Journal of Insect Science, (2018) 18(2): 41; 1–8
doi: 10.1093/jisesa/iey035
Research

OXFORD

The Complete Mitochondrial Genome of the Plant Bug *Lygus pratensis* Linnaeus (Hemiptera: Miridae)

Yao Tan,[1] Bing Jia,[1] Yuan-ming Chi,[1] Hai-bin Han,[2] Xiao-rong Zhou,[1] and Bao-ping Pang[1,3]

[1]Research Center for Grassland Entomology, Inner Mongolian Agricultural University, Hohhot 010019, China, [2]Institute of grassland research, Chinese Academy of Agricultural Science, Hohhot 010019, China, and [3]Corresponding author, e-mail: pangbp@imau.edu.cn

Subject Editor: Igor Sharakhov

Received 19 January 2018; Editorial decision 16 March 2018

Abstract

Lygus pratensis is a phytophagous pest responsible for yield losses in Bt alfalfa and other economic crops in Northwestern China. To better characterize Miridae at the genomic level, the complete mitochondrial (mt) genome of *L. pratensis* was sequenced and analyzed in this study. The mt genome was amplified via the polymerase chain reaction to generate overlapping fragments. These fragments were then sequenced, spliced, and analyzed to include the examination of nucleotide composition, codon usage, compositional biases, protein-coding genes (PCGs), and RNA secondary structures. Phylogenetic relationships between *L. pratensis* and other species in different Heteroptera families were also examined. The mt genome was found to be a typical circular genome with a length of 16,591 bp and a total AT content of 75.1%, encoded for 13 PCGs, 22 transfer RNAs (tRNAs), 2 ribosomal RNAs (lrRNA and srRNA), and a noncoding control region. The nucleotide composition of the entire mt genome was heavily biased toward A and T. All of the tRNAs were predicted to have classic clover leaf structures, but three of the tRNAs ($tRNA^{Asn}$, $tRNA^{His}$, $tRNA^{His}$) were missing the TΨC loop. The control region (2,017 bp), which was found to be located between 12S and $tRNA^{Ile}$, contained three tandem repeat elements. Phylogenetic analyses showed that *L. pratensis* is closely related to the other three examined *Lygus* bugs, and that it is a sister group to *Apolygus* and *Adelphocoris*. This study confirms the usability of the mt genome in phylogenesis studies pertaining to the *Lygus* genus, within Miridae.

Key words: *Lygus pratensis*, mitogenome, phylogenetic relationship, Miridae

Miridae, which belongs to the order Hemiptera, is one of the most diverse families of insects and covers over 1,200 genera and 11,000 species (Cassis and Schuh 2012). Some of species have been associated with serious economic losses of Bt crops in developed and developing countries, alike, including the United States (Bauer et al. 2006), Australia (Fitt et al. 1994), Canada (May et al. 2003), and China (Guo et al. 2005, Wu and Guo 2005, Lu et al. 2008). *Lygus pratensis* Linnaeus, which is the most common species found in Northwestern China, is one of five species affecting Bt crops in China (Lu et al. 2010). In recent years, *L. pratensis* has adversely affected alfalfa and other economic crops, and hindered the expansion of the dairy industry in Northwestern China. At present, very few studies have reported on the ecology, phenology, or managements of this pest (Li et al. 2007b, Lu et al. 2010). Furthermore, no molecular markers have been used to investigate *L. pratensis* population genetics structure or evolutionary patterns, with this insight potentially facilitating forecasting and management protocols.

In insects, mitochondrial (mt) genomes have been extensively studied for marker discovery (Cameron 2014). The insect mt genome consists of a circular, double-stranded molecule that is approximately 14–20 kb and contains a highly conserved exon arrangement comprising 37 genes. The insect mitogenome commonly includes 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), and 2 ribosomal RNA genes rRNAs, srRNA (12S), and lrRNA (16S) (Boore 1999, Taanman 1999, Cameron 2014). Additionally, the genome contains a major noncoding region known as the A+T-rich region that plays a role in the initiation of DNA replication and transcription (Clayton 1992, Wolstenholme 1992). In recent years, the numbers of complete insect mt genomes have increased. mt genomes are of interest due to their small size, rapid evolution, maternal inheritance, low rate or lack of recombination, and relatively conserved gene content and organization (Gissi et al. 2008, Timmermans et al. 2010). These genomes are widely used as a maternal molecular marker for taxonomy, molecular phylogenetics, population genetics, and phylogeography (Hwang et al. 2001, Ma et al. 2009, Chen et al. 2012). However, phylogenetic analysis based on the current taxon has been limited, thus potentially restricting the usefulness of an mtDNA genome based phylogenetic study (Zhao et al. 2015). Although researchers

© The Author(s) 2018. Published by Oxford University Press on behalf of Entomological Society of America.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
For commercial re-use, please contact journals.permissions@oup.com

have conducted some phylogenetic research on Miridae, studies have mainly focused on morphological analyses, with little molecular evidence presented (Schuh et al. 2009).

At present, the *L. pratensis* mt genome has not been previously reported. Herein, the complete *L. pratensis* mitogenome was cloned, sequenced, identified, and analyzed. Moreover, gene annotations, mitogenomic structural analysis, nucleotide composition, codon usage, tRNA secondary structure, noncoding and control region. Analyses were also performed. The research presented herein will offer a foundation for examining *L. pratensis* molecular evolution and provide further insight into the phylogenetic relationships among Hemiptera species.

