

PROOF COVER SHEET

Author(s): Dunyuan Huang, Tianjuan Su, Li Qu, Yupeng Wu, Ping Gu, Bo He, Xiaofeng Xu, and Chaodong Zhu

Article title: The complete mitochondrial genome of the *Colletes gigas* (Hymenoptera: Colletidae: Colletinae)

Article no: GDNA_A_987243

Enclosures: 1) Query sheet
2) Article proofs

Dear Author,

Please check these proofs carefully. It is the responsibility of the corresponding author to check against the original manuscript and approve or amend these proofs. A second proof is not normally provided. Informa Healthcare cannot be held responsible for uncorrected errors, even if introduced during the composition process. The journal reserves the right to charge for excessive author alterations, or for changes requested after the proofing stage has concluded.

The following queries have arisen during the editing of your manuscript and are marked in the margins of the proofs. Unless advised otherwise, submit all corrections using the CATS online correction form. Once you have added all your corrections, please ensure you press the “Submit All Corrections” button.

Please review the table of contributors below and confirm that the first and last names are structured correctly and that the authors are listed in the correct order of contribution.

Contrib. No.	Prefix	Given name(s)	Surname	Suffix
1		Dunyuan	Huang	
2		Tianjuan	Su	
3		Li	Qu	
4		Yupeng	Wu	
5		Ping	Gu	
6		Bo	He	
7		Xiaofeng	Xu	
8		Chaodong	Zhu	

AUTHOR QUERIES

Q1: Please update volume and page range for Huang et al. (2014) and Shi et al. (2013).

Q2: As per the journal's style, please provide seven authors before et al.,

MITOGENOME ANNOUNCEMENT

The complete mitochondrial genome of the *Colletes gigas*
(Hymenoptera: Colletidae: Colletinae)Dunyuan Huang^{1,2,3*}, Tianjuan Su^{2,3*}, Li Qu¹, Yupeng Wu⁴, Ping Gu², Bo He², Xiaofeng Xu¹, and Chaodong Zhu³¹College of Life Science, Institute of Genetic Resources, Nanjing Normal University, Nanjing, China, ²Forestry and Environmental Institute, Jiangxi Environmental Engineering Vocational College, Ganzhou, China, ³Key Laboratory of Zoological Systematics and Evolution (CAS), Institute of Zoology, Chinese Academy of Sciences, Beijing, China, and ⁴College of Environment and Safety, Taiyuan University of Science and Technology, Taiyuan, China

Abstract

The complete mitochondrial genome of the *Colletes gigas* (Hymenoptera: Colletidae) is determined to be 15,885 bp in length, containing 37 typical animal mitochondrial genes: 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes and a major non-coding AT-rich region. All PCGs initiate with typical ATN codons and end with the complete termination codon TAA, except for the *cob* gene, which harbors the stop codon of TAG. Twenty-four intergenic spacers (614 bp in total) and 7 overlapping regions (37 bp in total) are dispersed throughout the whole genome. The non-coding AT-rich region is 539 bp long and contains a tandem repeat region, which has also been reported in other insects. This is the first completely sequenced mitochondrial genome of the family Colletidae.

Keywords

Colletes gigas, Colletidae, Colletinae, complete mitochondrial genome, Hymenoptera

History

Received 29 October 2014

Revised 2 November 2014

Accepted 10 November 2014

Published online ■■■

In the past several decades, mitochondrial DNA has been extensively used as an informative molecular marker in various study areas, including species identification, molecular evolution, phylogenetic inference and resource conservation (Hu et al., 2010; Huang et al., 2014; Wei et al., 2014).

The Colletidae includes about 2000 species and is distributed nearly all around the world, with most of the species being found in South America and Australia (Danforth et al., 2006; Michener, 2007). Recently, Niu et al. (2013a, b, 2014a, b) reviewed different species group of the genus *Colletes* in China. They redescribed the male of the *C. gigas* from the collection of the Institute of Zoology, Chinese Academy of Sciences (Niu et al., 2013a). Besides, Zhao et al. (2010) studied the nesting biology of *C. gigas*. However, no complete mitochondrial genome has been reported from members of this family. Here, we present the first complete mitogenome sequence of the representative species of the family Colletidae, namely, *Colletes gigas*, with the aim of providing more useful information for the taxonomic and phylogenetic analysis of Hymenoptera.

