Parasitology Research (2007) 101(1):233-236.

Mitochondrial genomes of the human broad tapeworms Diphyllobothrium latum and Diphyllobothrium nihonkaiense (Cestoda : Diphyllobothriidae)

Minoru, Nakao ; Davaajav, Abmed ; Hiroshi, Yamasaki ; Akira, Ito

Submitted to *Parasitology Research* by Nakao M with 19 refs, 1 figure and 1 table

Short Communication

Mitochondrial genomes of the human broad tapeworms *Diphyllobothrium latum* and *Diphyllobothrium nihonkaiense* (Cestoda: Diphyllobothriidae)

Minoru Nakao, Davaajav Abmed, Hiroshi Yamasaki, Akira Ito

Minoru Nakao *, Hiroshi Yamasaki, Akira Ito

Department of Parasitology, Asahikawa Medical College, Midorigaoka-Higashi 2-1, Asahikawa, Hokkaido 078-8510, Japan

Davaajav Abmed

Parasitology Laboratory, National Center for Communicable Diseases, Ulaanbaatar, Mongolia

* Corresponding Author

nakao@asahikawa-med.ac.jp

Tel: +81-166-68-2422

Fax: +81-166-68-2429

Abstract Mitochondrial DNA (mtDNA) sequences of the human broad tapeworms *Diphyllobothrium latum* and *Diphyllobothrium nihonkaiense* have been totally determined. Both of them are closed circular molecules (total length 13720 bp in *D. latum* and 13747 bp in *D. nihonkaiense*) containing genes for 12 proteins, 22 transfer RNAs and two ribosomal RNAs. All the genes are coded on T-rich strand. The gene order of *Diphyllobothrium* mtDNAs is completely identical with that of *Taenia* and *Echinococcus* mtDNAs. The overall A+T contents of the genomes are 68.3% in *D. latum* and 67.8% in *D. nihonkaiense*. The pairwise divergence values of nucleotide sequences between these tapeworms ranged from 0.069 to 0.152 in protein-coding genes, demonstrating that *D. nihonkaiense* is a distinct species. The sequences determined in this study may provide useful marker systems for diagnostic, epidemiological and phylogeographical studies of human diphyllobothriasis.

The intestinal parasite of *Diphyllobothrium latum* or *Diphyllobothrium nihonkaiense* is the longest tapeworm infecting human. They are regarded as sibling species that are nearly indistinguishable morphologically. Two intermediate hosts (copepods and fishes) are required to maintain the life cycle of these tapeworms. Human infection occurs by ingesting the second stage larvae (plerocercoids) in raw or undercooked fishes. Freshwater fishes belonging to the genera *Esox*, *Perca* and *Lota* in the Holarctic region contain the plerocercoids of *D. latum*, whereas anadromous salmons of the genus *Oncorhynchus* in the North Pacific Ocean have the plerocercoids of *D. nihonkaiense* (Dick et al. 2001; Dupouy-Camet and Peduzzi 2004).

Formerly, the causative species of human diphyllobothriasis in Japan had been classified as *D. latum*. Based on morphological and ecological properties, Yamane et al (1986) revised the taxonomy of Japanese broad tapeworms with the description of *D. nihonkaiense*. The antigenicity (Fukumoto et al. 1988) and isozyme polymorphism (Fukumoto et al. 1990) of diphyllobothriid tapeworms made a distinction between *D. latum* and *D. nihonkaiense*, and the restriction fragment length polymorphism (RFLP) of ribosomal RNA gene supported the validity of *D. nihonkaiense* (Matsuura et al. 1992). As concerns the nucleotide sequence of mitochondrial DNA (mtDNA), genes of cytochorome *c* oxidase subunit 1 (*cox1*) and NADH dehydrogenase subunit 3 (*nad3*) have been sequenced in *D. nihonkaiense* (Kokaze et al. 1997; Miyadera et al. 2001). A subsequent comparative study has determined the *cox1* and *nad3* sequences of *D. latum* (Yera et al. 2006). The sequence information of mitochondrial genomes is required in diphyllobothrid tapeworms to make genetic markers for molecular identification. In this paper, we describe the characteristics of the complete mtDNA sequences of *D. latum* and *D. nihonkaiense*.

