Mar. Biotechnol. 1, 416–426, 1999

MARINE BIOTECHNOLOGY

Organization of the Mitochondrial Genome of a Deep-Sea Fish, *Gonostoma gracile* (Teleostei: Stomiiformes): First Example of Transfer RNA Gene Rearrangements in Bony Fishes

Masaki Miya[1,*] and Mutsumi Nishida[2]

[1]*Department of Zoology, Natural History Museum and Institute, Chiba, Aoba-cho, Chuo-ku, Chiba 260-8682, Japan*
[2]*Department of Marine Bioscience, Fukui Prefectural University, Obama, Fukui 917-0003, Japan*

Abstract: We determined the complete nucleotide sequence of the mitochondrial genome (except for a portion of the putative control region) for a deep-sea fish, *Gonostoma gracile*. The entire mitochondrial genome was purified by gene amplification using long polymerase chain reaction (long PCR), and the products were subsequently used as templates for PCR with 30 sets of newly designed, fish-universal primers that amplify contiguous, overlapping segments of the entire genome. Direct sequencing of the PCR products showed that the genome contained the same 37 mitochondrial structural genes as found in other vertebrates (two ribosomal RNA, 22 transfer RNA, and 13 protein-coding genes), with the order of all rRNA and protein-coding genes, and 19 tRNA genes being identical to that in typical vertebrates. The gene order of the three tRNAs ($tRNA^{Glu}$, $tRNA^{Thr}$, and $tRNA^{Pro}$) relative to cytochrome *b,* however, differed from that determined in other vertebrates. Two steps of tandem duplication of gene regions, each followed by deletions of genes, can be invoked as mechanisms generating such rearrangements of tRNAs. This is the first example of tRNA gene rearrangements in a bony fish mitochondrial genome.

Key words: long PCR, gene organization, transfer RNA, mitochondrial DNA, rearrangement

INTRODUCTION

Vertebrate mitochondrial gene order was initially considered conservative because the complete nucleotide sequences of the entire mitochondrial genome from 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 were subsequently 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), and marsupials (Pääbo et al., 1991; Janke et al., 1994), no such deviations were found among seven species of bony fish mitochondrial genomes that have been completely sequenced: loach (Tzeng et al., 1992), carp (Chang et al., 1994), trout (Zardoya et al., 1995), cod (Jo-

Received August 5, 1998; accepted February 19, 1999.
**Corresponding author; telephone +81-43-265-3111; fax +81-43-266-2481; e-mail miya@chiba-muse.or.jp*

hansen and Bakke, 1996), lungfish (Zardoya and Meyer, 1996), bichir (Noack et al., 1996), and coelacanth (Zardoya and Meyer, 1997).

During a series of molecular phylogenetic studies of the stomiiform fishes (Miya and Nishida, 1996, 1998, manuscript in preparation), we attempted to amplify the entire cytochrome *b* (cyt *b*) genes (approximately 1150 bp) from seven species of *Gonostoma* (family Gonostomatidae), using two primers designed on the two flanking transfer RNA genes ($tRNA^{Glu}$ and $tRNA^{Thr}$). The polymerase chain reaction (PCR) product from one species (*Gonostoma gracile*), however, was unexpectedly small (approximately 250 bp), and subsequent direct sequencing revealed that, instead of cyt *b,* $tRNA^{Pro}$ was found between these two tRNA genes. Further PCR and sequencing experiments demonstrated that the cluster of these three tRNAs ($tRNA^{Glu}$, $tRNA^{Pro}$, and $tRNA^{Thr}$) was adjacent to the 3′ end of the cyt *b* gene, suggesting that tRNA rearrangements relative to the latter have occurred in *G. gracile.* Such rearrangements have not been previously reported for any vertebrates (see Macey et al., 1997).

This article describes gene organization and tRNA rearrangements of the mitochondrial genome of *Gonostoma gracile,* a small deep-sea fish (<120 mm in standard length) endemic to the mesopelagic and bathypelagic zones (200–2000 m depth) of the western North Pacific (Kawaguchi, 1973). We used a long-PCR technique to purify the whole mitochondrial genome by gene amplification (Cheng et al., 1994a; Dowling et al., 1996), because it was difficult to obtain adequate amounts of appropriate tissue in good condition from such a small animal. We also designed 30 sets of primers (Table 1) based on highly conservative fish mitochondrial DNA regions, which were used for PCR and sequencing of contiguous, overlapping segments of the entire genome (Figure 1).

Materials and Methods

Fish Sample and DNA Extraction

A *Gonostoma gracile* specimen was collected from off the Pacific coast of Southern Japan and immediately preserved in 99.5% ethanol. Total genomic DNA was extracted from muscle tissue using the Qiagen QIAamp tissue kit following the manufacturer's protocol.

Mitochondrial DNA Purification by Long PCR

The entire mitochondrial genome of *Gonostoma gracile* was amplified using a long-PCR technique (Cheng et al., 1994a). Two sets of species-specific primers were designed in the previously determined partial sequences of the 16S rRNA and cyt *b* genes from *G. gracile* (Figure 1 and Table 1; Miya and Nishida, 1996; and unpublished data), so as to amplify the entire mitochondrial genome in two long-PCR reactions.