Materials and Methods

Specimen Collection

Adult *L. pratensis* species were collected from alfalfa fields in the Helin County Experimental Station at the Institute of Grassland Research (40°60′N, 111°80′E), Chinese Academy of Agricultural Science (Hohhut, Inner Mogolian Province, China) in May 2016. All of the samples were quickly preserved in 100% ethanol at −20°C until arrival at the Inner Mongolia Agricultural University (40°48′N, 111°42′E).

DNA Extraction, PCR Amplification, and Sequencing

Genomic DNA was extracted from eight adults using a DNeasy DNA Extraction kit (AiDeLai Biotech. Co. Ltd, Beijing, China) and mtDNA amplified via the polymerase chain reaction (PCR). Primers were designed to target generally conserved regions and amplified short fragments from *16s*, *12s*, *cox1*, *atp6*, *cox3*, *nad4*, *cytB*, and *nad* (Table 1). The PCR reaction was carried out with LA Taq polymerase (Takara, Japan) for 35 cycles at 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min/1 kb for the extension. The final $MgCl_2$ concentration in the PCR reaction was 2.0 mmol/liter. The obtained amplicons were then purified, cloned into pMD18-T vectors (Takara Dalian Co., Ltd., Dalian, China), and sequenced.

Sequence Analysis, Annotation, and Secondary Structure Predictions

The sequencing data were assembled using BioEdit software as previously described (Hall 1999). Protein-coding regions were identified using ORF Finder on the NCBI website and PCG nucleotide sequences were translated based on invertebrate mtDNA sequences. Sequence alignment was carried out using ClustalX (Thompson et al. 1997). Composition skew analysis was performed to examine base composition, with asymmetry measured according to the formulas (AT skew = [A−T]/[A+T] and GC skew = [G−C]/[G+C]) (Perna and Kocher 1995). MEGA 6.0 was used to calculate base composition and codon usage.

Transfer RNA genes and their predicted secondary structures were identified using tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE/), with a few of the tRNA genes identified by comparing sequence homologies with tRNAs from other published plant bugs. Two rRNA genes were identified by Blastn on the NCBI website and

Table 1. Primers used for amplifying the complete mt genome

Fragment	Primer name	Nucleotide sequences (5′→3′)	Sequence size (bp)	Starting site–stop site
1	LLF1	ACTAAATCCGTAGAGATC	973	82–1055
	LLR1	CGAAAGTGTCACTATTGATG		
2	LLF2	GAATAACAGAAAAGTTATC	1,069	891–1960
	LLR2	CGGTGATTCCCACCGATC		
3	LLF3	CATAATGGAGCATCAGTAG	1,108	1796–2904
	LLR3	CAGTATATGAATGTTCAGC		
4	LLF4	GATTACCCTGATAGATTC	1,300	2708–4008
	LLR4	ATGCTTGTTGCTGGGTCG		
5	LLF5	CCCAATTTGATGAACAAC	1,171	3823–4994
	LLR5	TCCTTTAGGTGGTCATAG		
6	LLF6	GATGACGAGATATCGTACG	1,051	4803–5854
	LLR6	GTTATTGTTAACTACATCC		
7	LLF7	TTCTAATCTTCGATATCG	1,242	5670–6911
	LLR7	TACCTTATACATCTTCTTG		
8	LLF8	CTCACAGTGAATATGACAG	970	6746–7716
	LLR8	TGCTGATTTCTTCAATGG		
9	LLF9	TTCGGTTAGTCATAACAG	1,494	7470–8964
	LLR9	GAATGTAGATTAATTCCTAC		
10	LLF10	TGCCATCAACATTGAAAG	1,261	8754–10015
	LLR10	GTAAGTCTTATCAATATCC		
11	LLF11	CAAACCCTACTTACATCAG	1,297	9833–11130
	LLR11	AATTGATGCAATCATGGC		
12	LLF12	GAATCAAGGATTATATAGG	1,322	10909–12231
	LLR12	TGCAACCATTTTCTGATGG		
13	LLR13	AGTCTGAGCAACAGCACG	1,091	11950–13041
	LLF13	GATCTAACCTGCCCTATG		
14	LLR14	GCTCTATAGGGTCTTCTCG	980	12842–13822
	LLF14	GTCGCTCCCAACATATGTTG		
15	LLR15	CTGATACAAAAGGTAAGC	994	13515–14509
	LLF15	GTTTTATTCATGCATATGG		
16	LLR16	GTATAACCGCAACTGCTGG	2,530	14328–1659181
	LLR16	CTAATAACAATAATGGTTC		

then validated by performing a homologous gene alignment against other plant bugs. Furthermore, the entire control region was examined for tandem repeats using the Tandem Repeats Finder program (Benson 1999).

