The mitochondrial genome of the land snail *Camaena cicatricosa* (Müller, 1774) (Stylommatophora, Camaenidae): the first complete sequence in the family Camaenidae

Pei Wang¹, Hai-Fan Yang², Wei-Chuan Zhou¹, Chung-Chi Hwang³, Wei-Hong Zhang⁴, Zhou-Xing Qian⁵

¹ Key Laboratory of Molluscan Quarantine and Identification of AQSIQ, Fujian Entry-Exit Inspection & Quarantine Bureau, Fuzhou, Fujian 350001, China ² National Wetland Museum of China, Hangzhou, Zhejiang 310013, China ³ Department of Life Sciences, National University of Kaohsiung, No.700, Kaohsiung University Road, Nan-Tzu District, Kaohsiung 81148, Taiwan ⁴ College of Life Science and Technology, Xinjiang University, Urumqi, Xinjiang 830046, China ⁵ Zhejiang Museum of Natural History, Hangzhou, Zhejiang 310014, China

Corresponding author: Wei-Chuan Zhou (wczhou@163.com); Chung-Chi Hwang (cchwang@nuk.edu.tw)

Academic editor: F. Köhler | Received 2 September 2014 | Accepted 1 October 2014 | Published 3 November 2014

http://zoobank.org/2CE08256-9EA9-4C68-9CCD-12DAE6CCB062

Citation: Wang P, Yang H-F, Zhou W-C, Hwang C-C, Zhang W-H, Qian Z-H (2014) The mitochondrial genome of the land snail *Camaena cicatricosa* (Müller, 1774) (Stylommatophora, Camaenidae): the first complete sequence in the family Camaenidae. ZooKeys 451: 33–48. doi: 10.3897/zookeys.451.8537

Abstract

The complete mitochondrial (mt) genome of the snail *Camaena cicatricosa* (Müller, 1774) has been sequenced and annotated in this study. The entire circular genome is 13,843 bp in size and represents the first camaenid mt genome, with content of 31.9%A, 37.9%T, 13.5%C and 16.7%G. Gene content, codon usage and base organization show similarity to a great extent to the sequenced mt genome from Stylommatophora, whereas, gene order is different from them, especially the positions of tRNA\(^{Cys}\), tRNA\(^{Phe}\), COII, tRNA\(^{Asp}\), tRNA\(^{Gly}\), tRNA\(^{His}\) and tRNA\(^{Trp}\). All protein coding genes use standard initiation codons ATN except for COII with GTG as start signal. Conventional stop codons TAA and TAG have been assigned to all protein coding genes. All tRNA genes possess the typical clover leaf structure, but the T\(^\psi\)C arm of tRNA\(^{Asp}\) and dihydrouridine arm of tRNA\(^{Ser(AGN)}\) only form a simple loop. Shorter intergenic spacers have been found in this mt genome. Phylogenetic study based on protein coding genes shows close relationship of Camaenidae and Bradybaenidae. The presented phylogeny is consistent with the monophyly of Stylommatophora.
Keywords
Camaena cicatricosa, Camaenidae, Stylommatophora, mitochondrial genome, secondary structure

Introduction

The mitochondrial (mt) genome of metazoa usually comprise 37 genes, including 13 protein coding genes (PCGs) (\textit{COI}–\textit{COIII}, \textit{Cytb}, \textit{ND1}–\textit{ND6}, \textit{ND4L}, \textit{ATP6} and \textit{ATP8}), two ribosomal RNA (rRNA) genes, and 22 transfer RNA (tRNA) genes (Boore 1999). Additionally, it also contains noncoding regions, such as the AT-rich region and short intergenic spacers (Wolstenholme 1992). The mt genome is characterized by small size (13–36 kb), maternal inheritance, lack of recombination, conserved genomic organization and rapid evolutionary rate compared with the nuclear genome (Avise 1994). It has been widely used in studies of systematics, phylogenetic analysis, phylogeography, population structure at diverse taxonomic groups (White et al. 2011; Gaitán-Espitia et al. 2013; Menegon et al. 2014). The mt genome is the most popular genetic marker though there are numerous debates on their utilization in systematic research (Delsuc et al. 2003; Cameron et al. 2004; Talavera and Vila 2011; Simon and Hadrys 2013; Cameron 2014). Over the last years, next generation sequencing technologies have accelerated further developments of mt genomics. The mt genomes of many vertebrates and insects are well sequenced and studied (Boore 1999; Hahn et al. 2013; Wang et al. 2014). However, studies on molluscan mt genomes are poor relatively (Kurabayashi and Ueshima 2000; Boore et al. 2004; Grande et al. 2008). Only 80 mt genomes of Gastropoda snails have been deposited in GenBank (up to 2014.9.20).

Camaenidae, one of the most diverse families, was erected by Pilsbry in 1893 (Pilsbry 1893–1895). The camaenids mainly feed on green plants and humus, and often harm a large number of crops, landscape plants and forest, leading to a depression in yield and a reduction in quality. Besides, they can spread zoonotic food borne parasitic disease and have great damage to human and animal health (Zhou et al. 2007). When humans are infected by ingesting snails, the nervous system can be injured (Liang and Pan 1992). The camaenids also play an important part in agricultural production and human activities as food, drug, arts, crafts, etc. (Chen and Gao 1987). \textit{Camaena cicatricosa} (Müller, 1774), the type species of the type genus \textit{Camaena} (Albers, 1850), occurs only in China, distributing in Guangdong, Guangxi, Guizhou, Yunnan and Hainan. Adult shell is large, thick and depressed conic. This snail usually feeds a broad range of fruits, vegetables, leaves and weeds (Xiao 1989).