Long PCR was done in a Perkin-Elmer Model 2400 thermal cycler, and reactions were carried out with 30 cycles of a 25-μl reaction volume containing 8.25 μl of sterile distilled H_2O, 2.5 μl of 10 × LA PCR buffer II (TaKaRa), 4.0 μl dNTP (4 mM), 2.5 μl each primer (5 μM), 0.25 μl of 2.5 U LA *Taq* (TaKaRa), and 5 μl of template. The thermal cycle profile was that of "shuttle PCR": denaturation at 98°C for 10 seconds, and annnealing and extension combined at the same temperature (68°C) for 5 or 12 minutes depending on the target sequence length. Long-PCR products were electrophoresed on a 0.6% SeaKem Gold agarose gel and later stained with ethidium bromide for band characterization via ultraviolet transillumination. The long-PCR products were diluted with TE buffer (1:100) for subsequent use as PCR templates.

PCR and Sequencing

Thirty sets of fish-universal primers were designed (Table 1) with reference to the aligned, complete nucleotide sequences from the mitochondrial genome of six species of bony fishes (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; and lungfish, Zardoya and Meyer, 1996).

PCR was done in a Perkin-Elmer Model 2400 thermal cycler, and reactions were carried out with 30 cycles of a 25-μl reaction volume containing 14.0 μl of sterile, distilled H_2O, 2.5 μl of 10× PCR buffer II (Perkin-Elmer), 2.5 μl of dNTP (4 mM), 2.5 μl of each primer (5 μM), 0.1 μl of 4 U *Taq* DNA polymerase (AmpliTaq, Perkin Elmer), and 1 μl of template. The thermal cycle profile was as follows: denaturation at 94°C for 15 seconds, annealing at 45°C for 15 seconds, and extension at 72°C for 45 seconds. PCR products were electrophoresed on a 2.0% NuSieve agarose gel and stained with ethidium bromide for band characterization via ultraviolet transillumination.

Double-stranded DNA products from PCR, purified by filtration through a Qiagen QIAquick column, were used for direct cycle sequencing with dye-labeled terminators (Applied Biosystems Inc.). Primers used were the same as those for PCR. All sequencing reactions were performed

Table 1. PCR and Sequencing Primers Used in the Analysis of *Gonostoma gracile* Mitochondrial Genome

Primers	Sequence (5′ → 3′)*
Long PCR primers	
Gogr-16S-L	TCG ACG AAA GGG TTT ACG ACC TCG ATG TTG
Gogr-Cb-H	GGT AGG AGC CGT AGT AAA GCC CTC GTC CGA
Gogr-Cb-L	AAC ACC CTC CAA CAT CTC TGC CTG ATG AAA
Gogr-16S-H	GAC CTG GAT TAC TCC GGT CTG AAC TCA GAT
PCR and sequencing primers†	
1. L613-Phe	TTA ACG CAA AGC ATA GCA CTG AA
H1358-12S	CGA CGG CGG TAT ATA GGC
2. L1083-12S	ACA AAC TGG GAT TAG ATA C
H1903-16S	GTA GCT CGT YTA GTT TCG GG
3. L1803-16S	AGT ACC GCA AGG GAA AGC TGA AA
H2590-16S	ACA AGT GAT TGC GCT ACC TT
4. L2510-16S	CGC CTG TTT ACC AAA AAC AT
H3084-16S	AGA TAG AAA CTG ACC TGG AT
5. L2949-16S	GGG ATA ACA GCG CAA TC
H3718-ND1	ACT TCG TAT GAA ATW GTT TG
6. L3437-ND1	AAT GTK GTM GGM CCT TAC GG
H3934-ND1	GCG TAT TCT ACG TTG AAT CC
7. L3737-ND1	CAA ACW ATT TCS TAT GAA GT
H4427-Met	CCG WCA TGT TTG GGG TAT GGG CCC
8. L4166-ND1	CGA TAT GAT CAA CTM ATK CA
H4866-ND2	AAK GGK GCK AGT TTT TGT CA
9. L4633-ND2	CAC CAC CCW CGA GCA GTT GA
H5334-ND2	CGK AGG TAG AAG TAH AGG CT
10. L5260-ND2	CTG GST TTA TGC CMA ART G
H5934-CO1	GGG TGC CAA TGT CTT TGT GGT T
11. L5644-Ala	GCA AMT CAG ACA CTT TAA TTA A
H6371-CO1	TTG ATT GCC CCK AGG ATW GA
12. L6199-CO1	GCC TTC CCW CGA ATA AAT AA
H6855-CO1	AGT CAG CTG AAK ACT TTT AC
13. L6730-CO1	TAT ATA GGA ATR GTM TGA GC
H7480-Ser	ATG TGG YTG GCT TGA AA
14. L7255-CO1	GAT GCC TAC ACM CTG TGA AA
H8168-CO2	CCG CAG ATT TCW GAG CAT TG
15. L7863-CO2	ATA GAC GAA ATT AAT GAC CC
H8589-ATP	AAG CTT AKT GTC ATG GTC AGT
16. L8329-Lys	AGC GTT GGC CTT TTA AGC
H9076-ATP	GGG CGG ATA AAK AGG CTA AT
17. L8894-ATP	TTG GAC TAC TWC CST ATA C
H9639-CO3	CTG TGG TGA GCY CAK GT
18. L9514-CO3	TTC TGA GCC TTC TAY CA
H10019-Gly	CAA GAC KGK GTG ATT GGA AG
19. L9916-CO3	CAC CAT TTT GGC TTT GAA GC
H10433-Arg	AAC CAT GGW TTT TTG AGC CGA AAT
20. L10267-ND3	TTT GAY CTA GAA ATY GC
H10970-ND4	GAT TAT WAG KGG GAG WAG TCA

