The complete mitochondrial genome of *Caprella scaura* (Crustacea, Amphipoda, Caprellidea), with emphasis on the unique gene order pattern and duplicated control region.

Ito Atsushi, Aoki Masakazu N., Yokobori Shin-ichi, Wada Hiroshi

*Mitochondrial DNA*

Volume 21

Number 5

Page range 183-190

Year 2010-10

(C) 2010 Informa Plc.

URL http://hdl.handle.net/2241/114281

doi: 10.3109/19401736.2010.517834
The complete mitochondrial genome of *Caprella scaura* (Crustacea, Amphipoda, Caprellidea), with emphasis on the unique gene order pattern and duplicated control region

**Short title**

*Complete mitochondrial genome of C. scaura*

**Atsushi Ito**

Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, 305-8572, Japan
Telephone and fax numbers: 81-29-853-4671
Email address: itats110@gmail.com

**Masakazu N. Aoki**

Shimoda Marine Research Center, University of Tsukuba, Shimoda, Shizuoka, 415-0025, Japan
Telephone number: 81-558-22-6618
Email address: m-aoki@kurofune.shimoda.tsukuba.ac.jp

**Shin-ichi Yokobori**

Department of Molecular Biology, School of Life Science, Tokyo University of Pharmacy and Life Science, Horinouchi, Hachioji, Tokyo 192-0392, Japan
Hiroshi Wada*
Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, 305-8572, Japan

Telephone and fax numbers: 81-29-853-4671
Email address: 98champ@msg.biglobe.ne.jp

*Corresponding author
ABSTRACT

The nucleotide and amino acid sequences, and the gene order of the mitochondrial genome make it highly informative for studying phylogeny, population genetics, and phylogeography. This study determined the complete mitochondrial genome of the caprellid species *Caprella scaura*. The mitochondrial genome of *C. scaura* has a total length of 15,079 bp, with an AT content of 66.43%.

The mitochondrial genome contains typical gene components, including 13 protein-coding genes, two ribosomal RNA genes, and 22 tRNA genes. In comparison with the mitochondrial genome of a gammarid, some distinct characteristics were found. For example, the order of the two conserved gene blocks is inverted between Gammaridea and *C. scaura*. In addition, two copies of almost identical control regions were found in the mitochondrial genome of *C. scaura*. These unique characteristics will be useful for determining the evolutionary history of the Caprellidea.

Keywords

Crustacea; Amphipoda; Entire mitochondrial genome; Gene rearrangement; Duplicate control regions
1. Introduction

The complete mitochondrial genome sequences of many crustacean species have been reported, and these data have provided useful information on population genetics, phylogeography, and the phylogenetic relationships among crustacean taxon. However, most complete mitochondrial genome sequence data come from the superorder Decapoda (Malacostraca), which includes animals such as crabs and shrimps (Hickerson and Cunningham 2000, Wilson et al. 2000, Yamaguchi et al. 2002, Miller et al. 2004, 2005, Cook 2005, Sun et al. 2005).

Peracarida have highly diverse morphologies and life histories. The suborder comprises nine orders, including Isopoda and Amphipoda, and is characterized by the brood pouch on thoracic segments. It includes an amphipod that inhabits the hadal zone (deeper than 10,000 m; France 1993), freshwater residents, and completely terrestrial species (e.g., wood lice). Some isopods are parasitic on fish or other crustaceans, and have a modified segmental structure. One group of amphipods, the Caprellidea, the focus of this study, has some degenerate thoracic limbs and posterior segments. From an evolutionary standpoint, this morphological and lifestyle variation makes the peracarids fascinating organisms to study (Ito et al. 2008).