The mt genome of land snail is similar to other invertebrates in containing 37 genes. Since the first mt genome of \textit{Albinaria caerulea} was obtained in 1995 (Hatzoglou et al. 1995), only ten mt genomes from eight species in the order Stylommatophora were determined prior to this study, consisting of three species in Helicidae (Terrett et al. 1995; Groenenberg et al. 2012; Gaitán-Espitia et al. 2013), two in Bradybaenidae (Yamazaki et al. 1997; Deng et al. 2014), one in Clausiliidae (Hatzoglou et al. 1995), one in Succineidae (White et al. 2011) and one in Achatinidae (He et al. 2014).
Although researchers have done some phylogenetic studies on Camaenidae, they often pay much attention to analyses of shell morphology or single gene fragment (Scott 1996; Wade et al. 2007). Complete mt genome evidence is still limited. We select *C. cicatricosa* as subject because of not only relatively wide distribution and varied morphology but also acting as type species of the type genus *Camaena*. We have analyzed nucleotide composition, codon usage, compositional biases, and constructing models of the secondary structure of tRNAs. Besides, we also discussed the phylogenetic relationships with other representative gastropods. This snail mt genome is the first model in the family Camaenidae, thus it can offer worthwhile information to other camaenids.

**Materials and methods**

**Genomic DNA extraction, PCR amplification and DNA sequencing**

Adults of *C. cicatricosa* were collected from Xishan Park in Guiping (23°23'58"N, 110°3'46"E), Guangxi, China in November 2, 2013. Specimens were initially preserved in 100% ethanol in the field, and then stored at -20 °C at Fujian Entry-Exit Inspection & Quarantine Bureau (FJCIQ). Total genomic DNA was extracted from the pedal muscle tissue of single individual using the DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions. Voucher specimen (FJCIQ 18483) is deposited at the Key Laboratory of Molluscan Quarantine and Identification of AQSIQ, Fujian Entry-Exit Inspection & Quarantine Bureau, Fuzhou, Fujian.

The entire genome was successfully amplified by polymerase chain reaction (PCR) in overlapping fragments with four pairs of mitochondrial universal primers from previous works (Palumbi et al. 1991; Folmer et al. 1994; Merritt et al. 1998; Hugall et al. 2002), and four pairs of perfectly matched specific primers designed from sequenced short fragments in this study (Table 1). Short PCRs (< 2 kb) were performed using Takara Taq DNA polymerase (TaKaRa, Dalian, China), with the following cycling conditions: 30 s at 94 °C, followed by 35 cycles of 10 s at 94 °C, 50 s at 40 °C or 45 °C, and 1 min at 72 °C. The final elongation step was continued for 10 min at 72 °C. Long range PCRs (> 4 kb) were performed using Takara Long Taq DNA polymerase (TaKaRa, Dalian, China) under the following cycling conditions: 1 min at 94 °C, followed by 40 cycles of 10s at 98 °C, 50 s at 60 °C, 4−8 min at 68 °C, and the final elongation step at 72 °C for 6 min. The PCR products were checked by spectrophotometry and 1.0% agarose gel electrophoresis.

Short fragments were sequenced from both directions after purification using the BigDye Terminator Sequencing Kit (Applied Biosystems, San Francisco, CA, USA) and the ABI PRIMER™3730XL DNA Analyzer (PE Applied Biosystems) with internal primers for primer walking. For the long fragments, the shotgun libraries of *C. cicatricosa* were constructed, and then the positive clones were sequenced using above kit and sequenator with vector-specific primers *Bea*Best primer M13-47 and *Bea*Best Primer RV-M.
Raw sequences were proof-read and aligned into contigs with BioEdit v.7.0.5.3 (Hall 1999). The tRNA genes were identified with tRNAscan-SE Search Server v.1.21 (Lowe and Eddy 1997) and DOGMA (Wyman et al. 2004), while others that could not be determined by these two tools were predicted by similarity comparison with other published land snails (Terrett et al. 1995; Yamazaki et al. 1997; Groenenberg et al. 2012; Gaitán-Espitia et al. 2013; He et al. 2014; Deng et al. 2014). The PCGs and rRNA genes were annotated by BLAST in Genbank with published available mitochondrial sequences of terrestrial snails.

PCGs were aligned with Clustal X (Thompson et al. 1997). The nucleotide composition and codon usage were analyzed with MEGA 5.0 (Tamura et al. 2011). Strand asymmetry was denoted by skew values, which were calculated according to the formulas: AT skew = (A−T)/(A+T) and GC skew = (G−C)/(G+C) (Perna and Kocher 1995).

