Agri Gene 5 (2017) 27–36

Contents lists available at ScienceDirect

Agri Gene

journal homepage: www.elsevier.com/locate/aggene

The mitogenome of the brown pod-sucking bug *Clavigralla tomentosicollis* Stäl (Hemiptera: Coreidae)

Laura D. Steele[a,*], Weilin Sun[b], M. Carmen Valero[a,c], James Adebayo Ojo[d], Keon Mook Seong[a], Brad S. Coates[e], Venu M. Margam[f], Manuele Tamò[g], Barry R. Pittendrigh[b]

[a] Department of Entomology, University of Illinois, Urbana-Champaign, IL 61801, USA
[b] Department of Entomology, Michigan State University, East Lansing, MI 48824, USA
[c] Carl R. Woese Institute of Genomic Biology, University of Illinois at Urbana-Champaign, IL 61801, USA.
[d] Department of Crop Production, Kwara State University, Malete, Ilorin, Nigeria
[e] United States Department of Agriculture – Agricultural Research Service, Corn Insect and Crop Genetics Research Unit, Genetics Laboratory, Iowa State University, USA
[f] Aspire Food Group, Kumasi, Ghana
[g] International Institute of Tropical Agriculture, Cotonou, Benin

ARTICLE INFO

Keywords:
Clavigralla tomentosicollis
Pod-sucking bug
Cowpea
Vigna unguiculata
mitogenome
Coreoidea

ABSTRACT

The brown pod-sucking bug, *Clavigralla tomentosicollis* Stäl (Hemiptera: Coreidae), causes significant damage to cultivated cowpea, *Vigna unguiculata* Walp, a staple crop in sub-Saharan Africa. *C. tomentosicollis* pierce and suck sap from cowpea pods, resulting in reduced grain yield and quality. The complete, 16,089 bp mitogenome of *C. tomentosicollis* encodes 13 protein-coding genes (PCGs), 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs) and an A + T rich control region, with gene order and orientation identical to that of the insect ancestral gene order. The initiation and termination codons for the PCGs used standard ATN codons and TAA or TAG codons respectively. All predicted tRNAs fold into a clover-leaf secondary structures with the exception of tRNA-Ser (AGN) with a semi-loop dihydrouridine arm. The 1509 bps A + T rich region contains a single 89 bp tandem repeat unit duplicated 3.7 times. When compared with other published Coreoidea mitogenomes, *C. tomentosicollis* was also highly A − T skewed, and similar in both size and A − T%; however, its longer tandem repeat within the A + T rich region was unique. The *C. tomentosicollis* mitogenome can serve as a foundation to combine molecular marker data with pest monitoring strategies to better understand the population dynamics of this species.

1. Introduction

Cowpea (*Vigna unguiculata* Walp.) is a legume crop that serves as a major food staple in sub-Saharan Africa, grown for both human and animal consumption (Singh and Singh, 2015). Cowpea, owing to its high protein content (23–25%), drought tolerance and nitrogen-fixing ability, is an important crop in many developing nations (Singh and Singh, 2015; Devi et al., 2015). A pest insect complex in Africa causes severe damage to cowpea that can result in yield losses of up to 70% (Adati et al., 2008; Aliyu et al., 2007). Notable pest species include the legume pod borer (*Marcua vitrata* Fabricius), the cowpea aphid (*Aphis craccivora* Koch), the flower thrips (*Megalurothrips sjostedti* Trybom), the cowpea weevil (*Callosobruchus maculatus* Fabricius) (a storage pest), and a complex of pod-sucking insects including the brown pod-sucking bug *Clavigralla tomentosicollis* Stäl (Coreidae), *Clavigralla shadabi* Dolling (Coreidae), and *Riptortus dentipes* Fabricius (Alydidae) (Koona et al., 2004; Soyelu et al., 2007; Dreyer and Baumgartner, 1994). The most damaging of the pod-sucking bugs, *C. tomentosicollis* (Hemiptera: Coreidae) can dramatically decrease yields in cowpea crops (Dreyer and Baumgartner, 1994; Jackai, 1990; Koona et al., 2002), with levels of damage ranging from 20–100% (Singh and Allen, 1980; Aliyu et al., 2007). Both nymphs and adults suck the sap from the pods causing premature pod drying and shriveling (Jackai et al., 2001). Pest management efforts have included the planting of resistant cowpea varieties (Olatunde et al., 2007; Dabire-Binso et al., 2010), the use of botanical pesticides (Oparaeke, 2006a, 2006b) and biocontrol methods among others.

The amount of genetic data available for *C. tomentosicollis* is sparse,

Abbreviations: mitogenomes, mitochondrial genomes; *C. tomentosicollis*, *Clavigralla tomentosicollis*; ORF, Open Reading Frame; PCG, protein coding genes; tRNA, transfer ribonucleic acid; rRNA, ribosomal ribonucleic acid; lrRNA, large rRNA; srRNA, small rRNA; cox, cytochrome oxidase c subunit; nad, NADH Dehydrogenase subunit; ATP, *atp* synthase subunit; cytb, cytochrome *b*; PCR, polymerase chain reaction

* Corresponding author.
E-mail address: steele11@illinois.edu (L.D. Steele).

http://dx.doi.org/10.1016/j.aggene.2017.07.002
Received 26 March 2017; Received in revised form 19 June 2017; Accepted 10 July 2017
Available online 16 July 2017
2352-2151/ © 2017 Published by Elsevier Inc.

with one study that generated an annotated transcriptome sequence which identified genes of interest for pest control and potential molecular genetic markers (Agunbiade et al., 2013). Insect mitochondrial genomes (mitogenomes) are circular, double-stranded DNA molecules that range in size from 14–20 kb (Kim et al., 2009; Liu et al., 2015; Zhang et al., 2015). Mitogenomes encode 22 transfer RNA (tRNA) genes, two ribosomal RNA (large and small rRNA) genes (*lrRNA* and *srRNA*), which play a role in the translation of seven NADH dehydrogenase subunits (*nad1-6* and *nad4L*), two ATPase genes (*atp6* and *atp8*), three cytochrome *c* oxidase subunits (*cox1-3*), and a cytochrome B subunit (*cytb*) (Moritz et al., 1987; Wolstenholme, 1992; Cameron, 2014a). Insect mitogenomes also contain a large, non-coding, A − T-rich region which serves as the origin for replication and transcription (Wolstenholme, 1992; Zhang and Hewitt, 1997; Cameron, 2014a). Mitogenomes are inherited maternally and undergo very little sequence recombination, all of which are considered advantageous for phylogenetic, comparative, and evolutionary genomics (Harrison, 1989; Boore, 1999; Whinnett et al., 2005; Timmermans et al., 2014; Yuan et al., 2015b). Additionally, mitogenomes serve as molecular markers for molecular evolution and population genetics studies (Harrison, 1989; Boore, 1999; Whinnett et al., 2005; Timmermans et al., 2014; Abascal et al., 2006). The cytochrome *c* oxidase I (*cox1*) gene in particular has aided in differentiating cryptic species of butterflies (Herbert et al., 2004) and subspecies of *M. vitrata* (Margam et al., 2011a, 2011b). Out of the 146 complete hemipteran mitogenomes available on NCBI, there are six from the superfamily Coreoidea: *Riptortus pedestris* (Alydidae), *Aeschyntelus notatus* (Rhopalidae), *Stictopleurus subviridis* (Rhopalidae), *Corizus tetraspilus* (Rhopalidae), *Leptocorisa* sp. (Alydidae), and *Hydaropsis longirostris* (Coreidae), the only one from the family Coreidae (Hua et al., 2008; Tang et al., 2014; Yuan et al., 2015a; Wang et al., 2015).

Herein we describe the sequencing, assembly and annotation of the complete mitogenome of *C. tomentosicollis*, including a comparative analysis with the six other currently available Coreoidea mitogenomes. A complete mitogenome for *C. tomentosicollis* has the potential to provide an important tool for insect comparative and evolutionary genomics, as well as be a tool for understanding population dynamics of this species that might ultimately lead to better pest management decisions.

