The mitochondrial genomes of *Macrocheraia grandis grandis* and *Myrmoplasta mira* (Hemiptera: Heteroptera: Pentatomomorpha) and the unique mitogenome rearrangement in Pyrrhocoroidea

Yu MEN[1, 2, 3], Petr KMENT[4], Frantisek STAHLAVSKY[5], Fei YE[2, 3], Yanhui WANG[2, 3], Qiang XIE[2, 3①]

1. *College of Life Sciences, Nankai University, Tianjin 300071, China*

2. *State Key Laboratory of Biocontrol, Sun Yat-sen University, Guangzhou, Guangdong 510275, China*

3. *Department of Ecology and Evolution, College of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510275, China*

4. *Department of Entomology, National Museum, Cirkusova 1740, 19300 Prague 9, Czech Republic*

5. *Department of Zoology, Charles University, Vinicna 7, CZ-128 44 Praha 2, Czech Republic*

Abstract: Sequencing technology has developed rapidly in recent years. Complete or nearly complete mitochondrial genomes (mitogenomes) of 155 species from 47 families in Heteroptera have been sequenced. However, the amounts of mitogenomes between those families are unbalanced, which makes it difficult to correctly discern the patterns of mitogenome rearrangement in Heteroptera. Among 21 species from ten families, ten variations in mitogenome rearrangement had been previously reported, among which the translocation between *tRNA-Thr* and *tRNA-Pro* was considered as a synapomorphy of Pyrrhocoroidea based on two mitogenomes. As only one mitogenome in each of Largidae and Pyrrhocoridae had been sequenced to conclude the synapomorphy, more mitogenomes of Pyrrhocoroidea need to be explored. In this study, additional two mitogenomes of Pyrrhocoroidea (*Macrocheraia grandis grandis* (Gray, 1832) and *Myrmoplasta mira* Gerstäcker, 1892) were sequenced. Both of them also possess the same translocation between *tRNA-Thr* and *tRNA-Pro*, which reaffirms that this kind of rearrangement is a molecular synapomorphy of Pyrrhocoroidea. Moreover, we discovered a more complex rearrangement in *Myrmoplasta mira*, in which six nearly identical duplications of *tRNA-Thr* were found located downstream of *tRNA-Pro*. Considering the high biodiversity of Heteroptera, more mitogenomic studies are needed to improve our knowledge about mitogenome rearrangements and the potential synapomorphies.

Key words: Pyrrhocoroidea; mitogenome; synapomorphy; phylogenomics; evolutionary history

巨红蝽和 *Myrmoplasta mira* 的线粒体基因组及红蝽总科线粒体基因特有的重排（半翅目：异翅亚目：蝽次目）

门宇[1, 2, 3]，Petr KMENT[4]，Frantisek STAHLAVSKY[5]，叶飞[2, 3]，王艳会[2, 3]，谢强[2, 3①]

1. 南开大学生命科学学院，天津 300071；2. 中山大学生物防治国家重点实验室，广东 广州 510275；3. 中山大学生命科学学院生态与进化学系，广东 广州 510275；4. Department of Entomology, National

Accepted 20 March 2019. Published 25 June 2019. Published online 6 June 2019.

① Corresponding author, E-mail: xieq8@mail.sysu.edu.cn

Museum, Cirkusova 1740, 19300 Prague 9, Czech Republic；5. Department of Zoology, Charles University, Vinicna 7, CZ-128 44 Praha 2, Czech Republic

摘要：近年来测序技术快速发展，已经测得了异翅亚目昆虫 47 科 155 种的完整或接近完整的线粒体基因组序列。然而，线粒体基因组的数量在这些科之间的分布并不均衡，影响了对异翅亚目昆虫线粒体基因组重排规律的探索。在已有的研究中，共在 10 科 21 个物种中发现了 10 类线粒体基因的重排方式，其中 *tRNA-Thr* 和 *tRNA-Pro* 的换位被认为是红蝽总科的共有衍征。但是仅用大红蝽科和红蝽科中各一个重排来总结共有衍征显得不够充分，因此需要获取更多红蝽总科的线粒体基因组序列。本研究补充了两个红蝽总科的线粒体基因组序列（巨红蝽（Gray, 1832）和 *Myrmoplasta mira* Gerstäcker, 1892）。它们在 *tRNA-Thr* 和 *tRNA-Pro* 之间存在相同的重排方式，支持这种重排是红蝽总科的共有衍征。此外，在 *Myrmoplasta mira* 中存在更为复杂的重排方式：在 *tRNA-Pro* 的下游存在 6 个几乎相同的 *tRNA-Thr* 拷贝。考虑到异翅亚目昆虫高度的生物多样性，未来需要获取更多的线粒体基因组序列，以完善对线粒体基因组重排和共有衍征的认识。

关键词：红蝽总科；线粒体基因组；共有衍征；种系基因组学；演化历史

Introduction

Heteroptera is the largest suborder of Hemiptera, containing 123 families and more than 45,200 described species (Henry 2017). Since the complete mitogenome of *Drosophila yakuba* (Diptera: Drosophilidae) was sequenced (Clary & Wolstenholme 1985), which was the first mitogenome sequenced in insects, complete or nearly complete mitogenomes of 155 species in 47 families of Heteroptera have been sequenced. However, the amounts of known mitogenomes in different families vary widely. More than half of described mitogenomes are concentrated in the Reduviidae (including Phymatinae), Miridae, Tingidae, Nabidae, Pentatomidae and Aradidae, and only one or two mitogenomes have been sequenced in each of the other 30 families (Table S1), making their patterns of rearrangement hard to summarize. The gene order of a mitogenome is generally conserved in insects, similar to what was found for *Drosophila yakuba*. But rearrangements still exist across the 17 orders of Insecta (Wang *et al.* 2016; Chen & Du 2016). Ten kinds of rearrangements in ten families of Heteroptera had been previously reported (Table 1), in which *tRNA-Ile* and *tRNA-Gln* switch their positions in all seven mitogenomes of Aradidae. Li *et al.* (2016) proposed that it was likely a synapomorphy of Aradidae.

Table S1. The number of sequenced mitogenome in each family of Heteroptera

Infraorder	Family	The number of sequenced mitogenome
Cimicomorpha	Anthocoridae	2
	Cimicidae	1
	Miridae	13
	Nabidae	7
	Reduviidae	28
	Tingidae	13
	Velocipedidae	1
Nepomorpha	Aphelocheiridae	2

Continued Table S1

Infraorder	Family	The number of sequenced mitogenome
	Belostomatidae	3
	Corixidae	1
	Gelastocoridae	1
	Helotrephidae	2
	Micronectidae	1
	Naucoridae	1
	Nepidae	2
	Notonectidae	3
	Ochteridae	1
	Pleidae	1
Pentatomomorpha	Aradidae	7
	Alydidae	2
	Coreidae	3
	Rhopalidae	3
	Stenocephalidae	1
	Berytidae	2
	Colobathristidae	1
	Lygaeidae	2
	Malcidae	2
	Rhyparochromidae	2
	Acanthosomatidae	2
	Cydnidae	1
	Dinidoridae	3
	Pentatomidae	14
	Plataspidae	2
	Scutelleridae	3
	Tessaratomidae	3
	Urostylididae	3
	Largidae	1
	Pyrrhocoridae	1
Gerromorpha	Gerridae	2
	Hydrometridae	1
	Veliidae	1
Leptopodomorpha	Leptopodidae	1
	Saldidae	3
Dipsocoromorpha	Ceratocombidae	2
	Schizopteridae	2
Enicocephalomorpha	Enicocephalidae	4