The whole mitogenome of *C. gigas* is a circular molecule of 15,885 bp in length, including typical 13 protein-coding genes

(PCGs), 22 transfer RNA genes (tRNAs), 2 ribosomal RNA genes (*rrnL* and *rrnS*) and an AT-rich region (Figure 1). Besides the AT-rich region, 24 intergenic spacers (614 bp in total) and 7 overlapping regions (37 bp in total) are dispersed throughout the whole genome (Figure 1). Like the majority of other insect groups, the nucleotide composition of the *C. gigas* mitogenome is significantly biased toward AT (86.2%). The overall base composition is A: 42.1%, T: 44.1%, C: 7.4% and G: 6.4%. The mitogenome nucleotide skewness (AT-skew = -0.023, GC-skew = -0.086) indicates slight T skew and C skew.

All PCGs begin with typical ATN codons (four ATA, five ATT and four ATG) and terminate with the complete stop codon TAA, except for the *cob* gene, which harbors the stop codon of TAG (Figure 1).

All tRNA genes have the typical cloverleaf structures, except for *trnS2*, whose dihydrouridine arm forms a simple loop, as observed in other determined insect mitogenomes (Sheffield et al., 2008; Wei et al., 2010). Furthermore, the large and small subunit rRNA genes were 1345 bp and 836 bp, with their A + T contents of 87.0% and 88.9%, respectively.

The major non-coding sequence of *C. gigas* mitochondrial genome is identified as the control region, which is located between *trnV* and *trnM* with a length of 539 bp. The base composition of the control region is A: 41.4%, T: 42.1%, C: 6.7% and G: 9.8%. And the high A + T content (83.5%) is common in the hymenopteran mitochondrial genome. With the similarity of 96%, this region harbors two nearly identical 153 bp long tandem repeats and one 122 bp long incomplete repeat. The presence of variable copy numbers of tandemly repeated sequences was considered as one of the characteristics of insect control region (Kim et al., 2010; Shi et al., 2013; Zhang et al., 2013).

*These authors contributed equally to this work.

Correspondence: Xiaofeng Xu, College of Life Science, Institute of Genetic Resources, Nanjing Normal University, No. 1 Wenyuan Road, Nanjing, China. E-mail: xuxiaofeng@njnu.edu.cn
Chaodong Zhu, Key Laboratory of Zoological Systematics and Evolution (CAS), Institute of Zoology, Chinese Academy of Sciences, Beijing, China. E-mail: zhucd@ioz.ac.cn

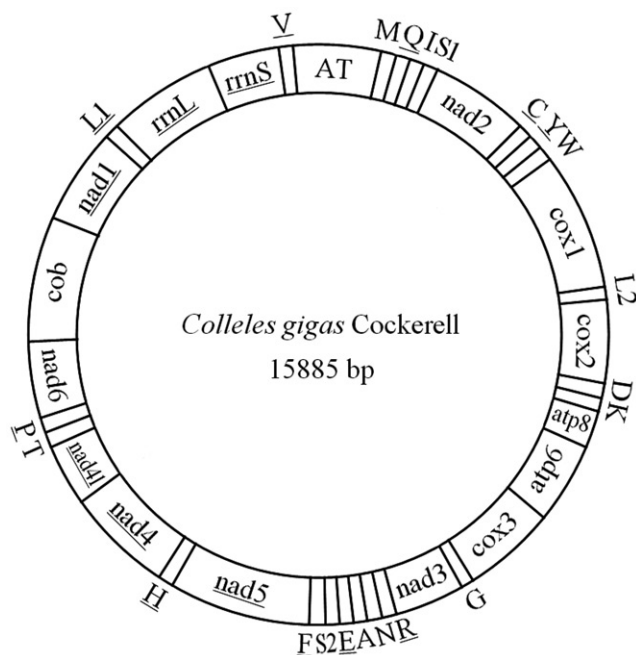


Figure 1. Organization of *Colleles gigas* mitochondrial genome. Transfer RNA genes are denoted according to the IPUC-IUB single-letter amino acid codes. AT indicates AT-rich region. All of the genes are transcribed in a clockwise direction, except those underlined.