2. Materials and methods

2.1. Sampling, DNA extraction and sequencing

C. tomentosicollis samples were collected from field grown cowpea plants in Benin, preserved in RNAlater and stored at − 4 °C. The abdomen from a single insect sample was rinsed with 100% ethanol and then ground to a powder in liquid nitrogen. The samples were suspended in Buffer BP1 (Qiaprep Spin Miniprep Kit, QIAGEN, Valencia, CA, USA), incubated while spinning for 20 min, and run through a QiaShredder column (QIAGEN). DNA was extracted using the Qiaprep Spin Miniprep Kit according to the manufacturer's instructions, yielding a mitochondria-enriched sample. DNA quantity was determined using NanoDrop 1000 UV/VIS Spectrophotometer (Thermo Scientific, Serial No. G642) and quality determined by 0.9% agarose gel electrophoresis.

2.2. Illumina HiSeq2500™ sequencing and assembly

A shotgun genomic library was constructed using the Hyper Library Construction Kit (Kapa Biosystems, Wilmington, MA, USA). The library was loaded on a single lane and DNA was sequenced at 2 × 160 nt using a HiSeq SBS Kit v4 (Illumina, San Diego, CA, USA). Illumina shotgun sequencing was completed on an Illumina HiSeq2500 at the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois Urbana-Champaign. Resulting fastq files were input to CLC Genomics Workbench 8.0.2 (Cambridge, MA, USA) as paired reads (minimum distance 160 bp, maximum distance 825 bp), and trimmed for reads with quality score (q) < 20 and the two adapter sequences (5′-AGA TCG GAA GAG CAC ACG TCT GAA CTC CAG TCA C-3′; 5′-AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT-3′) removed. A de novo assembly was performed using the following parameters: mapping mode-create simple contig sequences (fast) contig length ≥ 200; automatic word size = yes; perform scaffolding = yes; auto-detect paired distances = yes.

2.3. Mitogenome annotation and analysis

All resulting contigs from the assembly of between 10,001–30,005 bps were extracted and used as blast queries (blastn; Match/Mismatch and Gap Costs = Match 2 Mismatch 3 Existence 5 Extension 2; Expectation value = 1.0E-6; Word size 11) against the mitogenomes of three other Coreoidea insect species: *A. notatus* (Accession number: NC_012446.1), *H. longirostris* (Accession number: NC_012456.1), and *R. pedestris* (Accession number: NC_012462.1). The BLAST results for contigs with an *E*-value of 0.00 were retained for further PCR validation. The NCBI ORF Finder (Open Reading Frame Finder) with invertebrate mitochondrial genetic codes was used to identify the protein coding genes (PCGs), which were confirmed via protein BLAST tool on NCBI (http://www.ncbi.nlm.nih.gov) and alignment with corresponding PCGs from six other Coreoidea (*A. notatus*, *H. longirostris*, *R. pedestris*, *Stictopleurus subivridis* [Accession number: NC_012888.1], *Corizus tetraspilus* [Accession number: KM983397], *Leptocorisa* sp. [Accession number: KM244663.1]). Three PCGs (*nad6*, *nad4L*, and *cox3*) were verified using PCR and additional Sanger sequencing as portions of these genes contained ambiguity. The *cox1* gene was analyzed for species identification using the Barcode of Life Data System (BOLD) (Ratnasingham and Hebert, 2007).

Transfer RNAs (tRNAs) and tRNA secondary structures were identified using ARWEN v1.2 (Laslett and Canback, 2008). The two ribosomal RNAs (rRNAs), *lrRNA* and *srRNA*, were determined via sequence comparison, using CLUSTAL Omega (Goujon et al., 2010; Sievers et al., 2011) with the six other published Coreoidea insect mitogenomes. Nucleotide composition was determined using MEGA v6.06 (Tamura et al., 2011) and strand asymmetry was calculated using the formulae: AT skew = [A% − T%] / [A% + T%] and GC skew = [G% − C%] / [G% ± C%] (Perna and Kocher, 1995). Codon usage and the relative synonymous codon frequencies (RSCU) were calculated using DnaSP 5.10.01 (Librado and Rozas, 2009). The Tandem Repeats Finder program (https://tandem.bu.edu/trf/trf.html; Benson, 1999) was used to predict tandem repeats in the A + T-rich region. DNA sequence and protein sequence percentage identities were calculated using CLUSTAL Omega (Goujon et al., 2010; Sievers et al., 2011).

2.4. Sanger sequence confirmation of nad6, nad4L, cox3 and the A + T-rich region

Additional verification of the sequence was performed for the genes *nad6*, *nad4l*, *cox3*, as each gene contained a small region of ambiguities, and the A + T-rich region, with its' high A − T content and repeat units making Illumina sequencing alone insufficient. Genes were amplified via PCR using primers designed using CLC Genomics Workbench and SnapGene Viewer Version 3.2.1 (GSL BioTech) (Supplementary Table 1). PCR was performed with GoTaq® Green Master Mix (Promega Corporation, USA) with the following cycle conditions: initial denaturation step: 2 min at 95 °C, then 35 cycles of 30 s denaturation at 95 °C, 30 s annealing at 49 °C, 30 s extension at 68 °C, and final extension step 5 min at 72 °C. Cleaned PCR products were Sanger sequenced at the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois Urbana-Champaign. Sequencing

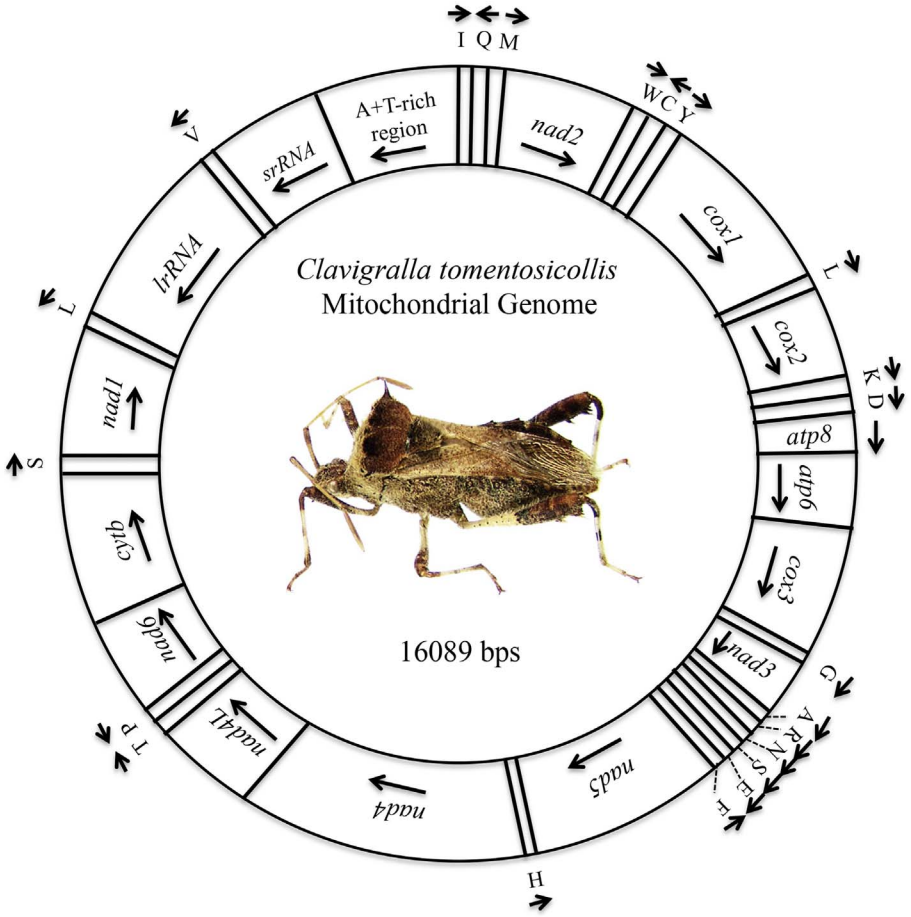


Fig. 1. *Clavigralla tomentosicollis* circular mitochondrial genome map, 16,089 bps in length, illustrating 13 protein coding genes, 22 transfer RNAs, two ribosomal RNAs, and an A + T-rich region. Gene abbreviations given on figure represent: *nad1-6* and *nad4L* indicate NADH dehydrogenase subunits, *cox1*, *cox2*, and *cox3* indicate cytochrome *c* oxidase subunits I-III, *atp6* and *atp8* indicate subunits 6 and 8 of F_0 ATPase, *lrRNA* and *srRNA* indicate the two ribosomal RNAs (16S and 12S, respectively), and transfer RNAs are labeled according to IUPAC-IUB single letter codes. Arrows indicate the gene direction (forward or reverse).

data was generated as .abi chromatogram files, imported into CLC Genomics Workbench, and a local BLAST (BLASTn) against the *C. tomentosicollis* mitogenome contig was performed.

