Zootaxa 4268 (1): 101–116
http://www.mapress.com/j/zt/

Article

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

https://doi.org/10.11646/zootaxa.4268.1.6
http://zoobank.org/urn:lsid:zoobank.org:pub:36BC08B9-DDFC-47C3-97E9-B92AC8FBB498

Complete mitochondrial genome and taxonomic revision of *Cardiodactylus muiri* Otte, 2007 (Gryllidae: Eneopterinae: Lebinthini)

JIAJIA DONG[1,2], NATALLIA VICENTE[2,3], IOANA C. CHINTAUAN-MARQUIER[2], CAHYO RAMADI[4], AGNÈS DETTAI[2] & TONY ROBILLARD[2]

[1]*College of Life Sciences, Shaanxi Normal University, 710062, Xi'an, P.R. China.*
[2]*Muséum national d'Histoire naturelle, Institut de Systématique, Evolution et Biodiversité ISYEB, UMR 7205 CNRS MNHN UPMC EPHE CP 50 (Entomologie), Sorbonne Universités, 75231 Paris Cedex 05, France. E-mail: jdong2@mnhn.fr*
[3]*Universidade Federal de Viçosa, Departamento de Biologia Animal. Av. PH Rolfs s/n. Viçosa, Minas Gerais, Brazil. CEP 36570-900.*
[4]*Division of Zoology, RC–Biology, Cibinong Science Center – LIPI, Jl. Raya Jakarta- Bogor Km. 46, Cibinong, 16911 Indonesia.*

Abstract

In the present study, we report the high-coverage complete mitochondrial genome (mitogenome) of the cricket *Cardiodactylus muiri* Otte, 2007. The mitogenome was sequenced using a long-PCR approach on an Ion Torrent Personal Genome Machine (PGM) for next generation sequencing technology. The total length of the amplified mitogenome is 16,328 bp, representing 13 protein-coding genes, 22 transfer RNA genes, two ribosomal RNA genes and one noncoding region (D-loop region). The new sets of long-PCR primers reported here are invaluable resources for future comparative evolutionary genomic studies in Orthopteran insects. The new mitogenome sequence is compared with published cricket mitogenomes. In the taxonomic part, we present new records for the species and describe life-history traits, habitat and male calling song of the species; based on observation of new material, the species *Cardiodactylus buru* Gorochov & Robillard, 2014 is synonymized under *C. muiri.*

Key words: mitogenome, long-PCR approach, next generation sequencing technology, New Guinea, Indonesia

Introduction

In the era of next generation sequencing technologies, complete mitochondrial genomes (mitogenomes) are becoming a common and valuable resource for phylogenetic studies and molecular systematics (Crampton-Platt *et al.* 2015, 2016; Gomez-Rodriguez *et al.* 2015). In a recent phylogeny of the diverse and ancient insect order Orthoptera, Song *et al.* (2015) used both mitogenomes and traditional molecular markers as sources of phylogenetic information. They showed that despite the long evolutionary history of these insects, their mitogenome has remained relatively stable, making it a powerful tool to resolve phylogenetic relationships. However, mitogenome sequences have been little investigated in Orthopteran insects compared with other animal groups (Zhou *et al.* 2010), and large areas of their phylogenetic tree are barely or not documented when considering these data. This is particularly true in the superfamily Grylloidea. Despite the diversity of crickets and their allies, only a handful of mitogenomes are available today in databases (four in Song *et al.* 2015). The families Gryllotalpidae, Myrmecophilidae and Schizodactylidae are each documented by one species, while the "true crickets", comprising three large families, namely Gryllidae, Phalangopsidae and Trigonidiidae (Chintauan-Marquier *et al.* 2016), are represented only by five species belonging to a single subfamily (Gryllinae): *Teleogryllus commodus* (Wolff *et al.* 2012), *Teleogryllus oceanicus* (Zhou *et al.* 2015), *Loxoblemmus equestris*, *Teleogryllus emma* and *Velarifictorus hemelytrus* (Yang *et al.* 2016a). The direct consequence of this is that the phylogenetic information of mitogenomes between and within cricket families has never been evaluated.

In this study, we present a new set of primers to obtain mitogenomes from long-range PCRs in next generation sequencing (NGS) (e.g., Jia *et al*. 2014, Hinsinger *et al*. 2015) and report the complete mitogenome of *Cardiodactylus muiri* Otte, 2007. This species belongs to the subfamily Eneopterinae, which is characterized by

Accepted by D. Rentz: 15 Mar. 2017; published: 15 May 2017

diverse shapes and original features related to acoustic communication (e.g., Robillard & Desutter-Grandcolas 2004a, b; Robillard *et al*. 2013; ter Hofstede *et al*. 2015). In the taxonomic part of the study, we revise *C. muiri* based on recent collections in New Guinea (Papua province, Indonesia) and *C. buru* Gorochov & Robillard, 2014 is synonymized under *C. muiri*. We also present new geographic records for the species and describe life-history traits, habitats and male calling song.

Materials and methods

Materials. The new material was collected in the "Lengguru 2014" expedition in Papua, Indonesia and will be deposited in the collections of the Muséum national d'Histoire naturelle Paris, France (MNHN) and in the Museum Zoologicum Bogoriense, Bogor, Indonesia (MZB). The crickets were collected mostly by sight, night and day, in order to observe the habitat and activity of the species. Square brackets in the list of specimens are used for additional information not mentioned on the specimen labels or translated from French labels.

Molecular techniques. Total genomic DNA was extracted from the median femoral muscles using the NucleoSpin® 96 Tissue Kit (Macherey-Nagel, Germany) and the automatic pipetting robot Eppendorf epMotion 5075 TMX following the manufacturer's manual. The complete mitochondrial genome sequence of *C. muiri* was obtained by overlapping three DNA fragments (> 5.5kb) amplified by long-range PCRs and the HotStart LongAmp® *Taq* DNA polymerase (New England BioLabs® Inc., Ipswich MA). The primers used for these PCRs are presented in Table 1. The amplicons were sequenced using the Ion Torrent Personal Genome Machine (PGM) sequencing platform with two level multiplexing (Hinsinger *et al.* 2015) at the "Service de Systématique Moléculaire" of the MNHN. The reads obtained from each long-range PCR were separately trimmed by error probability algorithm and assembled using the mapping to reference algorithm in GENEIOUS (8.1.7 version) (Table 1) (Kearse *et al.* 2012). The references used for mapping each long-range PCR were the partial mitochondrial gene markers of *C. muiri* obtained from a phylogenetic study of *Cardiodactylus* in preparation molecular data library (Dong *et al.* in prep., see details in Table 1). After this first assembly, the three contigs obtained were re-assembled into a circular mitochondrial genome with high coverage by the *de novo* assembly algorithm in GENEIOUS.

TABLE 1. List of primers used to amplify the mitogenome and mapping reference in assembly procedure of *Cardiodactylus muiri* Otte, 2007.

Set name	Primer name	Direction	Sequence (5'–3')	Mapping Reference	Length of the amplicon (bp)
SET-1	F146	Forward	GCTAATTAAAGCTAATGGGTTCATACC	Partial *cox1*, partial *cox2*	~ 5.8 kb
	R5927	Reverse	CAGTTTCGACCTGAAAGAAAGATTAA		
SET-2	F5346	Forward	GAAGCAGCAGCATGATACTGACA	Partial *cytb*	~ 7.7 kb
	R13080	Reverse	CTAATGATTATGCTACCTTTGCACGGTCAATATA		
SET-3	F12775	Forward	AACATTACGCTGTTATCCCTAAGGTA	Partial *16S*, partial *12S*	~ 5.8 kb
	R1281	Reverse	AGTAATGATNTAGACTGCAATTCTAAAG		

Sequence analysis. Annotation of the mitogenome was performed using the MITO webserver with invertebrate genetic code (Bernt *et al.* 2013) and modified after comparisons with other mitogenomes from Grylloidea species. The validation of tRNA sequences was performed in tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscanSE/) using the invertebrate mitogenome genetic codon (Schattner *et al.* 2005). The nucleotide base composition of the complete mitogenome was calculated in GENEIOUS.

Phylogenetic analysis. To validate the reliability of the mitogenome sequence and its identification, the *cytb* gene and *rrnS* sequences of this mitogenome were extracted and compared with other species of *Cardiodactylus*. In addition, a phylogenetic analysis of the family Gryllidae based on the sequences of *cytb* and *rrnS* was conducted using a Maximum Likelihood approach. The phylogenetic trees were reconstructed on the IQ-TREE web server (Nguyen *et al.* 2015).