To date, five complete or almost-complete mitochondrial genomes of peracarids (three isopods and two amphipods) have been made available (Cook et al. 2005, Kilper and Podsiadlowski 2006, Podsiadlowski and Bartolomaeus 2006, Marcade et al. 2007, Bauza-Ribot et al. 2009). In
Amphipoda, only the suborder Gammaridea (Figure 1A) has been the target of complete

*Metacrangonyx longipes*; Bauza-Ribot *et al*. 2009) or almost-complete (*Parhyale hawaiensis*; Cook *et al*. 2005, Bauza-Ribot *et al*. 2009) mitochondrial genome analysis. The mitochondrial genomes of amphipods show distinct gene arrangements that differ from the pancrustacean pattern or those seen in isopods. These observations suggest that the mitochondrial genome can provide useful information on peracarids. To better understand the molecular evolutionary history of the peracarid mitochondrial genomes, we report the complete mitochondrial genome sequence of another amphipod suborder: Caprellidea. Recently Kilpert and Podsiadlowski (2010) reported the complete mitochondrial genome of *Caprella mutica*. Referring to their data, we characterized the mitochondrial genome of the Caprellidea. The mitochondrial genome of *C. scaura* exhibits some unique characteristics that are not observed in gammarids. The complete mitochondrial genome data of *C. scaura* are useful for resolving the phylogeny of Caprellidea, as well as for providing new insight into the evolution of the mitochondrial genome in Metazoa.
2. Materials and Methods

2.1. Sampling and genomic DNA extraction

*Caprella scaura* was collected from Yokohama Bay, Kanagawa, Japan. Live samples were transported to the laboratory and preserved in 100% ethanol until use. Genomic DNA was extracted from the appendages and gills of a single individual using the DNeasy Tissue kit from Qiagen.

2.2. Long PCR and sequencing

Two primer sets were designed for long polymerase chain reaction (PCR) (Table I). One primer set amplified approximately half of the mitochondrial genome between *cox1* and *rnnL*, and the other amplified the remaining region. Long PCR was performed using LA *Taq* (Takara) and DNA Engine PTC-200 (MJ Research). The conditions for PCR cycling were as follows: 1 min at 95°C, 16 cycles of 15 s at 95°C and 12 min at 65°C, followed by 12 cycles of 15 s at 95°C and 12 min at 65°C, adding 15 s per cycle. Amplified DNA fragments were separated according to size by electrophoresis in 1% agarose gel and purified using a Sephaglas™ BandPrep Kit (Amersham). Purified DNA fragments were then subcloned using the TOPO XL PCR Cloning Kit (Invitrogen). Subcloned mitochondrial DNA was digested with *EcoRV*, *HindIII*, or *KpnI*. These restriction fragments were subcloned again and then sequenced by primer walking. Sequence analyses were
performed with an ABI PRISM 310 and 377 Genetic DNA sequencer (Applied Biosystems) using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

2.3. Mitochondrial genome structure

Protein-coding genes were identified by a BLAST search on DDBJ and by comparison to the mitochondrial genome of *Metacrangonyx longipes* (Bauza-Ribot *et al.* 2009). The borders of the protein-coding genes were determined by comparison with alignments of mitochondrial genes from crustacean species. The borders of the rRNAs were assumed to be the boundaries to the adjacent protein-coding genes or tRNA genes. Each tRNA gene was identified by constructing the tRNA cloverleaf structure by hand. The secondary structure of the noncoding region was predicted with mfold version 3.2 through the DNA mfold server (Zuker 2003). The default parameters were used, except under the ionic condition (50 mM NaCl, polymer).
3. Results and Discussion

3.1. Genome organization

The total length of the mitochondrial genome of *C. scaura* is 15,079 bp (DDBJ accession number AB539699; Figure 1, Table II), which is 348 bp shorter than the mitochondrial genome of *C. mutica* (Kilpert and Podsiadlowski 2010). Successful amplification of the complete mitochondrial DNA by the primer sets designed from *cox1* and *rnnL* in both directions (Table I) indicated that the *C. scaura* mitochondrial DNA is circular. The gene set is identical to that of *C. mutica* and consists of 13 protein-coding genes and two rRNA genes, as observed in most mesozoans (Table II). Unlike the mitochondrial genome of *M. longipes* (Gammaridea), which has only 21 tRNA genes (Bauza-Ribot *et al.* 2009), 22 tRNA genes were found in the *C. scaura* mitochondrial genome. In addition to the 37 genes, we identified eight intergenic spacers and two duplicated copies of control regions as described below (Table II). The AT content of the complete sequence is 66.43%, which is similar to that in *C. mutica* (68%), but is lower than the AT content of *M. longipes* (76.03%). Gene overlaps were found at 13 boundaries. The overlapping base length varied from one (in several genes) to 24 (5' end of *trnQ* and 5' end of *nad2*) bases.
3.2. Protein-coding and rRNA genes