2.5. Comparative analysis of Pentatomomorpha species

The A − T-rich region sequences for 28 complete Pentatomomorpha species were downloaded via NCBI (For Accession Numbers see Supplementary Table 6). Nucleotide composition was calculated using MEGA v6.06 (Tamura et al., 2011). Strand asymmetry was calculated using the formulae: AT skew = [A% − T%] / [A% + T%] and GC skew = [G% − C%] / [G% ± C%] (Perna and Kocher, 1995). The Tandem Repeats Finder program (https://tandem.bu.edu/trf/trf.html; Benson, 1999) was used to predict tandem repeats in the A + T-rich region, with tandem repeats requiring a 90% or higher percent matches result.

3. Results & discussion

3.1. C. tomentosicollis mitogenome organization and nucleotide composition

The complete 16,089 bp mitogenome of the brown pod-sucking bug *C. tomentosicollis* (GenBank Accession Number: KY274846) was assembled using a shotgun sequencing approach, combined with PCR amplification and Sanger sequencing of *nad6*, *nad4l*, *cox3* and the A + T-rich regions wherein ambiguity was initially observed. Importing of sequence files resulted in 467,469,646 paired-end Illumina raw reads with a mean length of 160.0 nucleotides. Raw reads were trimmed using Illumina adapter sequences, with quality score (q) > 20 and two ambiguous nucleotides allowed, resulting in 467,218,189 reads with a mean length of 140.1 nucleotides. de novo assembly resulted in a total of 712,688 contigs, with 10,230 contigs ranging in size from 10,001 to 30,005 bp (Supplementary Table 1). A BLAST of these 10,230 contigs against three Coreoidea species mitogenomes resulted in a contig with an *E*-value of 0.00 (best match with *R. pedestris*), which was utilized for further PCR confirmation and analysis.

The BOLD is a repository of sequence data commonly used for species identification (Ratnasingham and Hebert, 2007), and was applied as secondary confirmation of the morphologically-identified field collected samples used in this study. The resulting results from a BOLD search using the *cox1* gene from our *C. tomentosicollis* assembly as the query classified it to the Family Coreidae (99.39% similarity to an Early-Release specimen) [All Barcode Records on BOLD (4,944,211 sequences; current 06/07/2017)]. Additionally, the *cox1* gene from *C. tomentosicollis* also showed 89.71% similarity to another species of *Clavigralla*, *C. horrida*).

The *C. tomentosicollis* mitogenome is within the size range observed for other hemipteran mitogenomes; from 14,143 bps (*Paraplea frontalis*, Accession Number: NC_028629.1) to 18,414 bps (*Trialeurodes vaporariorum*, Accession Number: NC_006280.1). The *C. tomentosicollis* mitogenome is A + T biased [%A + T = 75.4%; Table 2], also within the range of the four complete Coreoidea mitogenomes (those that had the A + T-rich region sequence complete) (Table 3), as well as those of

Table 1
Annotation and organization of the *Clavigralla tomentosicollis* mitogenome.

Gene	Direction	Start	End	Length (bp)	Start codon	Stop codon	Anticodon	Intergenic nucleotides
tRNA-Ile	F	1	66	68			gat	0
tRNA-Gln	R	64	132	69			ttg	-3
tRNA-Met	F	135	203	69			cat	2
nad2	F	205	1206	1002	atg	taa		1
tRNA-Trp	F	1205	1272	68			tca	− 2
tRNA-Cys	R	1265	1326	62			gca	− 8
tRNA-Tyr	R	1327	1390	64			gta	0
cox1	F	1393	2926	1534	ttg	t (+ aa)		− 4
tRNA-Leu	F	2927	2992	66			taa	0
cox2	F	2993	3671	679	att	t (+ aa)		0
tRNA-Lys	F	3672	3743	72			ctt	0
tRNA-Asp	F	3744	3809	66			gtc	0
atp8	F	3811	3969	159	att	taa		1
atp6	F	3963	4634	672	atg	taa		− 7
cox3	F	4634	5420	787	atg	t (+ aa)		− 1
tRNA-Gly	F	5421	5483	63			tcc	0
nad3	F	5484	5835	352	att	t (+ aa)		0
tRNA-Ala	F	5836	5898	63			tgc	0
tRNA-Arg	F	5901	5965	65			tcg	2
tRNA-Asn	F	5965	6031	67			gtt	− 1
tRNA-Ser	F	6031	6099	69			gct	− 1
tRNA-Glu	F	6100	6164	65			ttc	0
tRNA-Phe	R	6165	6230	66			gaa	0
nad5	R	6232	7944	1713	atg	taa		1
tRNA-His	R	7946	8009	64			gtg	1
nad4	R	8010	9327	1319	atg	t (+ aa)		0
nad4L	R	9321	9602	291	ata	taa		− 9
tRNA-Thr	F	9611	9675	65			tgt	− 1
tRNA-Pro	R	9676	9740	65			tgg	0
nad6	F	9743	10,234	492	ata	taa		2
cytb	F	10,231	11,370	1140	ata	tag		− 4
tRNA-Ser	F	11,369	11,437	69			tga	− 2
nad1	R	11,456	12,388	933	ata	taa		18
tRNA-Leu	R	12,383	12,449	67			tag	− 6
lrRNA	R	12,450	13,714	1265				0
tRNA-Val	R	13,715	13,782	68			tac	0
srRNA	R	13,783	14,580	798				0
A + T-rich region		14,581	16,089	1509				0

other insect mitogenomes that have a broad range from 65.6% in *Reticulitermes* spp. (Isoptera) to 89.3% in *Apis mellifera* (Hymenoptera) (Cameron and Whiting, 2007; Gillespie et al., 2006). Additionally, nucleotide skew shows a positive AT skew (0.144) and a negative GC skew (− 0.187) for the *C. tomentosicollis* mitogenome, which is similar to the pattern in other hemipteran species (Wang et al., 2015; Zhang et al., 2015) including Coreoidea (Table 3).

The *C. tomentosicollis* mitogenome contained the typical 13 PCGs, 22

Table 2
Nucleotide composition of the *Clavigralla tomentosicollis* whole mitogenome, combined protein coding genes (PCGs), individual PCGs, combined tRNAs, combined rRNAs, individual rRNAs (*lrRNA* and *srRNA*), and the A + T-rich region.

	Size (bps)	A%	T%	G%	C%	%A + T	%G + C	AT skew	GC skew
Whole Genome	16,089	42.0	33.4	10.0	14.6	75.35	24.65	0.11	− 0.19
PCGs	10,130	41.1	34.6	10.6	13.7	75.72	24.28	0.09	− 0.13
nad2	1002	39.4	40.6	9.0	11.0	80.04	19.96	− 0.02	− 0.10
cox1	1534	32.5	37.9	14.7	14.9	70.40	29.60	− 0.08	− 0.01
cox2	679	37.1	35.6	13.3	14.0	72.75	27.25	0.02	− 0.03
atp8	159	44.0	37.7	5.7	12.6	81.76	18.24	0.08	− 0.38
atp6	672	37.6	39.4	9.4	13.5	77.08	22.92	− 0.02	− 0.18
cox3	787	34.7	37.1	13.9	14.4	71.79	28.21	− 0.03	− 0.02
nad3	352	36.4	39.5	9.1	15.1	75.85	24.15	− 0.04	− 0.25
nad5	1713	50.4	28.2	8.3	13.1	78.58	21.42	0.28	− 0.23
nad4	1318	50.6	27.2	8.3	13.9	77.77	22.23	0.30	− 0.25
nad4L	282	53.5	23.4	7.4	15.6	76.95	23.05	0.39	− 0.35
nad6	492	41.9	41.5	7.1	9.6	83.33	16.67	0.01	− 0.15
cytb	1140	35.9	35.7	13.0	15.4	71.58	28.42	0.00	− 0.09
nad1	933	51.8	25.4	7.9	14.9	77.17	22.83	0.34	− 0.31
tRNAs	1458	39.8	35.3	11.5	13.4	75.10	24.90	0.06	− 0.08
rRNAs	2063	44.6	31.8	8.5	15.1	76.39	23.61	0.17	− 0.28
lrRNA (16S)	1265	44.6	32.3	8.7	14.4	76.92	23.08	0.16	− 0.25
srRNA (12S)	798	44.6	31.0	8.1	16.3	75.56	24.44	0.18	− 0.34
A + T rich region	1509	40.0	30.6	8.2	21.1	70.64	29.36	0.13	− 0.44

AT skew = [A% − T%] / [A% + T%] GC skew = [G% − C%] / [G% ± C%].