The whole mitogenomes nucleotide base composition of *C. muiri* was calculated in GENEIOUS. The

nucleotide compositional skew was calculated following the formula (Konstantionv *et al.*, 2016): AT-skew=(A-T)/(A+T) and GC-skew=(G-C)/(G+C), where A, T, C and G are the frequencies of the four bases. Six grylline mitogenomes are currently present in GenBank: *Teleogryllus commodus* (JQ686193), *Teleogryllus emma* (EU557269), *Teleogryllus oceanicus* (KT824636), *Loxoblemmus equestris* (KU562919), *Teleogryllus emma* (KU562917) and *Velarifictorus hemelytrus* (KU562918)). Some of these sequences however include nuclear mitochondrial DNA sequences (numts), as reported from Wolff (2012) and Yang (2016) and were thus discarded here. We only considered the sequences of *L. equestris* (KU562919), *T. emma* (KU562917) *V. hemelytrus* (KU562918) and *C. muiri* which was obtained in this study. The nucleotide base composition comparison analysis for the whole mitogenomes, two ribosomal genes (the large and small ribosomal subunit, *rrnL* and *rrnS*) and the A+T-rich regions in grylline crickets were also calculated.

The percentage of pairwise identical bases (% identity) for the whole mitogenomes, each protein coding genes (PCGs) and two ribosomal genes (*rrnL* and *rrnS*) of grylline crickets were calculated in GENEIOUS. The average percentage of identical bases were then calculated between two subfamilies (Eneopterinae and Gryllinae) and within Gryllinae.

Besides, the comparison analysis of synonymous and non-synonymous substitutions (Ka/Ks) for all PCGs in grylline crickets were calculated using MEGA 7.0 (Kumar *et al.* 2016).

Acoustic analysis. The basic cricket song terminology follows Ragge & Reynolds (1998). One song unit is called a syllable and corresponds to one opening-closure cycle of the male forewings. Three males of *C. muiri* have been recorded in the field, in their natural habitat or in semi-captivity in a net cage hanged to a low tree. The recordings were made with a modified Condenser Microphone Capsule CM16 (Avisoft Bioacoustics, Berlin), with a flat frequency response from 3 to 150 kHz (R. Specht pers. comm.), connected to a Fostex FR-2LE digital field recorder (96 kilo-samples per second sampling frequency, 16 bit). Acoustic analyses were performed using the computer software Avisoft-SASLab Pro version 4.40 (Specht 2009). All recording files are deposited in the Sound Library of the MNHN.

Taxonomic part

Subfamily Eneopterinae Saussure, 1874

Tribe Lebinthini Robillard, 2004

Genus *Cardiodactylus* Saussure, 1878

Type species: *Cardiodactylus novaeguineae* (Haan, 1842)

Cardiodactylus muiri Otte, 2007a
(Figs 1–4)

Cardiodactylus muiri Otte, 2007a: 349–2007b: 30 (confirmation of depository)—Robillard *et al.* 2014: 68 (redescription).

Synonym name. *Cardiodactylus buru* Gorochov & Robillard, 2014, in Robillard *et al.* 2014: 25. New synonymy. The differences between *C. buru* and *C. muiri* in male genitalia described in Robillard *et al.* (2014) are clearly present in specimens of the same populations in the newly examined material from New Guinea and Kai Islands. These differences must then be considered as intraspecific variation of *C. muiri*; *C. buru* is consequently a junior synonym of *C. muiri*.

Type material. Male holotype: **Indonesia: Malaka Province,** Ceram, Piroe [Seram I., Piru], II.1909, F. Muir (BPBM) [not examined].

Type locality. Indonesia, Seram Island, Piru.

Distribution. Indonesia, Maluku islands (Ambon, Seram, Buru), Kai Islands, South-west part of New Guinea (surroundings of Kumawa and Kaimana).

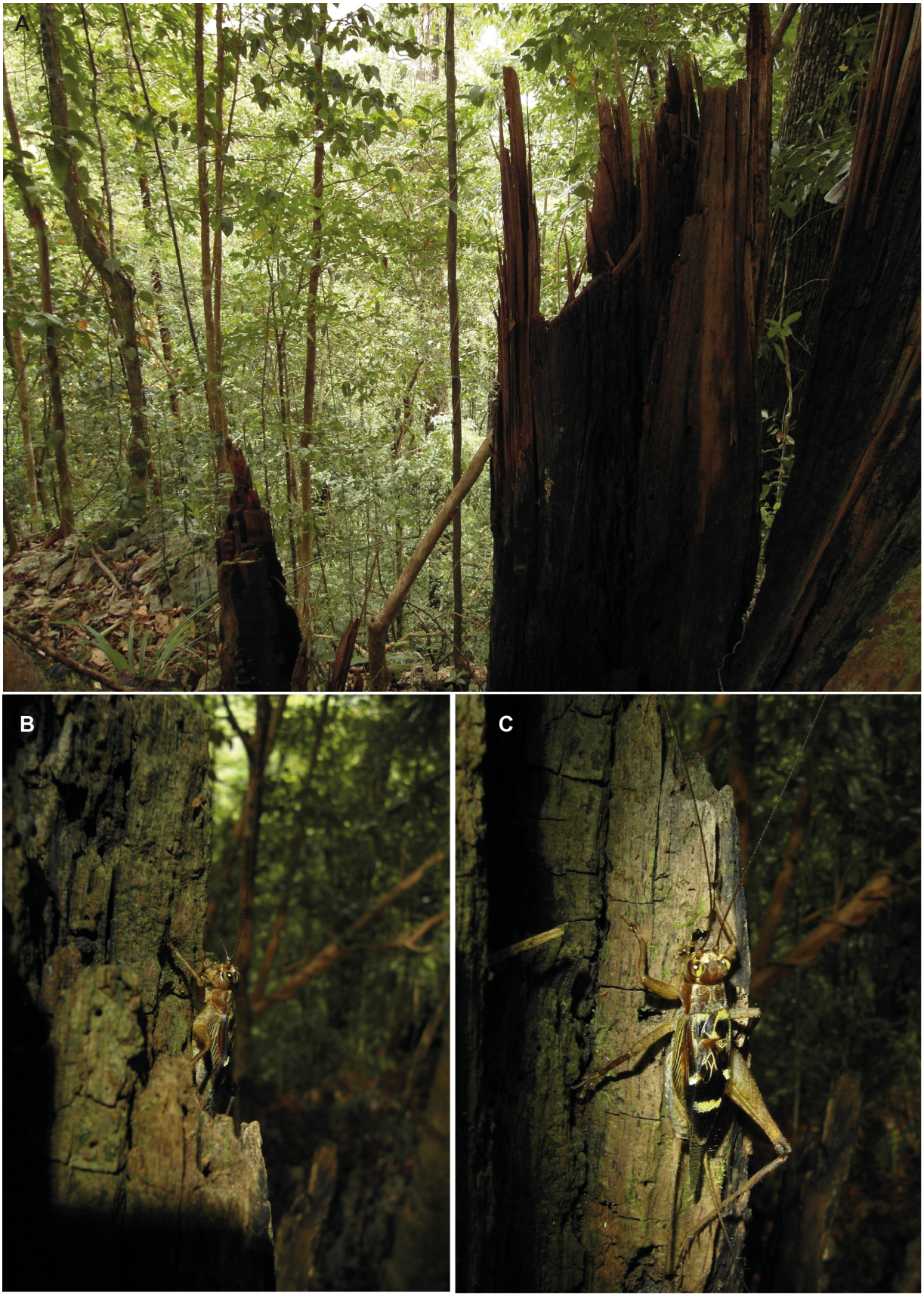

FIGURE 1. Habitat of *Cardiodactylus muiri* Otte, 2007. A, dead tree trunk in Kumawa (Papua, Indonesia); B, C, male sitting on dead tree (A) during the day.

FIGURE 2. Tree trunk with males of *Cardiodactylus muiri* Otte, 2007 engaged in singing activity during late afternoon in Kumawa forest (Papua, Indonesia). Red circles mark males' positions; height ca. 4 m.