Table 1. Rearrangements in Heteroptera

Family	Species	Rearrangements as potential synapomorphies	Other rearrangements
Urostylididae	*Urochela quadrinotata*	Absence of *I* and *Q*	
Largidae +	*Physopelta gutta*	*T-P* → *P-T*	
Pyrrhocoridae	*Dysdercus cingulatus*	(Hua *et al.* 2008)	
Aradidae	*Aneurus similis*	*I-Q* → *Q-I*	
	Aneurus sublobatus	(Li *et al.* 2016)	
	Brachyrhynchus hsiaoi		
	Libiocoris heissi		
	Neuroctenus parus		
	Aradacanthia heissi		*W-C* → *C-W*
	Aradus compar		*M-ND2-W-C-Y* → *C-Y-M-ND2-W*
Nabidae	*Nabis apicalis*	Absence of *I*, *Q* and *M*	
Reduviidae	*Phymata americana*		Absence of *I*
	Brontostoma colossus		*ND3-A* → *ND3-R-A*
	Reduvius tenebrosus		*I-Q* → *Q-I*
Enicocephalidae	*Henschiella* sp. PJ-2015	*T-P-ND6-CytB-S-ND1-L*-lrRNA-*V*-srRNA-*CR* → *CytB-S-CR*-lrRNA-*V*-srRNA-*ND1-L-P-T-ND6*	
	Oncylocotis sp. PJ-2015		
	Stenopirates sp. HL-2011		
	Stenopirates sp. PJ-2015		
Ceratocombidae	*Ceratocombus* sp. HL-2012	Translocation of *R* and *E* & *V*-srRNA → srRNA-*V*	
	Ceratocombus japonicus		
Belostomatidae	*Lethocerus deyrollei*	Absence of *I*	

A similar opinion was proposed by Hua *et al.* (2008) while studying the two mitogenomes in Largidae and Pyrrhocoridae, which are the only two families of the superfamily Pyrrhocoroidea. Each contains one mitogenome with the translocation between *tRNA-Thr* and *tRNA-Pro*, which was considered as a synapomorphy of Pyrrhocoroidea. But it is still unknown whether this synapomorphy will be supported when more mitogenomes in Pyrrhocoroidea are revealed. To secure more mitogenome sequences and check the reliability of such a synapomorphy, mitogenomes of *Macrocheraia grandis grandis* (Gray, 1832) (Largidae: Physopeltinae) and *Myrmoplasta mira* Gerstäcker, 1892 (Pyrrhocoridae) in Pyrrhocoroidea were sequenced and analyzed in this study. *Macrocheraia grandis grandis* is the species with the largest body size in Largidae, and *Myrmoplasta mira* is an ant-like Pyrrhocoridae.

Material and Methods

Samples and DNA extraction

The specimen of *Macrocheraia grandis grandis* was collected by Shasha YU in China, Hainan, Nankai Township on July 24, 2013 and preserved in absolute ethyl alcohol at -20°C. The total genomic DNA was extracted from thorax tissue and the whole head using the

CTAB-based method (Reineke *et al*. 1998). The specimen of *Myrmoplasta mira* was collected by Frantisek STAHLAVSKY, Jan A NEETHLING, Vera OPATOVA and Pavel JUST in South Africa, Eastern Cape on October 15, 2013. This sample was preserved in absolute ethyl alcohol at -20°C. Total genomic DNA was extracted from a fore leg using the TIANamp Micro DNA Kit (TIANGEN).

PCR amplification, sequencing and assembling

The mitogenome of *Myrmoplasta mira* was amplified in several overlapping PCR fragments. The thermal cycling program was set as follows: an initial denaturation at 95°C for 2 min, denaturation of 35 cycles at 95°C for 20 s, annealing at 45–62°C for 1 min, elongation at 72°C for 1–2 min, and a final elongation at 72°C for 10 min. PCR products were electrophoresed in 1% agarose gel and then sequenced by the Sanger method. The results were assembled by SeqMan in the DNAStar v7.1 program package. The *Macrocheraia grandis grandis* sample was quite large and had a high concentration of DNA, therefore its mitogenome sequence was obtained by high-throughput-sequencing. A DNA library was constructed with an insert size of 250 bp, and then sequenced with a 150 bp paired end on the Illumina platform. *De novo* assembly was performed by SOAPdenovo-Trans v1.03 (K-mer = 61 and 71) (Xie *et al.* 2014), and the assemblies were BLASTed using BLAST+ (Camacho *et al.* 2009). A second method was used with MITObim v1.9 to bait and assemble the mitogenome of *Macrocheraia grandis grandis* directly referring to that of *Dysdercus cingulatus*.

Analysis and annotation

The boundaries of tRNA genes and the secondary structures of tRNAs were identified by the Mitos WebServer (Bernt *et al.* 2013). Protein coding genes (PCGs) and the start and stop codons were determined by ORF Finder using invertebrate mitochondrial genetic codes which were implemented by NCBI website (https://www.ncbi.nlm.nih.gov/orffinder/). The boundaries of 12S rRNA and 16S rRNA genes and the secondary structures of those rRNAs were determined through comparison with other insects, including *Drosophila melanogaster* (Diptera: Drosophilidae) (Cannone *et al*. 2002), *Chauliops fallax* (Hemiptera: Malcidae) (Li *et al*. 2013), and *Neolethaeus assamensis* (Hemiptera: Rhyparochromidae) (Guo *et al.* 2017). Codon usage and nucleotide composition were analyzed using MEGA v7.0.

Results

Genome organization and structure

The mitogenome of *Macrocheraia grandis grandis* is a double-stranded circular DNA molecule. A nearly complete mitogenome 14,917 bp in length was sequenced without partial sequence of the control region (CR). The mitogenome of *Macrocheraia grandis grandis* contains the typical 37 genes, including 22 tRNA genes, 13 PCGs and 2 rRNA genes (Fig. 1, Table 2). This mitogenome is relatively compact with 16 gene overlaps. The longest overlap (8 bp) is located between *tRNA-Trp* and *tRNA-Cys*. Another two overlaps between *ATP6*/*ATP8* and *ND4*/*ND4L* consist of the same seven nucleotides (ATGATAA). In addition, eight intergenic spacers with a total of 127 nucleotides were found, ranging in length from 1 bp to 54 bp.

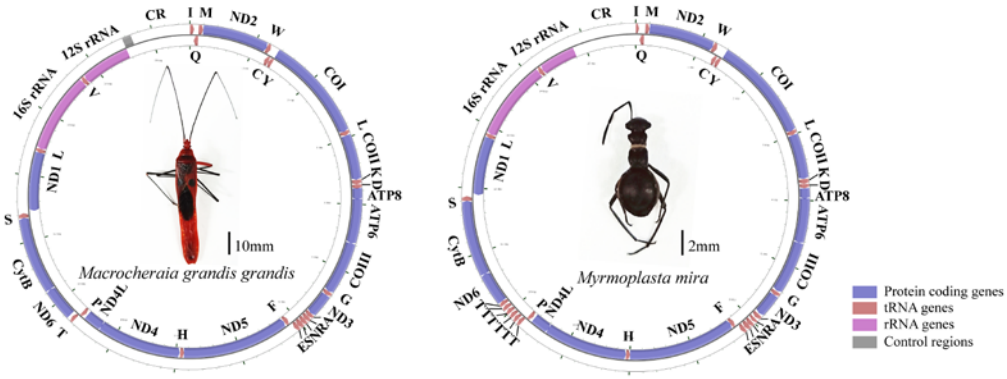

Figure 1. Mitochondrial genome maps of *Macrocheraia grandis grandis* (left) and *Myrmoplasta mira* (right).