Table 1. Continued

Primers	Sequence (5′ → 3′)*
21. L10765-ND4	TTA AAT CTC CTM CAA TGT TA
H11534-ND4	GCT AGK GTA ATA AWK GGG TA
22. L11424-ND4	TGA CTT CCW AAA GCC CAT GTA GA
H12145-His	CTA GTG TTT TKG TTA AAC TA
23. L11895-ND4	CCT AAC CTW ATG GGR GAA CT
H12632-ND5	GAT CAG GTT ACG TAK AGK GC
24. L12329-Leu	CTC TTG GTG CAA MTC CAA GT
H13069-ND5	GTG CTG GAG TGK AGT AGG GC
25. L12942-ND5	GAA ATT CAA CAA ATM YTT
H13727-ND5	GCG ATK ATG CTT CCT CAG GC
26. L13562-ND5	CTW AAC GCC TGA GCC CT
H14080-ND5	AGG TAK GTT TTG ATT AKK CC
27. L13940-ND5	TTC TTT CCK ACT ATT ATW CAC CG
H14834-CYB	GAG CCA AAG TTT CAT CA
28. L14850-CYB	GCC TGA TGA AAC TTT GGC TC
H15560-CYB	TAG GCR AAT AGG AAR TAT CA
29. L15411-CYB	GAT AAA ATT YCA TTC CAC CC
H15957-Pro	GAG TTG AAG TCT CTT CAY TYT G
30. L15777-CYB	TGA ATT GGC GGC ATA CCW GTA GA
H690-12S	GCG GAG GCT TGC ATG TGT A

†Primers are designated by their 3′ ends, which correspond to the position of the human mitochondrial genome (Anderson et al., 1981) by convention. L and H denote heavy strand and light strand, respectively. For relative positions of primers in the mitochondrial genome, see Figure 1.
*Positions with mixed bases are labeled with their IUB codes: R indicates A or G; Y, C or T; K, G or T; M, A or C; S, G or C; W, A or T.

according to the manufacturer's instructions. Labeled fragments were analyzed on a Model 373S DNA sequencer (Applied Biosystems Inc.).

Sequence Analyses

DNA sequences were analyzed using the computer software package program DNASIS version 3.2 (Hitachi Software Engineering Co. Ltd.). The locations of 13 protein-coding genes were determined by comparisons of DNA or amino acid sequences of other bony fish mitochondrial genomes. The 22 tRNA genes were identified by their proposed cloverleaf secondary structures (Kumazawa and Nishida, 1993) and anticodon sequences. The two rRNA genes were identified by sequence homology and proposed secondary structures (Gutell et al., 1993).

Results and Discussion

Long PCR and Sequencing Strategy

Recent development of a long-PCR technique enabled us to amplify up to a 35-kb target sequence with high fidelity (Barnes, 1994; Cheng et al., 1994a). Although Cheng et al., (1994b) successfully amplified 16.3 kb of the 16.6-kb human mitochondrial genome in a single long PCR, our preliminary experiments demonstrated that such a single long PCR was not feasible for the *Gonostoma gracile* mitochondrial genome because of GC-rich regions in the putative control region and the 5′ half of the 16S rRNA gene. Strings of C and G bases often inhibit PCR reactions.

Alternatively, we decided to divide the circular mitochondrial genome into two segments that overlapped by approximately 400 bp (Figure 1): one long segment was expected to cover all protein-coding genes and most tRNA genes, spanning from near the 3′ end of the 16S rRNA to the 5′ end of the cyt *b* genes, and another short segment was expected to cover the two rRNA genes and the entire putative control region, spanning from the 5′ end of the cyt *b* to the 3′ end of the 16S rRNA genes. Two sets of species-specific primers were designed (Table 1) on the basis of previously determined, partial sequences of the 16S rRNA and cyt *b* genes (Miya and Nishida, 1996, and unpublished data). These primers successfully amplified the two seg-

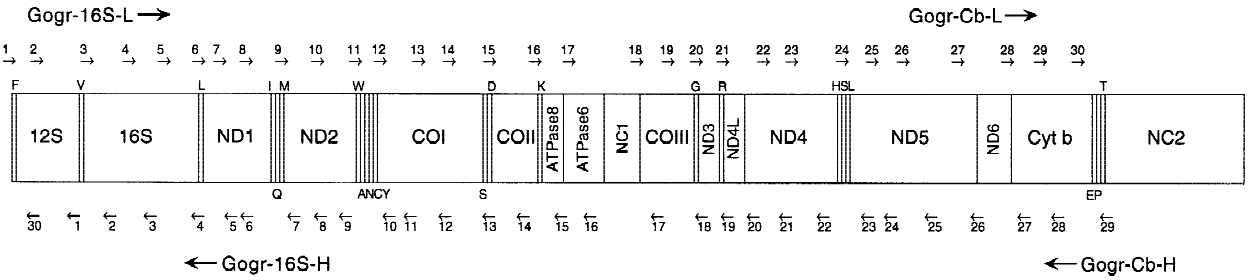

Figure 1. Gene organization and sequencing strategy for the *Gonostoma gracile* mitochondrial genome. All protein-coding genes are encoded by the H strand with the exception of ND6, which is coded by the L strand. Transfer RNA genes are designated by single-letter amino acid codes, those encoded by the H strand and L strand are shown above and below the gene map, respectively. Two pairs of long-PCR primers (Gogr-16S-L+Gogr-Cb-H and Gogr-Cb-L+Gogr-16S-H) amplify two segments of the entire mitochondrial genome, which overlap by approximately 400 bp in total. Relative positions of other primers are shown with numerals designated in Table 1. ND1–4L indicates NADH dehydrogenase subunits 1–4L; COI–III, cytochrome *c* oxidase subunits I–III; ATPase 6 and 8, ATPase subunits 6 and 8; Cyt *b,* cytochrome *b;* NC1 and NC2, major noncoding regions 1 and 2.

ments of the mitochondrial genome. Consequently, the mitochondrial genome of *G. gracile* was purified by gene amplification (Dowling et al., 1996), providing templates for subsequent amplifications and direct sequencings with contiguous, overlapping segments of the entire genome using 30 sets of newly designed, fish-universal primers (Figure 1; Table 1).