New records: Indonesia, West Papua: Nouvelle-Guinée [New Guinea], Kumawa, -4.0646111 133.037111 (KUM3), 82 m, forêt littorale sur karst [littoral forest on karst], ligne de crête, 13–15.XI.2014, T. Robillard: 1 male (LEN2014-TR606), jour [day], tronc chandelle [trunk of dead tree], molecular sample C158 (MNHN-E0-ENSIF4179); 1 female (LEN2014-TR582), nuit [night], litière [leaf litter], vidéo ponte dans racine 8 PM [video of oviposition in root] (MNHN); 1 female (LEN2014-TR583), nuit, litière, vidéo ponte dans racine 8 PM (MNHN); 1 juvenile (LEN2014-TR569), nuit (MNHN). Nouvelle-Guinée, Kumawa, -4.0688333 133.036444 (KUM2), 27 m, forêt littorale karstique sur pente, jour, tronc chandelle, 11–17.XI.2014, T. Robillard: 1 male (LEN2014-TR625) (MZB); 1 male (LEN2014-TR587), enregistrement appel en captivité [recording of calling song in captivity] Take Pap122 (MNHN-EO-ENSIF4403); 1 juvenile (LEN2014-TR595), jour, litière (MNHN); Nouvelle-Guinée, Kumawa, -4.0518611 133.066083 (KUM5), 87 m, forêt sur pente en amont de la rivière, 16–17.XI.2014, T. Robillard: 1 male (LEN2014-TR657), enregistrement appel en semi-captivité Takes Pap142-143 (MNHN-EO-ENSIF4404); 1 male (LEN2014-TR654), enregistrement appel en semi-captivité Takes Pap140-141 (MNHN-EO-ENSIF4405); Nouvelle-Guinée, Kumawa, -4.0555 133.066333 (KUM4), 47 m, forêt littorale karstique sur pente, 16.XI.2014, T. Robillard: 1 male (LEN2014-TR617), molecular sample C160 (MZB); 1 female (LEN2014-TR655), mort élevage (MZB); 1 female (LEN2014-TR615), nuit, tronc de pandanus h=1.80 m (MZB); 1 female (LEN2014-TR616), nuit, tronc de pandanus h=1.80 m (MNHN); Nouvelle-Guinée, Kaimana, -3.643667 133.757028 (KAI3), 200 m, forêt littorale sur pente, 19.X.2014, 1 female (LEN2014-TR54), jour, plante, molecular sample C150 (MNHN-EO-ENSIF4164). Indonesia, Kei Eil. [Kai Islands], Gn. [Gunung] Daab [Pulau Kai-besar], 1922, H. C. Siebers: 4 males, #79, #131, #112, #121; 1 female #138 (MNHN); 1 male, #131, identified *Cardiodactylus haani* Ss [Saussure] by Lucien Chopard (MZB-ORTH10475).

Life history traits and habitats. *C. muiri* lives in dense forested habitats on tree trunks of various sizes or on dead trees still standing (Fig. 1). Groups of singing males (three or more) distant by ca. 50 cm to one meter, are commonly found during afternoon on large tree trunks above three meters high (Fig. 2). They sit on the bark surface and sometimes in crevices of the bark or of epiphytes. Songs were not heard during night, even when habitats with singing males heard during the day were visited after sunset. It could indicate that the species is diurnal or crepuscular and that the formed couples hide at night for mating. However, the species seems to have a

low-light preference, given that light intensity under forest cover is limited during afternoon and in the confined habitats (Fig. 1). Females are found ovipositing during early night in the leaf litter or in roots near the ground (Fig. 3).

Calling song. The calling song of *C. muiri* (Fig. 4) consists of only one long syllable. The song bouts are rather irregular, which disqualifies them as echemes. At 26 °C (MNHN-EO-ENSIF4405, measurement of 28 syllables) the call duration is 38 ± 11.4 ms (mean ± SD), with a period of 4.78 s ± 3.83 s, giving a syllable duty cycle of 8%. The dominant frequency is 11.98 ± 0.24 kHz, which is rather low for the genus and corresponds to the third peak of the frequency spectrum, the two first peaks being little marked.

FIGURE 3. Females of *Cardiodactylus muiri* Otte, 2007. A, walking on leaf litter during early night; B, ovipositing in a mossy tree root at night; in Kumawa (Papua, Indonesia).

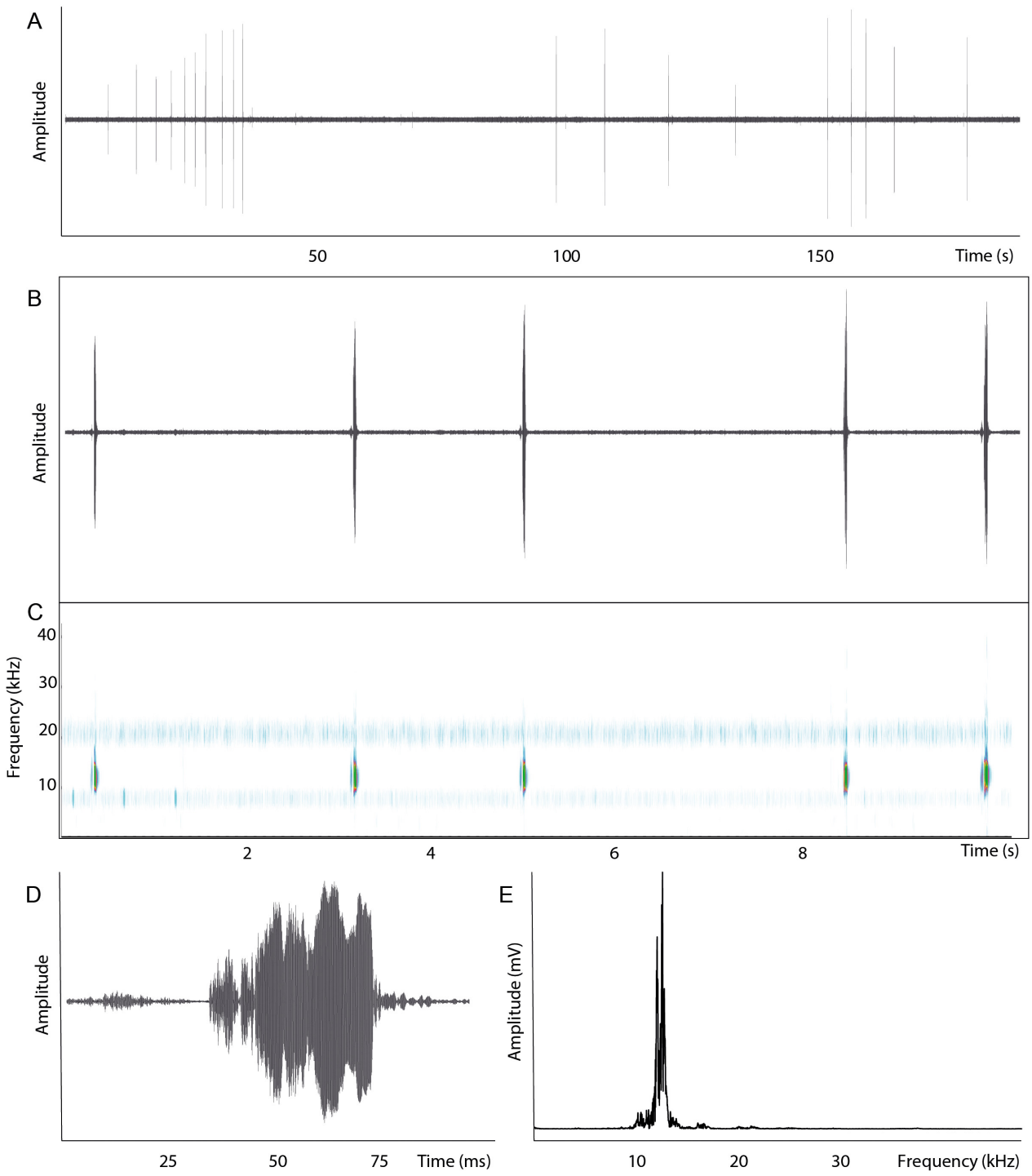

FIGURE 4. Calling song of *Cardiodactylus muiri* Otte, 2007. **A,** oscillogram of 19 syllables (= echemes); **B–C,** detailled oscillogram (B) and sonogram (C) of five syllables; **D,** detailed oscillogram of one syllable; **E,** frequency spectrum of one syllable.