Table 2. Organization of *Macrocheraia grandis grandis* mitochondrial genome

Name	Position	Strand	Length (bp)	Anticodon	Start codon	Stop codon	Intergenic nucleotides
tRNA-Ile	1-64	J	64	GAT			
tRNA-Gln	62-130	N	69	TTG			-3
tRNA-Met	130-198	J	69	CAT			-1
ND2	199-1179	J	981		ATA	TAA	0
tRNA-Trp	1178-1243	J	66	TCA			-2
tRNA-Cys	1236-1299	N	64	GCA			-8
tRNA-Tyr	1300-1363	N	64	GTA			0
COI	1365-2903	J	1539		TTG	TAA	1
tRNA-Leu(*UAA*)	2899-2963	J	65	TAA			-5
COII	2965-3637	J	673		ATG	T	1
tRNA-Lys	3638-3708	J	71	CTT			0
tRNA-Asp	3709-3771	J	63	GTC			0
ATPase8	3772-3927	J	156		ATA	TAA	0
ATPase6	3921-4586	J	666		ATG	TAA	-7
COIII	4586-5372	J	787		ATG	T	-1
tRNA-Gly	5373-5435	J	63	TCC			0
ND3	5436-5789	J	354		ATA	TAG	0
tRNA-Ala	5788-5849	J	62	TGC			-2
tRNA-Arg	5850-5914	J	65	TCG			0
tRNA-Asn	5915-5980	J	66	GTT			0
tRNA-Ser(*GCU*)	5980-6050	J	71	GCT			-1
tRNA-Glu	6050-6113	J	64	TTC			-1
tRNA-Phe	6112-6178	N	67	GAA			-2
ND5	6182-7873	N	1692		ATT	TAA	3
tRNA-His	7871-7934	N	64	GTG			-3
ND4	7940-9262	N	1323		ATG	TAG	5
ND4L	9256-9546	N	291		ATT	TAA	-7

Continued Table 2

Name	Position	Strand	Length (bp)	Anticodon	Start codon	Stop codon	Intergenic nucleotides
tRNA-Pro	9577-9638	N	62	TGG			30
tRNA-Thr	9649-9706	J	58	TGT			10
ND6	9761-10225	J	465		ATA	TAA	54
CytB	10225-11358	J	1134		ATG	TAA	-1
tRNA-Ser(*UGA*)	11357-11426	J	70	TGA			-2
ND1	11450-12376	N	927		ATA	TAA	23
tRNA-Leu(*UAG*)	12371-12437	N	67	TAG			-6
16S rRNA	12438-13696	N	1259				0
tRNA-Val	13697-13763	N	67	TAC			0
12S rRNA	13764-14554	N	791				0
Control region	14555-14917		363				0

The mitogenome of *Myrmoplasta mira* is a double-stranded circular DNA molecule. The nearly complete mitogenome 14,806 bp in length was sequenced without the sequence of the 5' terminal of 12S rRNA, control region and 5' terminal of *tRNA-Ile*. The mitogenome of *Myrmoplasta mira* contains 42 genes, including the typical 37 genes and five more duplications of *tRNA-Thr* (Fig. 1, Table 3). Ten gene overlaps were found. The longest two overlaps (7 bp) are same (ATGATAA) located between *ATP6*/*ATP8* and *ND4*/*ND4L*. Twelve intergenic spacers with a total of 211 nucleotides were found, ranging in size from 1 bp to 27 bp. Between the six duplications of *tRNA-Thr*, the five intergenic sequences were identical (CAAACAACCCTA).

Table 3. Organization of *Myrmoplasta mira* mitochondrial genome

Name	Position	Strand	Length (bp)	Anticodon	Start codon	Stop codon	Intergenic nucleotides
tRNA-Ile	1-49	J	49	GAT			
tRNA-Gln	61-129	N	69	TTG			11
tRNA-Met	142-207	J	66	CAT			12
ND2	208-1191	J	984		ATC	TAA	0
tRNA-Trp	1190-1258	J	69	TCA			-2
tRNA-Cys	1262-1324	N	63	GCA			3
tRNA-Tyr	1330-1394	N	65	GTA			5
COI	1403-2941	J	1539		TTG	TAA	8
tRNA-Leu(*UAA*)	2937-3002	J	66	TAA			-5
COII	3003-3678	J	676		ATG	T	0
tRNA-Lys	3682-3753	J	72	CTT			3
tRNA-Asp	3756-3818	J	63	GTC			2
ATPase8	3819-3974	J	156		ATT	TAA	0
ATPase6	3968-4636	J	669		ATG	TAA	-7
COIII	4636-5424	J	789		ATG	TAA	-1
tRNA-Gly	5434-5497	J	64	TCC			9

Continued Table 3

Name	Position	Strand	Length (bp)	Anticodon	Start codon	Stop codon	Intergenic nucleotides
ND3	5498-5851	J	354		ATA	TAA	0
tRNA-Ala	5866-5931	J	66	TGC			14
tRNA-Arg	5948-6011	J	64	TCG			16
tRNA-Asn	6015-6080	J	66	GTT			3
tRNA-Ser(*GCU*)	6080-6149	J	70	GCT			-1
tRNA-Glu	6156-6220	J	65	TTC			6
tRNA-Phe	6222-6288	N	67	GAA			1
ND5	6286-7996	N	1711		ATT	T	-3
tRNA-His	7997-8060	N	64	GTG			0
ND4	8061-9381	N	1321		ATG	T	0
ND4L	9375-9659	N	285		ATT	TAA	-7
tRNA-Pro	9687-9751	N	65	TGG			27
tRNA-Thr	9766-9827	J	62	TGT			14
tRNA-Thr	9840-9901	J	62	TGT			12
tRNA-Thr	9914-9975	J	62	TGT			12
tRNA-Thr	9988-10049	J	62	TGT			12
tRNA-Thr	10062-10123	J	62	TGT			12
tRNA-Thr	10136-10197	J	62	TGT			12
ND6	10198-10680	J	483		ATA	TAA	0
CytB	10680-11807	J	1128		ATG	TAG	-1
tRNA-Ser(*UGA*)	11806-11875	J	70	TGA			-2
ND1	11893-12823	N	931		ATA	T	17
tRNA-Leu(*UAG*)	12818-12883	N	66	TAG			-6
16S rRNA	12884-14148	N	1265				0
tRNA-Val	14149-14214	N	66	TAC			0
12S rRNA	14215-14806	N	592				0

Rearrangements were found in mitogenomes of both *Myrmoplasta mira* and *Macrocheraia grandis grandis*. The *tRNA-Thr* and *tRNA-Pro* switch their positions, and this is the same as in *Physopelta gutta* and *Dysdercus cingulatus*, the two other species in Pyrrhocoroidea reported by Hua *et al.* (2008). Moreover, in the mitogenome of *Myrmoplasta mira*, six duplications of *tRNA-Thr* were found located downstream of *tRNA-Pro*. The gene orders between *ND4* and *CytB* of the four sequenced mitogenomes in Pyrrhocoroidea are shown in Fig. 2.

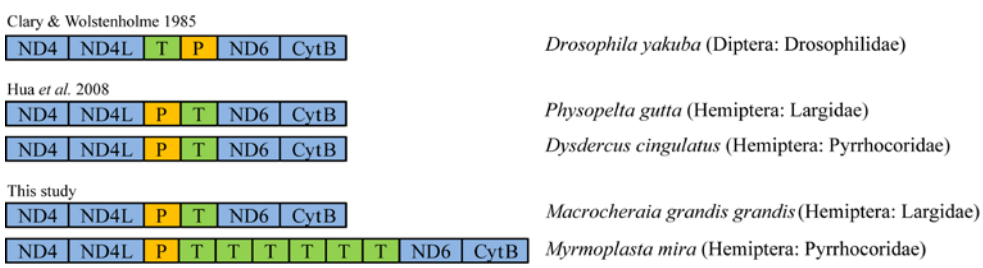

Figure 2. The rearrangements and duplications of *tRNA-Thr*(*T*) and *tRNA-Pro*(*P*) in Pyrrhocoroidea.