Although direct isolation of mitochondria from tissue is desirable (Zardoya and Meyer, 1997), a long-PCR approach should be useful for small, rare, or endangered species. Such an approach should also greatly reduce the possibility of amplification of mitochondrial pseudogenes in the nuclear genome (Dowling et al., 1996), which is a problem with direct PCR of the total genomic DNA.

Genome Content and Base Composition

The complete L-strand nucleotide sequence of *Gonostoma gracile,* except for a portion of the putative control region, is shown in Figure 2. The genome content of *G. gracile* included two rRNA, 22 tRNA, and 13 protein-coding genes, as found in other vertebrates, with two noncoding regions (Figures 1 and 2; Table 2). As in other vertebrates, most genes were encoded on the H strand, except for ND6 and eight tRNA genes, and all genes were similar in length to those in other bony fishes (Table 3).

Outside the putative control region (unable to be sequenced owing to numerous strings of C/G), the total length of the *G. gracile* mitochondrial genome was 16,144 bp, approximately 550 bp longer than those in other bony fishes (15,462–15,662 bp). As all of the genes were very similar in length to those in other bony fishes (Table 3), this difference was due apparently to the noncoding region (NC1) between ATPase 6 and COIII (304 bp), and intergenic spacers around the cyt *b* gene (239 bp in total) where tRNA rearrangements occurred (see below). Considering the exceptionally long putative control region (NC2, approximately 3 kb estimated from the PCR product) compared with those in other bony fishes (896–1184 bp), the *G. gracile* mitochondrial genome is not as economical as in other bony fishes. Such a long putative control region, however, has been observed among other stomiiform species, such as *Cyclothone alba* and *Chauliodus sloani* (M. Miya and M. Nishida, unpublished observations), suggesting that it is a common characteristic among stomiiform fishes.

The base composition of *G. gracile* was analyzed separately for the rRNA, tRNA, and protein-coding genes (Table 4). In the protein-coding gene, strong anti-G bias was observed in the third codon positions (9.6%) and pyrimidines were overrepresented in the second codon positions (68.4%), as has been noted for other vertebrate mitochondrial genomes, owing to the hydrophobic character of the proteins (Naylor et al., 1995). *Gonostoma gracile* tRNA genes were A + T-rich (61.6%), as in other vertebrates, while rRNA genes were slightly A + C-rich (52.9%), as in other bony fishes (Zardoya and Meyer, 1997).

Gene Order

The gene order of *G. gracile* was largely identical to that of typical vertebrates, although the positions of three tRNA