Phylogenetic analyses were performed based on 11 representative gastropod mt genomes from GenBank (Table 2) using maximum likelihood (ML) and maximum parsimony (MP) methods. One species of Opisthobranchia was selected as outgroup. A DNA alignment with 9,892bp length was inferred from the amino acid alignment of 13 PCGs using MEGA 5.0 (Tamura et al. 2011). The selection of best-fit-substitution model for ML estimation was performed using MEGA 5.0 with corrected Akaike information criterion (AIC). Node supports for ML and MP analyses were calculated through 1000 bootstrap replicates. All other settings were kept as default.

| No. of fragment | Primer name | Nucleotide sequence (5’–3’) and location | Size (bp) | Reference |
|-----------------|-------------|------------------------------------------|----------|-----------|
| 1               | LCO-1490    | GGTCAACAAATCATAAAGATATTGG                |          | Folmer et al. 1994 |
|                 | HCO-2198    | TAAACTGAGGTGACCAAATAATCA                |          | Folmer et al. 1994 |
| 2               | Fcoi        | TGAACATGTTATCCTCCAC (364–382)           | 1908     | Present study |
|                 | RL          | TAGGGTCTTCGTCCTTTT (2254–2271)          |          | Present study |
| 3               | 16Sar-L     | CGCCTGTGTTATCAAAAACAT                   |          | Palumbi et al. 1991 |
|                 | 16Sbr-H     | CCGGTCTGAACTCAGATACACGT                 |          | Palumbi et al. 1991 |
| 4               | FL2         | CGATGTTGGATTAGGAAGTGA (2415–2436)       | 4267     | Present study |
|                 | Rch2        | TAAAGGATTTTGGACCCAG (6661–6681)         |          | Present study |
| 5               | 144F        | TGAGSNCAATGTCNTWYTG                     |          | Merritt et al. 1998 |
|                 | 272R        | GCRAANAGRAARTACCAYTC                    |          | Merritt et al. 1998 |
| 6               | Fcb         | GTGGGTCAACAAATCTCTT (6662–6679)         | 816      | Present study |
|                 | Rcoii       | ATGAACACCTCGGGTAGTT (7460–7477)         |          | Present study |
| 7               | FCOII       | AAAATAATGCTATTTCATGAYCAY                 |          | Hugall et al. 2002 |
|                 | RCOII       | GCCTCGCAAATCTCTGARCTY                  |          | Hugall et al. 2002 |
| 8               | SF1F        | AAATTCCATTAGAGGGCTTATACGCGGC           | 6957     | Present study |
|                 | SF1R        | CAAGAGATAGTCCCGTGACATACTGCGC           | (68–79)  | Present study |

**Table 1.** Primer pairs used for PCR amplification.

Genome annotation and inference of secondary structure

Phylogenetic analyses were performed based on 11 representative gastropod mt genomes from GenBank (Table 2) using maximum likelihood (ML) and maximum parsimony (MP) methods. One species of Opisthobranchia was selected as outgroup. A DNA alignment with 9,892bp length was inferred from the amino acid alignment of 13 PCGs using MEGA 5.0 (Tamura et al. 2011). The selection of best-fit-substitution model for ML estimation was performed using MEGA 5.0 with corrected Akaike information criterion (AIC). Node supports for ML and MP analyses were calculated through 1000 bootstrap replicates. All other settings were kept as default.
The mitochondrial genome of the land snail Camaena cicatricosa (Müller, 1774)...

Results and discussion

The complete mt genome of C. cicatricosa was a double-stranded circular molecule of 13,843 bp in length (GenBank: KM365408). It contained 13 PCGs, 22 tRNA genes, two rRNA genes, similar to other mt genomes of land snails from the order Stylommatophora. All genes were divided into two groups, encompassing 24 genes on the majority coding strand (J strand) and others on the minority coding strand (N strand) (Fig. 1). However, the gene arrangement differed from that of the known land snails in the order Stylommatophora, specially the locations of tRNACys, tRNA Phe, COII, tRNA Asp, tRNA Gly, tRNA His and tRNA Trp (Fig. 2). Gene overlaps with a total of 242 bp were found at 16 gene junctions, and the longest overlap (50 bp) existed between ND6 and ND5. Besides, there were 144 nucleotides dispersed in 14 intergenic spacers with the shortest 1 bp and the longest 29 bp. The 29 bp long noncoding region was situated between COIII and tRNAIle; the shortest 1bp in three gene spacers (Table 3).

Protein coding genes

The length of PCGs was 10,941bp, accounting for 79.04% of the whole mt genome (Table 4). Most PCGs started with ATN as initiation codons (four with ATG, three with ATT, and five with ATA) except for COII gene with GTG (Table 3), while ATC, TTA, TTG, CTT and TCG as unconventional start signals have been found in other invertebrates (Raay and Crease 1994; Crease 1999; Yamazaki et al. 1997; Yu et al. 2007; Groenenberg et al. 2012). Conventional stop codons TAA and TAG had been assigned to all PCGs (Table 3). However, an incomplete terminator signal (T) has been found in other snails (Terrett et al. 1995; Hatzoglou et al. 1995; Yamazaki et al. 1997; White et al. 2011; Groenenberg et al. 2012; Gaitán-Espitia et al. 2013).