Phylogenetic Analyses

Phylogenetic analyses were performed based on 14 complete representative bug mt genomes from GenBank, and 11 Miridae species as follows: *Lygus hesperus* (NC_024641), *Lygus lineolaris* (NC_021975), *Lygus rugulipennis* (KJ170898), *L. pratensis* (MF188255), *Lygus lucorum* (NC_023083), *Adelphocoris fasciaticollis* (KJ001714), *Adelphocoris lineolatus* (KJ020286), *Adelphocoris nigritylus* (KJ020287), *Adelphocoris suturalis* (KJ020288), *Trigonotylus caelestialium* (KJ170899), and *Nesidiocoris tenuis* (NC_022677). The remaining species belonged to the Anthocoridae, Reduviidae, and Nabidae families in Hemiptera and included *Orius niger* (NC_012429), *Agriosphodrus dohrni* (NC_015842), and *Gorpis annulatus* (NC_019595). DNA alignments were inferred based on sequences from all 14 of the PCGs and concatenated using ClustalX. Phylogenetic relationships were examined using the neighbor joining (NJ), maximum likelihood (ML), and Bayesian inference (BI) methods. An NJ tree was constructed using MEGA 5.0 with 1,000 bootstrap replicates. ML analysis was conducted using RAxML 7.0 (Stamatakis 2006), with an MtREV model for amino acid sequences used to optimize the topology. Node support was calculated by acquiring bootstrap values from heuristic searches of 1,000 resampled data sets, using the rapid bootstrap feature in RAxML (Stamatakis et al. 2008). BI analysis was carried out using MrBayes 3.2.2 (Ronquist and Huelsenbeck 2003), with two simultaneous runs of 1,500,000 generations conducted for the matrix, with each one sampled every 100 generations with a burn-in of 25%. Stationarity was considered to be reached when the average SD of split frequencies was below 0.01 (Huelsenbeck et al. 2001).

Results

Features of the *L. pratensis* mt Genome

The complete *L. pratensis* mt genome is 16,591 bp double-stranded circular molecule (Fig. 1), containing 37 genes (13 PCGs, 22 tRNAs, and 2 rRNAs) and a noncoding region (control region). Twenty-three of the genes were located on the majority strand (J strand, direction: F), while the remaining were on the minority strand (N strand, direction: R; Table 2). In the mitogenome, genes were observed to be overlapped, contiguous, or separated by a few nucleotides. Gene overlaps were observed at 11 gene junctions, with the longest overlap (14 bp) present between the termination of ATPase6 and the initiation of COIII. There were seven spacer sequences in the *L. pratensis* mitogenome, with the longest intergenic spacer sequence (7 bp) located between $tRNA^{Ser}$ and ND1. The entire *L. pratensis* mt genome sequence was deposited in GenBank under accession number MF188255.

Protein-Coding Genes

The PCGs (11,087 bp) comprised 66.83% of the whole mt genome. Thirteen of the genes were strictly initiated with an ATN start codon to include ATT (6), ATG (5), ATA (1), and ATC (1). The conventional stop codon TAN was also observed to include TAA (9), TAG (1). Furthermore, three of the PCGs (COI, COIII, and ND3) were terminated with a single T as the stop codon (Table 2).

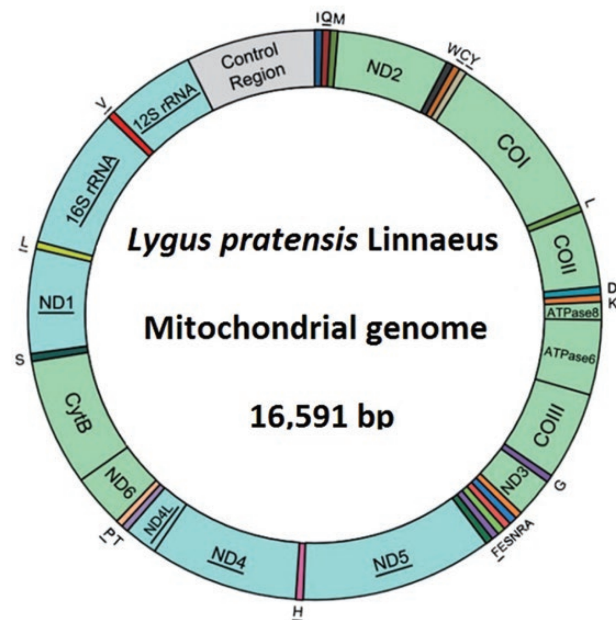

Fig. 1. Structure of the *Lygus pratensis* mt genome. The color blocks contain tRNAs that are labeled according to IUPAC-IUB single-letter amino acid codes. Genes that are underlined undergo transcription from left to right, while those without underlines proceed from right to left.

Base Composition and Codon Usage

The nucleotide composition of the *L. pratensis* mt genome was heavily biased toward A (42.7%) and T (32.4%), while much lower levels were seen for G (10.3%) and C (14.6%). The mt genome had an overall high A+T content of 75.1% with content levels of 74.7% in PCGs, 77.5% in tRNA genes, 79.1% in rRNA genes, and 72.2% in the control region (Table 3). When performing skew statistics on the PCGs, it was clearly shown that PCGs on the J strand were TA-skewed, while the N strand was AT-skewed.

The nucleotide AT bias was also reflected in the codon usage. When analyzing the base composition of 13 PCGs, the relative synonymous codon usage (RSCU) values showed that at the third codon position, A or T appeared substantially more often than did G or C (Fig. 2). While the four most used codons included TTT (F), TTA (L), ATT (I), and ATA (I).

tRNA and rRNA Genes

Within the *L. pratensis* mt genome, 22 tRNA genes were identified, with 21 identified using tRNAscan-SE and 1 by sequence comparison with other Hemipterans. All of the identified tRNAs had lengths between 62 bp ($tRNA^{Gly}$, $tRNA^{Ala}$, and $tRNA^{Phe}$) and 70 bp ($tRNA^{Lys}$; Table 2) and were predicted to have classic clover leaf structures. Furthermore, TΨC loops were found to be missing in $tRNA^{Asn}$, $tRNA^{His}$, and $tRNA^{His}$. Additionally, 22 unmatched base pairs were predicted within 11 of the tRNAs, with 20 of them being located in the AA arm (9 bp), the TΨC stem (3 bp), or DHU stem (8 bp). The remaining 2 unmatched base pairs included a U-U mismatch in the AA stem of $tRNA^{Ala}$, and a U-C mismatch in the AA stem of $tRNA^{Arg}$ (Fig. 3).