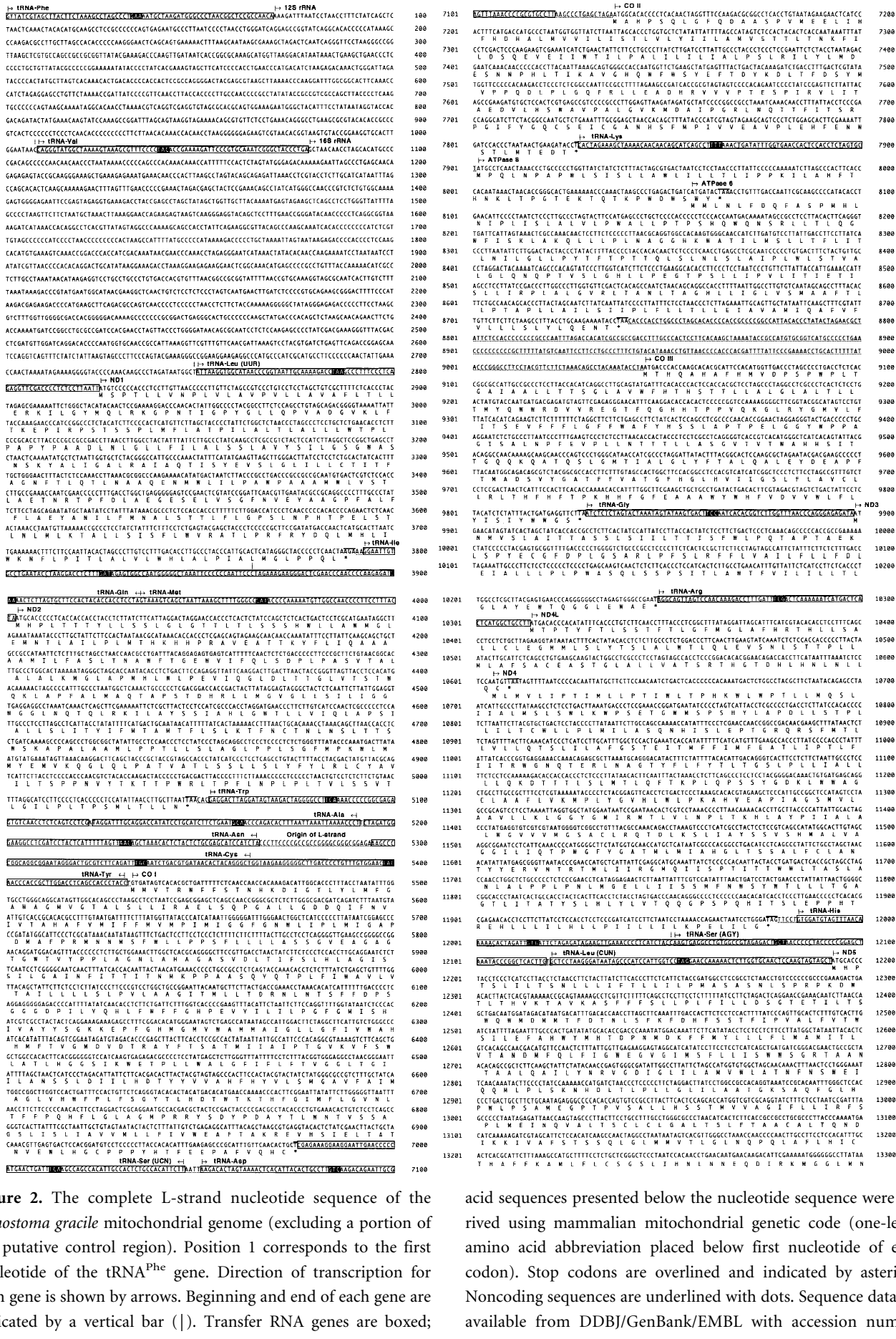

Figure 2. The complete L-strand nucleotide sequence of the *Gonostoma gracile* mitochondrial genome (excluding a portion of the putative control region). Position 1 corresponds to the first nucleotide of the tRNA^{Phe} gene. Direction of transcription for each gene is shown by arrows. Beginning and end of each gene are indicated by a vertical bar (|). Transfer RNA genes are boxed; corresponding anticodons are indicated in black boxes. Amino acid sequences presented below the nucleotide sequence were derived using mammalian mitochondrial genetic code (one-letter amino acid abbreviation placed below first nucleotide of each codon). Stop codons are overlined and indicated by asterisks. Noncoding sequences are underlined with dots. Sequence data are available from DDBJ/GenBank/EMBL with accession number AB016274.

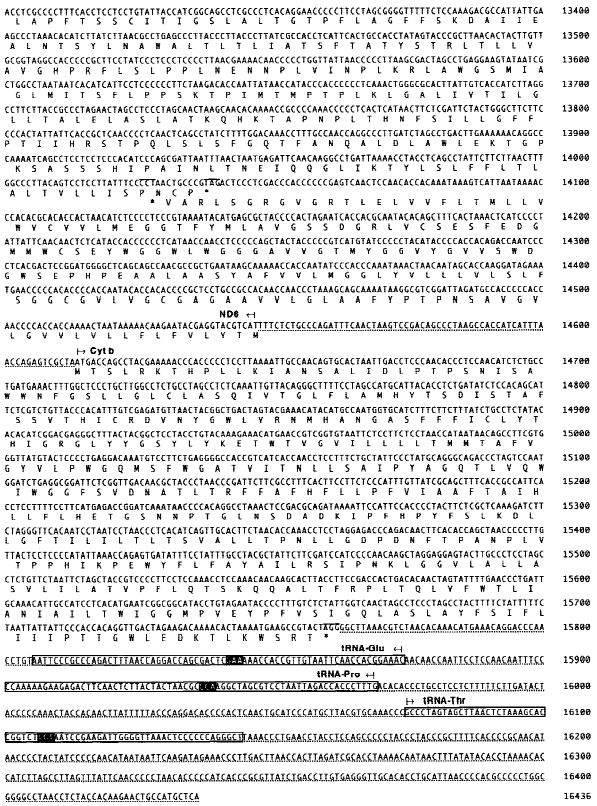

Figure 2. *Continued*

genes (tRNA^{Glu}, tRNA^{Thr}, and tRNA^{Pro}) relative to the cyt *b* gene were unique among known vertebrate gene orders. In typical vertebrates, cyt *b* is flanked by tRNA^{Glu} at the 5′ end and a cluster of tRNA^{Thr} and tRNA^{Pro} at the 3′ end (Figure 3A). In *G. gracile,* however, all three tRNA genes were adjacent to the 3′ end of cyt *b,* in a cluster in which tRNA^{Glu} preceded the other two tRNA genes, with a further switch in the order of tRNA^{Thr} and tRNA^{Pro}.

Gene rearrangements have been proposed as occurring by tandem duplication of gene regions as a result of slipped-strand mispairing, followed by deletions of genes (Levinson and Gutman, 1987; Moritz and Brown, 1987). The present gene order and associated intergenic spacers could have resulted from a double occurrence of this process (Figure 3): the first tandem duplication occurring in the tRNA^{Thr}–tRNA^{Pro} region (and a portion of the control region [CR]), followed by deletions of redundant genes (Figure 3B), and the second tandem duplication occurring in the tRNA^{Glu}–cyt *b* segment (plus an intergenic spacer between cyt *b* and tRNA^{Pro}), again followed by deletions of redundant genes (Figure 3C). This two-step tandem duplication and subsequent deletions apparently resulted in the observed gene order and associated intergenic spacers in *G. gracile* (Figure 3C).