The mature proglottid of adult *D. latum* was obtained from a Russian patient, who was treated in Mongolia. The proglottid of *D. nihonkaiense* was derived from a Japanese patient living in Hokkaido. Both of the species were morphologically identified and confirmed subsequently by the careful checking of their mtDNA sequences. Genomic DNA purified from each of the proglottids was used as a template for polymerase chain reaction (PCR). As reported previously (Nakao et al. 2003), the initial amplifications of mtDNA were carried out using primers designed from the conserved regions of *Echinococcus multilocularis* mtDNA (Nakao et al. 2002). The sequence data of these amplicons allowed us to design specific primers, whereby the remaining unknown regions were amplified. Each of the PCR products was directly read using BigDye terminator and ABI PRISM 377 genetic analyzer (Applied Biosystems). Large DNA templates were sequenced by primer walking. The resultant sequences covering a complete mitochondrial genome were compiled into a total sequence.

The diphyllobothriid mtDNAs determined in this study were covalently closed-circular molecules of 13720 base pairs (bp) in *D. latum* and 13747 bp in *D. nihonkaiense* (DDBJ/EMBL/GenBank accession nos. AB269325 and AB268585). Open reading frames (ORFs) of ATPase subunit 6 (*atp6*), cytochrome *b* (*cob*), cytochrome *c* oxidase subunits 1 to 3 (*cox1* to *3*) and NADH dehydrogenase subunits 1 to 6 and 4L (*nad1* to *6* and *nad4L*) were inferred by using the echinoderm mitochondrial genetic code (Nakao et al. 2000; Telford et al. 2000). Genes for ribosomal RNA (rRNA) and transfer RNA (tRNA) were detected in the 5'- and 3'-flanking regions of ORFs, based on their sequence motifs. As shown in Table 1, the genomes consisted of 12 protein-coding genes, 2 rRNA genes (*rrnL* and *rrnS*, large and small subunit rRNA), 22 tRNA genes (*trnA*, *trnC, trnD, trnE, trnF, trnG, trnH, trnI, trnK, trnL* (CUN)*, trnL* (UUR)*, trnM, trnN, trnP, trnQ, trnR, trnS* (AGN)*, trnS* (UCN)*, trnT, trnV*, *trnW* and *trnY,* one for each of 18 amino acids and two each for leucine and serine) and two short noncoding regions involved in replication and translation. A gene for ATPase subunit 8 (*atp8*) was absent from the genomes. Most of the genes were separated by a few bases, and the protein genes contained no introns. All the genes were located on the same strand, which was highly biased toward thymine and against cytosine. The nucleotide A+T contents were 68.3% in *D. latum* mtDNA and 67.8% in *D. nihonkaiense* mtDNA. The protein-coding genes were initiated by either an ATG or GTG codon and terminated by a TAG or TAA codon. The abbreviated stop codon T or TA, which is modified into a complete TAA codon by posttranscriptional polyadenylation, was found in *cox3*, *nad1* and *nad3* (Table 1). The gene order of *Diphyllobothrium* mtDNAs was completely identical with that of *Taenia* and *Echinococcus* mtDNAs (Le et al. 2002; Nakao et al. 2003). In tapeworm mitochondrial genomes hitherto examined, the translocation of *trnL* (UCN) and *trnS* (CUN) has been found only in *Hymenolepis diminuta* (Nickisch-Rosenegk et al. 2001).