Table 2. Summary of samples information used in this study.

| Subclass / order | Family     | Species                  | Accession number | Reference              |
|------------------|------------|--------------------------|------------------|------------------------|
| Stylommatophora  | Camaenidae | Camaena cicatricosa      | KM365408         | Present study          |
|                  | Bradybaenidae | Eubadra berkloesi  | Z71693–Z71701   | Yamazaki et al. 1997   |
|                  | Helicidae | Cornu asperum            | JQ417195         | Gaitán-Espitia et al. 2013 |
|                  | Cepaea nemoralis | CMU23045   |                  | Terrett et al. 1995    |
|                  | Cylindrus obtusus | JN107636   |                  | Groenenberg et al. 2012 |
|                  | Succineidae | Succinea putris          | JN627206         | White et al. 2011      |
|                  | Clausiliidae | Albinaria caerulea       | X83390           | White et al. 2011      |
|                  | Achatinidae | Achatina fulica          | NC024601         | He et al. 2014         |
| Basommatophora   | Lymnaeidae | Galba pervaia            | JN564796         | Liu et al. 2012        |
|                  | Aplysiidae | Aplysia californica      | AY569552         | Knudsen et al. 2006    |

Table 3. Summary of the protein coding genes of C. cicatricosa.

| Gene | Start codon | Stop codon |
|------|-------------|------------|
| COI  | ATG         | TAG/TAA    |
| COII | GTG         | TAA/TAAG   |
| COIII| ATG         | TAG/TAG    |
| ND1  | ATN         | TAG/TAG    |
| ND2  | ATN         | TAG/TAG    |
| ND3  | ATN         | TAG/TAG    |
| ND4  | ATN         | TAG/TAG    |
| ND5  | ATN         | TAG/TAG    |
| ND6  | ATN         | TAA/TAAG   |
| ND7  | ATN         | TAG/TAG    |
| ND8  | ATN         | TAG/TAG    |
| tRNAAsp | ATN          |                |
| tRNAVal | ATN          |                |
| tRNAIle | ATN         |                |
| tRNAMet | ATN         |                |
| tRNATrp | ATN         |                |
Transfer RNA genes

The 22 tRNA genes typically found in metazoan mt genomes were also discovered in *Camaena cicatricosa*, and 18 of them were determined by tRNAscan-SE (Lowe and Eddy 1997) and DOGMA (Wyman et al. 2004). The other four tRNA genes that could not be detected by the two programs were identified and drawn through comparison with known patterns of previous researches (Terrett et al. 1995; Grande et al. 2002; Groenenberg et al. 2012; Gaitán-Espitia et al. 2013). Fourteen tRNA genes were encoded on the J strand and the remainings on the N strand. Most
The mitochondrial genome of the land snail *Camaena cicatricosa* (Müller, 1774)...

**Camaena cicatricosa**

**Cornu aspersum**

**Cepaea nemoralis**

**Cylindrus obtusus**

**Euhadra herklotsi**

**Mastigeulota kiangsinsensis**

**Succinea putris**

**Albinaria caerulea**

**Achatina fulica**

**Figure 2.** Gene arrangement of nine mt genomes in the order Stylommatophora.

tRNA genes could be folded into classic clover leaf structures exclusive of \(tRNA^{Asn}\) and \(tRNA^{Ser(AGN)}\), in which their T\(\psi\)C arm and dihydrouridine (DHU) arm simply formed a loop (Fig. 3).
The length of tRNA genes ranged from 53 to 65 bp (Table 3). All amino acid acceptor (AA) arms (7 bp), anticodon (AC) loops (7 bp) and arms (5 bp) were almost invariant. However, other arms and loops changed considerably in size. Additionally, in some tRNA genes, non-Watson-Crick matches and aberrant loops...
The mitochondrial genome of the land snail *Camaena cicatricosa* (Müller, 1774)...