Three minor strand protein genes (nad1, nad4, nad4L) are found in the mitochondrial genome of *C. scaura* (Figure 1, Table II). This pattern is identical to that of *C. mutica*, but is distinct from that of two gammarids: *M. longipes*, which has five minor strand genes (nad1, nad4, nad4L, nad5, cob) and *P. hawaiensis*, which has four (nad1, nad4, nad4L, nad5). Most protein-coding genes had ATA or ATG as their start codon, which encodes methionine, while the start codon of *cox1* was ATT and that of *nad1* and *nad4L* was CTG (Table II). These unusual start codons are also found in *C. mutica* and *M. longipes* ([Bauza-Ribot et al. 2009; Kilpert and Podsiadlowski 2010]). *cox1* also starts from ATT in *C. mutica*, while *nad1* and *nad4L* start from ATA and TTG, respectively (Kilpert and Podsiadlowski 2010). The stop codon of six protein-coding genes was TAA, while one was TAG, and that of the remaining six genes was TA or a single T (Table II). We postulate that adenines are added to the single T by polyadenylation during mRNA processing ([Ojala et al. 1981]). The *rrnS* and *rrnL* genes of *C. scaura* were encoded on the minor strand (Table II), which is also seen in *C. mutica*, *M. longipes*, and *P. hawaiensis*.

3.3. tRNA-coding genes

The mitochondrial genome of *C. scaura* contains the usual 22 tRNA genes (Table II). Exactly the same set of tRNAs was identified in the *C. mutica* mitochondrial genome. Of these, 14 genes
are encoded on the major strand, while the remaining eight are encoded on the negative strand (Table II).

The secondary structure of tRNA-Ser (UGA) and tRNA-Ser (UCU) deviates from the typical cloverleaf structure; these genes lack the D-arm (Figure 2). The loss of the D-arm in tRNA-Ser (UCU) that specifies AGN codons is also observed in other crustaceans such as Ligia oceanica (Isopoda; Kilpert and Podsiadlowski 2006) and most metazoan mitochondria, although tRNA-Ser (UCU) in the mitochondrial genome of M. longipes does not show these characteristics (Bauza-Ribot et al. 2009). The loss of the D-arm of tRNA-Ser (UGA) is not observed in L. oceanica (Kilpert and Podsiadlowski 2006). Note that most metazoan mitochondrial trnS genes for AGN codons have the anticodon GCT at the DNA level. Modification at the first anticodon position might enable this tRNA to encode all four AGN box codons (Yokobori et al. 2001). The trnS(ucu) gene has been found in various lineages of metazoan mitochondria including Apis mellifera (honey bee) mitochondria (Crozier and Crozier 1993). The anticodon UCU can encode all AGN codons, and the anticodon may have originated independently from the GCU anticodon during the evolution of different lineages.

Three tRNA genes, trnC, trnQ, and trnE, seem to have a TV-replacement loop instead of a T-arm and a variable loop (Figure 2). The TV-replacement loops in tRNAs were first discovered in
nematode mitochondria (Okimoto and Wolstenholme 1990). Since then, different metazoan mitochondrial tRNA genes of various lineages have been found to have TV-replacement loops.

The inferred secondary structure of tRNA-Pro is also unusual (Figure 2). Between the acceptor stem and D-stem, there are three nucleotides instead of two. If we assume that one of four As from the fourth to seventh bulges out, then the secondary structure of this tRNA becomes more similar to the canonical tRNA secondary structure (see alternative structure of tRNA-Pro in Figure 2).

Four tRNA genes [trnN, trnQ, trnK, and trnS(ucu)] overlap by from one to four nucleotides with downstream tRNA genes in the same direction (Figure 2). From these tRNA genes, mature transcripts might be synthesized according to the pathway suggested by Yokobori and Pääbo (1995). In these genes, the addition of As at the “overlapped” regions retains the Watson–Crick base-pairing at the top of the acceptor stems of the trnN and trnQ transcripts.