In addition to the identified tRNAs, two rRNAs were also identified. The lrRNA gene (1,250 bp) was found to be located between the $tRNA^{Val}$ and $tRNA^{Leu}$ genes, while the srRNA gene (1,250 bp) was located between the $tRNA^{Val}$ gene and the control region. The

Table 2. Organization of the *Lygus pratensis* mt genome

Gene	Direction	Strand	Position (bp)	Anticodon	Size (bp)	Start codon	Stop codon	Intergenic nucleotide (bp)
tRNA-Ile (I)	F	J	1–65	GAT	65	–	–	0
tRNA-Gln (Q)	R	N	63–131	TTG	69	–	–	–3
tRNA-Met (M)	F	J	132–199	CAT	68	–	–	0
ND2	F	J	200–1,207	–	1,008	ATT	TAA	0
tRNA-Trp (W)	F	J	1,206–1,270	TCA	65	–	–	–2
tRNA-Cys (C)	R	N	1,262–1,325	GCA	63	–	–	–8
tRNA-Tyr (Y)	R	N	1,325–1,391	GTA	66	–	–	0
COI	F	J	1,392–2,925	–	1,534	ATG	T-	0
tRNA-Leu (L)	F	J	2,926–2,991	TAA	66	–	–	0
COII	F	J	2,992–3,672	–	681	ATC	TAA	0
tRNA-Lys (K)	F	J	3,673–3,743	CTT	70	–	–	0
tRNA-Asp (D)	F	J	3,745–3,810	GTC	65	–	–	2
ATPase8	F	J	3,809–3,971	–	162	ATT	TAG	0
ATPase6	F	J	3,965–4,648	–	684	ATG	TAA	–7
COIII	F	J	4,634–5,418	–	784	ATG	T-	–14
tRNA-Gly (G)	F	J	5,419–5,481	TCC	62	–	–	0
ND3	F	J	5,482–5,832	–	352	ATA	T-	0
tRNA-Ala (A)	F	J	5,834–5,895	TGC	62	–	–	1
tRNA-Arg (R)	F	J	5,899–5,965	TCG	67	–	–	3
tRNA-Asn (N)	F	J	5,964–6,028	GTT	65	–	–	–2
tRNA-Ser (S)	F	J	6,028–6,096	GCT	69	–	–	–1
tRNA-Glu (E)	F	J	6,096–6,159	TTC	64	–	–	–1
tRNA-Phe (F)	R	N	6,158–6,219	GAA	62	–	–	–2
ND5	R	N	6,220–7,918	–	1699	ATT	TAA	0
tRNA-His (H)	R	N	7,919–7,983	GTG	64	–	–	0
ND4	R	N	7,984–9,311	–	1,327	ATG	TAA	0
ND4L	R	N	9,303–9,608	–	306	ATT	TAA	–7
tRNA-Thr (T)	F	J	9,611–9,678	TGT	68	–	–	2
tRNA-Pro (P)	R	N	9,679–9,742	TGG	64	–	–	0
ND6	F	J	9,744–10,232	–	489	ATT	TAA	1
CytB	F	J	10,234–11,367	–	1134	ATG	TAA	1
tRNA-Ser (S)	F	J	11,368–11,433	TGA	67	–	–	–1
ND1	R	N	11,441–12,367	–	927	ATT	TAA	7
tRNA-Leu (L)	R	N	12,368–12,431	TAG	64	–	–	0
16S-rRNA (lrRNA)	R	N	12,432–13,681	–	1250	–	–	0
tRNA-Val (V)	R	N	13,682–13,750	TAC	69	–	–	0
12S-rRNA (srRNA)	R	N	13,751–14,575	–	824	–	–	0
Control region	–	–	14,576–16,591	–	2,017	–	–	0

Table 3. *Lygus pratensis* mt genome nucleotide composition and skew

Feature	%A	%C	T%	G%	A+T%	AT[a] skew	GC[b] skew	No. of nucleotide
Whole genome	42.7	14.6	32.4	10.3	75.1	0.14	–0.17	16591
PCGs	43.8	10.4	30.9	14.9	74.7	0.17	0.18	11087
PCGs (J)	39.2	15.2	33.8	11.8	73.0	0.07	–0.13	6828
PCGs (N)	43.6	7.0	22.4	12.5	66.0	0.32	0.28	4979
tRNA genes	41.3	10.1	36.2	12.4	77.5	0.07	0.10	1427
tRNA genes (J)	41.1	10.8	36.5	11.6	77.6	0.06	0.10	919
tRNA genes (N)	41.7	15.2	35.9	7.3	77.6	0.07	–0.35	521
rRNA genes	44.2	12.8	34.9	8.1	79.1	0.12	–0.22	2074
Control region	36.8	16.2	35.4	11.7	72.2	0.02	–0.16	2017

[a]AT skew = (A–T)/(A+T).
[b]GC skew = (G–C)/(G+C).

AT content was 79.1% for these rRNA genes, with lrRNA having a content level of 79.8% and srRNA a level of 78.2%.

