Kazuo Umetsu[1]
Naruki Iwabuchi[2]
Isao Yuasa[3]
Naruya Saitou[4]
Paul F. Clark[5]
Geoff Boxshall[5]
Motoki Osawa[1]
Keiji Igarashi[6]

[1]Department of Forensic Medicine, Yamagata University School of Medicine, Yamagata, Japan
[2]Sendai Science Museum, Sendai, Japan
[3]Department of Legal Medicine, Tottori University School of Medicine, Yonago, Japan
[4]Laboratory of Evolutionary Genetics, National Institute of Genetics, Mishima, Japan
[5]Department of Zoology, The Natural History Museum, London, UK
[6]Faculty of Agriculture, Yamagata University, Tsuruoka, Japan

Complete mitochondrial DNA sequence of a tadpole shrimp (*Triops cancriformis*) and analysis of museum samples

The complete mitochondrial DNA (mtNDA) of the tadpole shrimp *Triops cancriformis* was sequenced. The sequence consisted of 15 101 bp with an A+T content of 69%. Its gene arrangement was identical with those sequences of the water flea (*Daphnia pulex*) and giant tiger prawn (*Penaeus monodon*), whereas it differed from that of the brine shrimp (*Artemia franciscana*) in the arrangement of its genes for tRNAs. Phylogenetic analysis revealed *T. cancriformis* to be more closely related to the water flea than to the brine shrimp and giant tiger prawn. We also compared the 16S rRNA sequences of five formalin-fixed tadpole shrimps that had been collected in five different locations and stored in a museum. The sequence divergence was in the range of 0–1.51%, suggesting that those samples were closely related to each other.

Keywords: Formalin-fixed sample / Mitochondrial DNA / Molecular phylogeny / Notostraca / Tadpole shrimp
EL 5169

1 Introduction

Fossils of the tadpole shrimp from approximately 200 million years ago are indistinguishable morphologically from the present species. Consequently, the tadpole shrimp has been regarded as a "living fossil" [1, 2]. Three species of tadpole shrimps are found in Japan: *Triops longicaudatus, Triops granarius,* and *Triops cancriformis. T. cancriformis* is widely distributed in North Africa, Europe, the Middle East, India, and Siberia [2] and has been classified as an endangered species in some regions. However, it is rather common in Japan, where it was reported for the first time in 1951 [3]. Several habitats have been verified in Yamagata Prefecture. The origin of the *T. cancriformis* in Japan was uncertain, because the species has not been recorded to exist in neighboring countries such as Korea and China. Mitochondrial DNA (mtDNA) is one of the best markers for revealing phylogenetic relationships among related groups of animals [4–6]. But information on mtDNA of the tadpole shrimp is scarce [7, 8]. Because museum-stored samples collected from various locations are available, such samples would be useful if DNA information can be extracted from those formalin-fixed samples. The aim of the present study was to determine the complete sequence of the *T. cancriformis* mitochondiral genome and to clarify the usefulness of the formalin-fixed samples for examining phylogenetic relationships in *T. cancriformis*.

Correspondence: Dr. Kazuo Umetsu, Department of Forensic Medicine, Yamagata University School of Medicine, 2-2-2 Iida-Nishi, Yamagata 990-9585, Japan
E-mail: kumetsu@med.id.yamagata-u.ac.jp
Fax: +81-23-628-5273

Abbreviations: AA, amino acid; **NJ**, neighbor-joining; **nt**, nucleotide

2 Materials and methods

2.1 mtDNA cloning and sequencing

Samples of *T. cancriformis* were collected from rice fields in Yamagata Prefecture, Japan. The mtDNA was prepared from tissues and eggs according to standard protocols [9]. The first primers for cloning (TaKaRa LA PCR *in vitro* cloning kit; TaKaRa, Tokyo, Japan) were designed from the nucleotide (nt) sequence of the COI gene of *Triops australiensis* mtDNA (DDBJ/ENBL/GenBank accession number AF218284). Sequencing was performed by primer walking (primers were synthesized on the basis of the newly clarfied mtDNA sequence of *T. cancriformis*). The PCR products were separated on 1% agarose gels, and then purified by using a Gel Extraction Kit (Qiagen, Hilden, Germany). PCR products were subjected to direct DNA sequencing using an ABI PRISM 310 sequencer and

0173-0835/02/2412–4080 $17.50+.50/0

BigDye Terminator Cycle Sequencing FS Kit (Applied Biosystems, Foster City, CA, USA). The 5'- and 3'-ends of genes and amino acid (aa) sequences of protein-coding genes of *T. cancriformis* mtDNA were inferred from alignments with the *Daphnia pulex* sequences [10] and by BLAST searches [11].

