J Mol Evol (2001) 52:311–320
DOI: 10.1007/s002390010161

JOURNAL OF MOLECULAR EVOLUTION

Complete Mitochondrial DNA Sequence of *Conger myriaster* (Teleostei: Anguilliformes): Novel Gene Order for Vertebrate Mitochondrial Genomes and the Phylogenetic Implications for Anguilliform Families

Jun G. Inoue,[1] Masaki Miya,[2] Katsumi Tsukamoto,[1] Mutsumi Nishida[1]

[1] Ocean Research Institute, University of Tokyo, 1-15-1 Minamidai, Nakano-ku, Tokyo 164-8639, Japan
[2] Department of Zoology, Natural History Museum & Institute, Chiba, 955-2 Aoba-cho, Chuo-ku, Chiba 260-8682, Japan

Received: 13 July 2000 / Accepted: 30 November 2000

Abstract. The complete nucleotide sequence of the mitochondrial genome was determined for a conger eel, *Conger myriaster* (Elopomorpha: Anguilliformes), using a PCR-based approach that employs a long PCR technique and many fish-versatile primers. Although the genome [18,705 base pairs (bp)] contained the same set of 37 mitochondrial genes [two ribosomal RNA (rRNA), 22 transfer RNA (tRNA), and 13 protein-coding genes] as found in other vertebrates, the gene order differed from that recorded for any other vertebrates. In typical vertebrates, the ND6, $tRNA^{Glu}$, and $tRNA^{Pro}$ genes are located between the ND5 gene and the control region, whereas the former three genes, in *C. myriaster*, have been translocated to a position between the control region and the $tRNA^{Phe}$ gene that are contiguously located at the 5′ end of the 12S rRNA gene in typical vertebrates. This gene order is similar to the recently reported gene order in four lineages of birds in that the latter lack the ND6, $tRNA^{Glu}$, and $tRNA^{Pro}$ genes between the ND5 gene and the control region; however, the relative position of the $tRNA^{Pro}$ to the ND6–$tRNA^{Glu}$ genes in *C. myriaster* was different from that in the four birds, which presumably resulted from different patterns of tandem duplication of gene regions followed by gene deletions in two distantly related groups of organisms. Sequencing of the ND5–cyt *b* region in 11 other anguilliform species, representing 11 families, plus one outgroup species, revealed that the same gene order as *C. myriaster* was shared by another 4 families, belonging to the suborder Congroidei. Although the novel gene orders of four lineages of birds were indicated to have multiple independent origins, phylogenetic analyses using nucleotide sequences from the mitochondrial 12S rRNA and cyt *b* genes suggested that the novel gene orders of the five anguilliform families had originated in a single ancestral species.

Key words: Gene rearrangement — Long PCR — Complete mtDNA sequence — Phylogenetic implications — Anguilliformes — Congroidei — Eels

Introduction

The vertebrate mitochondrial gene order was initially considered conservative because the complete nucleotide sequences of the entire mitochondrial genome of mammals (Anderson et al. 1981, 1982; Bibb et al. 1981) and the African clawed frog (Roe et al. 1985) showed a common gene order. Although deviations from this gene order have subsequently been identified in various vertebrate lineages, including lampreys (Lee and Kocher 1995), amphibians (Yoneyama 1987; Macey et al. 1997), reptiles (Kumazawa and Nishida 1995; Quinn and Mindell 1996; Macey et al. 1997), birds (Desjardins and Morais 1990, 1991; Quinn and Wilson 1993; Mindell et al. 1998), and marsupials (Pääbo et al. 1991; Janke et al. 1994), none have been found to date among bony fish, in which the mitochondrial genomes have been completely sequenced [loach (Tzeng et al. 1992), carp (Chang et al. 1994), trout (Zardoya et al. 1995), cod (Johansen and

Correspondence to: J.G. Inoue; *e-mail:* jinoue@ori.u-tokyo.ac.jp

Bakke 1996), lungfish (Zardoya and Meyer 1996), bichir (Noack et al. 1996), coelacanth (Zardoya and Meyer 1997), ginbuna (Murakami et al. 1998), Atlantic salmon (Hurst et al. 1999)]. Miya and Nishida (1999), however, recently found the first example of transfer RNA (tRNA) gene rearrangements in the mitochondrial genome of a bony fish, *Gonostoma gracile* (Stomiiformes), in which the positions of three tRNA genes ($tRNA^{Glu}$, $tRNA^{Thr}$, and $tRNA^{Pro}$) relative to the cytochrome *b* (cyt *b*) gene differed from that determined in other vertebrates. Unlike mitochondrial gene rearrangements in other vertebrate lineages, such a gene order appears to occur only in the above species, even following comparison with seven other, currently recognized species of *Gonostoma* (Miya and Nishida 1999, 2000a, unpublished data).

During our molecular phylogenetic studies of eels (the Anguilliformes), one of the basal teleost orders, we attempted to amplify the entire cyt *b* gene sequence (approximately 1150 bp) from various anguilliform species, using several primers designed on the cyt *b* and three flanking tRNA genes ($tRNA^{Glu}$, $tRNA^{Thr}$, and $tRNA^{Pro}$). However, the combinations of these primers for the $tRNA^{Glu}$ gene and the putative downstream genes (cyt *b*, $tRNA^{Thr}$, and $tRNA^{Pro}$) consistently failed to yield polymerase chain reaction (PCR) products for several specific anguilliform species. Accordingly, instead of using $tRNA^{Glu}$ gene-based primers, we employed primers designed on a more upstream gene (ND5) for the anguilliform, *Conger myriaster* (Congridae). The resulting PCR products from this species were unexpectedly small, subsequent sequencing experiments demonstrating that the ND6 and $tRNA^{Glu}$ genes did not occur between the ND5 and the cyt *b* genes in this species. Therefore, we determined the complete mitochondrial DNA (mtDNA) sequence for *C. myriaster* so as to locate the translocated ND6 and $tRNA^{Glu}$ genes.

This paper describes gene organization and gene rearrangements in the mitochondrial genome of *Conger myriaster*, a common conger eel in Japanese shallow waters. We employed a PCR-based approach developed by Miya and Nishida (1999) for sequencing the complete mitochondrial genomes of fishes. In addition, we sequenced the ND5–cyt *b* region for another 11 anguilliform species (representing 11 families) and one outgroup species (*Megalops cyprinoides*) to determine whether or not they share the same gene order as *C. myriaster* and plotted these gene order characters on the phylogenetic tree derived from analyses of nucleotide sequences from the mitochondrial 12S rRNA and cyt *b* genes.

Materials and Methods

Fish Samples and DNA Extraction

Mitochondrial DNA (mtDNA) sequences were obtained from examples of 12 species, representing 12 of 15 families in the order Anguilliformes (Robins 1989). *Megalops cyprinoides* (Elopiformes: Megalopidae) was used as an outgroup, on the basis of currently accepted teleost phylogeny (Greenwood et al. 1966; J. Nelson 1994; Forey et al. 1996), so as to assess the ancestral gene order among the anguilliform families. All specimens were preserved in 99.5% ethanol immediately after collection. Total genomic DNA was extracted using the Qiagen QIAamp tissue kit following the manufacturer's protocol. Voucher specimens were deposited in the Fish Collection, Natural History Museum & Institute, Chiba (CBM-ZF), and the United States National Museum of Natural History (USNM).