Results of sequencing and phylogenetic analyses

The mitogenome is 16,328 bp long (Fig. 5), organized in the typical set of 37 genes (13 protein-coding genes, 22 transfer RNA genes and two ribosomal RNA genes) and one A+T-rich control region including two 279 bp repeated regions (Table 2).

TABLE 2. Annotation of the mitochondrial genome of *Cardiodactylus muiri* Otte, 2007.

Gene	Strand	Anticodon	Start codon	Stop codon	Position
tRNA-Ile	+	AUC			1–63 (0)
tRNA-Gln	-	UUG			61–129 (0)
tRNA-Met	+	AUG			143–211 (0)
nd2	+		ATT	T	212–1229 (0)
tRNA-Trp	+	UCA			1230–1294 (-8)
tRNA-Cys	-	GCA			1287–1348 (0)
tRNA-Tyr	-	GUA			1361–1425 (-8)
cox1	+		ATC	TAA	1418–2965 (0)
tRNA-Leu^2	+	UUA			2972–3036 (0)
cox2	+		ATG	T	3042–3720 (+3)
tRNA-Lys	+	AAG			3718–3788 (0)
tRNA-Asp	+	GAC			3789–3854 (0)
atp8	+		ATT	TAA	3855–4016 (+7)
atp6	+		ATG	TAA	4010–4693 (+1)
cox3	+		ATG	T	4693–5479 (0)
tRNA-Gly	+	GGA			5480–5542 (0)
nd3	+		ATA	TAA	5543–5896 (0)
tRNA-Ala	+	GCA			5904–5968 (0)
tRNA-Arg	+	CGA			5979–6041 (-1)
tRNA-Glu	-	UUC			6041–6105 (0)
tRNA-Ser^1	-	GCU			6110–6176 (0)
tRNA-Asn	-	GUU			6177–6244 (0)
tRNA-Phe	-	GAA			6246–6307 (0)
nd5	-		ATT	T	6308–8021 (0)
tRNA-His	-	GUG			8040–8103 (-1)
nd4	-		ATG	T	8103–9445 (+7)
nd4l	-		ATG	TAA	9439–9735 (0)
tRNA-Thr	+	ACA			9747–9810 (0)
tRNA-Pro	-	UGG			9811–9877 (0)
nd6	+		ATC	TAA	9880–10404 (+1)
cytb	+		ATG	TAA	10404–11537 (0)
tRNA-Ser^2	+	UCA			11562–11624 (0)
nd1	-		ATA	T	11625–12594 (+6)
tRNA-Leu^1	-	UAG			12589–12656 (+41)
rrnL	-				12616–13977 (+18)
tRNA-Val	-	UAC			13960–14027 (0)
rrnS	-				14028–14798 (0)
A+T-rich region	+	N/A			14857–16328 (0)

The ML gene trees, based on *cytb* and *rrnS* markers, are congruent and showed that the sequences issued from the present study clustered with published sequences of *Cardiodactylus* species and other representatives of the subfamily Eneopterinae (Fig. 6).

The nucleotide composition of *Cardiodactylus muiri* is typically insect A+T biased, with 76.5%, and is slightly

A skewed (AT-skew=0.08) and strongly C skewed (GC-skew= -0.29). The nucleotide compositional skew of the whole mitogenomes, *rrnL* genes, *rrnS* genes and the A+T-rich regions of Gryllinae crickets are also represented (Table 3, Fig. 7). The conventional start/stop codons and incomplete stop codons are represented in the PCGs (see details in Table 2). All the protein-coding genes (PCGs) in the mitogenome start with a typical ATN codon but end with different stop codons. The complete stop codon TAA are determined in genes *cox2*, *atp8*, *atp6*, *nd3*, *nd4l*, *nd6* and *cytb*, while the incomplete stop codons (T) are found in genes *nd2*, *cox1*, *cox3*, *nd5*, *nd4*, *nd1*. The lengths of 22 tRNA genes range from 62 to 71 bp. Two rRNA genes (*rrnL* and *rrnS*), separated by tRNA-Val, are located between tRNA-Leu and A+T-rich region. The lengths of *rrnL* and *rrnS* are 1,362 and 771 bp, respectively. Their A+T contents are 78.5% and 74.9%, respectively. The A+T-rich region of 1,472 bp, with an A+T content of 80.1%, is located between *rrnS* and trnI.

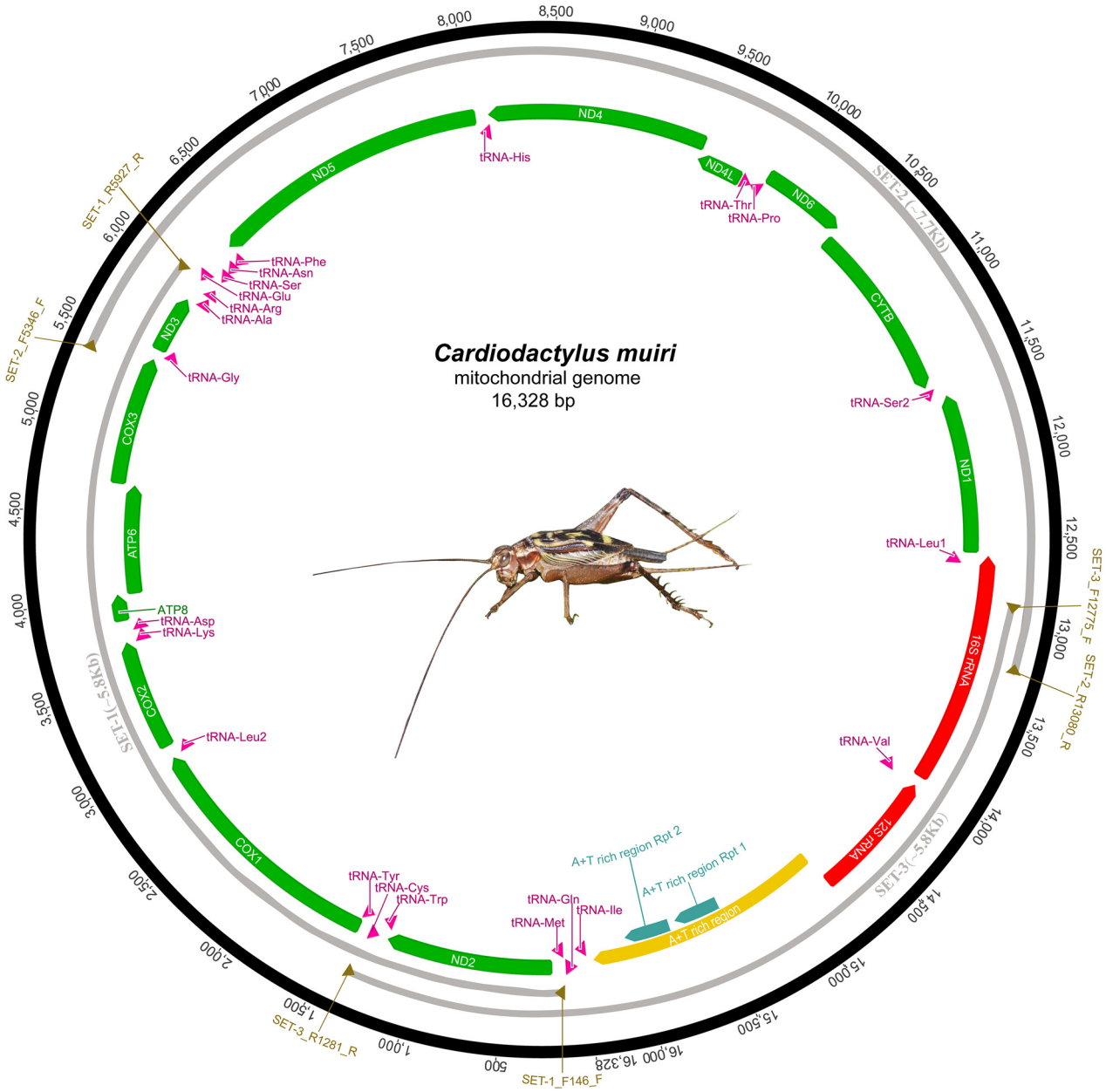

FIGURE 5. Map of the mitochondrial genome of *Cardiodactylus muiri* Otte, 2007.