Transfer RNAs

The secondary structures of tRNAs in *Macrocheraia grandis grandis* and *Myrmoplasta mira* are shown in Figs. 3, 4. According to the algorithm of the Mitos WebServer, most tRNAs could be folded into the classic cloverleaf secondary structure, except for *tRNA-Thr* in *Macrocheraia grandis grandis*, in which the TψC stem simply formed a loop. Furthermore, among the six duplications of *tRNA-Thr* in *Myrmoplasta mira*, the three nucleotides in the 3' terminal are different between the last one (ACU) and the previous five (CUC). The lengths of the amino acid acceptor stem, anticodon stem and anticodon loop are conserved. In contrast, the DHU arm and TψC arm are more variable in size.

Ribosomal RNAs

In the mitogenomes of *Macrocheraia grandis grandis* and *Myrmoplasta mira*, the 16S rRNA genes are located between *tRNA-Leu* and *tRNA-Val*. And the 12S rRNA genes are located between *tRNA-Val* and the control region. The secondary structures of 16S rRNA both consist of six domains (domain III absent in Arthropoda) and 43 helices (Figs. S1, S2). The secondary structure of 12S rRNA in *Macrocheraia grandis grandis* contains three domains and 28 helices(Fig. S3). The secondary structure of 12S rRNA in *Myrmoplasta mira* contains three domains and 19 helices (Fig. S4).

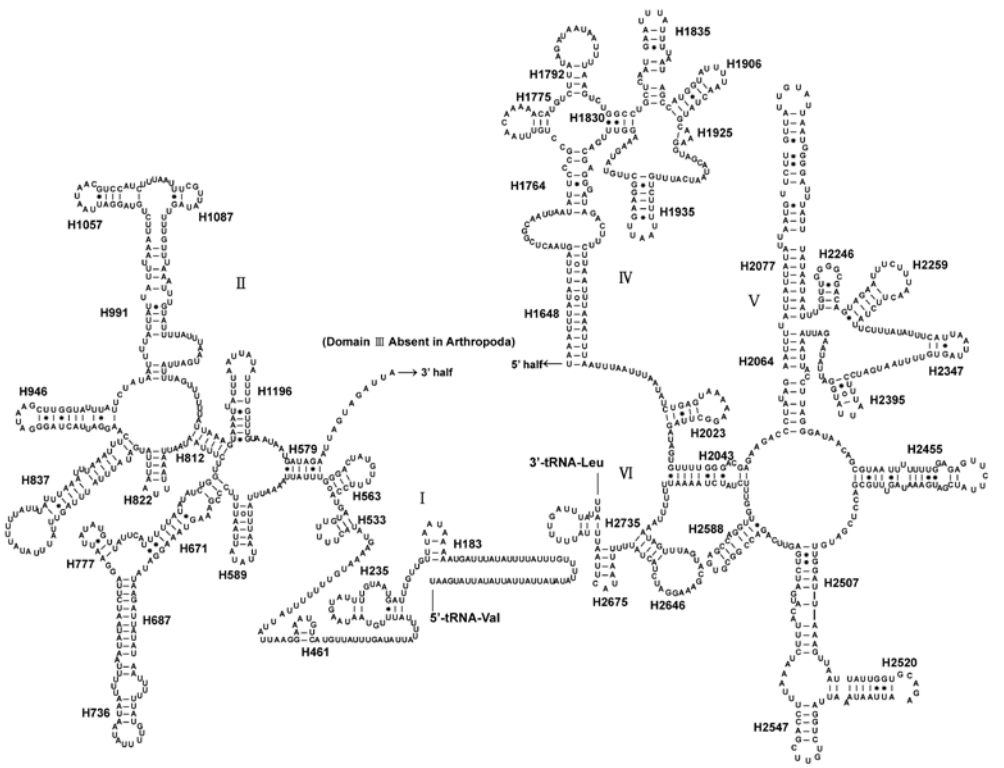

Figure S1. Putative secondary structure of 16S rRNA in *Macrocheraia grandis grandis*.

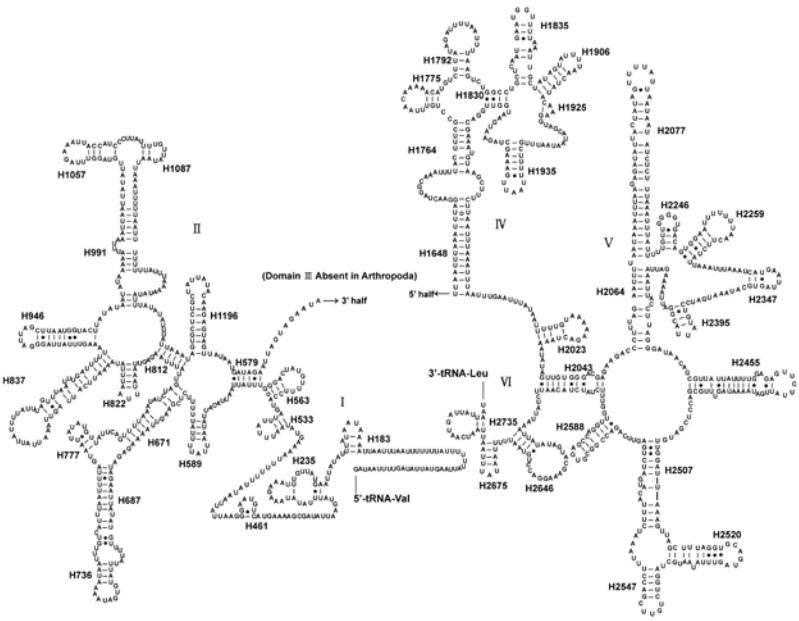

Figure S2. Putative secondary structure of 16S rRNA in *Myrmoplasta mira*.

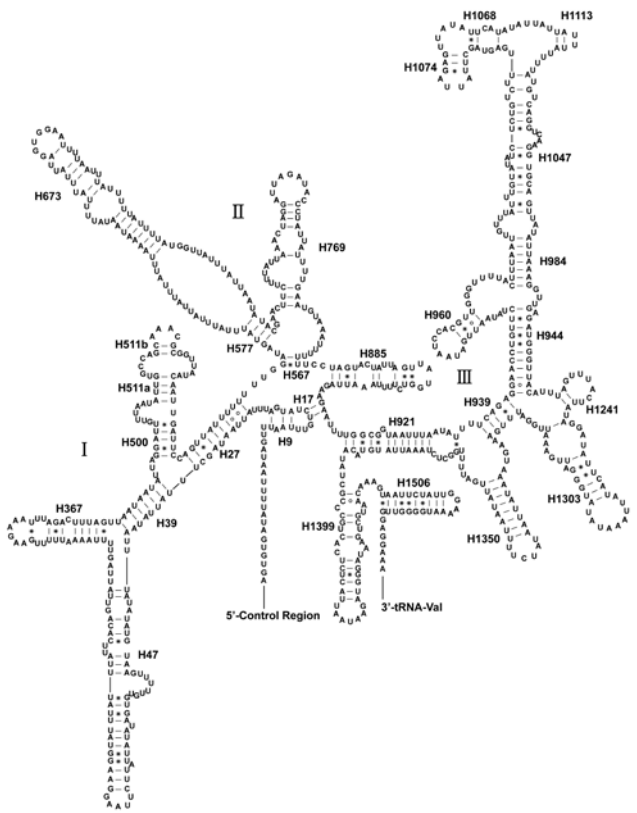

Figure S3. Putative secondary structure of 12S rRNA in *Macrocheraia grandis grandis*.