Table 2. Location of Genes in the Mitochondrial Genome of *Gonostoma gracile*

Gene	Position number From	To	Size (bp)	Codon Start	Stop
tRNA^{Phe}	1	69	69		
12S rRNA	70	1008	939		
tRNA^{Val}	1009	1078	70		
16S rRNA	1079	2749	1671		
tRNA$^{Leu(UUR)}$	2750	2824	75		
ND1	2825	3790	966	ATG	AGA
tRNA^{Ile}	3793	3864	72		
tRNA^{Gln}	3864	3934	71(L)		
tRNA^{Met}	3934	4002	69		
ND2	4003	5049	1047	ATG	TAA
tRNA^{Trp}	5053	5121	69		
tRNA^{Ala}	5123	5191	69 (L)		
tRNA^{Asn}	5193	5265	73 (L)		
tRNA^{Cys}	5295	5363	69 (L)		
tRNA^{Tyr}	5363	5431	69 (L)		
COI	5433	6975	1543	GTG	T—
tRNA$^{Ser\ (UCN)}$	6976	7046	71 (L)		
tRNA^{Asp}	7051	7122	72		
COII	7137	7827	691	ATG	T—
tRNA^{Lys}	7828	7900	73		
ATPase 8	7902	8066	165	ATG	TAA
ATPase 6	8057	8740	684	ATG	TAA
COIII	9045	9829	785	ATG	TA–
tRNA^{Gly}	9830	9898	69		
ND3	9899	10247	349	ATG	T—
tRNA^{Arg}	10248	10314	67		
ND4L	10315	10611	297	ATG	TAA
ND4	10605	11978	1374	ATG	TAG
tRNA^{His}	11984	12051	68		
tRNA$^{Ser(AGY)}$	12052	12118	67		
tRNA$^{Leu(CUN)}$	12120	12192	73		
ND5	12193	14040	1848	ATG	TAG
ND6	14026	14547	522 (L)	ATG	AGG
Cyt*b*	14614	15762	1149	ATG	AGG
tRNA^{Glu}	15806	15874	69 (L)		
tRNA^{Pro}	15901	15969	69 (L)		
tRNA^{Thr}	15975	16144	70		

Multiple deletions of redundant genes seemed to be incomplete in *G. gracile,* because four stretches of short noncoding sequences (26–104 nucleotides) occurred around the genes involved in the tRNA rearrangments (Figure 2). Although no homologous regions were identified for

Table 3. Comparisons of Lengths (bp) of Bony Fish Mitochondrial Genes

Gene	*G. gracile*	Cod	Trout	Carp	Loach	Bichir	Lungfish	Coelacanth
12S rRNA	939	950	944	951	937	950	933	983
16S rRNA	1671	1669	1680	1681	1680	1655	1591	1665
ND1	966	972	972	975	975	958	966	972
ND2	1047	1047	1050	1047	1047	1036	1028	1047
COI	1543	1551	1551	1551	1551	1557	1548	1548
COII	691	691	691	691	691	688	691	691
ATPase 8	165	167	168	165	168	168	168	168
ATPase 6	684	684	670	684	684	682	682	662
COIII	786	786	784	786	768	784	784	786
ND3	349	351	349	351	351	346	346	349
ND4L	297	297	297	297	297	297	297	297
ND4	1374	1381	1381	1383	1383	1378	1384	1381
ND5	1848	1839	1839	1824	1837	1842	1836	1836
ND6	522	523	522	519	522	504	513	519
Cyt *b*	1149	1141	1141	1141	1144	1141	1144	1142
Control region	≈3,000	997	1003	927	896	1068	1184	781
Total*	≈19,000	16,696	16,642	16,575	16,558	16,624	16,646	16,407

*Total lengths of mitochondrial genome including intergenic spacers.

Table 4. Base Composition of the Mitochondrial Genome of *Gonostoma gracile*

	A	C	G	T
Proteins				
1st	25.7	29.1	24.5	20.6
2nd	18.0	29.3	13.6	39.1
3rd	28.8	39.0	9.6	22.5
Total	24.2	32.5	15.9	27.4
tRNAs	26.0	23.4	23.8	26.9
rRNAs	31.4	30.2	19.6	18.8

three of these sequences, the remaining sequence (66 nucleotides), located adjacent to the 5′ end of cyt *b,* showed some similarity (63.9%) to that of $tRNA^{Glu}$ within the aligned sequences (Figure 4). Because this noncoding sequence was located at the original position of $tRNA^{Glu}$ and apparently does not form a stable secondary structure common among tRNA genes, it was likely to be a pseudogene of $tRNA^{Glu}$ that had not been completely deleted after tandem duplication of the $tRNA^{Glu}$–cyt *b* region (Figure 3B). This observation supported the concept of the most recent tandem duplication and subsequent deletion events having occurred in the $tRNA^{Glu}$–cyt *b* region (Figure 3C), rather than in the $tRNA^{Thr}$–$tRNA^{Pro}$ region (Figure 3B).

Protein-Coding Genes

Among 13 protein-coding genes of *G. gracile,* there were two reading-frame overlaps on the same strand (ATPases 8 and 6 shared 10 nucleotides; ND4L and ND4 shared 7 nucleotides) (Figure 2). As in other bony fishes, all the mitochondrial protein-coding genes began with an ATG start codon, except COI, which started with GTG (Table 2). Open reading frames of *G. gracile* ended with TAA (ND2, ATPase 8, ATPase 6, and ND4L), AGA (ND1), TAG (ND4 and ND5), and AGG (ND6 and cyt *b*), and the remainder had incomplete stop codons, either T (COI, COII, and ND3) or TA (COIII) (Table 2).