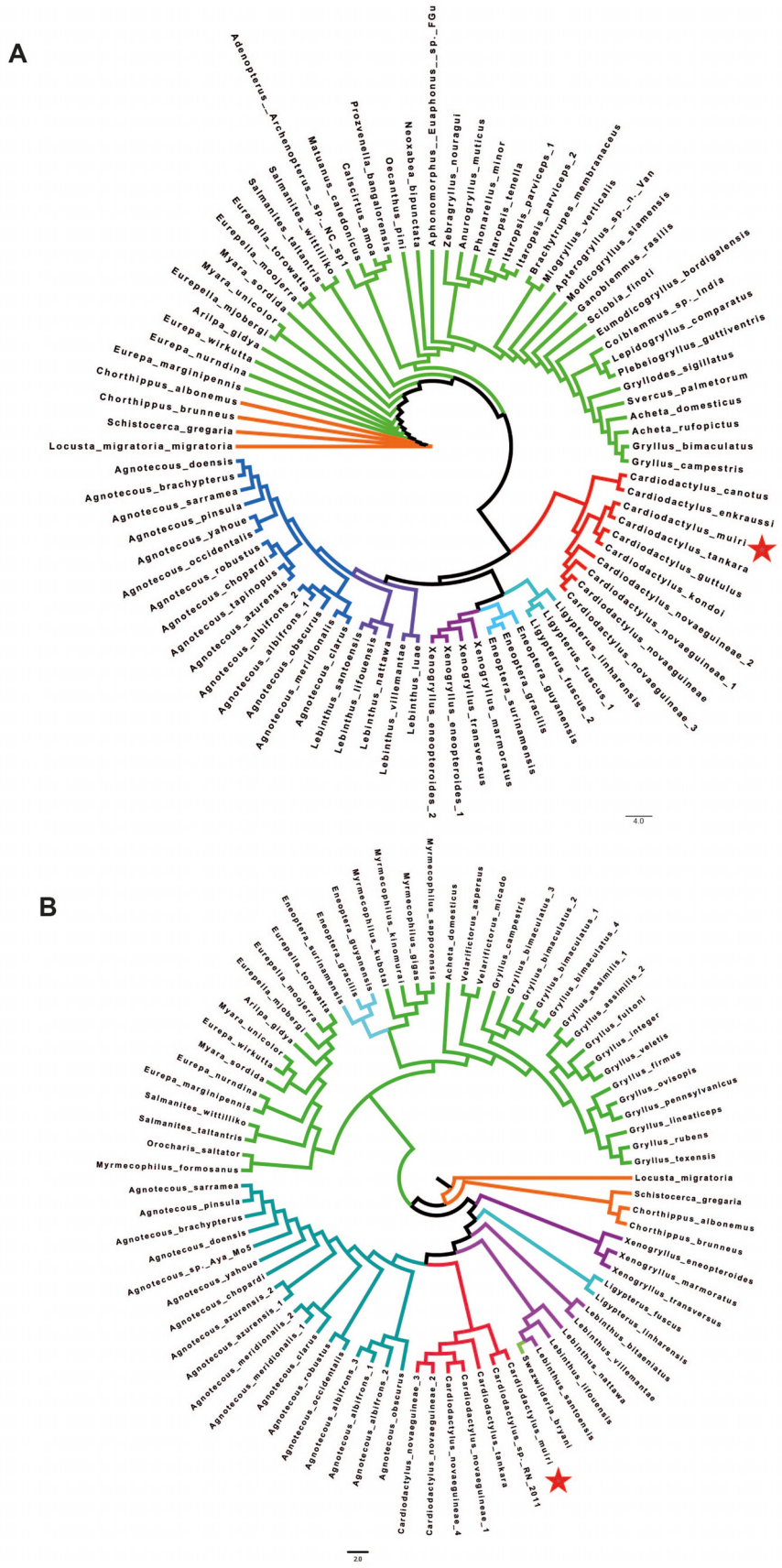

FIGURE 6. A, Maximum likelihood phylogeny of the family Gryllidae inferred from *rrnS* partial sequences available in GenBank. **B,** Maximum likelihood phylogeny of the family Gryllidae inferred from *cytb* partial sequences available in GenBank. The red star indicates the species under study.

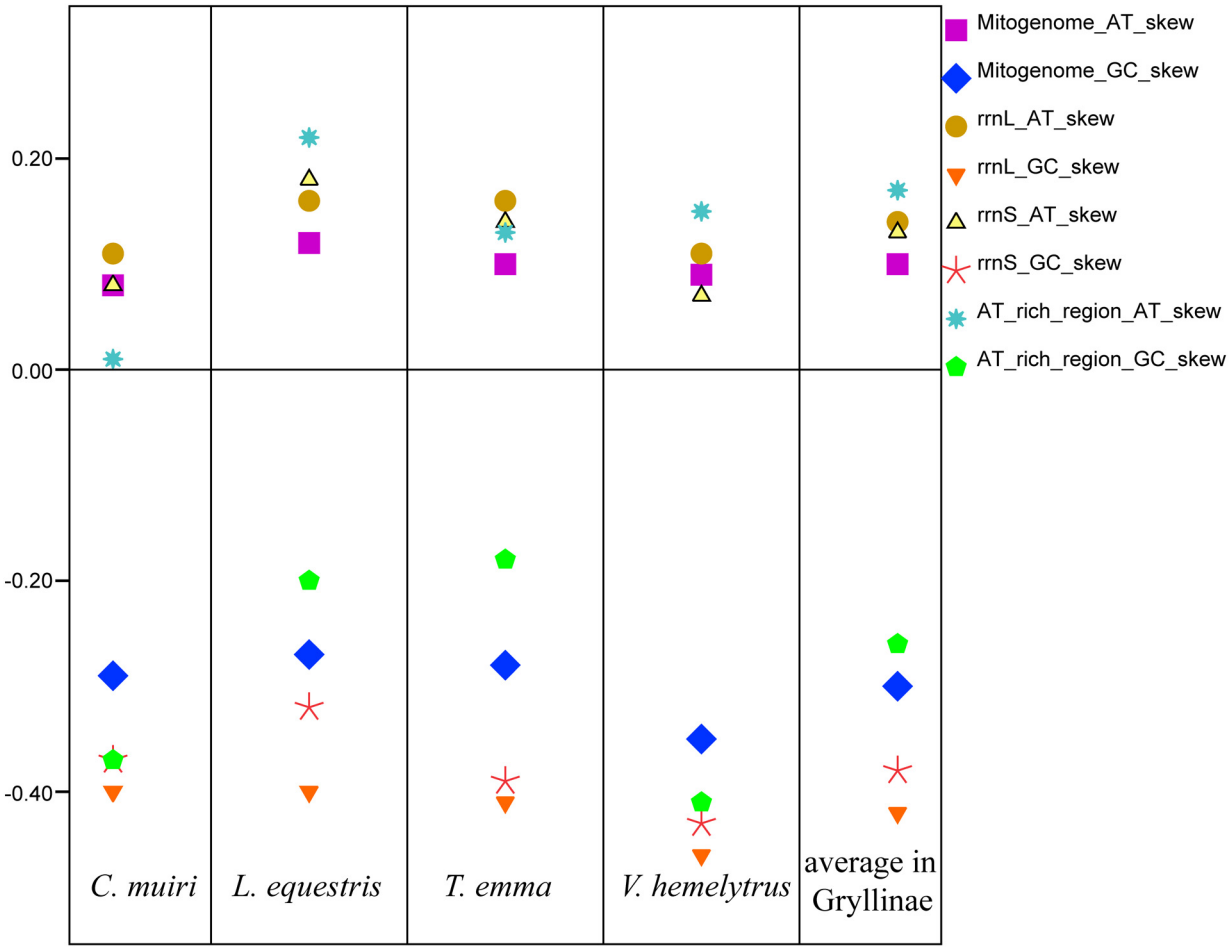

FIGURE 7. Comparison of AT-skews and GC-skews of Eneopterinae and Gryllinae.

TABLE 3. General nucleotide composition of complete mitogenome and nucleotide compositional skew of the Gryllidae mitogenomes, *rrnL* genes, *rrnS* genes and the A+T-rich regions in grylline crickets. A (%), T (%), G (%) and C (%) mean the percentage of adenine, thymine, guanine and cytosine. The number with asterisk (*) means the average nucleotide compositional skew in the subfamily Gryllinae.