Table 3
Comparative analysis of seven Coreoidea mitogenomes, including mitogenome size, nucleotide composition of the whole mitogenome and the A + T-rich region, and overall mitogenome completeness.

Family	Species	Size (bp)	Total %A + T	Total AT-skew	Total GC-skew	A + T-rich region %A + T	Completeness	NCBI accession	Reference
Coreidae	*Clavigralla tomentosicollis*	16,089	75.35	0.14	− 0.19	70.64	Complete	Current study	Current study
	Hydaropsis longirostris	16,521	75.46[a]	0.10	− 0.25	73.78[a]	Complete	NC_012456	Hua et al., 2008
Alydidae	*Riptortus pedestris*	17,191	76.59[a]	0.09	− 0.19	76.42[a]	Complete	NC_012462	Hua et al., 2008
	Leptocorisa sp.	15,322	73.62[b]	0.20	− 0.16	Incomplete	Partial	KM244663	Tang et al., 2014
Rhopalidae	*Aeschyntelus notatus*	14,532[b]	75.71[a,b]	0.13	0.50	Incomplete	Partial	NC_012446	Hua et al., 2008
	Stictopleurus subviridis	15,319	75.71	0.12	− 0.16	77.08	Complete	NC_012888	NCBI Direct Submission
	Corizus tetraspilus	14,989	74.88	0.140	− 0.19	60.68[a]	Complete	KM983397	Yuan et al., 2015a

AT skew = [A% − T%] / [A% + T%] and GC skew = [G% − C%] / [G% ± C].
[a] Values obtained from corresponding reference.
[b] Genome size is missing A − T region.

Table 4
Comparative analysis of seven Coreoidea mitogenomes, including gene size and the nucleotide composition of the combined protein coding genes (PCGs) and the two individual rRNAs (*lrRNA* and *srRNA*).

	Size (bps)	A%	T%	G%	C%	%A + T	%G + C	AT skew	GC skew
PCGs									
Clavigralla tomentosicollis	10,130	41.1	34.6	10.6	13.7	75.72	24.28	0.09	− 0.13
Hydaropsis longirostris	11,048	41.3	34.0	9.7	15.0	75.31	24.69	0.10	− 0.21
Riptortus pedestris	11,056	40.9	35.2	10.2	13.8	76.09	23.91	0.08	− 0.15
Aeschyntelus notatus	11,038	42.5	32.6	10.5	14.4	75.10	24.9	0.13	− 0.16
Stictopleurus subviridis	11,068	42.3	32.8	10.6	14.4	75.03	24.97	0.13	− 0.15
Corizus tetraspilus	11,047	42.8	31.9	10.4	14.8	74.73	25.27	0.15	− 0.17
Leptocorisa sp.	11,053	43.7	29.2	11.5	15.6	72.89	27.11	0.20	− 0.15
rRNAs									
lrRNA									
Clavigralla tomentosicollis	1265	44.6	32.3	8.7	14.4	76.92	23.08	0.16	− 0.25
Hydaropsis longirostris	1251	44.3	33.4	7.8	14.5	77.70	22.30	0.14	− 0.30
Riptortus pedestris	1260	44.7	34.8	7.9	12.7	79.44	20.56	0.13	− 0.24
Aeschyntelus notatus	1269	46.3	33.2	7.2	13.2	79.51	59.57	0.17	0.56
Stictopleurus subviridis	1263	45.9	32.9	7.8	13.3	78.86	21.14	0.17	− 0.26
Corizus tetraspilus	1266	47.1	31.1	7.4	14.4	78.20	21.80	0.20	− 0.32
Leptocorisa sp.	783	48.9	27.6	8.2	15.3	76.50	23.50	0.28	− 0.30
srRNA									
Clavigralla tomentosicollis	798	44.6	31.0	8.1	16.3	75.56	24.44	0.18	− 0.34
Hydaropsis longirostris	779	46.0	31.5	7.4	15.1	77.41	22.59	0.19	− 0.34
Riptortus pedestris	785	44.6	32.2	7.6	15.5	76.82	23.18	0.16	− 0.34
Aeschyntelus notatus	782	46.7	31.7	7.5	14.1	78.39	21.61	0.19	− 0.30
Stictopleurus subviridis	776	44.8	32.6	8.5	14.0	77.45	22.55	0.16	− 0.25
Corizus tetraspilus	786	45.5	31.3	7.5	15.6	76.84	23.16	0.19	− 0.35
Leptocorisa sp.	654	41.6	28.1	12.4	17.9	69.72	30.28	0.19	− 0.18

AT skew = [A% − T%] / [A% + T%]; GC skew = [G% − C%] / [G% ± C].

tRNAs, 2 rRNAs and an A + T-rich control region, and order and orientation of the 37 genes is identical to the ancestral gene order of insects, suggesting evolutionary conservation across the superfamily Coreoidea (Boore, 1999; Cameron et al., 2006; Wang et al., 2015). This ancestral gene order is also fairly conserved across the order Hemiptera, including being found in the six other currently published Coreoidea superfamily mitogenomes (*A. notatus*, *H. longirostris*, *R. pedestris*, *S. subivridis*, *Leptocorisa* sp., and *C. tetraspilus*) (Fig. 1; Table 1). The hemipteran exceptions are gene rearrangements identified in four heteropteran families plus one family in each of the suborders Auchenorrhyncha and Sternorrhyncha (Wang et al., 2015).

The majority strand (J strand) encodes 9 PCGs and 14 tRNAs while the opposite strand (N strand) encodes four PCGs, eight tRNAs, and two rRNAs. There are 13 locations where genes overlap a total of 47 bp. The longest overlap is 8 bp in length between the genes *tRNA-Trp* and *tRNA-Cys*. The next longest overlaps were both 7 bp in length, with the first between the genes *atp8* and *atp6* and the second between *nad4* and *nad4L*. In addition to the A + T-rich region, there were eight intergenic spacers totaling 28 bases. All intergenic spacers were one or two bp long with the exception of between *tRNA-Leu* and *nad1* at 18 bp.

3.2. Protein coding genes

The 13 PCGs span 10,130 bp or 63.0% of the total *C. tomentosicollis* mitogenome. At 75.72%, the overall A − T content of the PCGs was 0.37% higher than that of the entire mitogenome (Table 2), in contrast to the four complete Coreoidea mitogenomes (Tables 3 and 4) but consistent with other Coreoidea species (Table 4). The *C. tomentosicollis* mitogenome PCGs contain 3678 codons (excluding stop codons), with the most utilized codons Leu* (UUA; n = 423), Ile (AUU; n = 340), Phe (UUU; n = 291), and Met (AUA; n = 246) (Table 5). Similarly, five of the six Coreoidea species have UUA, AUU, UUU, and AUA as the most frequent used codons; the exception is *C. tetraspilus* (Fig. 2; Supplementary Table 5) (Yuan et al., 2015a). Relative synonymous codon

Table 5
Codon usage and the relative synonymous codon frequencies (RSCU) in the 13 protein coding genes of *Clavigralla tomentosicollis*. A total of 3681 codons were analyzed, excluding stop codons.