Transfer RNA Genes

The *G. gracile* mitochondrial genome contained 22 tRNA genes interspersed between the rRNA and protein-coding genes (Figures 1 and 2). The tRNA genes range in size from 67 to 75 nucleotides (Table 2), large enough so that the encoded tRNAs can fold into the cloverleaf secondary structure characteristic of tRNAs (data not shown). This is pos-

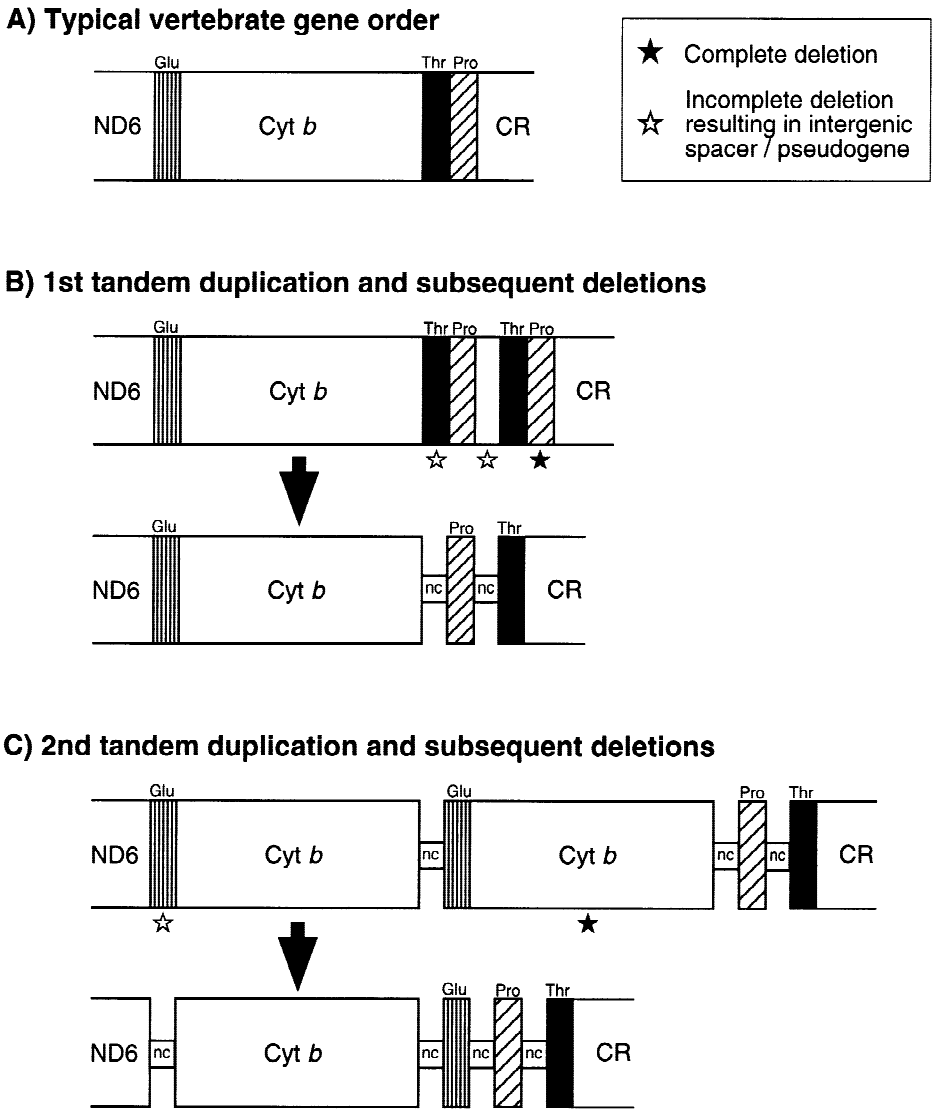

Figure 3. Proposed mechanism of tRNA gene rearrangements in *Gonostoma gracile* under a model of tandem duplication of gene regions and subsequent gene deletions. **A:** Typical vertebrate gene order of $tRNA^{Glu}$, cyt *b*, $tRNA^{Thr}$, and $tRNA^{Pro}$. **B:** First tandem duplication in the $tRNA^{Thr}$–$tRNA^{Pro}$ region (plus a portion of the control region [CR]) and subsequent deletions of redundant genes. **C:** Second tandem duplication in the $tRNA^{Glu}$–cyt *b* region (plus an intergenic spacer between cyt *b* and $tRNA^{Pro}$) and subsequent deletions of redundant genes, resulting in the observed gene order in *G. gracile*. Note that the four observed intergenic spacers around the four genes (nc indicates noncoding region; see also Figure 2) can result from incomplete deletion events. Also note that the intergenic spacer between ND6 and cyt *b* appears to be a functionless copy (or pseudogene) of $tRNA^{Glu}$, resulting from the most recent tandem duplication and subsequent deletion (see text and Figure 4).

sible provided that formation of the G-U wobble and other atypical pairings were allowed in the stem regions. All postulated cloverleaf structures, with the exception of $tRNA^{Ser(AGY)}$, contain 7 bp in the amino acid stem, 5 bp in the TYC stem, 5 bp in the anticodon stem, and 4 bp in the DHU stem. The $tRNA^{Ser(AGY)}$ of *G. gracile* had no recognizable DHU stem and loop, as found in the bichir (Noack et al., 1996) and lamprey (Lee and Kocher, 1995).