### Table 3. Organization of the *Camaena cicatricosa* mt genome.

| Gene    | Direction | Location | Size (bp) | Anticodon | Start codon | Stop codon | Intergenic nucleotides |
|---------|-----------|----------|-----------|-----------|-------------|------------|------------------------|
| COI     | F         | 1–1527   | 1527      | ATG       | TAG         |            |                        |
| tRNA<sup>cyt</sup> | F       | 1527–1585 | 59        | 1557–1559 TAC |             | –1         |                        |
| tRNA<sup>Lys(UCN)</sup> | F       | 1586–2582 | 997       |           |             | 0          |                        |
| tRNA<sup>Leu(GUG)</sup> | F       | 2583–2642 | 60        | 2611–2613 TAG |             | 0          |                        |
| tRNA<sup>Leu(UUR)</sup> | F       | 2640–2702 | 63        | 2669–2671 TGG |             | –3         |                        |
| tRNA<sup>Ala</sup> | F       | 2705–2764 | 60        | 2735–2737 TGC |             | 2          |                        |
| ND6     | F         | 2784–3251 | 468       | ATA       | TAA         | 19         |                        |
| ND5     | F         | 3202–4893 | 1692      | ATT       | TAA         | –50        |                        |
| ND1     | F         | 4914–5786 | 873       | ATA       | TAA         | 20         |                        |
| ND4L    | F         | 5797–6072 | 276       | ATA       | TAA         | 10         |                        |
| CytB    | F         | 6076–7188 | 1113      | ATG       | TAA         | 3          |                        |
| tRNA<sup>Val</sup> | F       | 7185–7246 | 62        | 7215–7217 GCA |             | –4         |                        |
| tRNA<sup>Trp</sup> | F       | 7249–7309 | 61        | 7279–7281 GAA |             | 2          |                        |
| COII    | F         | 7310–7984 | 675       | GTG       | TAG         | 0          |                        |
| tRNA<sup>Asn</sup> | F       | 7989–8048 | 60        | 8019–8021 GTC |             | 4          |                        |
| tRNA<sup>Asp</sup> | F       | 8075–8136 | 62        | 8105–8107 GTA |             | 26         |                        |
| tRNA<sup>Tyr</sup> | F       | 8132–8191 | 60        | 8162–8164 TCC |             | –5         |                        |
| tRNA<sup>Pro</sup> | F       | 8188–8246 | 59        | 8218–8220 GTC |             | –4         |                        |
| tRNA<sup>Ser(UCN)</sup> | F       | 8255–8314 | 60        | 8282–8284 TCA |             | 8          |                        |
| tRNA<sup>Ser(AGN)</sup> | F       | 8311–8369 | 59        | 8338–8340 TTG |             | –4         |                        |
| tRNA<sup>Glu</sup> | R       | 8366–8429 | 64        | 8398–8400 TAA |             | –4         |                        |
| ATP8    | R         | 8431–8595 | 165       | ATG       | TAA         | 1          |                        |
| tRNA<sup>Arg</sup> | R       | 8597–8652 | 56        | 8620–8622 GTT |             | 1          |                        |
| ATP6    | R         | 8652–9332 | 681       | ATT       | TAA         | –1         |                        |
| tRNA<sup>Gln</sup> | R       | 9309–9366 | 58        | 9339–9341 TCG |             | –24        |                        |
| tRNA<sup>His</sup> | R       | 9366–9430 | 65        | 9393–9395 TTC |             | –1         |                        |
| SrRNA   | R         | 9431–10112 | 682      |          |             | 0          |                        |
| tRNA<sup>Lys(UCN)</sup> | R       | 10113–10174 | 62     | 10140–10142 CAT |             | 0          |                        |
| ND3     | R         | 10165–10524 | 360   |          |             | –10        |                        |
| tRNA<sup>Ser(UUR)</sup> | R       | 10517–10569 | 53      | 10548–10550 TGA |             | –8         |                        |
| tRNA<sup>Ser(AGN)</sup> | R       | 10570–10629 | 61      | 10594–10596 GCT |             | 0          |                        |
| ND4     | F         | 10648–11988 | 1341   |          |             | 18         |                        |
| tRNA<sup>Thr</sup> | R       | 11940–11999 | 60      | 11967–11969 TGT |             | –49        |                        |
| ND3     | R         | 11965–12792 | 828    |          |             | –35        |                        |
| tRNA<sup>Val</sup> | R       | 12989–13843 | 54     | 13819–13821 TTT |             | –39        |                        |

Note: Negative numbers indicate adjacent gene overlap.

had been found. For example, a total of 73 unmatched base pairs existed in some tRNAs, and 38 of them were G-U pairs, situated in the AA stem (13 bp), the T stem (10 bp), the AA stem (8 bp) and the DHU stem (7 bp). The remaining five base pairs included U-U mismatches, U-C mismatches, A-C mismatches, A-G
mismatches and A-A mismatches (Fig. 3). Nevertheless, the post-transcriptional RNA-editing mechanism can rectify these mismatches to maintain tRNA functions (Tomita et al. 2001).

**Ribosomal RNA genes**

The rRNA genes comprising large rRNA subunit (lrRNA) and small rRNA subunit (srRNA) are presumed to block in the spaces of flanking genes (Boore 2001; 2006). The lrRNA gene was situated between tRNAVal and tRNALeu(CUN) revealing 78.23% consistency with Euhadra herklotsi and Mastigeulota kiangsinensis. The srRNA gene was located between tRNAGlu and tRNAMet (Fig. 1). The length of them were determined to be 997 bp and 682 bp respectively (Table 3).

**Base composition and codon usage**

Like other snail mt genomes, the nucleotide composition of the C. cicatricosa mt genome was obviously biased toward adenine and thymine (A = 31.90%, T = 37.90%, C = 13.50%, G = 16.70%). The entire mt genome had a high A+T content of 69.80%, by the composition of 69.32% in PCGs, 71.41% in tRNA genes, 72.42% in rRNA genes. Nucleotide bias can also be reflected by codon usage. Evidently, we can see that NNA and NNU were applied frequently in most PCGs. Furthermore, codons TTT (phenylalanine), TTA (leucine), ATT (isoleucine) and ATA (methionine) which were used widely were all composed of A and T. Especially, more and more codons were biased in favor of those codons with A or T in the third position (Fig.4).

The nucleotide composition of metazoan mt genomes usually demonstrate an obvious strand bias (Hassanin et al. 2005; Hassanin 2006) that can be described as AT and GC skews (Perna and Kocher 1995). The PCGs skew statistics of C. cicatricosa showed a great TA skew and nearly equal G and C on the N strand, whereas a great GC skew on the J strand. The nucleotide composition of tRNAs on the J strand were GC and TA skews, evidently exceeding values on the N strand (Table 4). AT and GC skews of C. cicatricosa mt genome differ from the strand biases of metazoan mtDNA (generally positive AT skew and negative GC skew for the J strand, contrary to the N strand for most metazons).

**Noncoding regions**

The noncoding regions of C. cicatricosa mt genome contained some short intergenic spacers. These short sequences possibly acted as splicing recognition sites during the process of transcription (He et al. 2005). In the sequenced complete mt genome of the order Stylommatophora, the short intergenic spacers range from 1 bp to 65 bp (Hatzoglou et al. 1995; Terrett et al. 1995; Yamazaki et al. 1997; White et al. 2011;
The mitochondrial genome of the land snail *Camaena cicatricosa* (Müller, 1774)...