Amino acid	Codon	#	RSCU	Amino acid	Codon	#	RSCU
Phenylalanine – Phe (F)	UUU	291	1.72	Tyrosine – Tyr (Y)	UAU	150	1.83
	UUC	47	0.28		UAC	14	0.17
Leucine – Leu (L)[a]	UUA	423	4.69	Stop	UAA	0	0.00
tRNA-Leu(UUR)	UUG	43	0.48		UAG	0	0.00
Leucine – Leu (L)	CUU	36	0.40	Histidine – His (H)	CAU	59	1.57
tRNA-Leu(CUN)	CUC	1	0.01		CAC	16	0.43
	CUA	35	0.39	Glutamine – Gln (Q)	CAA	50	1.85
	CUG	3	0.03		CAG	4	0.15
Isoleucine – Ile (I)	AUU	340	1.87	Asparagine – Asn (N)	AAU	158	1.79
	AUC	24	0.13		AAC	19	0.21
Methionine – Met (M)	AUA	246	1.84	Lysine – Lys (K)	AAA	86	1.55
	AUG	21	0.16		AAG	25	0.45
Valine – Val (V)	GUU	90	1.86	Aspartic acid – Asp (D)	GAU	54	1.66
	GUC	7	0.14		GAC	11	0.34
	GUA	93	1.92	Glutamic acid – Glu (E)	GAA	81	1.86
	GUG	4	0.08		GAG	6	0.14
Serine – Ser (S)	UCU	101	2.39	Cysteine – Cys (C)	UGU	46	1.80
tRNA-Ser(UCN)	UCC	6	0.14		UGC	5	0.20
	UCA	106	2.51	Tryptophan – Trp (W)	UGA	94	1.88
	UCG	0	0.00		UGG	6	0.12
Proline – Pro (P)	CCU	82	2.43	Arginine – Arg (R)	CGU	17	1.26
	CCC	6	0.18		CGC	2	0.15
	CCA	46	1.36		CGA	34	2.52
	CCG	1	0.03		CGG	1	0.07
Threonine – Thr (T)	ACU	91	1.85	Serine – Ser (S)[a]	AGU	39	0.92
	ACC	11	0.22	tRNA-Ser(AGN)	AGC	5	0.12
	ACA	93	1.91		AGA	81	1.92
	ACG	1	0.02		AGG	0	0.00
Alanine – Ala (A)	GCU	75	1.99	Glycine – Gly (G)	GGU	83	1.54
	GCC	16	0.42		GGC	3	0.06
	GCA	60	1.59		GGA	113	2.09
	GCG	0	0.00		GGG	17	0.31

[a] Two amino acids, Leucine (CUN or UUR) and Serine (AGN and UCN), are coded for by two separate codon formats with each format associated with a different tRNA, as indicated in the table.

usage (RSCU) in *C. tomentosicollis* revealed that three codons (UCG, GCG, AGG) were unused (Table 5), and comparative analysis indicated the absence of GCG from *C. tetrapilus* and AGG from *S. subviridis* (Fig. 3). Across all Coreoidea mitogenomes, codons more frequently had A or U in the third codon position that synonymous G/C codons (Fig. 3).

Twelve PCGs have ATN start codons (Table 1), which encode Met (ATG ATA) or Ile ATT. One PCG had an atypical start codon TTG (*cox1*), commonly observed in other published hemipteran mitogenomes (Supplementary Table 3) (Hua et al., 2008; Li et al., 2012; Yuan et al., 2015b).

Eight of the termination codons were typical TAG/TAA stop codons (Table 1). The remaining five genes (*cox1*, *cox2*, *cox3*, *nad3*, and *nad4*) ended in a single T, with post-transcription polyadenylation proposed to add the AA necessary to complete a functional TAA stop codon (Ojala et al., 1981). This post-transcriptional polyadenylation is frequently seen in insect mitogenomes and all published Coreoidea mitogenomes; ranging from 4/13 genes in *S. subviridis* to 9/13 genes in *C. tetraspilus* (Supplementary Table 3; Wei et al., 2010; Yuan et al., 2015b, 2015a; Wang et al., 2015; Du et al., 2016). Unlike *C. tomentosicollis*, both *A. notatus* and *H. longirostris* had TA partial stop codons. There are four locations where two PCGs overlap (*atp8* with *atp6*, *atp6* with *cox3*, *nad4* with *nad4L*, and *nad6* with *cytb*), and five locations where a PCG and a tRNA overlap (*nad2* with tRNA-Trp, *cox1* with tRNA-Tyr, *nad4L* with tRNA-Thr, *cytb* with tRNA-Ser, and *nad1* with tRNA-Leu).

3.3. Transfer RNA genes

The typical 22 tRNAs found in metazoan mitogenomes were identified in *C. tomentosicollis*; the J strand contained fourteen of the tRNAs while the remaining eight were located on the N strand (Table 1; Fig. 1). Transfer RNAs ranged in length from 62 bp to 72 bp. There are four locations where two tRNAs overlap (tRNA-Ile and tRNA-Gln; tRNA-Trp and tRNA-Cys; tRNA-Arg and tRNA-Asn; tRNA-Asn and tRNA-Ser). All tRNAs have the canonical clover-leaf structure except for tRNA-Ser (AGN) in which the dihydrouridine (DHU) arm lacks a stem (Supplementary Fig. 1). This variant tRNA-Ser (AGN) has been observed across insect species and is considered a typical characteristic of metazoan mitogenomes (Cameron, 2014b; Zhang et al., 2015; Du et al., 2016), including other Hemipterans (Wang et al., 2015). The identified anticodons for each tRNA isotype were identical to those in other hemipteran mitogenomes published previously (Li et al., 2012; Teng et al.,

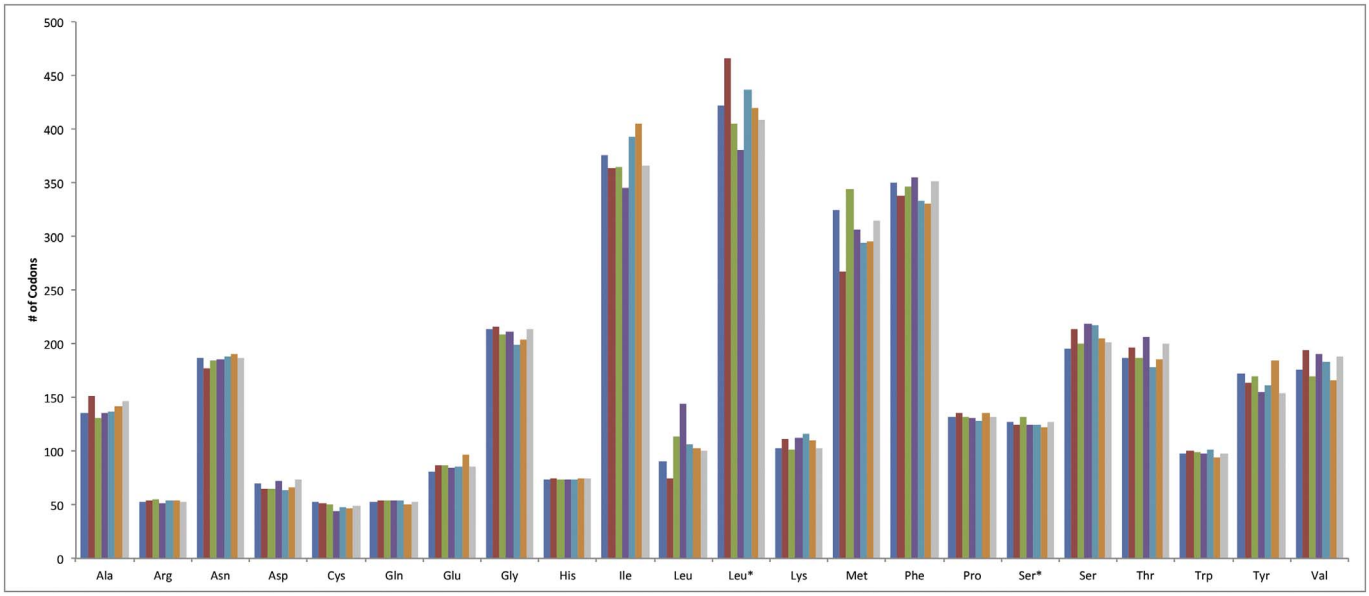


Fig. 2. Codon usage in the protein codon genes of the *Clavigralla tomentosicollis* genome compared to six other Coreoidea species. Codon usage calculations, excluding start codons and stop codons, were completed using DNA sP program. The y-axis represents the total number of codons. Codons are listed on the x-axis. Colored bars represent each species: Blue = *Aeschyntelus notatus*; Red = *Clavigralla tomentosicollis*; Green = *Corizus tetraspilus*; Purple = *Leptocorisa* spp.; Teal = *Hydaropsis longirostris*; Orange = *Riptortus pedestris*; Gray = *Stictopleurus subviridis*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2013; Yuan et al., 2015b, 2015a). The tRNAs contain a total of 1458 bps (9.1%), a nucleotide composition with a positive A − T skew (0.060) and a negative G-C skew (− 0.076), and a total A − T content of 75.1%.