2.2 Phylogenetic analysis

The nt and aa sequences were aligned by using the CLUSTAL-W program [12]. Evolutionary distances were computed by Kimura's two-parameter method [13], and the phylogenetic analysis was performed by the neighbor-joining (NJ) method [14] and by the use of TreeView [15]. Comparisons were made between the complete sequence of tadpole shrimp (*T. cancriformis*) mtDNA and those sequences of eight other arthropods. The species and common names, and DDBJ/ENBL/GenBank accession numbers of the sequences are as follows: *Daphnia pulex* (water flea, AF117817) [10], *Artemia franciscana* (brine shrimp, X69067) [16], *Penaeus monodon* (giant tiger prawn, AF217843) [17], *Ixodes hexagonus* (tick, AF081828) [18], *Limulus polyphemus* (horseshoe crab, AF216203) [19], *Drosophila yakuba* (fruit fly, X03240) [20], *Anopheles gambiae* (mosquito, L20934) [21], and *Locusta migratoria* (grasshopper, S80245) [22]. *Crassostrea gigas* (Pacific oyster, AF177226) was included as an outgroup.

2.3 Analysis of mtDNA from formalin-fixed samples

As listed in Table 1, five specimens of *T. cancriformis*, which had been kept at The Natural History Museum, London, were analyzed for the 16S rRNA gene of their mtDNA. They were most probably fixed initially in formalin and then transferred into industrial-methylated spirit (90% ethanol, 3–10% methanol, 0–7% water) for long-term storage with the reference collections. DNA was prepared from about 10 eggs of the tadpole shrimp. The eggs of each sample were crushed in a 1.5 mL tube, and 50 μL of proteinase K solution (25 ng of proteinase K; 20 mM Tris-HCl, pH 8.0; 20 mM EDTA; 0.1 M 2-mercaptoethanol; 1% sodium *N*-lauroyl sarcosinate) was added; and the mixture was then incubated for 1 h at 60°C. Subsequently, 150 μL of guanidine-detergent lysing solution (4.5 M guanidine thiocyanate; 20 mM Tris-HCl, pH 8.0; 20 mM EDTA; 0.4 M NaCl; 0.1 M 2-mercaptoethanol; 0.5% sodium *N*-lauroyl sarcosinate; 20 ng of glycogen) was added. The supernatant was precipitated with 2.5 volumes of ethanol. Since DNA extracted from formalin-fixed samples is decomposed into very small molecular sizes [23], small fragments ranging from about 100 to 200 bp were the target of the PCR. Three pairs of primers (16SF1: 5'-TCCCTAGGGTAACTTGATCTG-3' and 16SR1: 5'-ACTCTAAGTTATTGGGTTGGG-3'; 16SF2: 5'-GATCCATAGGGTCTTATCGTC-3' and 16SR2: 5'-TTGAGGTATAGCCTGCTCTATG-3'; 16SF3: 5'-TATATTTGCCGAGTTCCTTTG-3' and 16'SR3: 5'-AATTAGGGTTAGAGGGATTAGC-3') were used to amplify the 16S rRNA gene. The PCR products were separated on 1.5% agarose gels, and then purified by using a Gel Extraction Kit (Qiagen). Sequencing of the purified products and phylogenetic analysis were completed by using the methods described above.