	Length (bp)	whole mitogenome					
Species		A (%)	T (%)	G (%)	C (%)	AT-skew	GC-skew
Cardiodactylus muiri	16,328	41.4	35.1	8.4	15.2	0.08	-0.29
Loxoblemmus equestris	16,314	40.3	31.6	10.2	17.9	0.12	-0.27
Teleogryllus emma	15,697	40.2	33.0	9.6	17.2	0.10	-0.28
Velarifictorus hemelytrus	16,123	39.7	32.9	8.9	18.4	0.09	-0.35
						0.10*	0.30*

continued.

	rrnL genes		*rrnS* genes		A+T-rich region	
Species	AT-skew	GC-skew	AT-skew	GC-skew	AT-skew	GC-skew
Cardiodactylus muiri	0.11	-0.40	0.08	-0.37	0.01	-0.37
Loxoblemmus equestris	0.16	-0.40	0.18	-0.32	0.22	-0.20
Teleogryllus emma	0.16	-0.41	0.14	-0.39	0.13	-0.18
Velarifictorus hemelytrus	0.11	-0.46	0.07	-0.43	0.15	-0.41
	0.14*	0.42*	0.13*	0.38*	0.17*	-0.26*

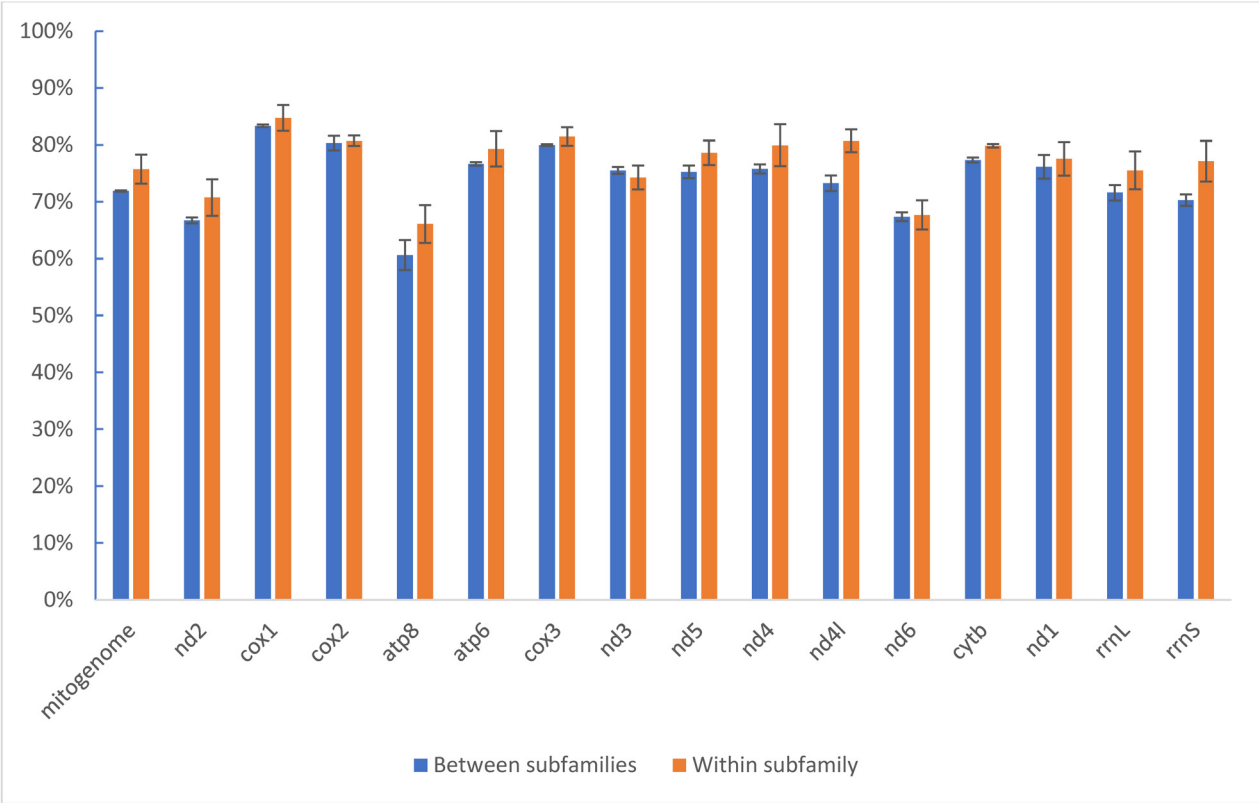

FIGURE 8. Percentage of pairwise identical bases (% identity) of mitogenome and each gene between the two subfamilies and within Gryllinae.

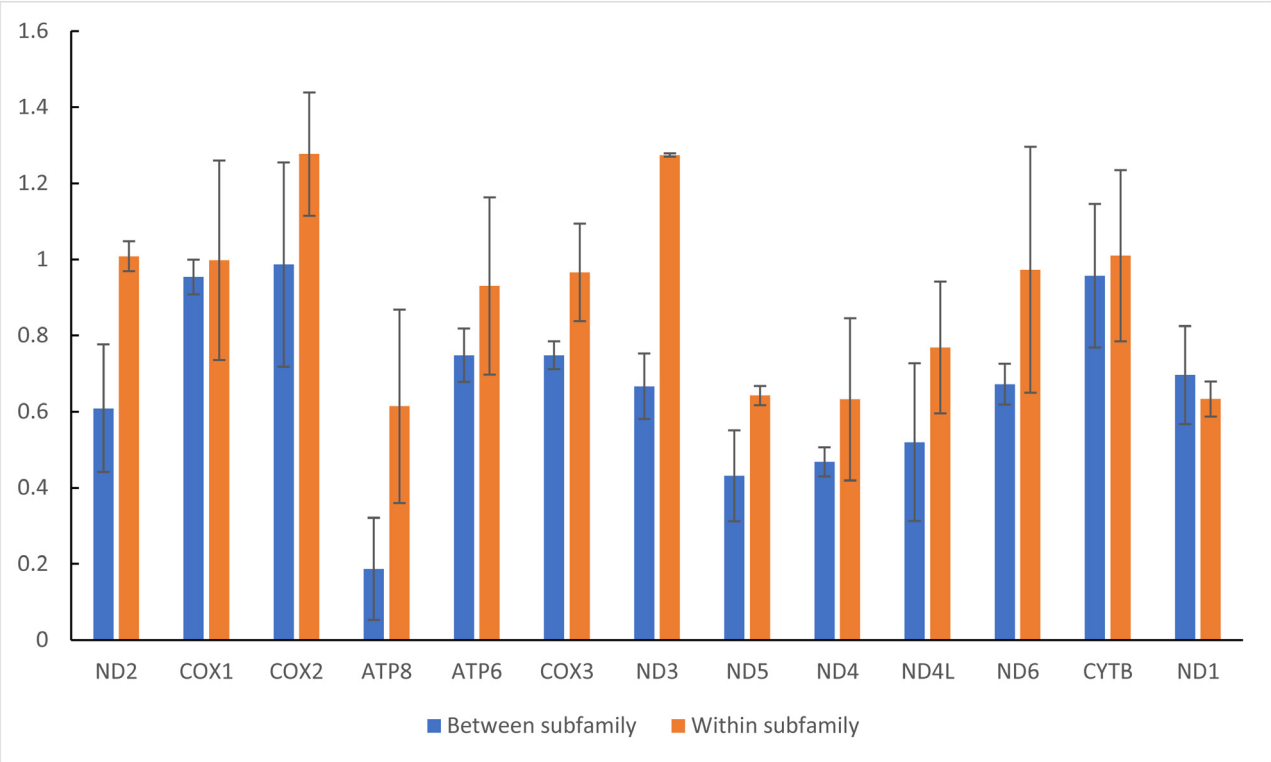

FIGURE 9. Evolutionary rate of each protein-coding (PCG) in the mitogenomes of Eneopterinae and Gryllinae.

The percentage of pairwise identical bases (% identity) for the whole mitogenomes, each PCG and two ribosomal genes, and the percentage of identity between the two subfamilies (Eneopterinae and Gryllinae) and within Gryllinae are presented in Table S1 and Fig. 8. Except for gene *nd3*, the percentage identities and the pairwise genetic distance within Gryllinae were higher than between the two subfamilies. Pairwise comparison of Ka/Ks for each PCG are also presented in Table S2 and Fig. 9. The overall Ka/Ks values are less than 1.0 between the two subfamilies, and higher than 1.0 in genes *nd1*, *cox1*, *cox2*, *nd3* and *cytb* within the subfamily Gryllinae. The values for Cytochrome Oxidase complex (*cox1*, *cox2* and *cox3*) and *cytb* are less than 1.0, but higher than for other genes between the two subfamilies. The smallest value between the two subfamilies is that for the *atp8*. The variability of Ka/Ks ratio within Gryllinae is similar with that between the two subfamilies.