### 3.4. Duplicated control region

As in the mitochondrial genome of C. mutica, C. scaura possesses two control regions (Figure 1, Table II). The AT content of these regions (approximately 74%) is higher than that of the total mitochondrial genome (66.43%). The nucleotide sequences of these regions are almost identical (533/532; Table II). A few arthropods, such as ticks (Shao et al. 2005), thrips (Hexapoda, Thysanoptera; Shao and Barker 2003), and sea fireflies (Crustacea, Ostracoda; Ogoh and Ohmiya

11
2004) exhibit a duplicated control region. The stomatopod *Squilla mantis* also has two AT-rich regions (Cook 2005), while *Geothelphusa dehaani* (Japanese freshwater crab) has a pseudo-control region in addition to a putative control region (Segawa and Aotsuka 2005). In both species, however, the length and nucleotide sequences are quite different from one another (230/861 in *S. mantis* and 69/514 in *G. dehaani*).

The high sequence similarity of the duplicated control regions of *C. scaura* may occur because duplication took place recently, or might be due to concerted evolution after duplication (Shao et al. 2005). Kumazawa *et al.* (1998) suggested that a mitochondrial genome with two control regions may be able to replicate more efficiently than a mitochondrial genome with one control region. The biological significance of multiple control regions is not completely clear. The mitochondrial genome of *C. scaura* can contribute to future studies on the evolution of duplicated control regions in mitochondrial genome of the Metazoa.

### 3.5. Gene order

We found that some conserved blocks consisted of more than two genes among the three amphipod mitochondrial genomes (Figure 3). However, the location of these blocks is different between gammarids and *Caprella*. 
We identified the following three conserved blocks: Block A, which includes \textit{nad2}, \textit{cox1}, \textit{cox2}, \textit{atp8}, \textit{atp6}, \textit{cox3}, \textit{nad3}, and some tRNA genes (shown in dark gray in Figure 3); Block B, which includes the four genes \textit{nad5}, \textit{trnH}, \textit{nad4}, and \textit{nad4L} (shown in an intermediate shade of gray); and Block C, which includes the five genes \textit{nad1}, \textit{trnL(uag)}, \textit{rrnL}, \textit{trnV}, and \textit{rrnS} (shown in light gray). All genes in Blocks B and C are encoded on the minor strand except for \textit{nad5} of \textit{Caprella} (Figure 3). These blocks are also found in the putative ancestral ground pattern of Pancrustacea (Crustacea + Hexapoda; Boore \textit{et al.} 1995, 1998). The three blocks are located in the order A, B, C as in the two gammarids. However, in \textit{Caprella}, the blocks are located in the order A, C, B (Figure 3). Transposition of the gene block must have occurred in the lineage leading to Caprellidea. In addition, inversion of \textit{nad5} within Block C must have occurred in the caprellid lineage.

In addition, the order of \textit{trnA}, \textit{trnS(ucu)}, \textit{trnN}, \textit{trnE}, \textit{trnR}, and \textit{trnF} is conserved in \textit{Caprella} and \textit{M. longipes}, but not in \textit{P. hawaiensis} (Figure 3). By contrast, the location of the control region is identical in \textit{M. longipes} and \textit{P. hawaiensis}, but not in \textit{Caprella} (Figure 3).

Comparing the mitochondrial genomes of amphipods, we estimated the rearrangement steps of protein-coding genes. The order and transcriptional direction of protein-coding genes are well conserved between \textit{P. hawaiensis} and the ancestral ground pattern of Pancrustacea (Boore \textit{et al.} 1995, 1998) (Figure 3). However, in \textit{M. longipes}, \textit{cob} seems to have translocated over Block C, and its transcriptional direction was inverted (Figure 3). By contrast, in \textit{Caprella}, the transcriptional
direction of *nad5* was inverted (Figure 3). In addition, as mentioned above, translocation between Blocks B and C and duplication of the control region occurred in *Caprella*. Translocation of tRNA genes frequently occurred in Amphipoda (Figure 3).