The pairwise divergence value of mtDNA between *D. latum* and *D. nihonkaiense* was determined for each gene. In the case of tRNA genes, all the short sequences were concatenated into a single sequence (*trn*) to calculate the value. Within the family Taeniidae, the sister species *Taenia saginata* and *Taenia asiatica* are recognized as a pair of recently evolved species (Hoberg et al. 2001), and their divergence values of *cox1* and *cob* sequences (accession nos. AB066494, AB066495, AB066580 and AB066581) were used as lower limits for inter-specific variation. The alignments of DNA sequences were achieved by the program ClustalX (Thompson et al. 1997), and Kimura's 2-parameter distance (Kimura, 1980) was calculated using a gamma shape parameter (a=1). As shown in Fig. 1, the divergence values between *D. latum* and *D. nihonkaiense* ranged from 0.069 (*cox2*) to 0.152 (*atp6*) in 12 protein-coding genes. These high values proved *D. nihonkaiense* to be a distinct species. The comparison of divergence values among mitochondrial genes clearly shows that protein genes are more variable than rRNA and tRNA genes. Consequently, protein-coding regions in mtDNA may be adequate to detect inter- and intra-specific polymorphisms.

The taxonomy of the genus *Diphyllobothrium* is totally dependent on phenetic classification, but the paucity of morphorogical characters causes misidentifications, particularly in larvae and immature adults. The adult tapeworms of *D. latum*, *D. nihonkaiense*, *D. klebanovskii* and *D. ursi* are parasitic in humans and their long strobilae are morphologically similar to each other (Dick et al. 2001; Rausch 1954; Yamane et al. 1988). The latter three species use Pacific salmons as second intermadiate hosts, and the morphological differentiation of their plerocercoids is largely impossible. The recent advance of food transportation by air is responsible for the occurrence of human diphyllobothriasis in nonendemic areas (Yera et al. 2006), indicating that the accurate identification of foreign parasites is difficult in local diagnostic laboratories. The molecular identification of parasites, therefore, is necessary especially for problematic species in morphological classification. The complete sequences of mitochondrial genomes determined in this study may provide useful marker systems for diagnostic, epidemiological and phylogeographical studies of human diphyllobothriasis. The laborious accumulation of sequence information in tapeworm mitochondrial genomes will clarify the phylogeny of the genus *Diphyllobothrium*.

Acknowledgements We thank local physicians who collected tapeworms. This study was supported in part by Grants-in-Aid for Scientific Research from the Japan Society for Promotion of Science (14256001 and 17256002) to Akira Ito.

References

Dick TA, Nelson PA, Choudhury A (2001) Diphyllobothriasis: update on human cases, foci, patterns and sources of human infections and future considerations. Southeast Asian J Trop Med Public Health 32 (Suppl 2): 59-76

Dupouy-Camet J, Peduzzi R (2004) Current situation of human diphyllobothriasis in Europe. Euro Surveill 9: 31-35

Fukumoto S, Yazaki S, Kamo H, Yamane Y, Tsuji M (1988) Distinction between *Diphyllobothrium nihonkaiense* and *Diphyllobothrium latum* by immunoelectrophoresis. Jpn J Parasitol 37: 91-95

Fukumoto S, Yazaki S, Maejima J, Tsuboi T, Hirai K (1990) Soluble protein profiles and isozyme patterns of *Diphyllobothrium pacificum* by isoelectric focusing: comparison with those of related diphyllobothriid species. Yonago Acta Med 33: 61-70

Hoberg EP, Alkire NL, de Queiroz A, Jones A (2001) Out of Africa: origins of the *Taenia* tapeworms in humans. Proc Biol Sci 268: 781-787

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120

Kokaze A, Miyadera H, Kita K, Machinami R, Noya O, Alarcon de Noya B, Okamoto M, Horii T, Kojima S (1997) Phylogenetic identification of

Sparganum proliferum as a pseudophyllidean cestode. Parasitol Int 46: 271-279

Le TH, Blair D, McManus DP (2002) Mitochondrial genomes of parasitic flatworms. Trends Parasitol 18: 206-213

Matsuura T, Bylund G, Sugane K (1992) Comparison of restriction fragment length polymorphisms of ribosomal DNA between *Diphyllobothrium nihonkaiense* and *D. latum*. J Helminthol 66: 261-266

Miyadera H, Kokaze A, Kuramochi T, Kita K, Machinami R, Noya O, Alarcon de Noya B, Okamoto M, Kojima S (2001) Phylogenetic identification of *Sparganum proliferum* as a pseudophyllidean cestode by the sequence analyses on mitochondrial COI and nuclear sdhB genes. Parasitol Int 50: 93-104