Ribosomal RNA Genes

The 12S and 16S ribosomal RNA genes of *G. gracile* were 939 and 1671 nucleotides long, respectively (Table 2). They were located, as in other vertebrates, between $tRNA^{Phe}$ and $tRNA^{Leu}$, being separated by $tRNA^{Val}$ (Figures 1 and 2). Preliminary assessment of the secondary structure of *G. gracile* indicated that the present sequence could be reason-

```
tRNA^Glu     TTCCC-GCCCAGACTTTAACCAGGACCAGCGACTCGAAAAACCACCGTTGTAATTCAACCA
Pseudogene  ...T.T.......T..C...T.A.T..GA.AG.C.T..--G.....A.C---...T.....
```

Figure 4. Aligned DNA sequences of a portion of $tRNA^{Glu}$ and its putative pseudogene located between ND6 and cyt *b* from *Gonostoma gracile* (see Figure 2). Identity with the first sequence is denoted by dots. Insertions/deletions of specific nucleotides are indicated by dashes (–).

ably superimposed on the proposed secondary structure of carp 12S rRNA and cow 16S rRNA (Gutell et al., 1993).

Noncoding Sequences

As in most vertebrates, the origin of light strand replication (O_L) in *G. gracile* was in a cluster of five tRNA genes (WANCY region, Figure 2) and comprised 44 nucleotides in length. This region has the potential to fold into a stable stem-loop secondary structure with 10 bp in the stem (including one bulge) and nine nucleotides in the loop. The conserved motif 5′-GCCGG-3′ was found at the base of the stem within the $tRNA^{Cys}$ gene.

Although two major noncoding regions were identified in *G. gracile* (NC1 and NC2), the control region could not be recognized within those regions sequenced (Figure 2). NC1, located between ATPase 6 and COIII, and 304 bp long, had no characteristics of the control region, such as a conserved sequence block (CSBs; Walberg and Clayton, 1981) or termination-associated sequence (TASs; Doda et al., 1981). NC2, located between $tRNA^{Pro}$ and $tRNA^{Phe}$, and being approximately 3 kb long, also had no such characteristics within the region sequenced.

Acknowledgments

We thank T. Komai for collection of the specimen of *Gonostoma gracile* and M. Yamaguchi and A. Kawaguchi for assistance in DNA sequencing. This work was supported in part by grants from the Ministry of Education, Science, Sports and Culture of Japan (No. 09740644) to M.M. and (No. 10660189) to M.N.

References

Anderson, S. Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, C.D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R., and Young, I.G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.

Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F., and Young, I.G. (1982). Complete sequence of bovine mitochondrial DNA: conserved features of the mammalian mitochondrial genome. *J Mol Biol* 156:683–717.

Barnes, W.M. (1994). PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc Natl Acad Sci USA* 91:2216–2220.

Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W., and Clayton, D.A. (1981). Sequence and gene organization of mouse mitochondrial DNA. *Cell* 26:167–180.

Chang, Y.-S., Huang, F.-I., and 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., Chang, S.-Y., Gravitt, P., and Respess, R. (1994a). Long PCR. *Nature* 369:684–685.

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

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

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

Doda, J.N., Wright, C.T., and Clayton, D.A. (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.

Dowling, T.E., Moritz, C., Palmer, J.D., and Rieseberg, L.H. (1996). Nucleic acids, III: analysis of fragments and restriction sites. In: *Molecular Systematics, 2nd ed.,* Hillis, D.M., Moritz, C., and Mable, B.K. (eds.). Sunderland, Mass.: Sinauer Associates, 249–320.

Gutell, R.R., Gray, M.W., and Schnare, M.N. (1993). A compilation of large subunit (23S and 23S-like) ribosomal RNA structures: 1993. *Nucleic Acids Res* 21:3055–3074.

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

Johansen, S., and 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.

Kawaguchi, K. (1973). Biology of *Gonostoma gracile* Günther (Gonostomatidae), II: geographical and vertical distribution. *J Oceanogr Soc Jpn* 29:113–120.

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

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

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

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

Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., and Papenfuss, T.J. (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.

Miya, M., and Nishida, M. (1996). Molecular phylogenetic perspective on the evolution of the deep-sea fish genus *Cyclothone* (Stomiiformes: Gonostomatidae). *Ichthyol Res* 43:375–398.

Miya, M., and Nishida, M. (1998). Molecular phylogeny and evolution of the deep-sea fish genus. *Sternoptyx. Mol Phylogenet Evol* 10:11–22.

Moritz C., and Brown, W.M. (1987). Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. *Proc Natl Acad Sci USA* 84:7183–7187.

Naylor, G.J., Collins, T.M., and Brown, W.M. (1996). Hydrophobicity and phylogeny. *Nature* 373:555–556.

Noack, K., Zardoya, R., and 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, W.K., Whitfield, K.M., Kumazawa, Y., and Wilson, A.C. (1991). Rearrangements of mitochondrial transfer RNA genes in marsupials. *J Mol Evol* 33:426–430.

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

Quinn, T.W., and Wilson, A.C. (1993). Sequence evolution in and around the mitochondrial control region in birds. *J Mol Evol* 37: 417–425.

Roe, B.A., Ma, D.-P., Wilson, R.K., and Wong, J.F.-H. (1985). The complete sequence of the *Xenopus laevis* mitochondrial genome. *J Biol Chem* 260:9759–9774.

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

Walberg, M.W., and Clayton, D.A. (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.

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