The ground pattern of the gene order of Pancrustacea is well conserved in several crustaceans such as Stomatopoda and Decapoda. However, each species of Isopoda and Amphipoda, including *C. scaura*, exhibits a gene order distinct from the ground pattern (Figure 3). The pattern between Amphipoda and Isopoda also diverged from each other (Figure 3). While the mitochondrial genomes of amphipods show a pattern relatively similar to the ancestral ground pattern, three isopod species have gene order patterns quite different from the ancestors of Amphipoda and Pancrustacea (Figure 3). Frequent gene rearrangements may help in inferring phylogenetic relationships within Peracarida.

### 3.6. Conclusion

We determined the complete nucleotide sequence of the mitochondrial genome of the caprellid *C. scaura*. The mitochondrial genome has a total length of 15,079 bp, with an AT content of 66.43%. The mitochondrial genome contains typical gene components, including 13 protein-coding genes, two ribosomal RNA genes, and 22 tRNA genes. In comparison with the gammarid mitochondrial genome, we found some conserved gene blocks among three amphipod
mitochondrial genomes. Between the Gammaridea and Caprella, the conserved gene blocks were inverted. The positions of the tRNA-coding genes also show variation. In addition, two copies of almost identical control regions were found in the mitochondrial genome of Caprella. These traceable gene order changes in the mitochondrial genome will be a useful source of information for phylogenetic analyses of peracarids and within Caprellidea.

Declaration of Interest

This work was supported by a research grant from the Research Institute of Marine Invertebrates (grant number: 07-RIMI-16).
References


Kilpert F, Podsiadlowski L. 2006. The complete mitochondrial genome of the common


Gene 327: 131–139.


Okimoto R, Wolstenholme DR, 1990. A set of tRNAs that lack either the T psi C arm or the dihydrouridine arm: towards a minimal tRNA adaptor. EMBO J 9: 3405–3411.


Wilson K, Cahill V, Ballment E, Benzie J. 2000. The complete sequence of the
mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan
crustaceans more closely related to insects than to branchiopods? Mol Biol Evol
17: 863-874.
Yamaguchi MM, Miya, MU, Nishida M. 2002. Complete mitochondrial DNA sequence
of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Dacapoda). Gene
295: 89-96.

**Figure captions**

**Figure 1.** Map of the complete mitochondrial genome of *Caprella scaura*. tRNA genes
are represented by their one-letter amino acid code. The genomic regions in which
genes are encoded on the minor stand are shown by thick lines.

**Figure 2.** Secondary structures of mitochondrial tRNA in *Caprella scaura*. Nucleotides
shared with downstream tRNA genes in the same direction are indicated by small letters.

Suggested nucleotides added posttranscriptionally are shown in parentheses. In the case of tRNA-Pro, an alternative structure is also shown. Watson–Crick base pairs are indicated by bars, whereas G–U base pairs are indicated by asterisks.

**Figure 3.** Schematic gene arrangement of peracarids (Amphipoda and Isopoda) and putative ancestral mitochondrial genome of Pancrustacea redrawn from that described by Bauza-Ribot *et al.* (2009). Genes with thick underlines are encoded on the minor strand. Three conserved gene blocks within Amphipoda are indicated by different brightnesses; Block A (12 genes from *nad*2 to *nad*3) is dark gray, Block B (four genes from *nad*5 to *nad*4L) is intermediate gray, and Block C (five genes from *nad*1 to *rrn*S) is light gray. The genes in the mitochondrial genomes of the putative ancestral ground pattern and Isopoda, which correspond to A, B, and C, are also hatched. Note that the location of *trnG* (light gray) in Block A of amphipods is different from that in the ancestor of Isopoda and Pancrustacea. Hypothetical protein gene translocation and inversion events are shown between *Parhyale hawaiiensis* and *Metacrangonyx longipes,*
such as translocation and inversion of \textit{cob}. In \textit{Caprella}, translocation of Blocks B and C (thick arrow), duplication of the control region (thin arrow), and inversion of \textit{nad5} (intermediate arrow) are seen.
Caprella scaura
mitochondrial genome
15,079 bp