Groenenberg et al. 2012; Gaitán-Espitia et al. 2013; Deng et al. 2014) except *Achatina fulica* with 551 bp long noncoding region (He et al. 2014). However, the longest noncoding region was only 29 bp in *C. cicatricosa*. The shorter lengths of noncoding regions indicated that the mt genome of stylommatophorans are quite compact.

A large noncoding region called control region or AT-rich region is commonly seen in metazoan mt genomes (Boore 1999). In fact, variation of size for the entire mt genome can be chalked up to the presence of a number of tandem repeats (Zhang and Hewitt 1997) in control region, which may be caused by replication slippage (Levinson and Gutman 1987; Fumagalli et al. 1996). Nevertheless, putative control region (POR) was not aligned confidently in gastropods (Groenenberg et al. 2012) except *A. fulica* having a 551 bp POR between *COI* and *tRNA*\(^{Val}\) (He et al. 2014). Other eight stylommatophoran species may possess short POR regions located adjacent to *COIII* (Hatzoglou et al. 1995; Terrett et al. 1995; Yamazaki et al. 1997; White et al. 2011; Groenenberg et al. 2012; Gaitán-Espitia et al. 2013; Deng et al. 2014). The POR regions of three helicid species and *M. kiangsinensis* were located between *COIII* and *tRNA*\(^{Ser}\) with lengths of 158–189 bp, whereas in the other three species were located between *COIII* and *tRNA*\(^{Ile}\) with lengths of 42–47 bp. The 29 bp noncoding region of *C. cicatricosa* was located between *COIII* and *tRNA*\(^{Ile}\), but its length was shorter than other stylommatophorans.

**Figure 4.** Relative synonymous codon usage (RSCU) in the *Camaena cicatricosa* mt genome. Codon families are provided on the x axis.
Table 4. Nucleotide composition and skew of the *Camaena cicatricosa* mt genome.

| Feature                  | %A | %T | %G | %C | %A+T | AT Skew | GC Skew | No. of nucleotides |
|--------------------------|----|----|----|----|------|---------|---------|-------------------|
| Whole genome             | 31.90 | 37.90 | 16.70 | 13.50 | 69.80 | -0.09   | 0.11    | 13843             |
| Protein coding genes     | 31.18 | 38.14 | 17.05 | 13.64 | 69.32 | -0.10   | 0.11    | 10941             |
| Protein coding genes (J) | 28.83 | 40.41 | 17.54 | 13.23 | 69.24 | -0.17   | 0.14    | 8907              |
| Protein coding genes (N) | 28.22 | 41.45 | 15.44 | 14.90 | 69.67 | -0.19   | 0.02    | 2034              |
| tRNA genes               | 34.95 | 36.46 | 15.81 | 12.78 | 71.41 | -0.02   | 0.11    | 1322              |
| tRNA genes (J)           | 33.96 | 36.80 | 17.51 | 11.72 | 70.77 | -0.04   | 0.20    | 845               |
| tRNA genes (N)           | 35.85 | 36.69 | 14.68 | 12.79 | 72.54 | -0.01   | 0.07    | 477               |
| rRNA genes               | 35.14 | 37.28 | 14.83 | 12.75 | 72.42 | -0.03   | 0.08    | 1679              |

Phylogenetic analysis

ML tree was estimated according to the GTR+G+I substitution model selected by AIC. The ML and MP trees (Fig. 5) displayed the same topologies and presented eight major clades corresponding to the families Bradybaenidae, Camaeinidae, Helicidae, Succineidae, Clausiliidae, Achatinidae, Lymnaeidae and Aplysiidae. The monophyly of Stylommatophora was approved. Species in Helicidae were sister groups and congruent with previous works (Gaitán-Espitia et al. 2013). *C. cicatricosa* and *M. kiangsinensis* from China and *E. herklotsi* from Japan are monophyletic. However, the systematics of the families Camaeinidae, Helicidae and Bradybaenidae are complicated and not fully resolved. Systematic and phylogenetic studies based on analyses of morphological versus molecular markers have produced inconsistent results (Scott 1996; Cuezzo 2003; Wade et al. 2007; Hirano et al. 2014). A final assessment of the systematic relationships of the three families is pending requiring a more complete taxon sampling.
Acknowledgments

We sincerely thank Professor De-Niu Chen of Institute of Zoology, Chinese Academy of Science for helpful comments on the manuscript. This research is supported by National Natural Science Foundation of China (31372162; 31040072), Natural Science Foundation in Fujian Province (2012J01063), Science and Technology Program in Fujian Province (2012I0001), National Science Council, Taiwan (NSC 98-2621-B-390-001) and Public Science and Technology Research Funds Projects of General Administration of Quality Supervision, Inspection and Quarantine of the People’s Republic of China (No. 201410076).

