

Author Contributions

W.D. and Y.M.D. conceived and designed the experiments. Y.M.D. performed the experiments. W.D. Y.M.D. and C.N.Z. analyzed the data. W.D., Y.M.D., C.D. and Y.L.Z. wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-14703-3>.

Competing Interests: The authors declare that they have no competing interests.

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



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017