Table 1. List of the *T. cancriformis* specimens used in this study (The Natural History Museum)

Locality	Collector/present	Date
Doybayazit, Turkey	–	March 22, 1961
Vienna, Austria	Vienna Museum	November 8, 1911
Dayat, Algeria	Mon L. W. Rothschild	March 16, 1912
Srinagar, India	Zool. Survey of India	October 7, 1954
Kirkcudbrigh, Scotland	F. Baljour Brown	October 17, 1907

3 Results and discussion

The mitochondrial genome of the tadpole shrimp (*T. cancriformis*) was found to be a circular one with 15 101 nts. It encoded information of 22 tRNAs, 13 protein-coding genes and 2 rRNAs as is characteristic of mtDNAs of other metazoans. The complete nt sequence of the mtDNA of *T. cancriformis* was registered in DDBJ/EMBL/GenBank nt sequence databases with the accession number AB084514. The organization of the genome is shown in Fig. 1. The overall base composition of the L-strand was A: 36%; C: 18%; T: 33%; G: 13%. The A+T content was similar to that of the giant tiger prawn [17], but somewhat lower than that of the water flea [10] or the brine shrimp [16].

The location and the putative initiation and termination codons of the protein-coding sequences are shown in Table 2. The three initiation codons (ATA, ATG, GTG) are commonly found in other animals. The use of GTG for the initiation of translation of the ND5 mRNA is also shared by other curstaceans [10]. A 4 bp initiation codon (ATTA) for ND2 in *T. cancriformis* has also been proposed for other arthropod taxa, though this has not been confirmed experimentally [10, 18]. In nine genes (ND2, COI, COII, ATP8, ATP6, ND4L, ND6, Cytb, and ND1), TAA is used as the stop codon; and TAG in one (ND3). Three genes (COIII, ND5, and ND4) are characterized by having an incomplete stop codon TA. The tRNA genes were 64–71 bp long. The control region comprised 467 nts, and was the smallest among the crustaceans tested so far.

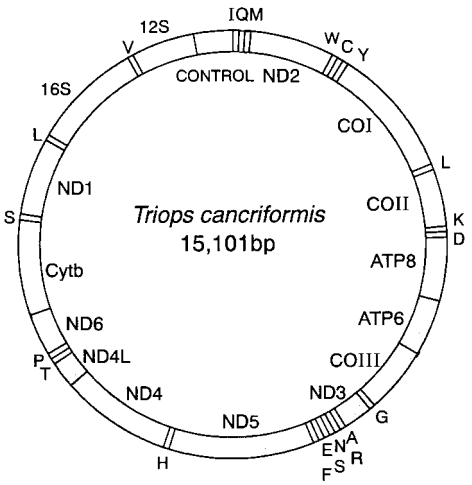

Figure 1. Genetic map of. *T. cancriformis* mtDNA. The tRNAs are identified by the 1-letter aa code.

This region was situated between the 12S rRNA and $tRNA^{Ile}$ genes. The aa sequences deduced from the protein-coding genes of tadpole shrimp were subjected to pairwise comparisons with the homologues of water flea, brine shrimp, giant tiger prawn, tick, horseshoe crab, fruit fly, mosquito, and grasshopper. The results showed that tadpole shrimp proteins had greater similarity in aa sequence to those of water fleas than to those of the brine shrimp and giant tiger prawn.

NJ trees (Fig. 2) were constructed from aa sequences of two well-conserved genes; cytochrome *c* oxidase subunit I (COI) and cytochrome *b* (Cytb), of one nonconserved gene, NADH dehydrogenase subunit 5 (ND5), and all protein-coding genes. Almost identical topologies were obtained in all four trees. *T. cancriformis* was more closely related to the water flea than to the brine shrimp and giant tiger prawn. However, the branching patterns of the brine shrimp were unstable in the mitochondrial trees. The subclass Branchiopoda of the class Crustacea is currently classifed into four orders, *i.e.*, Anostraca (brine shrimps), Notostraca (tadpole shrimps), Conchostraca (clam shrimps) and Cladocera (water fleas). mtDNA gene arrangements are generally well-conserved among arthropod groups, and suggest that gene arrangements could be used to deduce phylogenetic relationships [24]. The arrangement of *T. cancriformis* tRNA genes differs from that in the brine shrimp [16], but the arrangement of the protein-coding genes is the same as that in them. These results indicate that the tadpole shrimp is more closely related to the water flea than to the brine shrimp.

The use of mtDNA as a maternally inherited marker is now widespread to elucidate evolution and migrations. the partial sequence of the 16S rRNA gene of *T. cancriformis* has already been reported [7]. However, those data are probably erroneous, because there was no counterpart with the present mtDNA sequence data. Their PCR products would have been derived from the genomic DNA.