Noncoding and Overlapping Regions

The mt noncoding region was found to include short intergenic spacers and a long control region. The *L. pratensis* mitogenome harbors a total of 17 intergenic spacer sequences that vary from 1 to 7 bp in length and are distributed in seven regions. It also contains a long control region (2,017 bp) that is located between the 12S-rRNA and tRNA-Ile-Gln-Met gene cluster. The control region was found to have an AT content of 72.19%, with a positive AT skew and negative GC skew noted. Moreover, the complete control region contains three tandem repeats and includes a short 24 bp sequence tandemly repeated four times; a 25 bp sequence tandemly repeated three times,

Fig. 2. Relative synonymous codon usage (RSCU) within the *Lygus pratensis* mt genome.

with a partial forth; and a 61 bp sequence tandemly repeated eight times (Fig. 4). For the overlapping sequences, a total of 48 bp were identified, throughout 11 regions and ranged in size from 1 to 14 bp (Table 2).

Phylogenetic Analysis

Phylogenetic associations were examined among four Cimicomorpha families, with 14 mt genome sequences utilized. The NJ, ML, and BI analyses all basically generated the same tree topologies (Fig. 5). These findings showed that: 1) *L. pratensis* is clustered in the same clade as three of the *Lygus* species; 2) a sister relationship between *A. lucorum* and several *Lygus* species was highly supported; 3) among the 11 Miridae mtDNA sequences, 9 species formed a monophyletic; and 4) within Cimicomorpha, Miridae is a sister group to the Anthocoridae (*O. niger*), Reduviidae (*A. dohrni*), and Nabidae (*G. annulatus*) families.

Discussion

In the present study, the complete *L. pratensis* mt genome (16,591 bp) was cloned, assembled, and analyzed. Similar to most insects, this mt genome contained the standard 13 PCGs, 22 tRNA genes, 2 rRNA genes, and a control region (Zhang et al. 2013). It also shares the same gene arrangement as most of the sequenced Heteroptera mt genomes (Song et al. 2016). Among the *Lygus* mitogenomes reported in GenBank, *L. hesperus* (17,747 bp) is the largest, while *L. rugulipennis* (15,819 bp) is the smallest. All of the identified PCGs have conventional start codons (ATN), and nine have conventional stop codons (TAN). However, three had stops with a single T, with this phenomenon of an incomplete termination codon found in some insects, and sometimes at high frequencies (Liao et al. 2010), and this is presumed to be generated by posttranscriptional polyadenylation (Lee et al. 2006, Cha et al. 2007).

The nucleotide composition showed an AT bias, similar to that seen in the metazoan mt genome, and it is similar to the common strand biases of other bug mtDNA. In the *L. pratensis* mitogenome, 13 PCGs showed higher TA and GC skews in the N strands than in the J strands, while the J strand tRNAs had a higher GC skew than the N strands. This observation is most likely attributed to asymmetrical mutational constraints during replication (Hassanin et al. 2005, Min and Hickey 2007). Furthermore, the nucleotide AT bias was also reflected in the high AT usage, within codons, with TTT (F), TTA (L), ATT (I), and ATA (I) being the most prevalent.

All of the tRNAs were predicted to form classic clover leaf structures. While the TΨC stem, TΨC loop, DHU stem, and DHU loop were often variable, the AA arms (7 bp), anticodon loops (7 bp), and anticodon stems (5 bp) were invariant. In other insects, some tRNAs have been reported to form DHU arms with one typical loop and deficient DHU stem (Li et al. 2007a, Song and Liang 2009), but herein, the *L. pratensis* DHU arms were predicted with complete DHU stems and loops. However, mismatched and aberrant loops do exist in some tRNAs (Song et al. 2013). As to whether the function of these abnormal tRNAs is affected, some researchers have proposed that a posttranscriptional RNA editing mechanism might maintain the function of these tRNA genes (Tomita et al. 2002).

Within the mt genome, rRNAs are highly conserved and evolve slowly. It is difficult to precisely determine the ends of rRNA genes, and herein, the ends were assumed to extend to the boundaries of adjacent genes (Boore 2006); however, for the srRNA gene, the 3′ end was determined by comparison with other reported Hemipteran mitogenomes. As is seen in other Miridae species (Wang et al. 2014a), the *L. pratensis* mitogenome was found to harbor two rRNA genes (lrRNA and srRNA), but further examination into their secondary structures is still required.

The noncoding region located between 12S and *tRNA* ^{Ile} was annotated as the control region, which is involved in original replication and contains the presumed transcriptional initiation (Zhang et al. 1995). This is one of the most variable regions within the mt genome, and contains high rates of nucleotide substitution, insertion, and deletion, with some varied tandemly repeated elements also present at times. The control region is the largest noncoding region in the *L. pratensis* mt genome (2,017 bp), with the 5′ end adjoining the 12S rDNA and the 3′ end adjoining to *tRNA* ^{Ile}. The varying size of the control region is due to repeat unit copy number variations (Zhang and Hewitt 1997, Wang et al. 2014b). At present, intragenus variations in mtDNA control region, sequences appear to