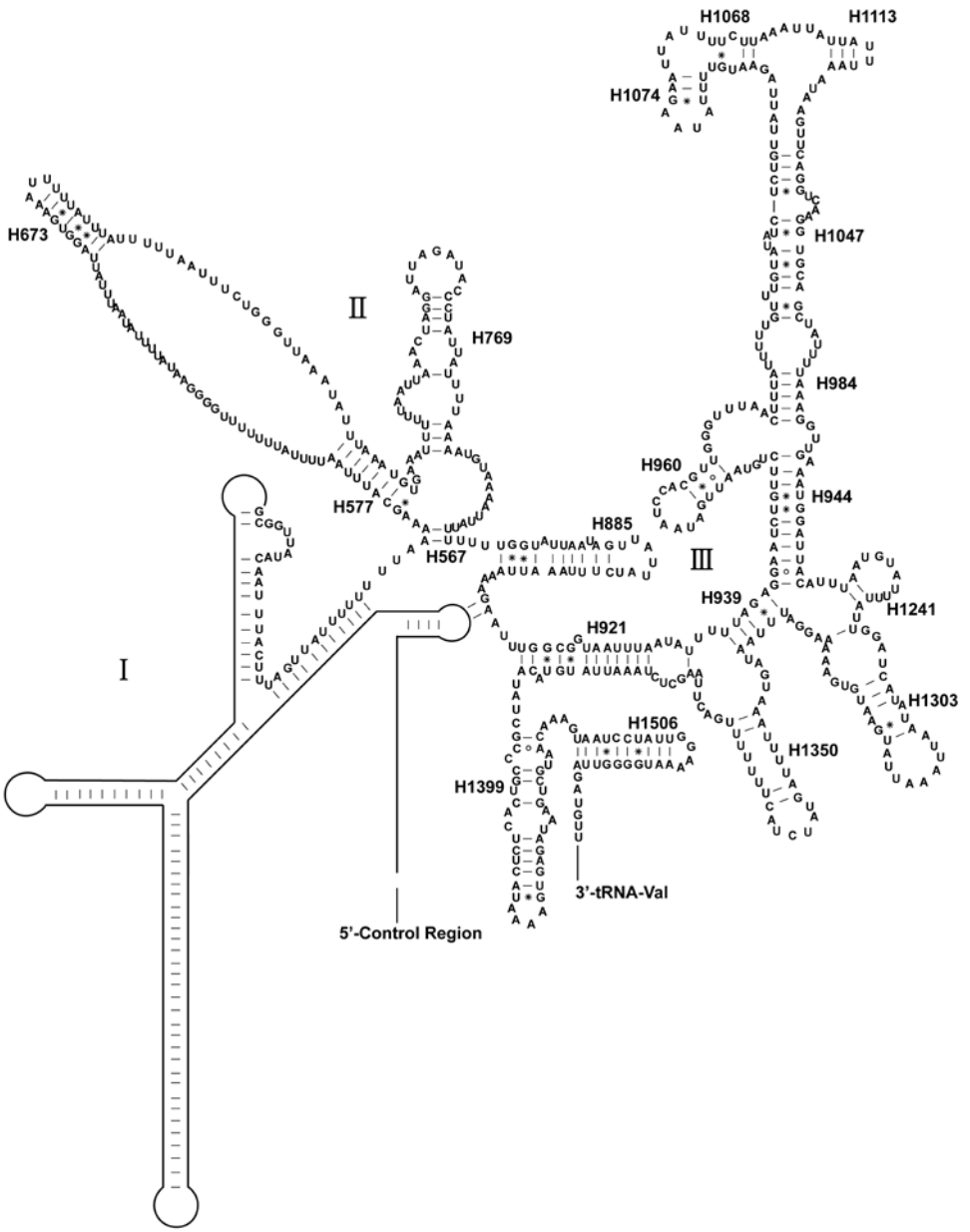

Figure S4. Putative secondary structure of 12S rRNA in *Myrmoplasta mira*.

Protein coding genes

In each mitogenome of *Macrocheraia grandis grandis* and *Myrmoplasta mira*, all 13 PCGs were found. Twelve of the 13 PCGs are initiated with start codons ATN, but the *COI* genes are initiated with TTG (Tables 2, 3). Most PCGs terminated with the common stop codons TAA or TAG, but *COII*, *ND5*, *ND4*, *ND1* in *Myrmoplasta mira* and *COII*, *COIII* in *Macrocheraia grandis grandis* are terminated with a single T. These unusual start codons and stop codons have also been observed in many other true bugs, such as *Stictopleurus subviridis*

(Hemiptera: Rhopalidae) (Hua *et al.* 2009), *Dimorphopterus japonicus* (Hemiptera: Blissidae) (Li 2014) and *Lethocerus indicus* (Hemiptera: Belostomatidae) (Devi *et al.* 2016).