Specimens analyzed were *Anguilla japonica* (Japanese eel, Anguillidae; CBM-ZF 10301), *Moringua edwardsi* (spaghetti eel, Moringuidae; USNM 326660), *Kaupichthys hyproroides* (false moray, Chlopsidae; USNM 327549), *Gymnothorax kidako* (moray eel, Muraenidae; CBM-ZF 10505), *Synaphobranchus kaupii* (northern cutthroat eel, Synaphobranchidae; CBM-ZF 10302), *Ophisurus macrorhynchus* (snake eel, Ophichthidae; CBM-ZF 10304), *Nessorhamphus danae* (longneck eel, Derichthyidae; CBM-ZF 10306), *Muraenesox bagio* (pike eel, Muraenesocidae; CBM-ZF 10307), *Nemichthys scolopaceus* (snipe eel, Nemichthyidae; CBM-ZF 10305), *Conger myriaster* (conger eel, Congridae; CBM-ZF 10309), *Nettastoma parviceps* (duckbill eel, Nettastomatidae; CBM-ZF 10308), *Stemonidium hypomelas* (sawtooth eel, Serrivomeridae; CBM-ZF 10303), and *Megalops cyprinoides* (tarpon, Megalopidae; CBM-ZF 10300).

Mitochondrial DNA Purification by Long PCR

The entire mitochondrial genome of *Conger myriaster* was amplified using a long-PCR technique (Cheng et al. 1994; Miya and Nishida 1999). Two sets of fish-versatile primer pairs (S-LA-16S-L + H15149-CYB and L12321-Leu + S-LA-16S-H; Fig. 1) were used so as to amplify almost the entire mitochondrial genome in two long-PCR reactions. Either the $tRNA^{Leu}$–cyt *b* or the $tRNA^{Leu}$–16S regions for other species were amplified using two sets of fish-versatile primer pairs (L12321-Leu + H15149-CYB or L12321-Leu + S-LA-16S-H; Figs. 1 and 2), respectively.

Long PCR was done in a Model 9700 thermal cycler (Perkin–Elmer) and reactions were carried out with 30 cycles of a 25-μl reaction volume containing 15.25 μl of sterile distilled H_2O, 2.5 μl of 10× LA PCR buffer II (Takara), 4.0 μl of dNTP (2.5 m*M*), 1.0 μl of each primer (5 μ*M*), 0.25 μl of 1.25-unit LA Taq (Takara), and 1.0 μl of template containing approximately 5 ng DNA. The thermal cycle profile was that of "shuttle PCR": denaturation at 98°C for 10 s and annealing and extension combined at the same temperature (68°C) for 16 min. Long-PCR products were electrophoresed on a 1.0% L 03 agarose gel (Takara) and later stained with ethidium bromide for band characterization via ultraviolet transillumination. Long-PCR products from the 13 species were diluted with TE buffer (1:20) for subsequent use as PCR templates, except for a region separating the two long PCR primers (S-LA-16S-L and S-LA-16S-H; Fig. 1) in *C. myriaster* and the 12S rRNA region in some other elopomorph fish, in which total genomic DNA was used instead of long-PCR products.

PCR and Sequencing

A total of 63 fish-versatile plus 7 species-specific primers was used to amplify contiguous, overlapping segments of the entire mitochondrial genome for *Conger myriaster* (Fig. 1). These primers were designed with reference to the aligned, complete nucleotide sequences from the mitochondrial genomes of six species of bony fish [loach (Tzeng et al. 1992), carp (Chang et al. 1994), trout (Zardoya et al. 1995), cod (Johansen and Bakke 1996), bichir (Noack et al. 1996), lungfish (Zardoya and Meyer 1996)]. Seven additional species-specific primers were used in regions where no appropriate fish-versatile primers were available (Fig. 1).

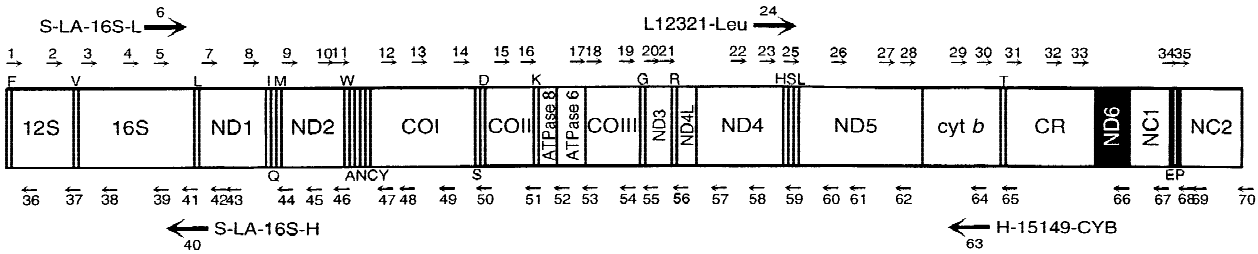

Fig. 1. Gene organization and sequencing strategy for the *Conger myriaster* mitochondrial genome. All protein-coding genes are encoded by the H-strand with the exception of ND6, which is coded by the L-strand. Transfer RNA (tRNA) genes are designated by single-letter amino acid codes, those encoded by the H-strand and L-strand being shown above and below the gene map, respectively. Translocated ND6, $tRNA^{Glu}$, and $tRNA^{Pro}$ genes are *highlighted*. Two pairs of long-PCR primers (S-LA-16S-L + L15149-CYB and L12321-Leu + S-LA-16S-H) amplify two segments that cover nearly the entire mitochondrial genome. 12S and 16S indicate 12S and 16S ribosomal RNA genes; ND1–6 and 4L, NADH dehydrogenase subunits 1–6 and 4L genes; COI–III, cytochrome *c* oxidase subunits I–III genes; ATPase 6 and 8, ATPase subunits 6 and 8 genes; cyt *b*, cytochrome *b* gene; CR, control region. Relative positions of other primers are shown by *small arrows* with numerals. For primer sequences, see Miya and Nishida (1999,[a] 2000b[b]) and Inoue et al. (2000,[c] 2001[d]): 1, L701-12S[d]; 2, L1374-12S[d]; 3, L1854-16S[c]; 4, L2510-16S[a]; 5, L2949-16S[a]; 6, L3074-16S (=S-LA-16S-L)[b]; 7, L3483-ND1[d]; 8, L4166-ND1[a]; 9, L4633-ND2[a]; 10, L5261-ND2[b]; 11, L5644-Ala[a]; 12, L6199-CO1[a]; 13, L6730-CO1[a]; 14, L7255-Co1[a]; 15, L7863-CO2[a]; 16, L8329-Lys[a]; 17, L8984-ATP[b]; 18, L9220-CO3[b]; 19, L9916-CO3[a]; 20, L10267-ND3[a]; 21, L10440-Arg[d]; 22, L11424-ND4[a]; 23, L11895-ND4[a]; 24, L12321-Leu[b]; 25, L12329-Leu[a]; 26, L12936-ND5[b]; 27, L13562-ND5[a]; 28, L13940-ND5[a]; 29, L15369-CYB[b]; 30, L15765-CYB[b]; 31, Comy-CR-L1 (5′-ATA ATA TGT AAA TAT TAC ATA CAC CTA TGG-3′); 32, Comy-CR-L2 (5′-GAC ATA AAA CAT ACA TTA GAA TAT ATC AGG-3′); 33, Comy-CR-L3 (5′-AAA GAT TTA GGT CGT ACA TGT C-3′); 34, L14734-Glu[c]; 35, Comy-NC-L (5′-ACC AAT TAA GAC GAG ACA TGA CG-3′); 36, H885-12S (5′-TAA CCG CGG YGG CTG GCA CGA-3′); 37, H1467-12S (5′-CGG TGT GTG CGC GCC TCA G-3′); 38, H2009-16S[d]; 39, H2582-16S[c]; 40, H2990-16S (=S-LA-16S-H)[b]; 41, H3058-16S[c]; 42, H3466-ND1[b]; 43, H3718-ND1[a]; 44, H4432-Met[b]; 45, H4866-ND2[a]; 46, H5334-ND2[a]; 47, H5937-CO1[b]; 48, H6371-CO1[a]; 49, H6855-CO1[a]; 50, H7480-Ser[a]; 51, H8168-CO2[a]; 52, H8589-ATP[a]; 53, H9076-ATP[a]; 54, H9639-CO3[a]; 55, H10035-Gly[c]; 56, H10433-Arg[a]; 57, H10970-ND4[a]; 58, H11618-ND4[b]; 59, H12145-His[a]; 60, H12632-ND5[a]; 61, H13069-ND5[a]; 62, H13727-ND5[a]; 63, H15149-CYB[b]; 64, H15557-CYB[c]; 65, H15915-Thr[d]; 66, H14473-ND6[b]; 67, Comy-NC-H1 (5′-CTA TGT CTT GTT CCT CAT GTG-3′); 68, H15973-Pro[b]; 69, Comy-CR-H2 (5′-CAA TAA TAA ACG TCA TGT CTC GTC-3′); 70, Comy-NC-H3 (5′-TGT TTT ATT TAT CCT GGG ATA GCG-3′).