Table 2. Organization of the *T. cancriformis* mitochondrial genome

Gene	From	To	Size (bp)	Start	Stop
$tRNA^{Ile}$	1	65	65		
$tRNA^{Gln}$	63	131	69 (L)		
$tRNA^{Met}$	128	192	65		
ND2	190	1191	1002	(A)TTA	TAA
$tRNA^{Trp}$	1188	1257	70		
$tRNA^{Cys}$	1252	1318	67 (L)		
$tRNA^{Tyr}$	1320	1383	64 (L)		
COI	1378	2919	1542	ATA	TAA
$tRNA^{Leu(UUR)}$	2915	2982	68		
COII	2987	3667	681	ATG	TAA
$tRNA^{Lys}$	3669	3738	70		
$tRNA^{Asp}$	3738	3804	67		
ATP8	3813	3962	150	ATA	TAA
ATP6	3956	4627	672	ATG	TAA
COIII	4627	5414	788	ATG	TA–
$tRNA^{Gly}$	5413	5476	64		
ND3	5480	5830	351	ATA	TAG
$tRNA^{Ala}$	5828	5896	69		
$tRNA^{Arg}$	5901	5964	64		
$tRNA^{Asn}$	5967	6036	70		
$tRNA^{Ser(AGN)}$	6037	6101	65		
$tRNA^{Glu}$	6102	6167	66		
$tRNA^{Phe}$	6169	6233	65 (L)		
ND5	6234	7960	1727 (L)	GTG	TA–
$tRNA^{His}$	7959	8025	67 (L)		
ND4	8026	9344	1319 (L)	ATG	TA–
ND4L	9338	9655	318 (L)	ATG	TAA
$tRNA^{Thr}$	9656	9723	68		
$tRNA^{Pro}$	9723	9786	64 (L)		
ND6	9788	10288	501	ATA	TAA
Cytb	10285	11421	1137	ATA	TAA
$tRNA^{Ser(UNC)}$	11417	11487	71		
ND1	11509	12426	918 (L)	ATG	TAA
$tRNA^{Leu/CUN)}$	12436	12500	65 (L)		
16S rRNA	12502	13807	1306 (L)		
$tRNA^{Val}$	13807	13877	71 (L)		
12S rRNA	13879	14634	756 (L)		
Control region	14635	15101	467		

L, indicated that the gene is transcribed from the light strand

To investigate the tadpole shrimp lineage, we examined the base sequence of 16S rRNA from specimens stored in The Natural Histrory Museum, London (Table 1). The majority of these specimens had been preserved in formalin to conserve their morphology, with the consequence that the DNA had decomposed into fragments of low molecular weight. Other studies have shown that the reduction of PCR product lengths can be used to improve