Nucleotide sequence accession number

The complete genome sequence of *C. gigas* has been assigned GenBank accession number KM978210.

Declaration of interest

This work was supported by the National Science Foundation of China (Grant No. J1211002) and the Science and Technology Research Project from Education Department of Jiangxi Province (Grant No. 14805). The authors alone are responsible for the content and writing of the article, and report no conflicts of interest.

References

Danforth BN, Sipes S, Fang J, Brady SG. (2006). The history of early bee diversification based on five genes plus morphology. *P Natl Acad Sci USA* 103:15118–23.

- Huang DY, Hao JS, Zhang W, Su TJ, Wang Y, Xu XF. (2014). The complete mitochondrial genome of *Melanargia asiatica* (Lepidoptera: Nymphalidae: Satyrinae). *Mitochondrial DNA*. doi: 10.3109/19401736.2014.919452.
- Hu J, Zhang DX, Hao JS, Huang DY, Cameron S, Zhu CD. (2010). The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini): Sequence, gene organization and a unique tRNA translocation event. *Mol Biol Rep* 37: 3431–8.
- Kim MJ, Wan X, Kim KG, Hwang JS, Kim I. (2010). Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *Afr J Biotechnol* 9:735–54.
- Michener CD. (2007). *The bees of the world*. Baltimore: The Johns Hopkins University Press.
- Niu ZQ, Khulman M, Zhu CD. (2013a). A review of the *Colletes succinctus*-group (Hymenoptera: Colletidae) from China with redescription of the male of *C. gigas*. *Zootaxa* 3626:173–87.
- Niu ZQ, Zhu CD, Khulman M. (2013b). Bees of the *Colletes clypearis*-group (Hymenoptera: Apoidea: Colletidae) from China with descriptions of seven new species. *Zootaxa* 3745:101–51.
- Niu ZQ, Zhu CD, Khulman M. (2014a). Bees of the *Colletes flavicornis*-group from China with description of one new species (Hymenoptera: Apoidea: Colletidae). *Zootaxa* 3780:534–46.
- Niu ZQ, Zhu CD, Khulman M. (2014b). The bees of the genus *Colletes* (Hymenoptera: Apoidea: Colletidae) from China. *Zootaxa* 3856: 451–83.
- Sheffield NC, Song H, Cameron SL, Whiting MF. (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Mol Biol Evol* 25: 2499–509.
- Shi QH, Huang DY, Wang YL, Hao JS. (2013). The complete mitochondrial genome of Blue Pansy, *Junonia orithya* (Lepidoptera: Nymphalidae: Nymphalinae). *Mitochondrial DNA*. doi: 10.3109/19401736.2013.823182.
- Wei SJ, Li Q, Achterberg K, Chen XX. (2014). Two mitochondrial genomes from the families Bethyridae and Mutillidae: Independent rearrangement of protein-coding genes and higher-level phylogeny of the Hymenoptera. *Mol Phylogenet Evol* 77:1–10.
- Wei SJ, Tang P, Zheng LH, Shi M, Chen XX. (2010). The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A + T content and a long intergenic spacer between *atp8* and *atp6*. *Mol Biol Rep* 37:1931–42.
- Zhao YH, Ding L, Yuan F, et al. (2010). Nesting biology of the *Colletes gigas* Cockerell (Hymenoptera: Colletidae). *Acta Entomologica Sinica* 53:1287–94.
- Zhang LL, Huang DY, Sun XY, Hao JS, Hao JJ, Peng CM, Yang Q. (2013). The complete mitochondrial genome of *Cupido argiades* (Lepidoptera: Lycaenidae). *Mitochondrial DNA* 24:475–7.

133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198

199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264