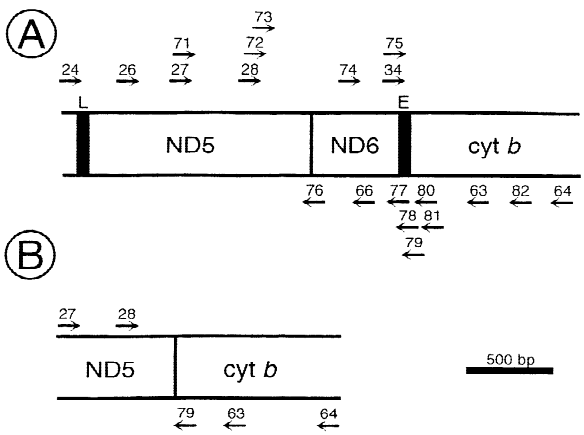

Fig. 2. Primer locations and gene orders in a ND5–cyt *b* region for 12 anguilliform species and an outgroup (*Megalops cyprinoides*). **A** Gene order of eight species representing Megalopidae, Anguillidae, Moringuidae, Chlopsidae, Muraenidae, Synaphobranchidae, Nemichthyidae, and Serrivomeridae. **B** Gene order of five species representing Ophichthidae, Derichthyidae, Muraenesocidae, Nettastomatidae, and Congridae. For abbreviations of genes, see the legend to Fig. 1. Relative positions of other primers are shown by *small arrows* with numerals. For primer sequences, see Miya and Nishida (1999,[a] 2000b[b]) and Inoue et al. (2000,[c] 2001[d]); 71, L13553-ND5[b]; 72, L13916-ND5 (5′-GCA CAA CTT CTC AAA TAT ACT TGG-3′); 73, Mlcy-ND6-L (5′-TCG CAA CTA ATT GAC CAA ACA TG-3′); 74, L14504-ND6[b]; 75, L14724-CYB[d]; 76, H14080-ND5[b]; 77, Mlcy-ND6-H (5′-TGT GGC TTC TAA TCC TGC ACC-3′); 78, Nepa-ND5-H (5′-GGT CAA GAA AAA CAT GGC TAG G-3′); 79, H14718-Glu[b]; 80, H14768-CYB (5′-TTK GCG ATT TTW AGK AGG GGG TG-3′); 81, H14834-CYB[a]; 82, H15341-CYB (5′-TTT GAT CCT GTT TCA TGG AGR AA-3′).

Table 1. Primers used to amplify and sequence the ND5–cyt *b* regions

Species	Long-PCR primers[a]	PCR and sequencing primers[a]
Megalops cyprinoides	24 + 64	27, 73, 75, 64, 77, 78
Anguilla japonica	24 + 64	27, 28, 75, 77, 82
Moringua edwardsi	24 + 63	71, 73, 74, 80, 63
Kaupichthys hyproroides	24 + 63	27, 72, 74, 63
Gymnothorax leucostigma	26 + 63	27, 74, 75, 77, 78, 81, 82
Synaphobranchus kaupii	24 + 40	27, 28, 34, 81, 82, 64
Ophisurus macrorhynchus	24 + 40	27, 28, 63, 64
Nessorhamphus danae	28 + 63	28, 63
Muraenesox bagio	27 + 63	28, 63
Nemichthys scolopaceus	24 + 40	28, 75, 82
Nettastoma parviceps	24 + 40	28, 79, 64
Stemonidium hypomelas	28 + 64	28, 82

[a] For primer sequences and locations, see Figs. 1 and 2.

For the ND5–cyt *b* regions in other elopomorph fish, an additional 12 fish-versatile plus 3 species-specific primers were used (Table 1, Fig. 2). Two species-specific primers for *M. cyprinoides* and one for *N. parviceps* were additionally used in regions where no appropriate fish-versatile primers were available (Fig. 2). For the 12S rRNA gene region, three fish-versatile primers were additionally used for another 12 elopomorph fish.

PCR was done in a Model 9700 thermal cycler (Perkin–Elmer) and reactions carried out with 30 cycles of a 25-μl reaction volume containing 14.4 μl of sterile, distilled H_2O, 2.5 μl of 10× PCR buffer (Perkin–Elmer), 2.0 μl of dNTP (2.5 m*M*), 2.5 μl of each primer (5 μ*M*), 0.1 μl of 0.5-unit Ex Taq (Takara), and 1.0 μl of template.