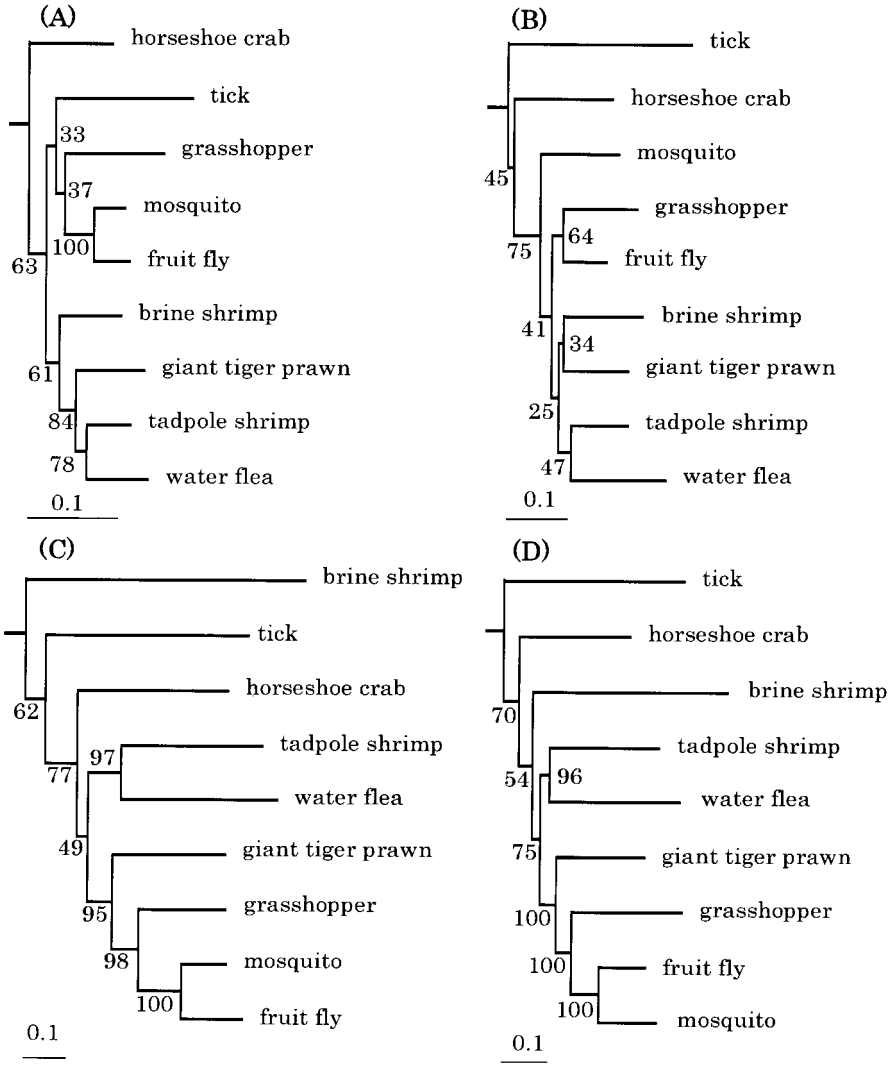

Figure 2. The phylogenetic tree for nine arthropods constructed by the NJ method, based on diffent aa sequences (A, COI; B, Cytb; C, ND5; D, all coding regions). Bootstrap values were computed for 1000 replications. The oyster was used as an outgroup for rooting trees.

the typing of highly degraded DNA [23]. In spite of the use of formalin-fixed samples, some sequence data (465 bp: nt 12 771–12 850, nt 12 937–13 096 and nt 13 181–13 405) revealed of mtDNA common to all samples (Table 3). Based on the observed sequence divergences of the 16S rRNA gene (0–1.51%), populations of *T. cancriforms* from six localities were more closely related to each other than to *T. granarius* (Fig. 3). The 16S rRNA sequence was suitable for estimating phyogenetic relationships in tadpole shrimps. *T. cancriformis* from high latitude regions in the Northern Hemisphere (Scotland, Austria and Japan) were relatively homogeneous for the nt sequence of 16S rRNA. These results do not indicate that *T. cancriformis* have been living in Japan for a long time, but suggest that they migrated recently from a high latitude in Eurasia. Since eggs of the tadpole shrimp are resistant to desication, passive dispersal by human activity, migratory birds or wind is assumed to be responsible [2].

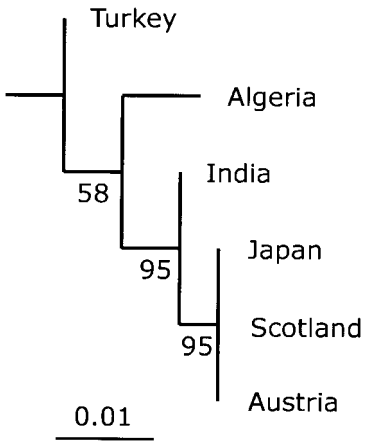

Figure 3. NJ trees based on 16S rRNA sequences of formalin-fixed samples. Bootstrap values were computed for 1000 replications. *T. granarius* (AF200963) was used as an outgroup for rooting the tree.

Table 3. Phylogenetically informative sites of 16S rRNA sequence data

nt No.	*T. cancriformis*						*T. granarius*[a)]
	Japan	Scotland	Austria	India	Turkey	Algeria	Japan
12771	C	C	C	C	C	T	T
12784	T	T	T	T	T	C	C
12798	A	A	A	A	A	T	T
12800	T	T	T	T	T	C	T
12804	T	T	T	T	T	C	C
12809	A	A	A	A	A	T	A
12835	C	C	C	T	T	T	T
13061	T	T	T	T	C	T	C
13369	A	A	A	A	T	A	T
13377	A	A	A	A	G	A	G
13384	A	A	A	A	G	A	T

a) Accession number, AF200963

Detailed elucidation of the phylogenetic relationships between subspecies of *T. cancriformis* [2] must await further studies, including morphological diversification. Furthermore, the role of morphological studies will become more important with the advancement of molecular studies. Thus, a more detailed study of combined morphological and molecular evolution in tadpole shrimps is planned.