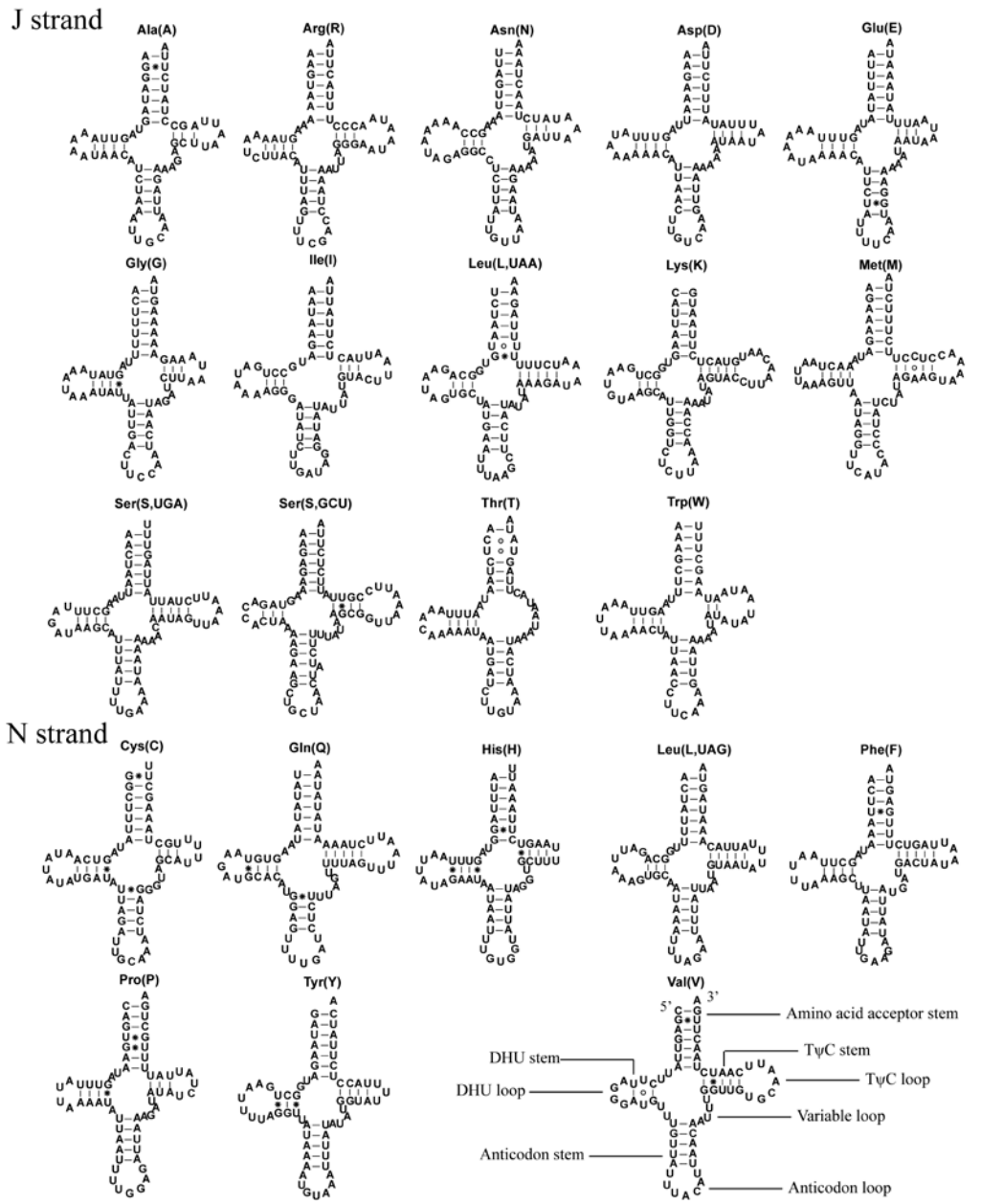

Figure 3. Putative secondary structures of tRNAs in the *Macrocheraia grandis grandis.* Typical Waston-Crick bonds are illustrated by shot lines; GU bonds by asterisk.

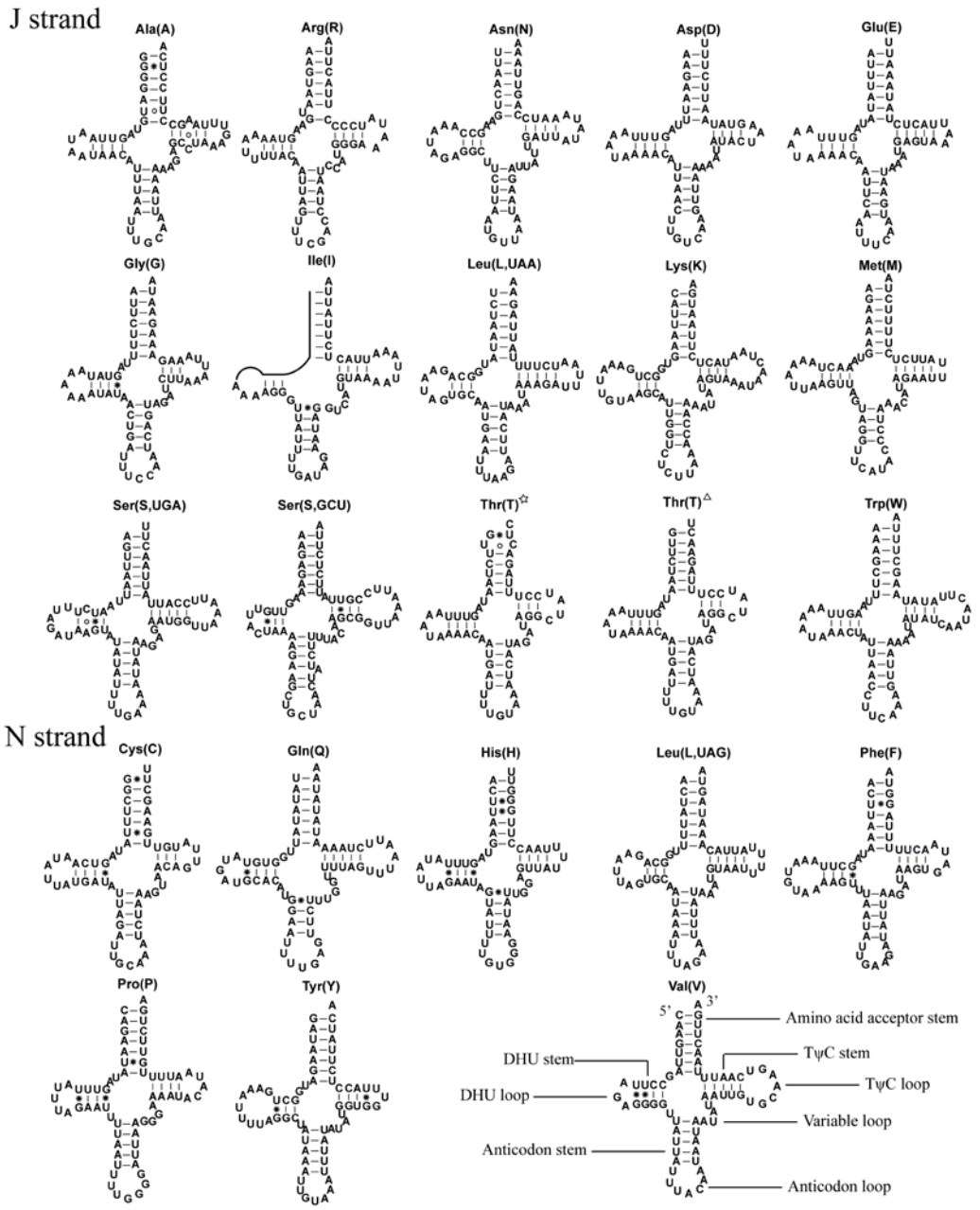

Figure 4. Putative secondary structures of tRNAs in the *Myrmoplasta mira* mitogenomes. The homologous nucleotides which were unsequenced in *Myrmoplasta mira* are shown by a line. Typical Waston-Crick bonds are illustrated by shot lines; GU bonds by asterisk. This kind of notation is also suitable for the secondary structures of 12S rRNAs and 16S rRNAs described in the supplementary materials.

Nucleotide composition and codon usage

For the mitogenomes of *Macrocheraia grandis grandis* and *Myrmoplasta mira*, the A+T

contents are 75.5% and 75.2% (Table 4). In *Macrocheraia grandis grandis*, the whole genome, rRNA genes, J strand PCGs and tRNA genes, and N strand tRNA genes are AT-skewed and CG-skewed. N strand PCGs are TA-skewed and GC-skewed. In *Myrmoplasta mira*, the whole genome, rRNA genes, J strand PCGs, N strand PCGs and tRNA genes are all AT-skewed and CG-skewed. J strand tRNA genes are AT-skewed and GC-skewed.

No significant differences of AT content were observed among the first (71.4% and 76.1%), second (77.6% and 70.1%) and third (75.8% and 77.1%) codon positions either in *Macrocheraia grandis grandis* or *Myrmoplasta mira*. Synonymous codon usage of twenty amino acids is shown in Fig. S5. For most amino acids, the most frequently used codons are NNA and NNU, which reflects the A/T bias of nucleotide composition.