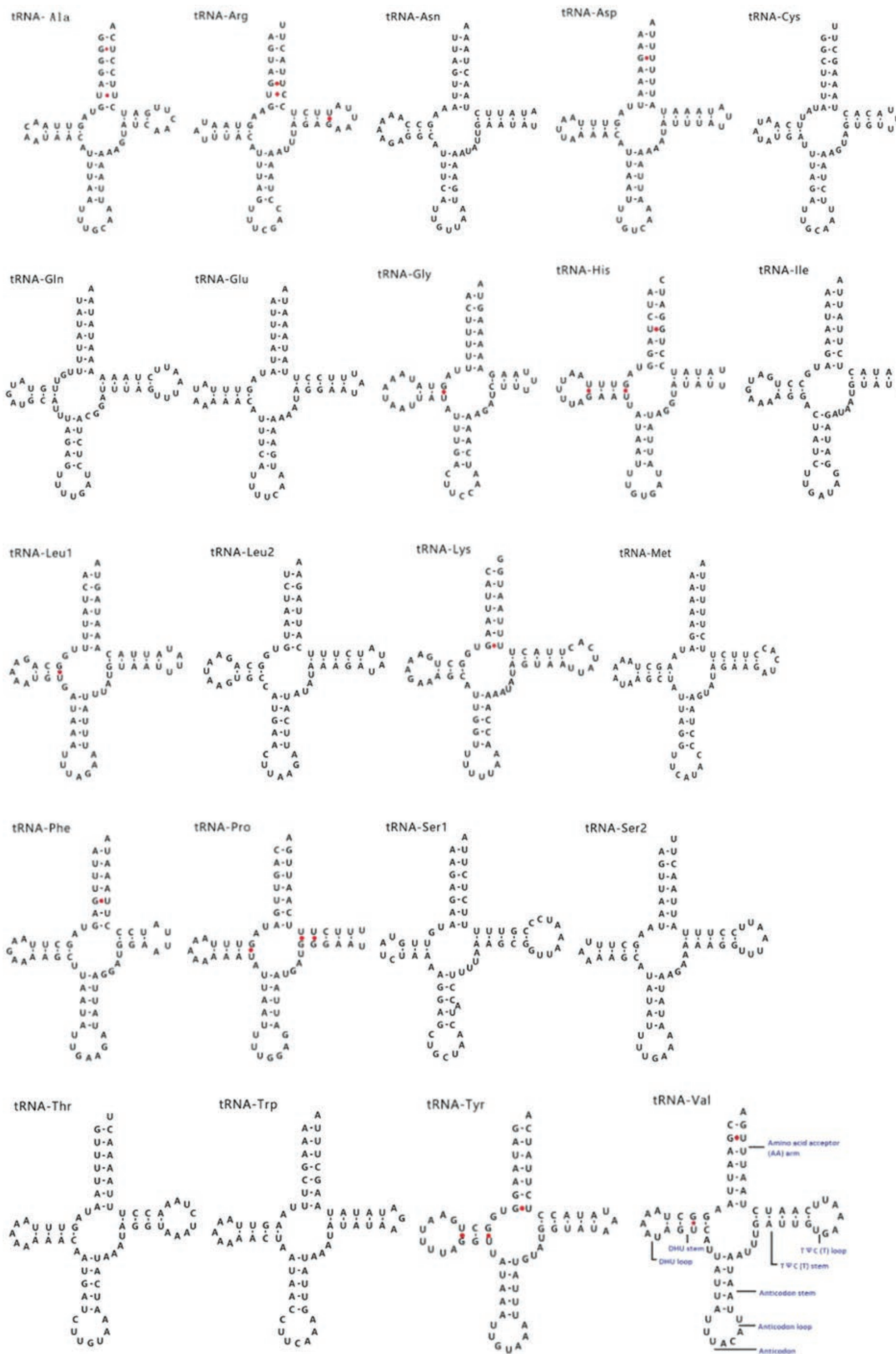

Fig. 3. Predicted tRNA secondary structures. Lines (-) indicate Watson–Crick base pairing, with unmatched base pairs indicated by red dots (●).

be attributed to distinct structural and evolutionary characteristics, including variable size, conserved structural elements, and tandem repeat sequences. This has been observed in the '*Adelphocoris*' and '*Lygus*' genera within the Miridae (Wang et al. 2014a), as well as other true bugs in Hemiptera (Dai et al. 2012; Li et al. 2012, 2013; Song et al. 2013). It has also been suggested that the existence of tandem repeats might be due to replication slippage (Fumagalli et al. 1996). These findings suggest that the variable control region sizes and abundance of tandem repeat elements may be helpful for evolutionary and population genetics studies of this genus.

Phylogenetic analyses clearly revealed a close relationship between *L. pratensis* and three other *Lygus* bugs, and it was identified

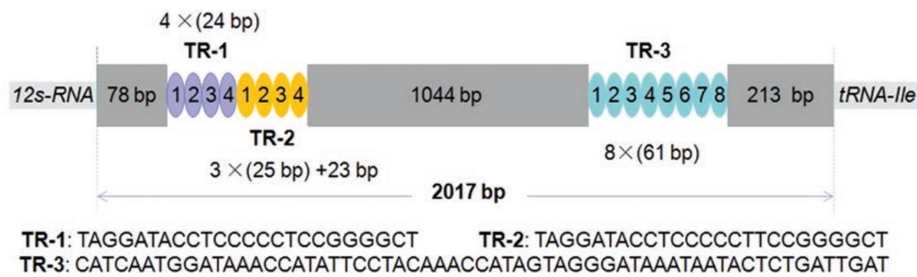

Fig. 4. *Lygus pratensis* control region organization. The locations and copy numbers of tandem repeats are indicated with colored ovals, while the remaining sequences are depicted with gray rectangles.