Nakao M, Sako Y, Ito A (2003) The mitochondrial genome of the tapeworm *Taenia solium*: a finding of the abbreviated stop codon U. J Parasitol 89: 633-635

Nakao M, Sako Y, Yokoyama N, Fukunaga M, Ito A (2000) Mitochondrial genetic code in cestodes. Mol Biochem Parasitol 111: 415-424

Nakao M, Yokoyama N, Sako Y, Fukunaga M, Ito A (2002) The complete mitochondrial DNA sequence of the cestode *Echinococcus multilocularis* (Cyclophyllidea: Taeniidae). Mitochondrion 1: 497-509

Rausch R (1954) Studies on the helminth fauna of Alaska. XXI. Taxonomy, morphological variation, and ecology of *Diphyllobothrium ursi* n. sp. provis. on Kodiak Island. J Parasitol 40: 540-563

Telford MJ, Herniou EA, Russell RB, Littlewood DT (2000) Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. Proc Natl Acad Sci USA 97: 11359-11364

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 Res 25: 4876-4882.

von Nickisch-Rosenegk M, Brown WM, Boore JL (2001) Complete sequence of the mitochondrial genome of the tapeworm *Hymenolepis diminuta*: gene arrangements indicate that Platyhelminths are Eutrochozoans. Mol Biol Evol 18: 721-730

Yamane Y, Kamo H, Bylund G, Wikgren BJ (1988) *Diphyllobothrium nihonkaiense* sp. nov. (Cestoda: Diphyllobothriidae) - Revised identification of Japanese broad tapeworms. Shimane J Med Sci 10: 29-48

Yera H, Estran C, Delaunay P, Gari-Toussaint M, Dupouy-Camet J, Marty P (2006) Putative *Diphyllobothrium nihonkaiense* acquired from a Pacific salmon (*Oncorhynchus keta*) eaten in France; genomic identification and case report. Parasitol Int 55: 45-49

Figure legend

Fig. 1. Pairwise divergence values of cestode mitochondrial genes. Closed bars show the values between *D. latum* and *D. nihonkaiense*, while open bars indicate the values between *T. saginata* and *T. asiatica*.