The thermal cycle profile was as follows: denaturation at 94°C for

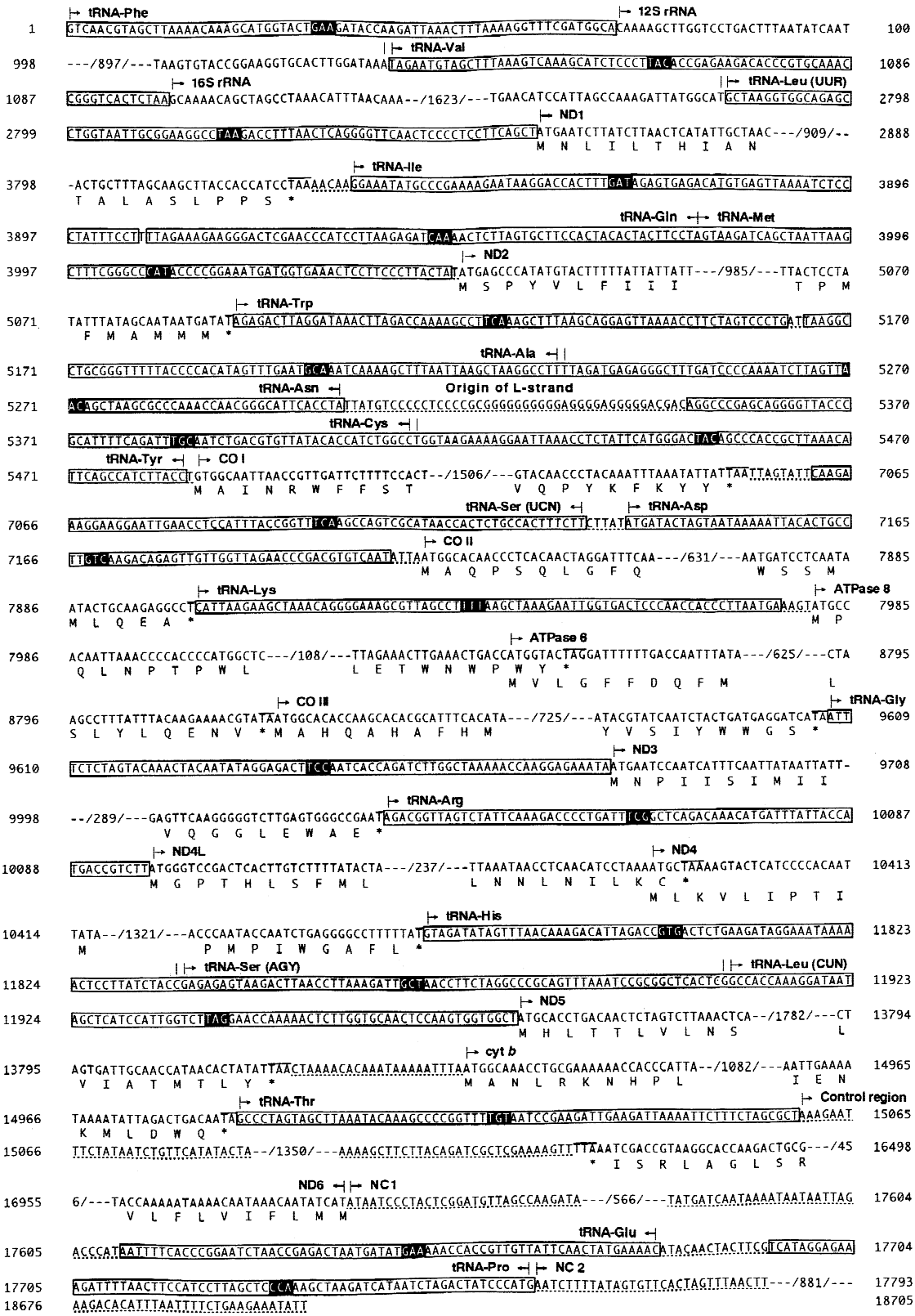

Fig. 3. Schematic representation of the complete L-strand nucleotide sequence of the *Conger myriaster* mitochondrial genome. Position 1 corresponds to the first nucleotide of the tRNA^{Phe} gene. Direction of transcription for each gene indicated by an *arrow.* Beginning and end of each gene indicated by *vertical lines.* Transfer RNA genes are *boxed;* corresponding anticodons are indicated by *block boxes.* Amino acid sequences presented below the nucleotide sequence were derived using the mammalian mitochondrial genetic code (one-letter amino acid abbreviations placed below the first nucleotide of each codon). Stop codons are *overlined* and indicated by *asterisks.* Noncoding sequences are *underlined with dots.* Numbers within slash marks indicate the numbers of nucleotides omitted from this figure. Sequence data are available from DDBJ/EMBL/GenBank under accession number AB038381. For abbreviations of genes, see the legend to Fig. 1.

15 s, annealing at 50°C for 15 s, and extension at 72°C for 20 s. PCR products were electrophoresed on a 1.0% L 03 agarose gel (Takara) and stained with ethidium bromide for band characterization via ultraviolet transillumination.

Double-stranded PCR products, purified by filtration through a Microcon-100 (Amicon Inc.) column, were subsequently used for direct cycle sequencing with dye-labeled terminators (Perkin–Elmer). Primers used were the same as those for PCR. All sequencing reactions were performed according to the manufacturer's instructions. Labeled fragments were analyzed on Model 310/377 DNA sequencers (Perkin–Elmer).

Sequence Analysis

DNA sequences were analyzed using the computer software package program DNASIS, version 3.2 (Hitachi Software Engineering Co. Ltd.). The location of 13 protein-coding genes were determined by comparisons of DNA or amino acid sequences of bony fish mitochondrial genomes. The 22 tRNA genes were identified by their cloverleaf-shaped secondary structures (Kumazawa and Nishida 1993) and anticodon sequences. The two rRNA genes were identified by sequence homology and secondary structure (Gutell et al. 1993). Sequence data are available from DDBJ/EMBL/GenBank under accession numbers AB038381 for *Conger myriaster*, AB038556 for *Anguilla japonica* (Inoue et al. 2001), and AB038410, AB038414–AB038420, and AB049978–AB049991 for the other 11 species.

Phylogenetic Analysis

The 12S rRNA gene sequence alignment for the 12 anguilliform and an outgroup species were initiated with Clustal X (Thompson et al. 1997), with default gap penalties, and the output was later improved by eye. Ambiguous alignment regions were excluded from the analyses, leaving 621 nucleotide positions from the 12S rRNA gene. The cyt *b* gene sequences were aligned by eye on the basis of inferred amino acid sequences. All third codon positions were excluded from the analyses, leaving 780 nucleotide positions from the cyt *b* gene. MacClade version 3.08 (Maddison and Maddison 1992) was used in various phases of the phylogenetic analyses, such as preparing data matrices in NEXUS format, exporting tree files, and exploring alternative tree topologies. Aligned sequence data in NEXUS format are available from J.G.I. upon request.

Maximum-parsimony (MP) analyses were performed with the branch-and-bound algorithm in PAUP 4.0b4a (Swofford 1998), using equal weighting for all substitutions. All phylogenetically uninformative sites were ignored. Gaps were considered as missing, rather than as fifth characters, to circumvent those longer than one or two bases being considered multiple events (Swofford 1993). To evaluate the robustness of the internal branches of the MP tree, 500 bootstrap replications (Felsenstein 1985) were executed, with 20 heuristic, random stepwise additions being performed at each replication.

Maximum-likelihood (ML) analyses were conducted, with 100 heuristic, random stepwise additions, using PAUP 4.0b4a to determine the statistically most likely phylogeny under the estimated transition/transversion bias and rates of base pair substitution at different codon positions. The parameter-rich HKY85 model of sequence evolution (Hasegawa et al. 1985) was used to accommodate unequal base frequency, with the following search option: use empirical base frequencies, estimate the transition/transversion ratio from the data set, consider rate heterogeneity among sites, and compute the shape parameter of the gamma distribution from the data.

Differences in tree topologies were compared statistically using the Templeton (1983) test and Kishino–Hasegawa (1989) test, implemented in PAUP version 4.0b4a.

Results and Discussion

Genome Content and Gene Order of Conger myriaster

The total length of the *Conger myriaster* mitochondrial genome was 18,705 bp (Fig. 3), the genome content of *C. myriaster* including 2 rRNA, 22 tRNA, and 13 protein-coding genes and a control region, as found in other vertebrates (Table 2, Figs. 1 and 3). Also, as in other vertebrates, most genes were encoded on the H-strand, except for the ND6 and eight tRNA genes, and all genes, except for the control region (1,410 bp), were similar in length to those in other bony fish.

As in most vertebrates, the origin of light-strand replication (O_L) in *C. myriaster* was located in a cluster of five tRNA genes (WANCY region; Fig. 3), being 59