Discussion

We generated the first mitogenome for the cricket subfamily Eneopterinae, which allows comparisons with existing mitogenomes from the subfamily Gryllinae. The gene order and orientation within the mitogenome of *C.muiri* are identical with those described for grylline crickets (*Loxoblemmus equestris*, *Teleogryllus emma* and *Velarifictorus hemelytrus*). The length of the whole mitogenome and the length range for tRNAs in *C. muiri* are also very close to Gryllinae (Yang *et al*. 2016b). The whole mitogenomes of these four crickets are weakly AT-skewed (0.08 to 0.10) and strongly GC-skewed (-0.35 to -0.27). The average AT- and GC-skew in the whole mitogenome of the Gryllinae are close to those of *C. muiri*. However, nucleotide compositional skew in *rrnL*, *rrnS* and A+T-rich region are stronger in Gryllinae than in *C. muiri*, with the exception of the GC-skew in A+T-rich region. It is logical that the nucleotide compositional skew in different genes of different subfamilies are slightly divergent, but more mitogenomes of other Eneopterinae crickets are necessary to test whether this tendency is consistent between the subfamilies.

There is no significant difference of the percentage identity, between the two subfamilies and within Gryllinae, for the mitogenome and the genes considered separately, except for the *atp8* and *nd6*. The percentage of pairwise identical bases for the mitogenome and for the individual genes are higher than 70%, both between the two subfamilies and within Gryllinae. The level of identity between the two subfamilies is higher or almost equal to that within Gryllinae in the genes *nd3* and *nd6*, respectively. According to the Ka/Ks ratio, the evolutionary rate of Cytochrome Oxidase complex (*cox1*, *cox2* and *cox3*) and *cytb* genes is high between the subfamilies and within Gryllinae. There is an obvious evolutionary rate discrepancy in NADH dehydrogenase complex genes and ATP synthase, since the genes *nd2* and *nd6* are less divergent than the other NADH dehydrogenase complex genes. In addition, the rate of evolution of *nd6* is almost the same, both between the two subfamilies and within Gryllinae. The gene *atp8* is found to have the slowest evolution rate between the two subfamilies, while the gene *cox2* has the highest rate of evolution within Gryllinae. Thus, the present study revealed these two genes as potential markers for future phylogenetic study at family and subfamily level in crickets. Our results lead to the conclusion that the mitogenomes in crickets have a relatively stable structural and functional evolution, as suggested by previous studies made in Orthoptera (Song *et al.* 2015) and in insects in general (e.g. Simon *et al*. (2006); Cameron (2014); Konstantinov *et al.* (2016)).

With the quick development of next generation sequencing (NGS) methods, the new sets of long-PCR primers defined here will allow to easily and cheaply acquire large amount of new data for large taxonomic samples. The new mitogenome described in the present study, combined with the existing datasets, arises as an invaluable resource for future comparative evolutionary genomic studies in crickets and insects, in general.

Acknowledgements

Fieldwork was supported by the Project Lengguru 2014 (www.lengguru.org), conducted by the French Institut de Recherche pour le Développement (IRD), the Indonesian Institute of Sciences (LIPI), the University of Papua (UNIPA), the University of Cendrawasih (UNCEN), the University of Musamus (UNMUS) and the Sorong Fisheries Academy (APSOR) with corporate sponsorship from COLAS Group, Veolia Water and the Total Foundation. Laboratory access and assistance was provided by the "Service de Systematique Moleculaire" of the

Muséum national d'Histoire naturelle, Paris (CNRS UMS 2700). This study was conducted in the context of the PhD thesis of JD, who was funded by China Scholarship Council (CSC), the Innovation Funds of Graduate Programs of Shaanxi Normal University (2012CXB019) and a supporting grant from "La Société des Amis du Muséum", and of NV, who was supported by the National Council for Scientific Development and Technological (CNPq—Brazil, PROTAX [158726/2013-8]), the Coordination for the Improvement of High Education Personnel (CAPES; PDSE—[BEX 2609/14-2]), and benefitted from financial support provided by CNPq and FAPEMIG, Process N. 563360/2010-0, Edital N. 47/2010 of the Ministério de Ciência e Tecnologia/CNPq/Ministério de Meio Ambiente/ Ministério da Educação/CAPES/Fundo Nacional de Desenvolvimento Científico e Tecnológico, Ação Transversal/Fundações de Apoio à Pesquisa, Sistema Nacional de Pesquisa em Biodiversidade, SISBIOTA Brazil. This work was also partially benefitted from the project of ICCM, who was supported by a grant from Agence Nationale de la Recherche under the LabEx ANR-10-LABX-0003-BCDiv, in the program "Investissements d'avenir" n ANR-11-IDEX-0004-02.

References

Bernt, M., Donath, A., Juhling, F., Externbrink, F., Florentz, C., Fritzsch, G., Putz, J., Middendorf, M. & Stadler, P.F. (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69, 313–319.
https://doi.org/10.1016/j.ympev.2012.08.023

Chintauan-Marquier, I.C., Legendre, F., Hugel, S., Robillard, T., Grandcolas, P., Nel, A., Zuccon, D. & Desutter-Grandcolas, L. (2016) Laying the foundations of evolutionary and systematic studies in crickets (Insecta, Orthoptera): a multilocus phylogenetic analysis. *Cladistics*, 32 (1), 54–81.
https://doi.org/10.1111/cla.12114

Cameron, S.L. (2014) Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny. *Annual Review of Entomology*, 59, 95–117.
https://doi.org/10.1146/annurev-ento-011613-162007

Crampton-Platt, A., Yu, D.W., Zhou, X. & Vogler, A.P. (2016) Mitochondrial metagenomics: letting the genes out of the bottle. *Gigascience*, 5, 15.
https://doi.org/10.1186/s13742-016-0120-y

Crampton-Platt, A., Timmermans, M.J., Gimmel, M.L., Kutty, S.N., Cockerill, T.D., Vun Khen, C. & Vogler, A.P. (2015) Soup to Tree: The Phylogeny of Beetles Inferred by Mitochondrial Metagenomics of a Bornean Rainforest Sample. *Molecular Biology and Evolution*, 32, 2302–2316.
https://doi.org/10.1093/molbev/msv111

Gomez-Rodriguez, C., Crampton-Platt, A., Timmermans, M.J.T.N., Baselga, A. & Vogler, A.P. (2015) Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution*, 6, 883–894.
https://doi.org/10.1111/2041-210X.12376

Hinsinger, D.D., Debruyne, R., Thomas, M., Denys, G.P.J., Mennesson, M., Utage, J. & Dettai, A. (2015) Fishing for barcodes in the Torrent: from COI to complete mitogenomes on NGS platforms. *DNA Barcodes*, 3, 170–186.
https://doi.org/10.1515/dna-2015-0019

Jia, H.Y., Guo, Y.F., Zhao, W.W. & Wang, K. (2014) Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. *Scientific Reports*, 4, 5737.
https://doi.org/10.1038/srep05737

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P. & Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28 (12), 1647–1649.
https://doi.org/10.1093/bioinformatics/bts199

Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33, 1870–1874.
https://doi.org/10.1093/molbev/msw054

Konstantinov, Y.M., Dietrich, A., Weber-Lotfi, F., Ibrahim, N., Klimenko, E.S., Tarasenko, V.I., Bolotova, T.A. & Koulintchenko, M.V. (2016) DNA import into mitochondria. *Biochemistry* (Moscow), 81, 1044–1056.
https://doi.org/10.1134/S0006297916100035

Nguyen, L.T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, 32, 268–274.
https://doi.org/10.1093/molbev/msu300

Ragge, D.R. & Reynolds, W.J. (1998) *The songs of the grasshoppers and crickets of Western Europe*. Harley Books, Colchester, 600 pp.

Robillard, T. & Desutter-Grandcolas, L. (2004a) High-frequency calling in Eneopterinae crickets (Orthoptera, Grylloidea, Eneopteridae): adaptive radiation revealed by phylogenetic analysis. *Biological Journal of the Linnean Society*, 83, 577–584.

Robillard, T. & Desutter-Grandcolas, L. (2004b) Phylogeny and the modalities of acoustic diversification in extant Eneopterinae (Insecta, Orthoptera, Grylloidea, Eneopteridae). *Cladistics,* 20, 271–293.