4 Concluding remarks

We determined the complete 15 101 bp sequence of the mtDNA of the tadpole shrimp (*T. cancriformis*). Phylogenetic analysis using the NJ method on the mtDNA revealed that the tadpole shrimp was more closely related to the water flea (*Daphnia pulex*) than to the brine shrimp (*Artemia franciscana*). Formalin-fixed samples, commonly stored in museums, were useful for the investigation of DNA. The analysis of mtDNA should become a good method for investigating both the geographical distribution and lineage relationships of tadpole shrimps. Some subspecies are morphologically distinguishable [2], and it is necessary to search for correlations between morphology and mtDNA data.

The authors would like to thank Dr. S. Chiba, Biological Institute, Tohoku University, Japan, for stimulating discussions.

Received June 17, 2002

5 References

[1] Trusheim, F., *Paleont. Z.* 1938, *19*, 198–216.

[2] Longhust, A. R., *Bull. Br. Mus. (Nat. Hist.) Zool.* 1955, *3*, 1–57.

[3] Abe, N., *Bull. Yamagata Uni. (Nat. Sci.)* 1951, *3*, 311–318.

[4] Boore, J. L., *Nucleic Acids Res.* 1999, *27*, 1767–1780.

[5] Zardoya, R., Meyer, A., *Genetics* 1999, *142*, 1249–1263.

[6] Takezaki, N., Gojobori, T., *Mol. Biol. Evol.* 1999, *16*, 590–601.

[7] Suno-uchi, N., Sasaki, F., Chiba, S., Kawata, M., *Biol. J. Linn. Soc.* 1997, *61*, 439–457.

[8] King, J. L., Hanner, R., *Mol. Phylogenet. Evol.* 1998, *10*, 23–36.

[9] Batucecas, B., Garesse, R., Calleja, M., Valverde, J. R., Marco, R., *Nucleic Acids Res.* 1988, *16*, 6515–6529.

[10] Crease, T. J., *Gene* 1999, *233*, 89–99.

[11] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., *J. Mol. Biol.* 1990, *215*, 403–410.

[12] Thompson, J. D., Higgins, D. G., Gibson, T. J., *Nucleic Acids Res.* 1994, *22*, 4573–4680.

[13] Kimura, M., *J. Biol. Evol.* 1980, *16*, 111–120.

[14] Saitou, N., Nei, M., *Mol. Viol. Evol.* 1987, *4*, 406–425.

[15] Page, R. D. M., *Comput. Appl. Biosci.* 1996, *12*, 357–358.

[16] Valverde, J. R., Batucecas, B., Moratilla, C., Marco, R., Garesse, R., *J. Mol. Evol.* 1994, *39*, 400–408.

[17] Wilson, K., Cahill, V., Ballment, E., Benzie, J., *Mol. Biol. Evol.* 2000, *17*, 863–874.

[18] Blach, I. V., Roehrdanz, R. L., *Mol. Biol. Evol.* 1998, *15*, 1772–1785.

[19] Lavrov, D. V., Boore, J. L., Brown, W. M., *J. Mol. Evol.* 2000, *17*, 813–824.

[20] Clary, D. O., Wolstenholme, D. R., *J. Mol. Evol.* 1985, *22*, 252–271.

[21] Beard, C. B., Hamm, D. M., Collins, F. H., *Insect Mol. Biol.* 1993, *2*, 103–104.

[22] Flook, P. K., Rowell, C. H. F., Gellissen, G., *J. Mol. Evol.* 1995, *41*, 928–941.

[23] Goelz, S. E., Hamilton, S. R., Vogelstein, B., *Biochem. Biophys. Res. Commun.* 1985, *130*, 118–126.

[24] Boore, J. L., Lavrov, D. V., Brown, W. M., *Nature* 1998, *376*, 667–668.