Table 4. Nucleotide composition of *Macrocheraia grandis grandis* and *Myrmoplasta mira*

Feature	T%	C%	A%	G%	Length (bp)	A+T%	AT-skew	GC-skew
Macrocheraia grandis grandis								
Whole genome	30.1	15.3	45.4	9.2	14917	75.5	0.20	-0.25
PCGs	41.6	12.5	33.4	12.6	10988	75.0	-0.11	0.00
PCGs-J	33.3	15.5	40.4	10.8	6755	73.7	0.10	-0.18
PCGs-N	54.8	7.7	22.3	15.3	4233	77.0	-0.42	0.33
First codon position	40.5	14.3	30.9	14.3	3663	71.4	-0.13	0.00
Second codon position	46.6	11.1	31.0	11.3	3663	77.6	-0.20	0.01
Third codon position	37.5	12.1	38.3	12.1	3662	75.8	0.01	0.00
tRNA genes	34.8	12.6	43.1	9.5	1441	77.9	0.11	-0.14
tRNA genes-J	34.8	11.0	43.5	10.7	917	78.3	0.11	-0.02
tRNA genes-N	34.7	15.5	42.4	7.4	524	77.1	0.10	-0.35
rRNA genes	31.5	14.5	46.5	7.6	2050	78.0	0.19	-0.31
Myrmoplasta mira								
Whole genome	32.5	15.2	42.7	9.6	14806	75.2	0.14	-0.23
PCGs	31.7	15.7	42.7	9.9	11026	74.4	0.15	-0.23
PCGs-J	36.1	15.6	37.2	11.1	6778	73.3	0.01	-0.17
PCGs-N	24.6	15.9	51.6	7.9	4248	76.2	0.36	-0.34
First codon position	28.8	13.8	47.3	10.0	3676	76.1	0.24	-0.16
Second codon position	31.2	17.9	38.9	12.1	3675	70.1	0.11	-0.19
Third codon position	35.1	15.5	42.0	7.5	3675	77.1	0.09	-0.35
tRNA genes	35.7	13.2	40.6	10.5	1747	76.3	0.06	-0.11
tRNA genes-J	35.9	11.5	40.5	12.1	1222	76.4	0.06	0.03
tRNA genes-N	35.2	17.1	40.8	6.9	525	76.0	0.07	-0.43
rRNA genes	34.3	13.8	44.3	7.5	1857	78.6	0.13	-0.29

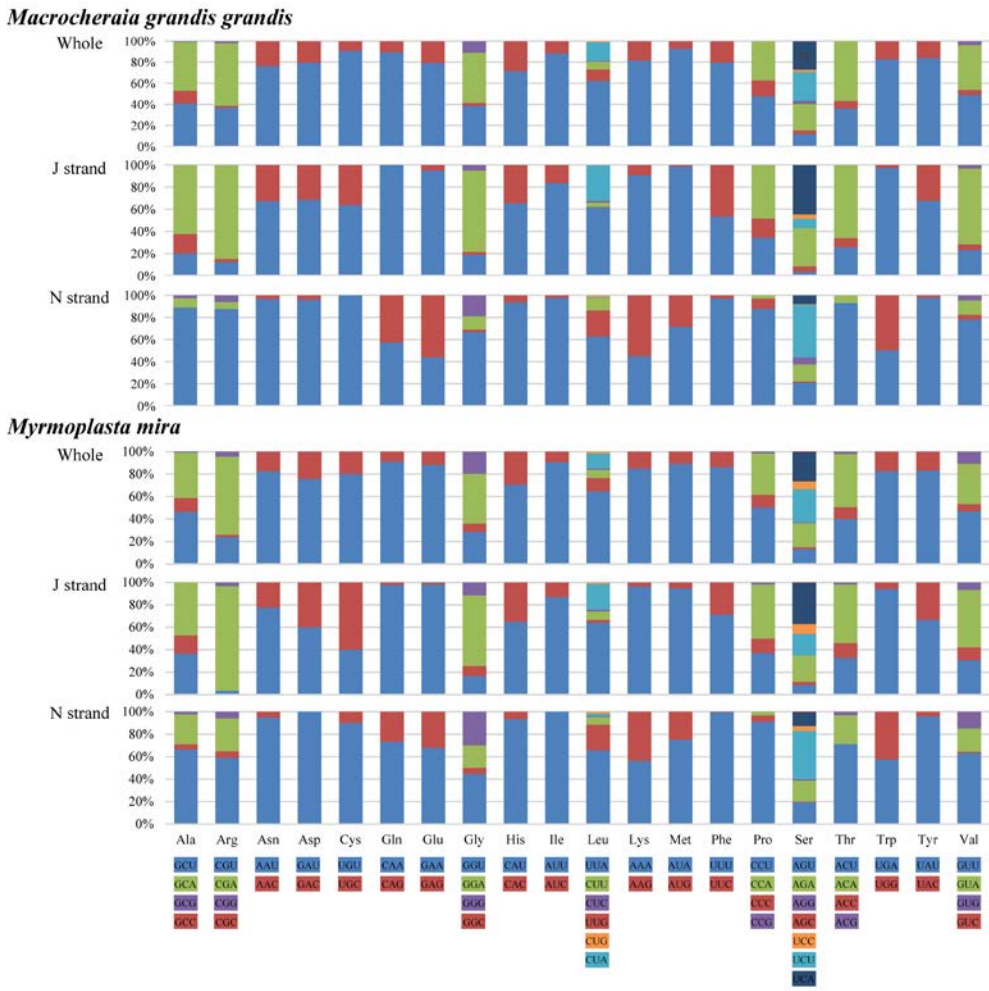

Figure S5. Percentage of synonymous codon usage of each amino acid.

Discussion

The nearly complete mitogenomes of *Macrocheraia grandis grandis* and *Myrmoplasta mira* were sequenced in this study. Both of them contain the translocation of *tRNA-Thr* and *tRNA-Pro*, which is the same as that of the other two mitogenomes in Pyrrhocoroidea (*Dysdercus cingulatus* and *Physopelta gutta*). In the mitogenome of *Myrmoplasta mira*, an undiscovered gene rearrangement in Heteroptera was found. It contains 27 tRNA genes with six duplications of *tRNA-Thr*, which are located downstream of *tRNA-Pro.*

Tandem duplication-random loss is the most commonly supposed model to explain the similar phenotypes in insects (Hua *et al.* 2008; Song *et al.* 2016; Ye *et al.* 2016). We assumed that in the mitogenome of the recent common ancestor of Pyrrhocoroidea (Fig. 5a), *tRNA-Thr* — *tRNA-Pro* duplicated to form an intermediate *tRNA-Thr* — *tRNA-Pro* — *tRNA-Thr* — *tRNA-Pro* (Fig. 5b), and then *tRNA-Thr* at the first set and *tRNA-Pro* at the last set were

randomly lost and form the current arrangement of the mitogenome in *Macrocheraia grandis grandis* (Fig. 5c). In *Myrmoplasta mira*, the generation of six nearly identical duplications of *tRNA-Thr* indicates that it is a recent event; the duplication of *tRNA-Thr* may have happen after the translocation of *tRNA-Thr* and *tRNA-Pro* (Fig. 5d).

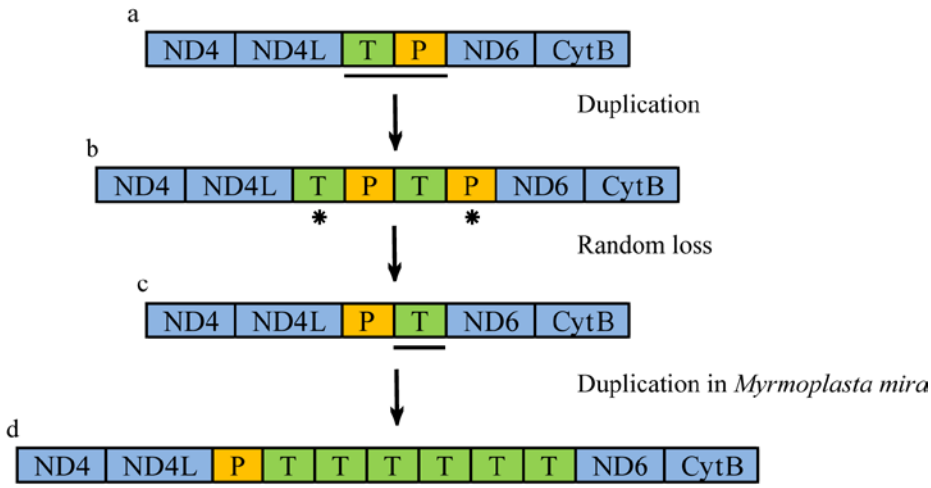

Figure 5. Putative process of rearrangements in the mitogenomes of Pyrrhocoroidea.