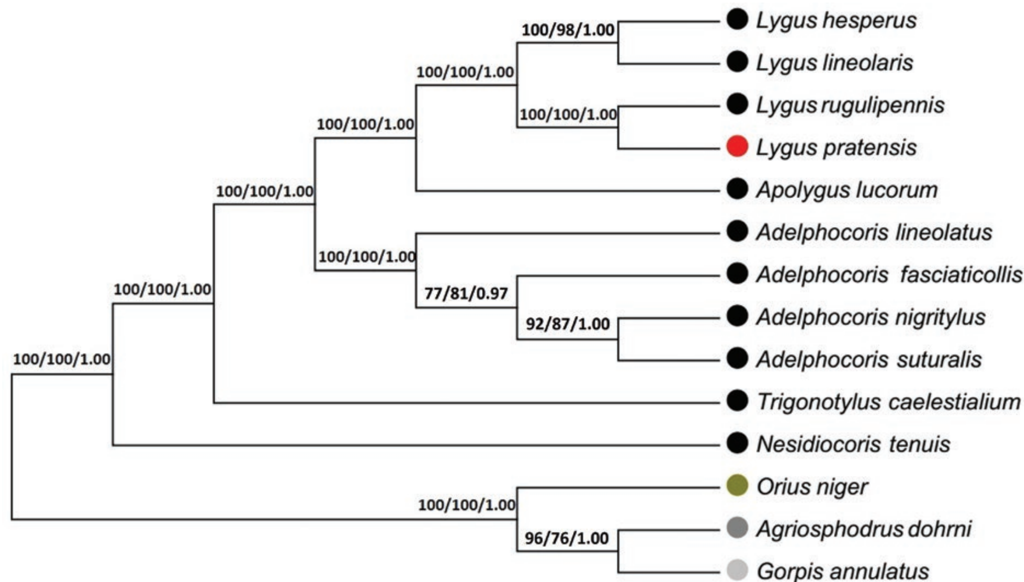

Fig. 5. Phylogenetic tree construction inferred from 14 PCGs from 11 Miridae species and 3 out-group Miridae. Numbers close to the nodes are bootstrap support values from the NJ analysis, ML analysis, and Bayesian posterior probability analyses. Miridae species ●; Anthocoridae species ●; Reduviidae species ●; Nabidae species ●; *Lygus pratensis* ●.

as a sister group to *Apolygus* and *Adelphocoris*. Moreover, Miridae was found to be a sister group to three other families within Cimicomorpha, but more taxonomic examination is still required. The findings presented herein are consistent with previous studies (Wang et al. 2014a, b), and contribute an additional mt genome sequence that can be utilized in Miridae genealogical classifications, and further shows why mtDNA is an ideal molecular marker for evolutionary and phylogenesis studies.

In conclusion, the present study confirmed the usability of mtDNA in a phylogenesis study of the *Lygus* genus, within Miridae.

Acknowledgments

This research was supported by grants from the National Science Foundation of China (31560519). We would also like to sincerely thank the anonymous reviewers for their comments and recommendations. We would like to thank LetPub (www.letpub.com) for providing linguistic assistance during the preparation of this manuscript. Y.T. and B.-P.P. conceived and designed the experiments. B.J. and H.-B.H. collected and maintained the specimens. Y.T., B.J., and Y.-M.C. performed the experimentation. Y.T. and Y.-M.C. analyzed the data. Y.T., X.-R.Z., and B.-P.P. wrote and revised the manuscript.

References Cited

Bauer, P. J., D. D. McAlister, III, and M. E. Roof. 2006. Evidence that light stink bug damage does not influence open end yarn processing performance. J. Cotton Sci. 3: 161–167.

Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27: 573–580.

Boore, J. L. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27: 1767–1780.

Boore, J. L. 2006. The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda). BMC Genomics. 7: 182.

Cameron, S. L. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59: 95–117.

Cassis, G., and R. T. Schuh. 2012. Systematics, biodiversity, biogeography, and host associations of the Miridae (Insecta: Hemiptera: Heteroptera: Cimicomorpha). Annu. Rev. Entomol. 57: 377–404.

Cha, S. Y., H. J. Yoon, E. M. Lee, M. H. Yoon, J. S. Hwang, B. R. Jin, Y. S. Han, and I. Kim. 2007. The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). Gene. 392: 206–220.

Chen, M., L. L. Tian, Q. H. Shi, T. W. Cao, and J. S. Hao. 2012. Complete mitogenome of the Lesser Purple Emperor *Apatura ilia* (Lepidoptera: Nymphalidae: Apaturinae) and comparison with other nymphalid butterflies. Zool. Res. 33: 191–201.

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

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

Fitt, G. P., C. L. Mares, and D. J. Llewellyn. 1994. Field evaluation and potential ecological impact of transgenic cottons (*Gossypium hirsutum*) in Australia. Biocontrol Sci. Tech. 4: 535–548.

Fumagalli, L., P. Taberlet, L. Favre, and J. Hausser. 1996. Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. Mol. Biol. Evol. 13: 31–46.

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

Guo, J. Y., H. X. Zhou, F. H. Wan, X. J. Liu, and Z. J. Han. 2005. Population dynamics and damage of *Lygus lucorum* in Bt cotton fields under two control measures. Chinese Bull. Entomol. 42: 424–428.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41: 95–98.

Hassanin, A., N. Léger, and J. Deutsch. 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inferences. Syst. Biol. 54: 277–298.

Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science. 294: 2310–2314.

Hwang, U. W., C. J. Park, T. S. Yong, and W. Kim. 2001. One-step PCR amplification of complete arthropod mitochondrial genomes. Mol. Phylogenet. Evol. 19: 345–352.

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

Li, X., K. Ogoh, N. Ohba, X. Liang, and Y. Ohmiya. 2007a. Mitochondrial genomes of two luminous beetles, *Rhagophthalmus lufengensis* and *R. ohbai* (Arthropoda, Insecta, Coleoptera). Gene. 392: 196–205.