Table 2. Location of features in the mitochondrial genome of *Conger myriaster*

Feature[a]	Position number		Size (bp)	Codon	
	From	To		Start	Stop
$tRNA^{Phe}$	1	70	70		
12S rRNA	71	1,027	957		
$tRNA^{Val}$	1,028	1,099	72		
16S rRNA	1,100	2,782	1,683		
$tRNA^{Leu(UUR)}$	2,783	2,858	76		
ND1	2,859	3,827	969	ATG	TAA
$tRNA^{Ile}$	3,833	3,905	73		
$tRNA^{Gln}$	3,907	3,977	71 (L)		
$tRNA^{Met}$	3,977	4,045	69		
ND2	4,047	5,091	1,045	ATG	T——
$tRNA^{Trp}$	5,092	5,162	71		
$tRNA^{Ala}$	5,165	5,233	69 (L)		
$tRNA^{Asn}$	5,234	5,305	72 (L)		
$tRNA^{Cys}$	5,350	5,414	65 (L)		
$tRNA^{Tyr}$	5,415	5,485	71 (L)		
COI	5,487	7,052	1,566	GTG	TAA
$tRNA^{Ser(UCN)}$	7,061	7,131	71 (L)		
$tRNA^{Asp}$	7,137	7,206	70		
COII	7,211	7,901	691	ATG	T——
$tRNA^{Lys}$	7,902	7,976	75		
ATPase 8	7,981	8,148	168	ATG	TAG
ATPase 6	8,139	8,821	683	ATG	TA—
COIII	8,822	9,606	785	ATG	TA—
$tRNA^{Gly}$	9,607	9,678	72		
ND3	9,679	10,027	349	ATG	T——
$tRNA^{Arg}$	10,028	10,097	70		
ND4L	10,098	10,394	297	ATG	TAA
ND4	10,388	11,768	1,381	ATG	T——
$tRNA^{His}$	11,769	11,837	69		
$tRNA^{Ser(AGY)}$	11,838	11,907	70		
$tRNA^{Leu(CUN)}$	11,908	11,980	73		
ND5	11,981	13,822	1,842	ATG	TAA
cyt *b*	13,845	14,986	1,142	ATG	TA—
$tRNA^{Thr}$	14,987	15,058	72		
Control region	15,059	16,468	1,410		
ND6	16,469	16,984	516 (L)	ATG	TAA
$tRNA^{Glu}$	17,611	17,679	69 (L)		
$tRNA^{Pro}$	17,694	17,763	70 (L)		

[a] For abbreviations of genes, see the legend to Fig. 1.

nucleotides in length. This region had the potential to fold into a stable stem–loop secondary structure, comprising 12 bp in the stem and 13 nucleotides in the loop. The conserved motif-like sequence 5′-GCTCGG-3′ was found at the base of the stem within the $tRNA^{Cys}$ gene (Fig. 3).

The *C. myriaster* mitochondrial genome contained three major noncoding regions [control region (CR), NC1, and NC2] separated by the ND6 and a cluster of $tRNA^{Glu}$ and $tRNA^{Pro}$ genes (Figs. 1 and 3). A noncoding region (1,410 bp) located between the $tRNA^{Thr}$ and the ND6 genes appeared to correspond to the control region, because it had conserved sequence blocks [CSB-D, -II, and -III (Walberg and Clayton 1981)] and a termination-associated sequence [TAS (Doda et al. 1981)] characteristic of that region. NC1 (629 bp), located between the ND6 and the $tRNA^{Glu}$ genes, had no control-region characteristics, such as CSBs and TASs. NC2 (942 bp) contained 54 copies of 12- to 13-bp strings (YTATYATA[A]RTAT). Multiple copies of the repeat in *C. myriaster* NC2 formed a stable secondary structure (data not shown), similar to that found for seven copies of a 27-bp string, occurring in the noncoding region adjacent to the 5′ end of the putative control region in lamprey (Lee and Kocher 1995).

The gene order of *C. myriaster* was largely identical to that of typical vertebrates, although the gene arrangement adjacent to the control region differed from all other vertebrate arrangements known to date (Macey et al. 1997; Boore 1999) (Figs. 1 and 3). In typical vertebrates, the ND5, $tRNA^{Glu}$, and $tRNA^{Pro}$ genes are located between the ND5 gene and the control region. In *C. myriaster*, however, the former three genes were translocated between the control region and the $tRNA^{Phe}$ gene (contiguously located at the 5′ end of the 12S rRNA gene in typical vertebrates).

Possible Mechanism for the Gene Rearrangement

Because the gene order in *Conger myriaster* is novel among vertebrates, it is reasonable to assume that the typical vertebrate gene order is more ancestral than that of the former (Fig. 4A). Gene rearrangements have been proposed as occurring by tandem duplication of gene regions as a result of slipped strand mispairing, followed

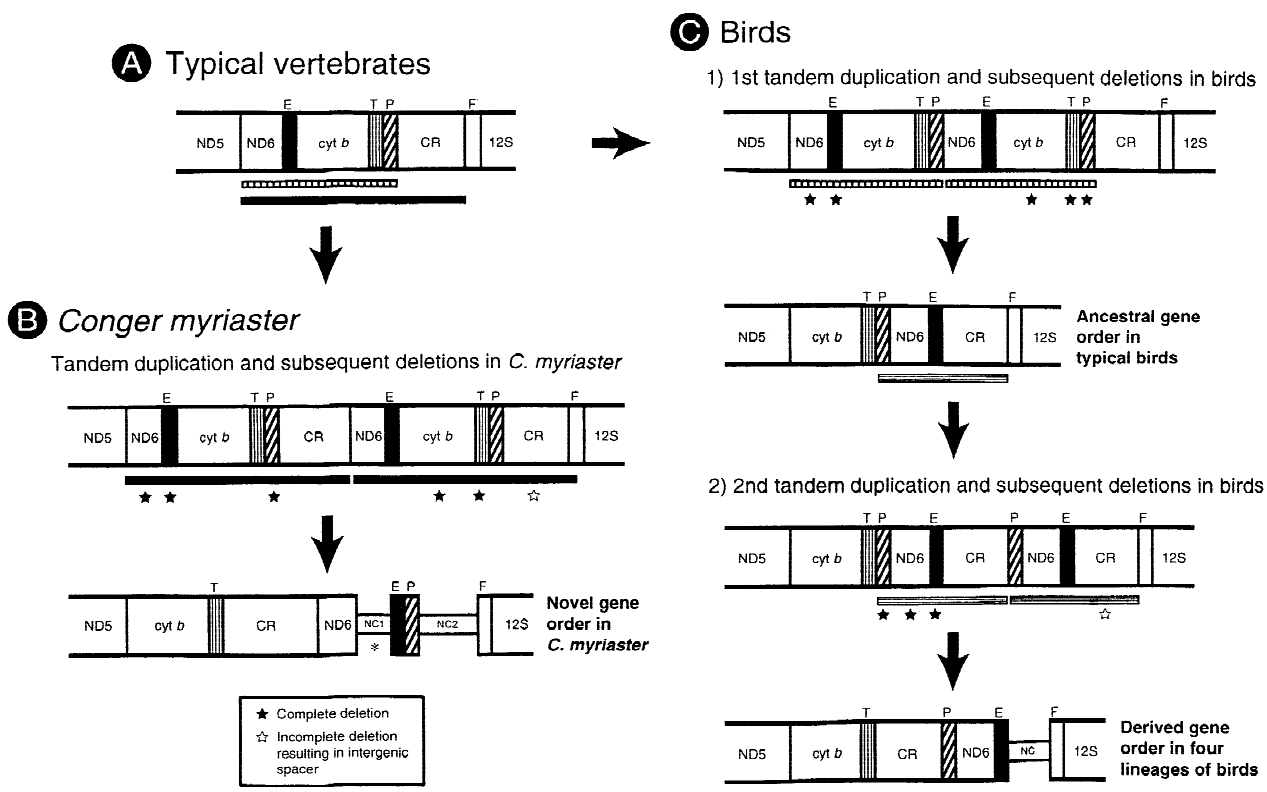

Fig. 4. Proposed mechanism of gene rearrangements in the ND5–$tRNA^{Phe}$ region in *Conger myriaster* and birds in a model of tandem duplication of gene regions and subsequent gene deletions. **A** Typical vertebrate gene order in a region of ND5–12S rRNA. **B** Tandem duplication in the ND6–CR region (*thick bar*) and subsequent deletions of redundant genes resulting in the observed gene order in *C. myriaster*. **C1** First tandem duplication in the ND6–$tRNA^{Pro}$ region (*vertical bar*) and subsequent deletions of redundant genes resulting in the ancestral gene order in birds (Quinn and Wilson 1993). **C2** Second tandem duplication in the $tRNA^{Pro}$–CR region (*horizontal bar*) and subsequent deletions of redundant genes resulting in the derived gene order in four divergent lineages of birds (Mindell et al. 1998; Bensch and Härlid 2000). In B, note that the observed intergenic spacer NC1 (*; noncoding region 1) may result from the duplication and degeneration of ND6, whereas NC2 may result from incomplete deletion events. It is notable that NC2 comprises 54 copies of 12- or 13-bp strings (YTATY-ATA[A]RTAT), which may be a degenerating vestige of the control region. For abbreviations of genes, see the legend to Fig. 1.

by deletions of genes (Levinson and Gutman 1987; Moritz and Brown 1986). The present gene order of *C. myriaster* and associated intergenic spacers could have resulted from an occurrence of this process, namely, tandem duplication occurring in the ND6–CR region, followed by deletions of redundant genes (Fig. 4B). Such tandem duplication and subsequent deletions most parsimoniously resulted in the observed gene order and associated intergenic spacers in *C. myriaster*.