The translocation of *tRNA-Thr* and *tRNA-Pro* is according to the available knowledge an autapomorphy of Pyrrhocoroidea. While monophyly of the superfamily Pyrrhocoroidea and the family Pyrrhocoridae is well supported by several morphological synapomorphies (Henry 1997; Hemala *et al*. a, b) as well as molecular evidence (e.g., Sudakaran *et al*. 2015; Gordon *et al*. 2016). However, the monophyly of the family Largidae, consisting of two subfamilies, the New World Larginae and Old World Physopeltinae, is morphologically supported only by symplesiomorhic characters (Stehlík 2013; Hemala *et al*. a,b) and also its molecular support is also not significant (Sudakaran *et al*. 2015; Gordon *et al*. 2016). The unique translocation of *tRNA-Thr* and *tRNA-Pro* thus confirm close relationship of Physopeltinae and Pyrrhocoridae, but obtaining a mitogenome of the subfamily Larginae is badly needed to generalize the situation in Pyrrhocoroidea and to evaluate the importance of the discovered translocations. Other synapomorphies and autapomorphies of groups in Heteroptera may be hidden by limited information insofar as only one or two mitogenomes have been sequenced in more than half of the families. More mitogenomes need to be explored to acquire a more comprehensive view on the evolution of mitogenomes in Heteroptera.

Acknowledgements

We sincerely thank Jan A NEETHLING (National Museum, South Africa), Vera OPATOVA (University of California, United States of America) and Pavel JUST (Charles University, Czech Republic) for collecting the specimen of *Myrmoplasta mira*. The work of P. Kment was financially supported by the Ministry of Culture of the Czech Republic (DKRVO 2019/5.I.a, National Museum, 00023272).

References

Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M & Stadler PF. 2013. Mitos: Improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics & Evolution*, 69: 31–39.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K & Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, 10(1): 421.

Cannone JJ, Subramanian S, Schnare MN, Collett JR, Souza LMD, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, Pande N, Shang Z, Yu N & Gutell RR. 2002. The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, 3(1): 2.

Chen ZT & Du YZ. 2016. Rearrangement of mitochondrial genome in insects. *Journal of Environmental Entomology*, 38(4): 843–851.

Clary DO & Wolstenholme DR. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *Journal of Molecular Evolution*, 22(3): 252–271.

Devi KM, Shantibala T & Debaraj H. 2016. The first complete mitochondrial genome of a Belostomatidae species, *Lethocerus indicus*, the giant water bug: An important edible insect. *Gene*, 591(1): 108–118.

Gordon ERL, McFrederick Q & Weirauch C. 2016. Phylogenetic evidence for ancient and persistent environmental symbiont reacquisition in Largidae (Hemiptera: Heteroptera). *Applied and Environmental Microbiology*, 82: 7123–7133. http://dx.doi.org/10.1128/AEM.02114-16

Guo PL, Wu YZ, Wang YH, Wu HY, Li T, Li JL, Bu WJ & Xie Q. 2017. The complete mitochondrial genome of *Neolethaeus assamensis* (Hemiptera: Heteroptera: Rhyparochromidae) and the implication for the phylogenetic position of Rhyparochromidae in Lygaeoidea. *Journal of Environmental Entomology*, 39(2): 314–331.

Hemala V, Kment P & Malenovský I. in prep. a. Comparative morphology of adult pregenital abdominal ventrites and trichobothria in Pyrrhocoroidea (Hemiptera: Heteroptera: Pentatomomorpha).

Hemala V, Kment P, Tihlaříková E, Neděla V & Malenovský I. in prep. b. External structures of metathoracic scent gland efferent system in the true bug superfamily Pyrrhocoroidea (Hemiptera: Heteroptera: Pentatomomorpha).

Henry TJ. 1997. Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea. *Annals of the Entomological Society of America*, 90(3): 275–301.

Henry TJ. 2017. Biodiversity of Heteroptera. *In*: Foottit RG & Adler PH (Eds.), *Insect biodiversity: Science and society. Volume I, Second Edition*. Wiley-Blackwell, WarnockPro by SPi Global, Chennai, India, pp. 279–335.

Hua JM, Dong PZ, Li M, Cui Y, Bu WJ & Xie Q. 2009. The analysis of mitochondrial genome of *Stictopleurus subviridis* Hsiao (Insecta, Hemiptera-Heteroptera, Rhopalidae). *Acta Zootaxonomica Sinica*, 34(1): 1–9.

Hua JM, Li M, Dong PZ, Cui Y, Xie Q & Bu WJ. 2008. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). *BMC Genomics*, 9(1): 610.

Li H, Shi A, Song F & Ca WZ. 2016. Complete mitochondrial genome of the flat bug *Brachyrhynchus hsiaoi* (Hemiptera: Aradidae). *DNA Sequence*, 27(1): 2.

Li T. 2014. *Phylogenomic Studies of Lygaeoidea (Hemiptera: Heteroptera) Based on the Mitochondrial Genomics*. Nankai University, Tianjin, 216 pp.

Li T, Gao CQ, Cui Y, Xie Q & Bu WJ. 2013. The complete mitochondrial genome of the stalk-eyed bug

Chauliops fallax Scott, and the monophyly of Malcidae (Hemiptera: Heteroptera). *PLoS ONE*, 8(2): e55381. DOI:10.1371/journal.pone.0055381.

Reineke A, Karlovsky P & Zebitz CPW. 1998. Preparation and purification of DNA from insects for AFLP analysis. *Insect Molecular Biology*, 7(1): 95–99.

Song F, Li H, Shao RF, Shi A, Bai XS, Zheng XR, Heiss E & Cai WZ. 2016. Rearrangement of mitochondrial tRNA genes in flat bugs (Hemiptera: Aradidae). *Scientific Reports*, 6(1): 25725.

Stehlík JL. 2013. Review and reclassification of the Old World genus *Physopelta* (Hemiptera: Heteroptera: Largidae). *Acta Entomologica Musei Nationalis Pragae*, 53(2): 505–584.

Sudakaran S, Retz F, Kikuchi Y, Kost C & Kaltenpoth M. 2015. Evolutionary transition in symbiotic syndromes enabled diversification of phytophagous insects on an imbalanced diet. *ISME Journal: Multidisciplinary Journal of Microbial Ecology*, 9(12): 2587–2604.

Wang YY, Liu XY, Garzon-Orduna IJ, Winterton SL, Yan Y, Aspock U, Aspock H & Yang D. 2016. Mitochondrial phylogenomics illuminates the evolutionary history of Neuropterida. *Cladistics*, DOI: 10.1111/cla.12186.

Xie YL, Wu GX, Tang JB, Luo RB, Patterson J, Liu SL, Huang WH, He GZ, Gu SC, Li SK, Zhou X, Lam TW, Li YR, Xu X, Wong GKS & Wang J. 2014. SOAPdenovo-Trans: *De novo* transcriptome assembly with short RNA-Seq reads. *Bioinformatics*, 30: 1660–1666.

Ye F, Lan X, Zhu WB & You P. 2016. Mitochondrial genomes of praying mantises (Dictyoptera, Mantodea): rearrangement, duplication, and reassignment of tRNA genes. *Scientific Reports*, DOI: 10.1038/srep25634.