Li, H. B., K. M. Wu, Y. Xu, X. R. Yang, J. Yao, and F. Wang. 2007b. Population dynamics of the pest mirids in cotton field in southern Xinjiang. Chinese Bull. Entomol. 44: 219–222.

Li, H., H. Liu, L. Cao, A. Shi, H. Yang, and W. Cai. 2012. The complete mitochondrial genome of the damsel bug *Alloeorhynchus bakeri* (Hemiptera: Nabidae). Int. J. Biol. Sci. 8: 93–107.

Li, T., C. Gao, Y. Cui, Q. Xie, and W. Bu. 2013. The complete mitochondrial genome of the stalk-eyed bug *Chauliops fallax* Scott, and the monophyly of Malcidae (Hemiptera: Heteroptera). PLoS One. 8: e55381.

Liao, F., L. Wang, S. Wu, Y. P. Li, L. Zhao, G. M. Huang, C. J. Niu, Y. Q. Liu, and M. G. Li. 2010. The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). Int. J. Biol. Sci. 6: 172–186.

Lu, Y. H., F. Qiu, H. Q. Feng, H. B. Li, Z. C. Yang, K. A. G. Wyckhuys, K. M. Wu. 2008. Species composition and seasonal abundance of pestiferous plant bugs (Hemiptera: Miridae) on Bt cotton in China. Crop Prot. 27: 465–472.

Lu, Y., K. Wu, Y. Jiang, B. Xia, P. Li, H. Feng, K. A. Wyckhuys, and Y. Guo. 2010. Mirid bug outbreaks in multiple crops correlated with wide-scale adoption of Bt cotton in China. Science. 328: 1151–1154.

Ma, C., C. Liu, P. Yang, and L. Kang. 2009. The complete mitochondrial genomes of two band-winged grasshoppers, *Gastrimargus marmoratus* and *Oedaleus asiaticus*. BMC Genomics. 10: 156.

May, W. E., J. J. Soroka, H. A. Loeppky, and D. C. Murrell. 2003. The effects of trichlorfon and deltamethrin on alfalfa plant bug and *Lygus* bug (Heteroptera: Miridae) populations in alfalfa grown in Canada. Crop Prot. 22: 883–889.

Min, X. J., and D. A. Hickey. 2007. DNA asymmetric strand bias affects the amino acid composition of mitochondrial proteins. DNA Res. 14: 201–206.

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

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

Schuh, R. T., C. Weirauch, and W. C. Wheeler. 2009. Phylogenetic relationships within the Cimicomorpha (Hemiptera: Heteroptera): a total evidence analysis. Syst. Entomol. 34: 15–38.

Song, N., and A. P. Liang. 2009. Complete mitochondrial genome of the small brown planthopper, *Laodelphax striatellus* (Delphacidae: Hemiptera), with a novel gene order. Zoolog. Sci. 26: 851–860.

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

Song, F., H. Li, R. Shao, A. Shi, X. Bai, X. Zheng, E. Heiss, and W. Cai. 2016. Rearrangement of mitochondrial tRNA genes in flat bugs (Hemiptera: Aradidae). Sci. Rep. 6: 25725.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22: 2688–2690.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. 57: 758–771.

Taanman, J. W. 1999. The mitochondrial genome: structure, transcription, translation and replication. Biochim. Biophys. Acta. 1410: 103–123.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882.

Timmermans, M. J., S. Dodsworth, C. L. Culverwell, L. Bocak, D. Ahrens, D. T. Littlewood, J. Pons, and A. P. Vogler. 2010. Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. Nucleic Acids Res. 38: e197.

Tomita, K., S. Yokobori, T. Oshima, T. Ueda, and K. Watanabe. 2002. The cephalopod *Loligo bleekeri* mitochondrial genome: multiplied noncoding regions and transposition of tRNA genes. J. Mol. Evol. 54: 486–500.

Wang, Y., H. Li, P. Wang, F. Song, and W. Z. Cai. 2014a. Comparative mitogenomics of plant bugs (Hemiptera: Miridae): identifying the AGG codon reassignments between Serine and Lysine. PLoS One. 9: 1–10.

Wang, P., H. Li, Y. Wang, J. H. Zhang, X. Dai, J. Chang, B. W. Hu, and W. Z. Cai. 2014b. The mitochondrial genome of the plant bug *Apolygus lucorum* (Hemiptera: Miridae): presently known as the smallest in Heteroptera. Insect Sci. 21: 159–173.

Wolstenholme, D. R. 1992. Genetic novelties in mitochondrial genomes of multicellular animals. Curr. Opin. Genet. Dev. 2: 918–925.

Wu, K. M., and Y. Y. Guo. 2005. The evolution of cotton pest management practices in China. Annu. Rev. Entomol. 50: 31–52.

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

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

Zhang, H. L., H. H. Zeng, Y. Huang, and Z. M. Zheng. 2013. The complete mitochondrial genomes of three grasshoppers, *Asiotmethis zacharjini*, *Filchnerella helanshanensis* and *Pseudotmethis rubimarginis* (Orthoptera: Pamphagidae). Gene. 517: 89–98.

Zhao, G., H. Li, P. Zhao, and W. Cai. 2015. Comparative mitogenomics of the assassin bug genus *Peirates* (Hemiptera: Reduviidae: Peiratinae) reveal conserved mitochondrial genome organization of *P. atromaculatus*, *P. fulvescens* and *P. turpis*. PLoS One. 10: e0117862.