3.4. Ribosomal RNA genes

The N strand-encoded both the 1265 bp 16S ribosomal RNA (*lrRNA*), and the 798 bp 12S ribosomal RNA (*srRNA*) genes, positioned between tRNA-Leu and tRNA-Val, and tRNA-Val and the A + T rich region respectively (Table 1). Conservation of rRNA sequence and position was determined via alignment against six Coreoidea species (Table 1; Supplementary Figs. 2 and 3). The combined A − T content for both rRNAs was 76.4%, while individually 76.9% for *lrRNA* and 75.6% for *srRNA* (Table 2). The A − T content for both *C. tomentosicollis* rRNAs was lower than other Coreoidea species (Table 4). As observed in other Coreoidea species, AT skew for both *lrRNA* and *srRNA* was positive (0.160 and 0.180, respectively) while GC skew was negative (− 0.247 and − 0.336, respectively) (Table 4).

3.5. A + T-rich region

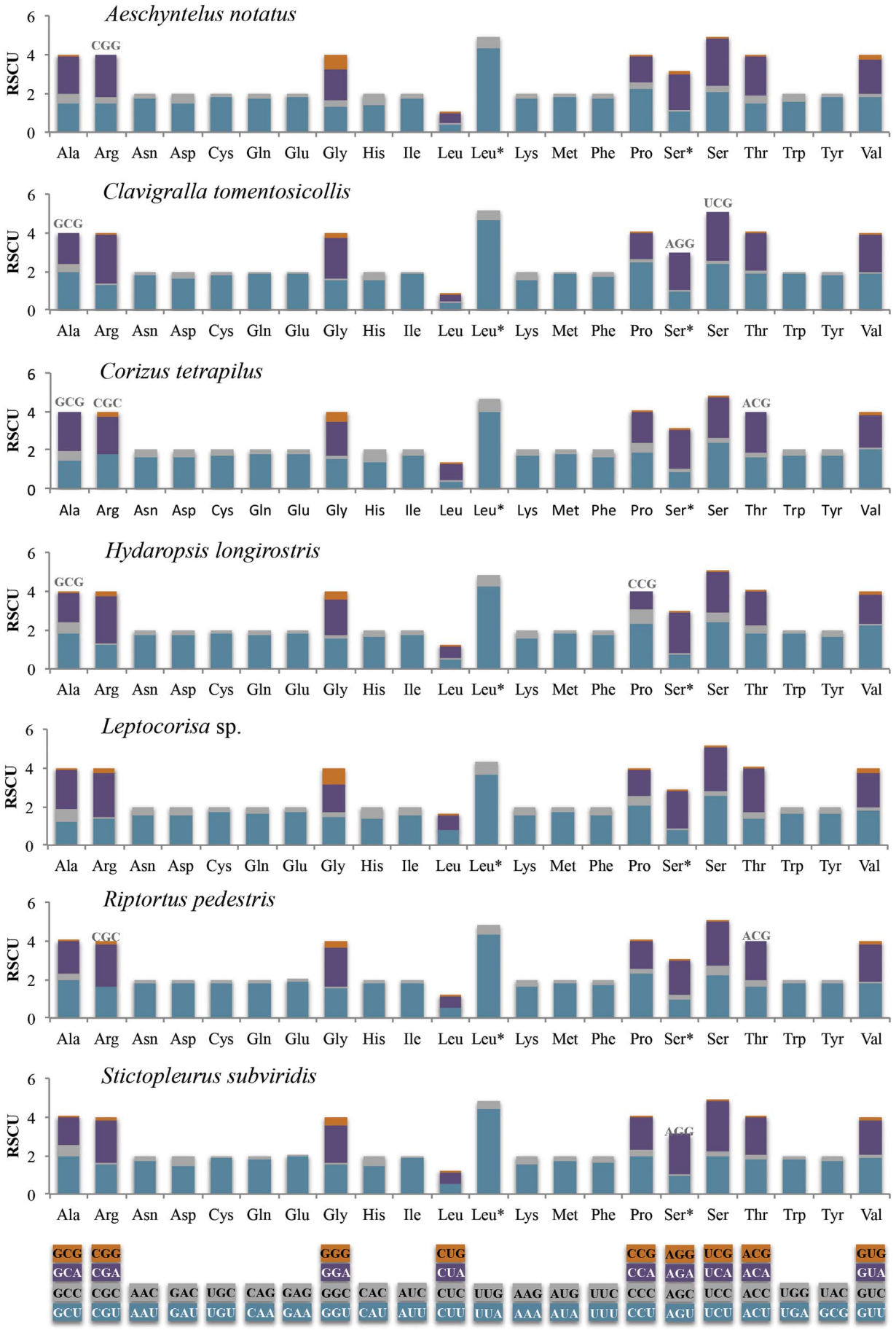
The *C. tomentosicollis* A + T-rich control region was 1509 bp and located between *srRNA* and tRNA-Ile. Sequencing this region was obtained by combining results from the Illumina sequence assembly and Sanger sequencing of PCR amplicons. The *C. tomentosicollis* A + T rich region has a nucleotide composition of 70.64% A − T, which is 4.71% lower than that of the overall genome, as well as of the combined PCGs, the combined tRNAs, and the combined rRNAs (Table 2).

A comparison of the A − T-rich region among 28 species within the Infraorder Pentatomomorpha with complete mitogenomes (representing five superfamilies and 18 families) showed that the %A − T in 17 of these was higher than *C. tomentosicollis* (Supplementary Table 6). This region of *C. tomentosicollis* had the second lowest %A − T known among Coreoidea, with only *C. tetraspilus* lower at 60.68% (Table 3). When compared to the entire *C. tomentosicollis* genome, the A − T region has a higher, positive AT skew (0.13) and a considerably higher, negative GC skew (− 0.44) (Table 2).

A single tandem repeat element was located within the A + T-rich region of the *C. tomentosicollis* mitogenome, which consisted of a 89 bp repeat unit repeated 3.7 times (Fig. 4). Comparative analysis across the Infraorder Pentatomomorpha found tandem repeats present in all five superfamilies, ranging from 4.1 repeats of a 9 bp tandem repeat in *Libiocoris heissi* to a 2.1 repeat of a 256 bp tandem repeat in *Coptosoma bifaria* (Supplementary Table 7). Along with *R. pedestris* (Hua et al., 2008; Yuan et al., 2015a) *C. tomentosicollis* is only the second Coreoidea species in which a tandem repeat has been identified. However, the tandem repeats in *R. pedestris* were considerably shorter, with two different sets of repeat units of 30 bp and 19 bp, and with fewer number of repeats, at 2.6 and 2.7, respectively. Analogous tandemly repetitive regions are described in other insect species (Monteforte et al., 1993; Lee et al., 2009; Liu et al., 2013; Liu et al., 2015; Ye et al., 2015), and are believe to evolve via the mechanism of slip strand mispairing (Levinson and Gutman, 1987). Although replication initiation and transcriptional start sites are well defined within vertebrate A + T-rich control regions, corresponding invertebrate mechanisms are less well understood (Saito et al., 2005). Therefore, the functional impact of repeat expansion within the control region of insects remains unknown, but in many vertebrate systems there appears to be selective constraints on repeat expansion (Casane et al., 1997).