Multiple deletions of redundant genes seemed to be incomplete in *C. myriaster*, owing to two stretches of noncoding sequences (NC1, 626 bp; NC2, 942 bp) occurring around the genes involved in the rearrangements. Although no homologous regions were identified for these two sequences, the latter (NC2), located between the $tRNA^{Pro}$ and the $tRNA^{Phe}$ genes, had a tandem repeat sequence characteristic of the control region of many vertebrates (Hoelzel et al. 1994; Lee et al. 1995; Zardoya and Meyer 1997). Because this noncoding sequence was located at the original position of the control region, it was likely to be a degenerating vestige of a duplicate control region. This observation supports the concept of the tandem duplication and subsequent deletion events as having occurred in the ND6–CR region (Fig. 4B).

Several mitochondrial gene rearrangements have been reported around the control region in other vertebrates (Macey et al. 1997; Boore 1999). It should be noted that the novel gene order in *C. myriaster* was similar to the recently reported gene order in four lineages of birds (Fig. 4C) (Mindell et al. 1998; Bensch and Härlid 2000) in that the ND6, $tRNA^{Glu}$, and $tRNA^{Pro}$ genes are translocated to a position between the CR and a supposedly noncoding region. However, a difference in the gene orders of *C. myriaster* and the four birds was found in the relative position of the $tRNA^{Pro}$ gene to the ND6–$tRNA^{Glu}$ genes (Figs. 4B and C). It seems plausible that a double occurrence of tandem duplication of gene regions followed by gene deletions in the four birds resulted in this difference as suggested by Mindell et al. (1998) (Fig. 4C). Also, it should be noted that the three genes (ND6, $tRNA^{Glu}$, and $tRNA^{Pro}$) involved in gene rearrangements in the two very distantly related lineages were all encoded by the light strand of vertebrate mitochondrial DNA. These findings suggest that some constraints on gene order mutation in vertebrates exist around the origin of heavy-strand replication within the mitochondrial control region (Mindell et al. 1998).

Gene Orders in Other Anguilliform Families and Phylogenetic Implications

The ND5–cyt *b* regions for 11 other anguilliform species, representing 11 families, plus an outgroup species, were sequenced so as to assess whether or not they shared the mitochondrial gene order found in *Conger myriaster*.

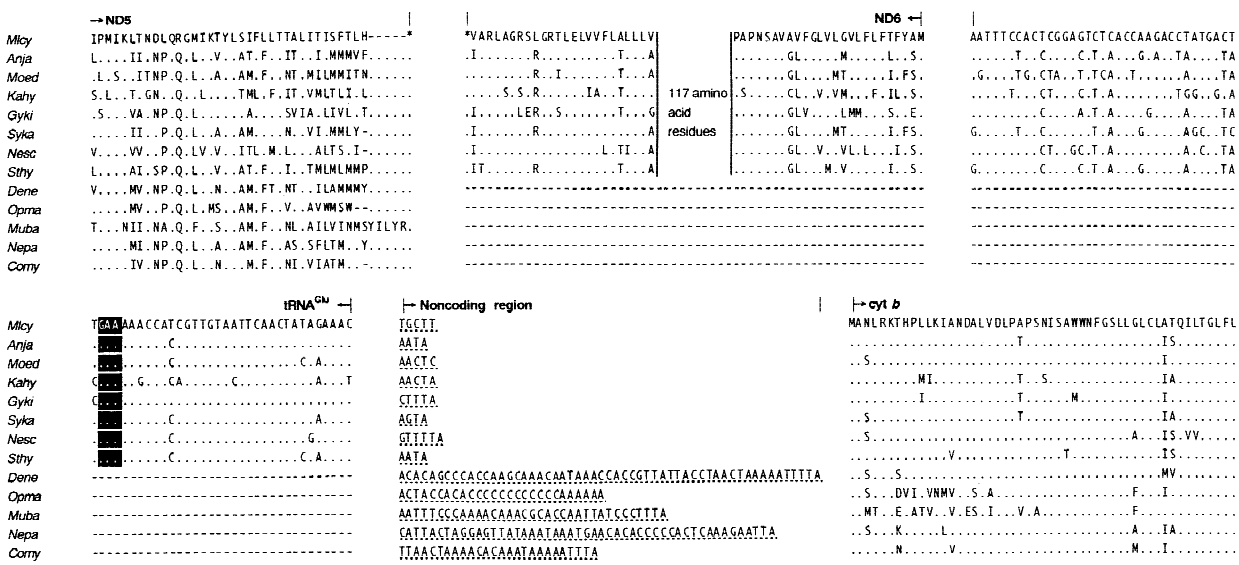

Fig. 5. Aligned nucleotide and amino acid sequences of the ND5–cyt *b* gene junctions for 12 anguilliform species and an outgroup (*Megalops cyprinoides*). Identity to the first sequence denoted by *dots.* Insertions/deletions of specific nucleotides indicated by *dashes.* Direction of transcription for each gene shown by an *arrow.* Beginning and end of each gene indicated by *vertical lines.* The $tRNA^{Glu}$ gene was identified by the cloverleaf secondary structure proposed by Kumazawa and Nishida (1993) and anticodon sequences, indicated by *black boxes.* Amino acid sequences derived using the mammalian mitochondrial genetic code. Stop codons indicated by *asterisks.* Noncoding sequences *underlined with dots.* Nucleotide sequence data are available from DDBJ/EMBL/GenBank under accession numbers AB038381 (*Conger myriaster*), AB038556 (*Anguilla japonica*), and AB038410, AB038414–AB038420, and AB049978–AB049980 (11 other species). For abbreviations of genes, see the legend to Fig. 1. *Mlcy, Megalops cyprinoides; Anja, Anguilla japonica; Moed, Moringua edwardsi; Kahy, Kaupichthys hyproroides; Gyki, Gymnothorax kidako; Syka, Synaphobranchus kaupii; Nesc, Nemichthys scolopaceus; Sthy, Stemonidium hypomelas; Dene, Nessorhamphus danae; Opma, Ophisurus macrorhynchus; Muba, Muraenesox bagio; Nepa, Nettastoma parviceps; Comy, Conger myriaster.*

The results demonstrated that representatives from four families (Ophichthidae, Derichthyidae, Muraenesocidae, and Nettastomatidae) all lacked the ND6 and $tRNA^{Glu}$ genes in this region, with short stretches of noncoding sequences (4–55 bp) being observed alternatively (Fig. 5). Although the two genes (ND6 and $tRNA^{Glu}$) were not located in the entire mitochondrial genomes from these four species, this observation suggests that they have a gene order identical to that of *C. myriaster*. In contrast, the gene order of the other eight species, including that of the outgroup, was the same as that of typical vertebrates, with the ND6 and $tRNA^{Glu}$ genes being located between the ND5 and the cyt *b* genes (Fig. 5).