Robillard, T., ter Hofstede, H.M., Orivel, J. & Vicente, N.M. (2015) Bioacoustics of the Neotropical Eneopterinae (Orthoptera, Grylloidea, Gryllidae). *Bioacoustics-the International Journal of Animal Sound and Its Recording*, 24, 123–143. https://doi.org/10.1080/09524622.2014.996915

Robillard, T. Gorochov, A.V., Poulain, S. & Suhardjono, Y.R. (2014) Revision of the cricket genus Cardiodactylus (Orthoptera, Eneopterinae, Lebinthini): the species from both sides of the Wallace line, with description of 25 new species. *Zootaxa*, 3854 (1), 1–104. https://doi.org/10.11646/zootaxa.3854.1.1

Robillard, T., Montealegre-Z, F., Desutter-Grandcolas, L., Grandcolas, P. & Robert, D. (2013) Mechanisms of high-frequency song generation in brachypterous crickets and the role of ghost frequencies. *Journal of Experimental Biology*, 216, 2001–2011. https://doi.org/10.1242/jeb.083964

Schattner, P., Brooks, A.N. & Lowe, T.M. (2005) The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research*, 33, 686–689. [W686–689] https://doi.org/10.1093/nar/gki366

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B. & Beckenbach, A.T. (2006) Incorporating Molecular Evolution into Phylogenetic Analysis, and a New Compilation of Conserved Polymerase Chain Reaction Primers for Animal Mitochondrial DNA. *Annual Review of Ecology*, *Evolution, and Systematics*, 37, 545–579. https://doi.org/10.1146/annurev.ecolsys.37.091305.110018

Specht R. (2009) *Avisoft-SASLab Pro: Sound Analysis and Synthesis Laboratory.* Avisoft Bioacoustics, Berlin. [software]

Song, H.J., Amedegnato, C., Cigliano, M.M., Desutter-Grandcolas, L., Heads, S.W., Huang, Y., Otte, D. & Whiting, M.F. (2015) 300 million years of diversification: elucidating the patterns of orthopteran evolution based on comprehensive taxon and gene sampling. *Cladistics*, 31, 621–651. https://doi.org/10.1111/cla.12116

ter Hofstede, H.M., Schoneich, S., Robillard, T. & Hedwig, B. (2015) Evolution of a Communication System by Sensory Exploitation of Startle Behavior. *Current Biology*, 25, 3245–3252. https://doi.org/10.1016/j.cub.2015.10.064

Wolff, J.N., Shearman, D.C., Brooks, R.C. & Ballard, J.W. (2012) Selective enrichment and sequencing of whole mitochondrial genomes in the presence of nuclear encoded mitochondrial pseudogenes (numts). *PLoS One*, 7, e37142. https://doi.org/10.1371/journal.pone.0037142

Yang, J., Ren, Q. & Huang, Y. (2016a) Complete mitochondrial genomes of three crickets (Orthoptera: Gryllidae) and comparative analyses within Ensifera mitogenomes. *Zootaxa*, 4092, 529–547.

Yang, J., Ye, F. & Huang, Y. (2016b) Mitochondrial genomes of four katydids (Orthoptera: Phaneropteridae): New gene rearrangements and their phylogenetic implications. *Gene*, 575, 702–711.

Zhou, J.X., Jia, Y.C., Yang, X.C. & Li, Q. (2015) The complete mitochondrial genome of the black field cricket, *Teleogryllus oceanicus*. *Mitochondrial DNA*, 2015, 1–2.

Zhou, Z., Ye, H., Huang, Y. & Shi, F. (2010) The phylogeny of Orthoptera inferred from mtDNA and description of *Elimaea cheni* (Tettigoniidae: Phaneropterinae) mitogenome. *Journal of Genetics and Genomics*, 37, 315–324. https://doi.org/10.1016/S1673-8527(09)60049-7

TABLE S1. Genetic distance for complete mitogenome, all PCGs and two ribosomal genes showed by the percentage of pairwise identical bases (% identity). The average genetic distances and their standard deviation between two subfamilies (Eneopterinae and Gryllinae) and within the subfamily Gryllinae are also shown in the table. The abbreviations of species as following: *C. muiri* (*C*), *L. equestris* (*L*), *T. emma* (*T*) and *V. hemelytrus* (*V*).

	Percentage of pairwise identical bases between subfamilies (%)				Percentage of pairwise identical bases within Gryllinae (%)			
	C / L	*C /T*	*C / V*	Mean ± SD	*L / T*	*L / V*	*T / V*	Mean ± SD
Mitogenome	71.80	72.01	71.80	71.87±0.12	74.82	73.76	78.62	75.73±2.56
nd2	66.76	66.08	67.36	66.73±0.52	68.44	69.35	74.40	70.73±3.21
cox1	83.1	83.63	83.27	83.34±0.22	84.41	82.65	87.15	84.74±2.27
cox2	78.96	82.01	79.94	80.30±1.27	81.04	79.70	81.42	80.72±0.90
atp8	59.88	64.15	57.86	60.63±2.62	64.15	64.15	69.93	66.08±3.34
atp6	76.51	77.09	76.36	76.65±0.31	80.47	75.77	0.816	79.29±3.11
cox3	79.77	80.15	80.03	79.98±0.16	80.03	81.17	83.21	81.47±1.61
nd3	75.21	74.93	76.35	75.50±0.61	73.50	72.65	76.64	74.26±2.10
nd5	73.73	75.60	76.42	75.25±1.13	77.51	77.21	81.09	78.60±2.16
nd4	75.02	76.88	75.39	75.76±0.80	78.82	76.96	84.04	79.94±3.67
nd4l	72.11	72.45	75.17	73.24±1.37	79.25	79.93	82.99	80.72±1.99
nd6	68.01	66.29	67.82	67.37±0.77	69.19	64.72	69.19	67.70±2.58
cytb	77.90	77.28	76.83	77.34±0.44	79.54	79.89	80.16	79.86±0.31
nd1	76.19	73.55	78.70	76.15±2.10	74.22	78.60	0.798	77.55±2.95
rrnL	71.73	73.13	69.90	71.59±1.32	75.78	72.06	78.71	75.52±3.33
rrnS	69.50	71.66	69.67	70.28±0.98	76.28	74.04	81.04	77.12±3.57

TABLE S2. Comparison of Ka/Ks ratios for all PCGs. The average Ka/Ks ratio and their standard deviation between subfamilies (Eneopterinae and Gryllinae) within the subfamily Gryllinae are also shown in the table. The abbreviations of species as following: *C. muiri* (*C*), *L. equestris* (*L*), *T. emma* (*T*) and *V. hemelytrus* (*V*).

	Ka/Ks ratio pairwise comparison between subfamilies				Ka/Ks ratio pairwise comparison within Gryllinae			
Gene	*C / L*	*C / T*	*C / V*	Mean ± SD	*L / T*	*L / V*	*T / V*	Mean ± SD
nd2	0.47	0.79	0.57	0.61±0.17	1.01	0.97	1.05	1.01±0.04
cox1	1.00	0.91	0.95	0.95±0.05	1.01	1.25	0.73	1.00±0.26
cox2	1.02	0.70	1.24	0.99±0.27	1.10	1.41	1.32	1.28±0.16
atp8	0.07	0.15	0.34	0.19±0.13	0.77	0.32	0.75	0.61±0.25
atp6	0.67	0.81	0.76	0.75±0.07	0.71	1.18	0.90	0.93±0.23
cox3	0.72	0.79	0.74	0.75±0.04	1.11	0.88	0.91	0.97±0.13
nd3	0.57	0.70	0.73	0.67±0.09	1.27	1.28	1.28	1.27±0.00
nd5	0.56	0.41	0.33	0.43±0.12	0.66	0.65	0.61	0.64±0.03
nd4	0.50	0.42	0.49	0.47±0.04	0.64	0.84	0.42	0.63±0.21
nd4l	0.67	0.61	0.28	0.52±0.21	0.96	0.71	0.63	0.77±0.17
nd6	0.68	0.72	0.61	0.67±0.05	0.76	1.35	0.81	0.97±0.32
cytb	0.75	0.99	1.13	0.96±0.19	0.93	0.83	1.26	1.01±0.22
nd1	0.73	0.81	0.56	0.70±0.13	0.66	0.66	0.58	0.63±0.05