4. Conclusions

The complete mitogenome described here provides a foundation for future studies that may lead to a better understanding of the population dynamics of this pest insect and the possibility better pest management decisions. Additional studies, including molecular markers development to examine genetic variability in *C. tomentosicollis* populations (Agunbiade et al., 2013) and assessing population movement from wild leguminous hosts (during insect breeding season) to cowpea plants undergoing pod development (Koona et al., 2002; Agunbiade et al., 2014), are underway. Future studies applying variation in the mitogenome may help to better understand *C. tomentosicollis* population structure across sub-Saharan Africa. Informed decisions by producers as to when control measures should be applied, e.g., release of biocontrol



(caption on next page)

Fig. 3. Relative synonymous codon usage (RSCU) of seven Coreoidea species (*Aeschyntelus notatus*, *Clavigralla tomentosicollis*, *Corizus tetraspilus*, *Hydaropsis longirostris*, *Leptocorisa* spp., *Riptortus pedestris*, and *Stictopleurus subviridis*). Codon families are displayed on the x-axis, with the third codon positions (G, A, C, or U) represented by orange, purple, gray, and teal, respectively. Codons not present in a species genome are designated on the corresponding graph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

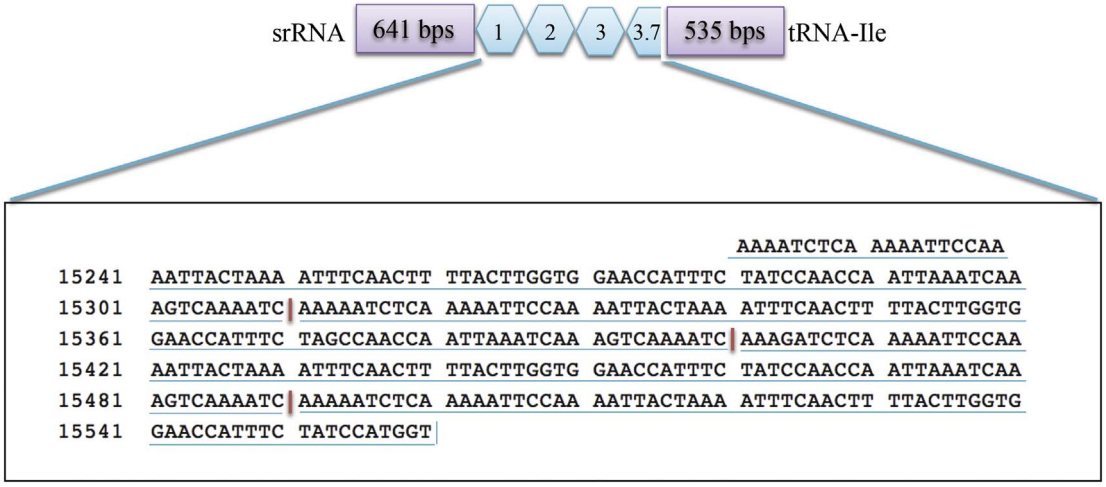


Fig. 4. A putative tandem repeat with three 89 bp repeat elements and one 66 bp incomplete repeat element was identified in the A − T rich region of the *Clavigralla tomentosicollis* mitogenome. Blue underline indicates individual repeats while red vertical lines separate repeats. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

agents where the pest is endemic during the dry season may ultimately help to minimize the pest population leading to potentially increased cowpea production.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.aggene.2017.07.002.

Acknowledgements

Thank you to Dr. George Goergen at IITA-Benin for the image of *Clavigralla tomentosicollis* in Fig. 1. Thank you to Susan Balfe for helping with formatting and editing of Fig. 3.

This study was supported by The Feed the Future Innovation Lab for Collaborative Research on Grain Legumes (formerly known as the Dry Grain Pulses Collaborative Research Support Program) and by the Bureau for Food Security, United States Agency for International Development under the terms of Grant No. EDH-A-00-07-00005-00 (to MT and BRP). The opinions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development or the U.S. government.

References

Abascal, F., Posada, D., Knight, R.D., Zardoya, R., 2006. Parallel evolution of the genetic code in arthropod mitochondrial genomes. PLoS Biol. 4 (5), e127.

Adati, T., Tamò, M., Yusuf, S.R., Downham, M.C.A., Singh, B.B., Hammond, W., 2008. Integrated pest management for cowpea-cereal cropping systems in the West African savannah. Int. J. Trop. Insect Sci. 27 (3/4), 123–137.

Agunbiade, T.A., Coates, B.S., Datinon, B., Djouaka, R., Sun, W., Tamò, M., Pittendrigh, B.R., 2014. Genetic differentiation among *Maruca vitrata* F. (Lepidoptera: Crambidae) populations on cultivated cowpea and wild host plants: implications for insect resistance management and biological control strategies. PLoS ONE 9 (3), e92072.

Agunbiade, T.A., Sun, W., Coates, B.S., Djouaka, R., Tamò, M., Ba, M.N., Binso-Dabire, C., Baoua, I., Olds, B.P., Pittendrigh, B.R., 2013. Development of reference transcriptomes for the major field insect pests of cowpea: a toolbox for insect pest management approaches in West Africa. PLoS ONE 8 (11), e79929.

Aliyu, M., Ladan, T., Ahmed, B.I., Abdullahi, J., 2007. Studies on the efficacy of black soap and kerosene mixture on the control of Pod sucking bugs (*Clavigralla tomentosicollis* stal.) on Cowpea (*Vigna unguiculata* (L.) Walp). Emir. J. Food Agric. 19 (2), 08–14.

Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27 (2), 573–580.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27 (8), 1767–1780.

Cameron, S.L., Whiting, M.F., 2007. Mitochondrial genomic comparisons of the subterranean termites from the genus *Reticulitermes* (Insecta: Isoptera: Rhinotermitidae). Genome 50 (2), 188–202.

Cameron, S.L., 2014a. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. Syst. Entomol. 39 (3), 400–411.

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

Cameron, S.L., Beckenbach, A.T., Dowton, M., Whiting, M.F., 2006. Evidence from mitochondrial genomics on interordinal relationships in insects. Arthropod Syst. Phylo. 64 (1), 27–34.

Casane, D., Dennebouy, N., De Rochambeau, H., Mounolou, J.C., Monnerot, M., 1997. Nonneutral evolution of tandem repeats in the mitochondrial DNA control region of lagomorphs. Mol. Biol. Evol. 14 (8), 779–789.

Dabire-Binso, C.L., Ba, N.M., Sanon, A., Drabo, I., Bi, K.F., 2010. Resistance mechanism to the pod-sucking bug *Clavigralla tomentosicollis* (Hemiptera: Coreidae) in the cowpea IT86D-716 variety. Int. J. Trop. Insect Sci. 30 (4), 192–199.

Devi, C.B., Kushwaha, A., Kumar, A., 2015. Sprouting characteristics and associated changes in nutritional composition of cowpea (*Vigna unguiculata*). J. Food Sci. Technol. 52 (10), 6821–6827.

Dreyer, H., Baumgartner, J., 1994. The influence of post-flowering pests on cowpea seed yield with particular reference to damage by Heteroptera in southern Benin. Agric. Ecosyst. Environ. 53, 137–149.

Du, C., Shilin, H., Xuhao, S., Liao, Q., Zhang, X., Yue, B., 2016. The complete mitochondrial genome of *Epicauta chinensis* (Coleoptera: Meloidae) and phylogenetic analysis among Coleopteran insects. Gene 578, 274–280.

Gillespie, J.J., Johnston, J.S., Cannone, J.J., Guttell, R.R., 2006. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRAN genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. Insect Mol. Biol. 15 (5), 657–686.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., Lopez, R., 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 38, W695–W699.

Harrison, R.G., 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4 (1), 6–11.

Herbert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. U. S. A. 101, 14812–14817.

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

Jackai, L.E.N., 1990. Screening of cowpeas for resistance to *Clavigralla tomentosicollis* Stal (Hemiptera: Coreidae). J. Econ. Entomol. 83 (2), 300–305.

Jackai, L.E.N., Nokoe, S., Tayo, B.O., Koona, P., 2001. Interferences on pod wall and seed defenses against the brown cowpea coreid bug, *Clavigralla tomentosicollis* Stal. (Hem., Coreidae) in wild and cultivated *Vigna* species. J. Appl. Entomol. 125, 277–286.

Kim, M.I., Baek, J.Y., Kim, M.J., Jeong, H.C., Kim, K.G., Bae, C.H., Han, Y.S., Jin, B.R., Kim, I., 2009. Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. Mol. Cell 28 (4), 347–363.