The presence/absence of this novel gene order and the current higher-level classification within the Anguilliformes (Robins 1989) are contrasted in Fig. 6. All species of the two suborders (Anguilloidei and Muraenoidei) and an outgroup shared the typical vertebrate gene order, whereas species of another suborder (Congroidei) exhibited two patterns of gene order: those with the typical vertebrate gene order (Synaphobranchidae, Serrivomeridae, and Nemichthyidae) vs those with a novel gene order (Derichthyidae, Muraenesocidae, Nettastomatidae, Congridae, and Ophichthidae). Although the superfamilial classification of the Congroidea based on the morphology of the hyomandibular (Robins 1989) was inconsistent with the pattern of presence/absence of the novel gene order, it should be noted that the five families with the novel gene order exhibit complete congruence with components of one of the three evolutionary lineages recognized by G. Nelson (1966), who examined gill-arch and other characters in the Anguilliformes. If the novel gene order originated in a single ancestral species in the suborder Congroidei, this pattern of presence/absence may be a good phylogenetic marker for identification of a monophyletic group as suggested by Kumazawa and Nishida (1995) and Macey et al. (1997) for higher vertebrate relationships. On the contrary, Mindell et al. (1998) and Bensch and Härlid (2000) demonstrated that the recently reported gene orders in four lineages of birds (Fig. 4C2) have multiple independent origins among avian lineages.

To evaluate alternative hypotheses on the origin of the novel gene order, partial nucleotide sequences from the mitochondrial 12S rRNA and cyt *b* genes were analyzed for the relationships among anguilliform families (Fig.

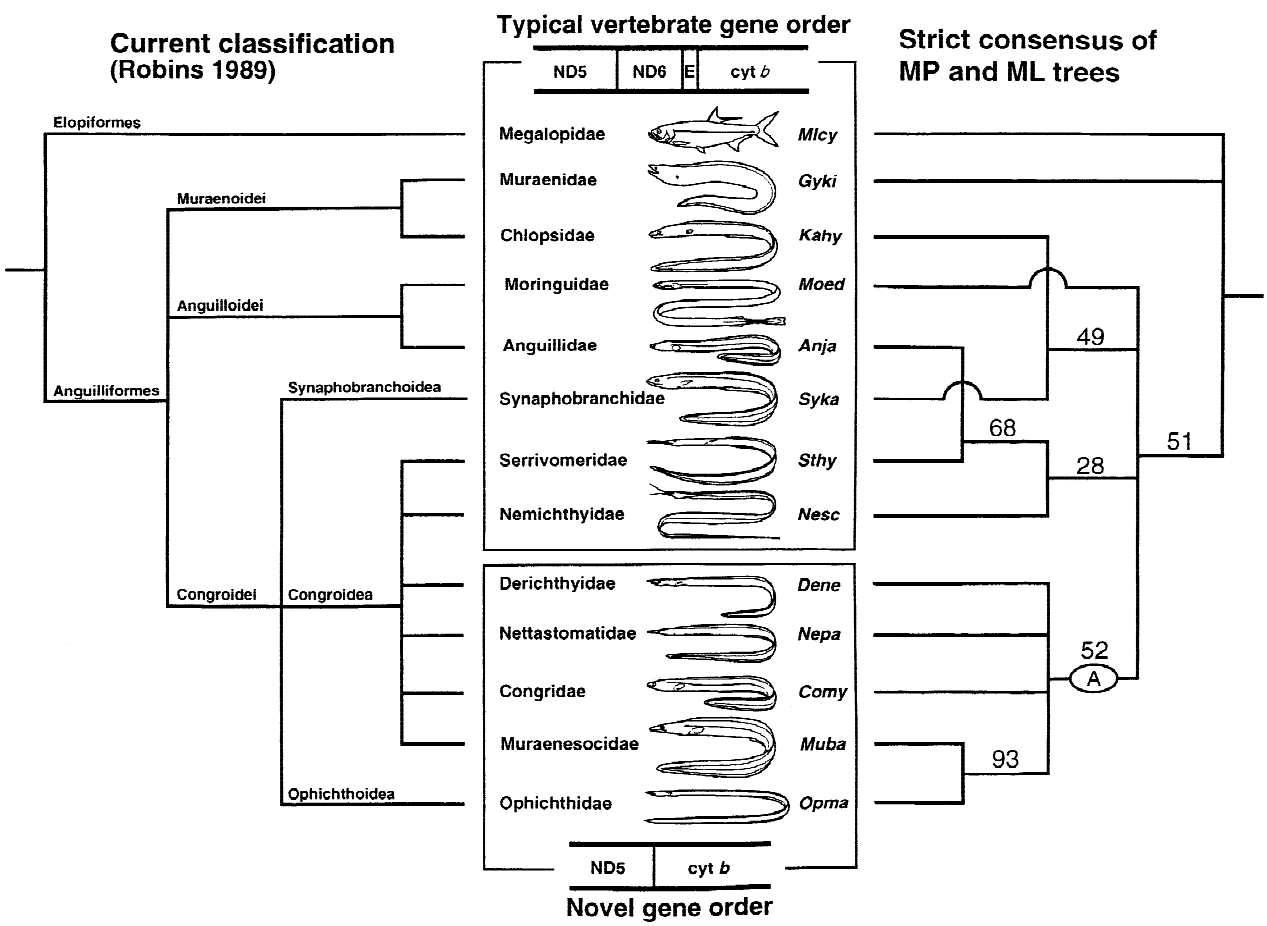

Fig. 6. **Left** Relationships between the current classification of eels (Robins 1989) and the gene organization of the ND5–cyt *b* region of 12 anguilliform and an outgroup species. **Right** Strict consensus of MP and ML trees derived from the partial nucleotide sequences from the mitochondrial 12S rRNA and cyt *b* genes, for 12 anguilliform and an outgroup species. Numbers above branches indicate bootstrap values obtained for 500 replicates for the MP tree. For abbreviations of specific names, see the legend to Fig. 5.

6). The maximum-parsimony (MP) analysis produced a single MP tree (tree length = 761, consistency index = 0.463, retention index = 0.288, and rescaled consistency index = 0.133). Although the MP and maximum-likelihood (ML) tree topologies differed somewhat, it should be noted that both analyses consistently reproduced monophyly of the five families with the novel gene order (Fig. 6, Clade A). In addition, when monophyly of Robins' (1989) "superfamily Congroidea" was enforced using the constraint option in PAUP, two minimum-length trees that required an additional 36 steps were found, the differences being highly significant in both Templeton ($z = -3.432$, $p = 0.0006$, and $z = -3.467$, $p = 0.0005$) and Kishino–Hasegawa ($\Delta\ln L \pm \text{SE} = 101.4 \pm 26.5$, $p = 0.0001$, and $\Delta\ln L \pm \text{SE} = 100.7 \pm 26.6$, $p = 0.0002$) tests.

On the basis of the pattern of the presence/absence of the novel gene order that exhibits congruence with the molecular phylogeny of the anguilliform families (Fig. 6), we conclude that the present study provides new insight into the evolution and phylogeny of this morphologically conservative group of fish, considering the controversial higher-level classification within the Anguilliformes (see Robins 1989; Forey et al. 1996). Also, we believe that longer DNA sequences from other anguilliform families, such as the complete mt DNA sequences, would resolve their complex higher-level relationships.