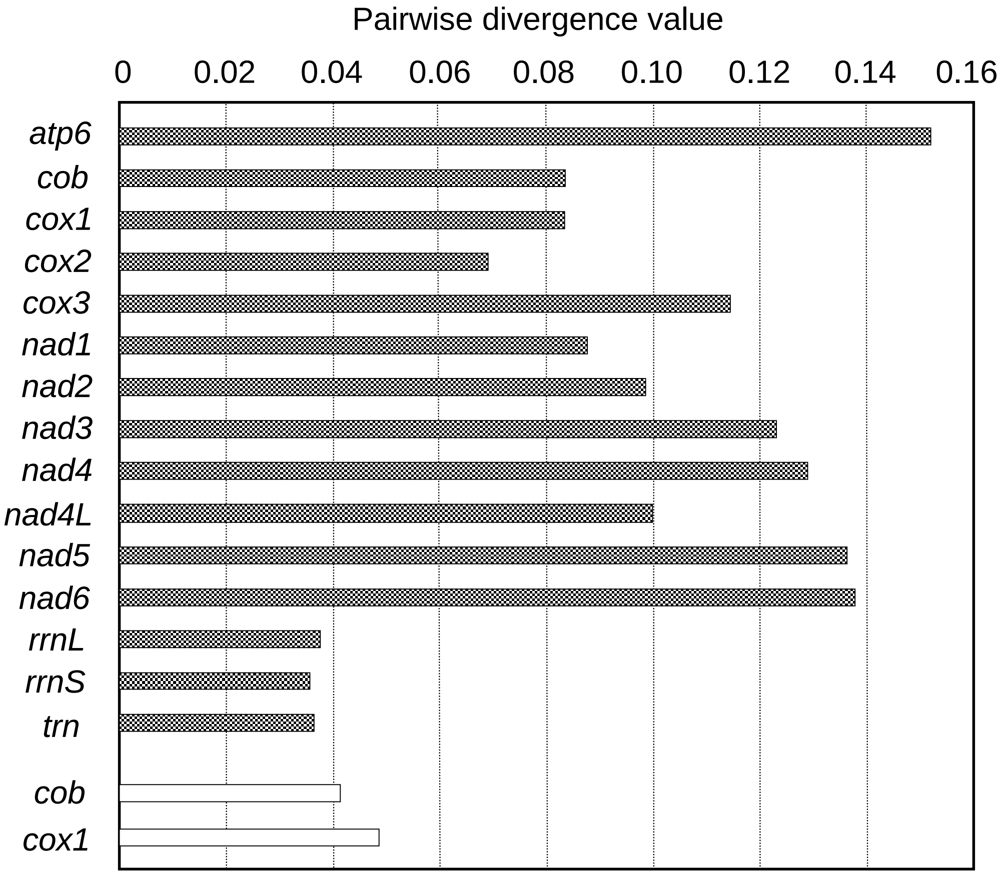

Table 1. Location of genes and noncoding regions (NCR) in the mitochondrial genomes of *Diphyllobothrium latum* and *Diphyllobothrium nihonkaiense*.

	D. latum (total 13720 bases)		*D. nihonkaiense* (total 13747 bases)	
Genes and NCR	Location 5'-3'	Sizes (Start/stop)	Location 5'-3'	Sizes (Start/stop)
trnY	1-65	65	1-65	65
1st NCR	66-286	221	66-289	224
trnL(CUN)	287-353	67	290-356	67
trnS(UCN)	370-426	57	369-425	57
trnL(UUR)	442-505	64	441-503	63
trnR	506-561	56	504-559	56
nad5	565-2133	1569 (ATG/TAA)	563-2131	1569 (ATG/TAA)
2nd NCR	2134-2428	295	2132-2462	331
trnG	2429-2494	66	2463-2528	66
cox3	2498-3140	643 (GTG/T)	2532-3174	643 (GTG/T)
trnH	3141-3207	67	3175-3240	66
cob	3211-4317	1107 (ATG/TAA)	3244-4350	1107 (ATG/TAA)
nad4L	4319-4579	261 (ATG/TAA)	4352-4612	261 (ATG/TAA)
nad4	4540-5790	1251 (ATG/TAG)	4573-5823	1251 (ATG/TAG)
trnQ	5791-5853	63	5824-5886	63
trnF	5850-5916	67	5883-5949	67
trnM	5913-5979	67	5946-6012	67
atp6	5983-6492	510 (ATG/TAG)	6016-6525	510 (ATG/TAG)
nad2	6495-7373	879 (ATG/TAG)	6528-7406	879 (ATG/TAG)
trnV	7375-7438	64	7408-7471	64
trnA	7447-7507	61	7480-7540	61
trnD	7512-7575	64	7544-7607	64
nad1	7576-8465	890 (ATG/TA)	7608-8497	890 (ATG/TA)
trnN	8466-8531	66	8498-8563	66
trnP	8539-8603	65	8571-8635	65
trnI	8613-8675	63	8645-8707	63
trnK	8683-8746	64	8715-8778	64
nad3	8748-9093	346 (ATG/T)	8780-9125	346 (ATG/T)
trnS(AGN)	9094-9152	59	9126-9184	59
trnW	9155-9217	63	9187-9249	63
cox1	9226-10791	1566 (ATG/TAG)	9258-10823	1566 (ATG/TAG)
trnT	10782-10843	62	10814-10875	62
rrnL	10844-11810	967	10876-11838	963
trnC	11811-11874	64	11839-11902	64
rrnS	11875-12618	744	11903-12645	743
cox2	12619-13188	570 (ATG/TAA)	12646-13215	570 (ATG/TAA)
trnE	13190-13253	64	13217-13280	64
nad6	13259-13717	459 (ATG/TAG)	13286-13744	459 (ATG/TAG)