References

Avise JC (1994) Molecular Markers, Natural History and Evolution. Chapman and Hall, New York. doi: 10.1007/978-1-4615-2381-9
Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Research 27: 1767–1780. doi: 10.1093/nar/27.8.1767
Boore JL (2001) Complete mitochondrial genome sequence of the polychaete annelid Platynereis dumerilii. Molecular Biology and Evolution 18: 1413–1416. doi: 10.1093/oxfordjournals.molbev.a003925
Boore JL (2006) The complete sequence of the mitochondrial genome of Nautilus macromphalus (Mollusca: Cephalopoda). BMC Genomics 7: 182. doi: 10.1186/1471-2164-7-182
Boore JL, Medina M, Rosenberg LA (2004) Complete sequences of the highly rearranged molluscan mitochondrial genomes of the Scaphopod Graptacme eborea and the bivalve Mytilus edulis. Molecular Biology and Evolution 21: 1492–1503. doi: 10.1093/molbev/msh090
Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. Annual Review of Entomology 59: 95–117. doi: 10.1146/annurev-ento-011613-162007
Cameron SL, Miller KB, Haese CA, Whiting MF, Barker SC (2004) Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea sensu lato (Arthropoda). Cladistics 20: 534–557. doi: 10.1111/j.1096-0031.2004.00040.x
Chen DN, Gao JX (1987) Economic fauna sinica of China. Science Press, Beijing, China.
Crease TJ (1999) The complete sequence of the mitochondrial genome of Daphnia pulex (Cladocera: Crustacea). Gene 233: 89–99. doi: 10.1016/S0378-1119(99)00151-1
Delsuc F, Phillips MJ, Penny D (2003) Comment on “Hexapod origins: monophyletic or paraphyletic?”. Science 301: 1482. doi: 10.1126/science.1086558
Deng PJ, Wang WM, Huang XC, Wu XP, Xie GL, Ouyang S (2014) The complete mitochondrial genome of Chinese land snail Mastigeulota kiangsinensis (Gastropoda: Pulmonata: Bradybaenidae). Mitochondrial DNA. doi: 10.3109/19401736.2014.953083
Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.
Fumagalli L, Taberlet P, Favre L, Haussler J (1996) Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. Molecular Biology and Evolution 13: 31–46. doi: 10.1093/oxfordjournals.molbev.a025568

Gaitán-Espitia JD, Scheihing R, Poulin E, Artacho P, Nespolo RF (2013) Mitochondrial phylogeography of the land snail Cornu aspersum: tracing population history and the impact of human-mediated invasion in austral South America. Evolutionary Ecology Research 15: 1–18.

Grande C, Templado J, Cervera JL, Zardoya R (2002) The complete mitochondrial genome of the nudibranch Roboastra europaea (Mollusca: Gastropoda) supports the monophyly of opisthobranchs. Molecular Biology and Evolution 19(10): 1672–1685. doi: 10.1093/oxfordjournals.molbev.a003990

Grande C, Templado J, Zardoya R (2008) Evolution of gastropod mitochondrial genome arrangements. BMC Evolutionary Biology 8: 61. doi: 10.1186/1471-2148-8-61

Groenenberg DSJ, Pirovano W, Gittenberger E, Schilthuizen M (2012) The complete mitogenome of Cylindrus obtusus (Helicidae, Ariantinae) using Illumina next generation sequencing. BMC Genomics 13: 114. doi: 10.1186/1471-2164-13-114

Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads-a baiting and iterative mapping approach. Nucleic Acids Research, 371.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.

Hassanin A (2006) Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. Molecular Phylogenetics and Evolution 38: 100–116. doi: 10.1016/j.ympev.2005.09.012

Hassanin A, Léger N, Deutsch J (2005) Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoan, and consequences for phylogenetic inferences. Systematic Biology 54: 277–298. doi: 10.1080/10635150590947843

Hatzoglou E, Rodakis GC, Lecanidou R (1995) Complete sequence and gene organization of the mitochondrial genome of the land snail Albinaria coerulea. Genetics 140(4): 1353–1366.

He Y, Jones J, Armstrong M, Lamberti F, Moens M (2005) The mitochondrial genome of Xiphinema americanum sensu stricto (Nematoda: Enoplea): Considerable economization in the length and structural features of encoded genes. Journal of Molecular Evolution 61: 819–833. doi: 10.1007/s00239-005-0102-7

He ZP, Dai XB, Zhang S, Zhi TT, Lun ZR, Wu ZD, Yang TB (2014) Complete mitochondrial genome of the giant African snail, Achatina fulica (Mollusca: Achatinidae): a novel location of putative control regions (CR) in the mitogenome within Pulmonate species. Mitochondrial DNA. doi: 10.3109/19401736.2014.930833

Hirano T, Kameda Y, Kimura K, Chiba S (2014) Substantial incongruence among the morphology, taxonomy, and molecular phylogeny of the land snails Aegista, Landouria, Trishoplita, and Pseudobuliminus (Pulmonata: Bradybaenidae) occurring in East Asia. Molecular Phylogenetics and Evolution 70: 171–181. doi: 10.1016/j.ympev.2013.09.020
The mitochondrial genome of the land snail Camaena cicatricosa (Müller, 1774)...

Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail Gnarosophia bellendenkerensis (Brazier 1875). PNAS 99(9): 6112– 6117. doi: 10.1073/pnas.092538699

Knudsen B, Kohn AB, Nahir B, McFadden CS, Moroz LL (2006) Complete DNA sequence of the mitochondrial genome of the sea-slug, Aplysia californica: conservation of the gene order in Euthyneura. Molecular Phylogenetics and Evolution 38(2): 459–469. doi: 10.1016/j.ympev.2005.08.017

Kurabayashi A, Ueshima R (2000) Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod Pupa striogosa: systematic implication of the genome organization. Molecular Biology and Evolution 17: 266–277. doi: 10.1093/oxfordjournals.molbev.a026306

Liu GH, Wang SY, Huang WY, Zhao GH, Wei SJ, Song HQ, Xu MJ, Lin RQ, Zhou DH, Zhu XQ (2012) The complete mitochondrial genome of Galba pervia (Gastropoda: Mollusca), an intermediate host snail of Fasciola spp. PLoS ONE 7(7): e42172. doi: 10.1371/journal.pone.0042172

Lowe TM, Eddy SR (1997) TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research 25: 955–964. doi: 10.1093/nar/25.5.0955

Menegon M, Loaderb SP, Marsdenc SJ, Branchd WR, Davenporte TRB, Ursenbacherf S (2014) The genus Atheris (Serpentes: Viperidae) in East Africa: phylogeny and the role of rifting and climate in shaping the current pattern of species diversity. Molecular Phylogenetics and Evolution 79: 12–22. doi: 10.1016/j.ympev.2014.06.007

Merritt TJS, Shi L, Chase MC, Rex MA, Etter RJ, Quattro JM (1998) Universal cytochrome b primers facilitate intraspecific studies in molluscan taxa. Molecular Marine Biology and Biotechnology 7: 7–11.

Palumbi S, Martin A, Romano S, Mcmillan WO, Stice L, Grabowwski G (1991) The Simple Fool’s Guide to PCR. Department of Zoology, University of Hawaii, Honolulu.

Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. Journal of Molecular Evolution 41: 353–358. doi: 10.1007/BF01215182

Pilsbury HA (1893–1895) Manual of Conchology, ser.2, vol.9. (Helicidae, vol.7). Guide to the study of Helices, 366 pp.

Raay TJ, Crease TJ (1994) Partial mitochondrial DNA sequence of the crustacean Daphnia pulex. Current Genetics 25: 66–72. doi: 10.1007/BF00712970

Scott B (1996) Phylogenetic relationships of the Camaenidae. Journal of Molluscan Studies 62: 65–73. doi: 10.1093/mollus/62.1.65

Simon S, Hadrys H (2013) A comparative analysis of complete mitochondrial genomes among Hexapoda. Molecular Phylogenetics and Evolution 69: 393–403. doi: 10.1016/j.ympev.2013.03.033

Talavera G, Vila R (2011) What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. BMC Evolutionary Biology 11: 315. doi: 10.1186/1471-2148-11-315
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. Molecular Biology and Evolution 28: 2731–2739. doi: 10.1093/molbev/msr121
Terrett JA, Miles S, Thomas RH (1995) Complete DNA sequence of the mitochondrial genome of Cepaea nemoralis (Gastropoda: Pulmonata). Journal of Molecular Evolution 42: 160–168. doi: 10.1007/BF02198842
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25: 4876–4882. doi: 10.1093/nar/25.24.4876
Tomita K, Yokobori S, Oshima T, Ueda T, Watanabe K (2001) The cephalopod Loligo bleekeri mitochondrial genome: multiplied noncoding regions and transposition of tRNA genes. Journal of Molecular Evolution 54: 486–500. doi: 10.1007/s00239-001-0039-4
Wade CM, Hudelot C, Davison A, Naggs F, Mordan PB (2007) Molecular phylogeny of the helicoid land snails (Pulmonata: Stylommatophora: Helicoidea), with special emphasis on the Camaenidae. Journal of Molluscan Studies 73: 411–415. doi: 10.1093/mollus/eym030
Wang P, Li H, Wang Y, Zhang JH, Dai X, Chang J, Hu BW, Cai WZ (2014) The mitochondrial genome of the plant bug Apolygus lucorum (Hemiptera: Miridae): presently known as the smallest in Heteroptera. Insect Science 21: 159–173. doi: 10.1111/1744-7917.12029
White TR, Conrad MM, Tseng R, Balayan S, Golding R, Martins AM, Dayrat BA (2011) Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. BMC Evolutionary Biology 11(1): 295. doi: 10.1186/1471-2148-11-295
Wolstenholme DR (1992) Genetic novelties in mitochondrial genomes of multicellular animals. Current Opinion in Genetics & Development 2: 918–925. doi: 10.1016/S0959-437X(05)80116-9
Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20(17): 3252–3255. doi: 10.1093/bioinformatics/bth352
Xiao WL (1989) Study on ecology of Camaena cicatricosa (Muller). Journal of Jinan University 1: 9.
Yamazaki N, Ueshima R, Terrett JA, Yokobori S, Kaifu M, Segawa R, Kobayashi T, Numachi K, Ueda T, Nishikawa K, Watanabe K, Thomas RH (1997) Evolution of Pulmonate Gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Cepaea and Albinaria and implications of unusual tRNA secondary structures. Genetics 145: 749–758.
Yu DJ, Xu L, Nardi F, Li JG, Zhang RJ (2007) The complete nucleotide sequence of the mitochondrial genome of the oriental fruit fly, Bactrocera dorsalis (Diptera: Tephritidae). Gene 396: 66–74. doi: 10.1016/j.gene.2007.02.023
Zhang DX, Hewitt GM (1997) Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochemical Systematics and Ecology 25: 99–120. doi: 10.1016/S0305-1978(96)00042-7
Zhou WC, She SS, Chen DN, Lin J, Guo YH, Chen SL (2007) The intermediate host of Angiostrongylus cantonensis-Molluscan. Chinese Journal of Zoonoses 23(4): 401–408.