Acknowledgments. We thank the captains, officers, crew, scientists, and students onboard for their assistance in collecting samples during the KH95-2 cruise of the R/V *Hakuho Maru* and the KT96-1, 96-19, 97-10, 99-6, and 99-10 cruises of the R/V *Tansei Maru.* This study would not have been possible without generous donations of tissue materials, for which we sincerely thank E.O. Wiley, S. Ohta, and A. Torii. Helpful suggestions regarding the identification of specimens were provided by K. Hatooka, T. Maeda, and H. Yeh. Thanks are also due to Y. Fukuyo, K. Furuya, K. Nanba, and graduate students at Fisheries Oceanography Laboratory, University of Tokyo, for generously allowing us to use their experimental facilities. We also thank G.S. Hardy for reading the manuscript. The helpful comments of an anonymous reviewer are also greatly appreciated. This study was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture of Japan (Nos. 10460081, 10660189, 11691177, 12460083, and 12NP0201) and "Research for the Future" Program No. JSPS-RFTF 97L00901 from the Japan Society for the Promotion of Science. K.T. was supported by the Research Foundation from Touwa Shokuhin Shinkoukai and the Eel Research Foundation from Noborikai.

References

Anderson S, Bankier AT, Barrell BG, et al. (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Anderson S, de Bruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG (1982) Complete sequence of bovine mitochondrial DNA: Conserved features of the mammalian mitochondrial genome. J Mol Biol 156:683–717

Bensch S, Härlid A (2000) Mitochondrial genomic rearrangements in songbirds. Mol Biol Evol 17:107–113

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. Cell 26:167–180

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27:1767–1780

Chang Y-S, Huang F-I, Lo T-B (1994) The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. J Mol Evol 38:138–155

Cheng S, Higuchi R, Stoneking M (1994) Complete mitochondrial genome amplification. Nature Genet 7:350–351

Desjardins P, Morais R (1990) Sequence and gene organization of the chicken mitochondrial genome. J Mol Biol 212:599–634

Desjardins P, Morais R (1991) Nucleotide sequence and evolution of coding and noncoding regions of a quail mitochondrial genome. J Mol Evol 32:153–161

Doda JN, Wright CT, Clayton DA (1981) Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc Natl Acad Sci USA 78: 6116–6120

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791

Forey PL, Littlewood DTJ, Ritchie P, Meyer A (1996) Interrelationships of elopomorph fishes. In: Stiassny MLJ, Parenti LR, Johnson GD (eds) Interrelationships of fishes. Academic Press, San Diego, pp 175–191

Greenwood PH, Rosen DE, Weitzman SH, Myers GS (1966) Phyletic studies of teleostean fishes, with a provisional classification of living forms. Bull Am Mus Nat Hist 131:339–456

Gutell RR, Gray MW, Schnare MN (1993) A compilation of large subunit (23S and 23S-like) ribosome RNA structures. Nucleic Acids Res 21:3055–3074

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Hoelzel AR, Lopez JV, Dover GA, O'Brien SJ (1994) Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. J Mol Evol 39:191–199

Hurst CD, Bartlett SE, Davidson WS, Bruce IJ (1999) The complete mitochondrial DNA sequence of the Atlantic salmon, *Salmo salar.* Gene 239:237–242

Inoue JG, Miya M, Tsukamoto K, Nishida M (2000) Complete mitochondrial DNA sequence of the Japanese sardine, *Sardinops melanostictus.* Fish Sci 66:924–932

Inoue JG, Miya M, Aoyama J, Ishikawa S, Tsukamoto K, Nishida M (2001) Complete mitochondrial DNA sequence of the Japanese eel, *Anguilla japonica.* Fish Sci (in press)

Janke A, Feldmaier-Fuchs G, Thomas WK, von Haeseler A, Pääbo S (1994) The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137:243–256

Johansen S, Bakke I (1996) The complete mitochondrial DNA sequence of Atlantic cod (*Gadus morhua*): Relevance to taxonomic studies among codfishes. Mol Mar Biol Biotechnol 5:203–214

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J Mol Evol 29:170–179

Kumazawa Y, Nishida M (1993) Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J Mol Evol 37:380–398

Kumazawa Y, Nishida M (1995) Variation in mitochondrial tRNA gene organization of reptiles as phylogenetic markers. Mol Biol Evol 12:759–772

Lee W-J, Kocher TD (1995) Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: Early establishment of the vertebrate genome organization. Genetics 139:873–887

Lee W-J, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. J Mol Evol 41: 54–66

Levinson G, Gutman GA (1987) Slipped-strand mispairing: A major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol Biol Evol 14:91–104

Maddison WP, Maddison DR (1992) MacClade version 3. Sinauer, Sunderland, MA

Mindell DP, Sorenson MD, Dimcheff DE (1998) Multiple independent origins of mitochondrial gene order in birds. Proc Natl Acad Sci USA 95:10693–10697

Miya M, Nishida M (1999) Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): First example of transter RNA gene rearrangements in bony fishes. Mar Biotechnol 1:416–426

Miya M, Nishida M (2000a) Molecular phylogeny of the deep-sea fish genus *Gonostoma* (Teleostei: Stomiiformes): Two paraphyletic clades and resurrection of *Sigmops*. Copeia 2000:378–389

Miya M, Nishida M (2000b) Use of mitogenic information in teleostean molecular phylogenetics: A tree-based exploration under the maximum-parsimony optimality criterion. Mol Phylogenet Evol 17:437–455

Moritz C, Brown WM (1986) Tandem duplications of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. Science 233:1425–1427

Murakami M, Yamashita Y, Fujitani H (1998) The complete sequence of mitochondrial genome from a gynogenetic triploid "ginbuna" (*Carassius auratus langsdorfi*). Zool Sci 15:335–337

Nelson GJ (1966) Gill arches of teleostean fishes of the order Anguilliformes. Pacific Sci 20:391–408

Nelson JS (1994) Fishes of the world, 3rd ed. John Wiley and Sons, New York

Noack K, Zardoya R, Meyer A (1996) The complete mitochondrial DNA sequence of the bichir (*Polypterus ornatipinnis*), a basal ray-finned fish: Ancient establishment of the consensus vertebrate gene order. Genetics 144:1165–1180

Pääbo S, Thomas WK, Whitfield KM, Kumazawa Y, Wilson AC (1991) Rearrangements of mitochondrial transfer RNA genes in marsupials. J Mol Evol 33:426–430

Quinn TW, Mindell DP (1996) Mitochondrial gene order adjacent to the control region in crocodile, turtle, and tuatara. Mol Phylogenet Evol 5:344–351

Quinn TW, Wilson AC (1993) Sequence evolution in and around the mitochondrial control region in birds. J Mol Evol 37:417–425

Robins CR (1989) The phylogenetic relationships of the anguilliform fishes. In: Böhlke EB (ed) Fishes of the western North Atlantic. Part 9. Vol 1. Orders Anguilliformes and Saccopharyngiformes. Sears Foundation for Marine Research, New Haven, CT, pp 9–23

Roe BA, Ma D-P, Wilson RK, Wong JF-H (1985) The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. J Biol Chem 260:9759–9774

Swofford DL (1993) PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Smithsonian Institute, Washington DC

Swofford DL (1998) PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0. Sinauer, Sunderland, MA

Templeton AR (1983) Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of human and the apes. Evolution 37:221–244

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

Tzeng C-S, Hui C-F, Shen S-C, Huan PC (1992) The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: Conservation and variations among vertebrates. Nucleic Acids Res 20:4853–4858

Walberg MW, Clayton DA (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res 9:5411–5421

Yoneyama Y (1987) The nucleotide sequence of the heavy and light strand replication origins of the *Rana catesbeiana* mitochondrial genome. Nippon Ika Daigaku Zasshi 54:429–440 (in Japanese)

Zardoya R, Meyer A (1996) The complete nucleotide sequence of the mitochondrial genome of the lungfish (*Protopterus dolli*) supports its phylogenetic position as a close relative of land vertebrates. Genetics 142:1249–1263

Zardoya R, Meyer A (1997) The complete DNA sequence of the mitochondrial genome of a "living fossil," the coelacanth (*Latimeria chalumnae*). Genetics 146:995–1010

Zardoya R, Garrido-Pertierra A, Bautista JM (1995) The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. J Mol Evol 41:942–951