Koona, P., Osisanya, E.O., Jackai, L., Tamò, M., Markham, R.H., 2002. Resistance in accessions of cowpea to the coreid pod-bug *Clavigralla tomentosicollis* (Hemiptera: Coreidae). J. Econ. Entomol. 95 (6), 1281–1288.

Koona, P., Osisanya, E.O., Jackai, L.E.N., Tonye, J., 2004. Infestation and damage by

Clavigralla tomentosicollis and *Anoplocnemis curvipes* (Hemiptera: Coreidae) in cowpea plants with modified leaf structure and pods in different positions relative to the canopy. Environ. Entomol. 33 (3), 471–476.
Laslett, D., Canback, B., 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24 (2), 172–175.
Lee, E.M., Hong, M.Y., Kim, M.I., Kim, M.J., Park, H.C., Kim, K.Y., Lee, I.H., Bae, C.H., Jin, B.R., Kim, I., 2009. The complete mitogenome sequences of the palaeopteran insects *Ephemera orientalis* (Ephemeroptera: Ephemeridae) and *Davidius lunatus* (Odonata: Gomphidae). Genome 52, 810–817.
Levinson, G., Gutman, G.A., 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4, 203–221.
Li, H., Liu, H., Cao, L., Shi, A., Yang, H., Cai, W., 2012. The complete mitochondrial genome of the damsel bug *Alloeorhynchus bakeri* (Hemiptera: Nabidae). Int. J. Biol. Sci. 8 (1), 93–107.
Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.
Liu, Q.-N., Zhu, B.-J., Dai, L.-S., Liu, C.-L., 2013. The complete mitogenome of *Bombyx mori* strain Dazao (Lepidoptera: Bombycidae) and comparison with other lepidopteran insects. Genomics 101, 64–73.
Liu, Q.N., Bian, D.D., Jiang, S.H., Ge, B.M., Zhou, C.L., Tang, B.P., 2015. Characterization of the complete mitochondrial genome of the Oriental armyworm, *Mythimna separata* (Lepidoptera: Noctuidae). Eur. J. Entomol. 112 (3), 399–408.
Margam, V.M., Coates, B.S., Ba, M.N., Sun, W., Binso-Dabire, C.L., Baoua, I., Ishiyaku, M.F., Shukle, J.T., Hellmich, R.L., Covas, F.G., Ramasamy, S., Armstrong, J., Pittendrigh, B.R., Murdock, L.L., 2011a. Geographic distribution of phylogenetically-distinct legume pod borer, *Maruca vitrata* (Lepidoptera: Pyraloidea: Crambidae). Mol. Biol. Rep. 38, 893–903.
Margam, V.M., Coates, B.S., Hellmich, R.L., Agunbiade, T., Seufferheld, M.J., Sun, W., et al., 2011b. Mitochondrial genome sequence and expression profiling for the legume pod borer *Maruca vitrata* (Lepidoptera: Crambidae). PLoS ONE 6 (2), e16444.
Monteforte, A., Barrio, E., Latorre, A., 1993. Characterization of the length polymorphism in the A + T-rich region of the *Drosophila obscura* group species. J. Mol. Evol. 36, 214–223.
Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA: relevance for population genetics and systematics. Annu. Rev. Ecol. Syst. 18, 269–292.
Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474.
Olatunde, G.O., Biobaku, I.A., Ojo, D.K., Pitan, O.O.R., Adegbite, E.A., 2007. Inheritance of resistance in cowpea (*Vigna unguiculata*) to the pod-sucking bug *Clavigralla tomentosicollis* (Hemiptera: Coreidae). Trop. Sci. 47 (3), 128–133.
Oparaeke, A.M., 2006a. The potential for controlling *Maruca vitrata* Fab. and *Clavigralla tomentosicollis* Stal. using different concentrations and spraying schedules of *Syzigium aromaticum* (L.) Merr and Perr on cowpea plants. J. Plant Sci. 1 (2), 132–137.
Oparaeke, A.M., 2006b. Effect of aqueous extracts of tropical plants for management of *Maruca vitrata* Fab. and *Clavigralla tomentosicollis* Stal. on cowpea, *Vigna unguiculata* (L.) Walp. plants. J. Entomol. 3 (1), 70–75.
Perna, N.T., Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41 (3), 353–358.
Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: the barcode of life data system. (www.barcodinglife.org) Mol. Ecol. Notes 7 (3), 355–364.
Saito, S., Tamura, K., Aostsuka, T., 2005. Replication origin of mitochondrial DNA in insects. Genetics 171, 1695–1705.
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.
Singh, S.R., Allen, D.J., 1980. Pests, diseases, resistance and protection of *Vigna unguiculata* (L.) Walp. In: Summerfield, R.J., Bunting, H.A. (Eds.), Advances in Legumes Science. London Royal Botanical Garden, Kew, and Ministry of Agriculture, Fisheries and Food, London, pp. 419–443.
Singh, C., Singh, N.N., 2015. Bioefficacy and economics of different insecticides again legume pod borer (*Maruca vitrata* Fabricius) infesting cowpea (*Vigna unguiculata* Walpers). Indian J. Agric. Res. 49 (4), 358–362.
Soyelu, O.L., Akingbohungbe, A.E., Okonji, R.E., 2007. Salivary glands and their digestive enzymes in pod-sucking bugs (Hemiptera: Coreoidea) associated with cowpea *Vigna unguiculata* ssp. *unguiculata* in Nigeria. Int. J. Trop. Insect Sci. 27 (1), 40–47.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Tang, M., Tan, M., Meng, G., Yang, S., Su, X., Liu, S., Song, W., Li, Y., Wu, Q., Zhang, A., Zhou, X., 2014. Multiplex sequencing of pooled mitochondrial genomes-a crucial step toward biodiversity analysis using mito-metagenomics. Nucleic Acids Res. 42 (22), e166.
Teng, L., Gao, C., Cui, Y., Xie, Q., Bu, W., 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.
Timmermans, M.J.T.N., Lees, D.C., Simonsen, T.J., 2014. Towards a mitogenomic phylogeny of Lepidoptera. Mol. Phylogenet. Evol. 79, 169–178.
Wang, Y., Chen, J., Jiang, L.Y., Qiao, G.X., 2015. Hemipteran mitochondrial genomes: features, structures and implications for phylogeny. Int. J. Mol. Sci. 16 (6), 12382–12404.
Wei, S.-J., Shi, M., Sharkey, M.J., van Achterberg, C., X-x, Chen, 2010. Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects. BMC Genomics 11, 371.
Whinnett, A., Willmott, K.R., Brower, A.V.Z., Simpson, F., Zimmermann, M., Lamas, G., Mallet, J., 2005. Mitochondrial DNA provides an insight into the mechanism driving diversification in the ithomiine butterfly *Hyposcada anchiala* (Lepidoptera: Nymphalidae: Ithomiinae). Eur. J. Entomol. 102, 633–639.
Wolstenholme, D.R., 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. 141, 173–216.
Ye, F., Liu, T., King, S.D., You, P., 2015. Mitochondrial genomes of two phlebotomine sand flies, *Phlebotomus chinensis* and *Phlebotomus papatasi* (Diptera: Nematocera), the first representatives from the family Psychodidae. Parasit Vectors 8, 472.
Yuan, M.-L., Zhang, Q.-L., Gou, Z.-L., Wang, J., Shen, Y.-Y., 2015a. The complete mitochondrial genome of *Corizus tetraspilus* (Hemiptera: Rhopalidae) and phylogenetic analysis of Pentatomomorpha. PLoS ONE 10 (6), e0129003.
Yuan, M.-L., Zhang, Q.-L., Guo, Z.-L., Wang, J., Shen, Y.-Y., 2015b. Comparative mitogenomic analysis of the superfamily Pentatomoidea (Insecta: Hemiptera: Heteroptera) and phylogenetic implications. BMC Genomics 16, 460.
Zhang, D.X., Hewitt, G.M., 1997. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. Biochem. Syst. Ecol. 25, 99–120.
Zhang, N.-X., Yu, G., Li, T.-J., He, Q.-Y., Zhou, Y., Si, F.-L., Ren, S., Chen, B., 2015. The complete mitochondrial genome of *Delie antiqua* and its implications in Dipteran phylogenetics. PLoS ONE 10 (10), e0139736.