J Mol Evol (1995) 41:942–951

JOURNAL OF MOLECULAR EVOLUTION

The Complete Nucleotide Sequence of the Mitochondrial DNA Genome of the Rainbow Trout, *Oncorhynchus mykiss*

Rafael Zardoya, Amando Garrido-Pertierra, José M. Bautista

Departamento de Bioquímica y Biología Molecular IV, Universidad Complutense de Madrid, Facultad de Veterinaria, Cuidad Universitaria, 28040 Madrid, Spain

Received: 22 November 1994 / Revised: 26 January 1995

Abstract. The complete nucleotide sequence of the mitochondrial DNA of the rainbow trout, *Onchorynchus mykiss,* has been determined. The total length of the molecule is 16,660 bp. The rainbow trout mitochondrial DNA has the same organization described in eutherian mammals, the clawed frog (*Xenopus laevis*), and the two fish species, Oriental stream loach (*Crossotoma lacustre*) and carp (*Cyprinus carpio*). Alignment and comparison of the deduced amino acid sequences of the 13 proteins encoded by rainbow trout and other vertebrate mitochondrial genomes allowed us to estimate that COI is the most conserved mitochondrial subunit (amino acid identity ranging from 85.6% to 94.8%) whereas ATPase 8 is the most variable one (amino acid identity ranging from 30.8% to 70.4%). Putative secondary structures for the 22 tRNAs found in the molecule are given along with an extensive comparison of tRNA sequences among representative species of each major group of vertebrates. In this sense, an unusual cloverleaf structure for the $tRNA^{Ser(AGY)}$ is proposed. A stem-loop structure inferred for the origin of the L-strand replication (O_L) and the presence of a large polycytidine tract in the O_L loop is described. The existence of this stretch instead of the usual T-rich sequence reported so far in mammal mtDNAs is explained in terms of a less-strict template dependence of the RNA primase involved in the initiation of L-strand replication.

Key words: Mitochondrial DNA — Rainbow trout — Molecular phylogeny

Correspondence to: J.M. Bautista

Introduction

The organization of vertebrate mitochondrial genomes is highly conserved and extremely compact (Attardi 1985). A total of 22 tRNAs, two rRNAs, and 13 protein genes that are required for mitochondrial function (Chomyn et al. 1985) are coded in a covalently closed circular molecule of merely 15–17 kb. Despite its highly conserved gene organization, mtDNAs evolve very rapidly, accumulating differences in their nucleotide sequences, and therefore are particularly valuable for estimating genetic distances among species and for other phylogenetic studies (Brown et al. 1979). The complete mtDNA sequences of 14 vertebrate species have been reported. Most of them are mammals [human (Anderson et al. 1981) mouse (Bibb et al. 1981), cow (Anderson et al. 1982), rat (Gadaleta et al. 1989), fin whale (Árnason et al. 1991), harbor seal (Árnason and Johnsson 1992), grey seal (Árnason et al. 1993), blue whale (Árnason and Gullberg 1993), opossum (Janke et al. 1994), and horse (Xu and Árnason 1994)]; the others are an amphibian [clawed frog (Roe et al. 1985)], two fish [oriental stream loach (Tzeng et al. 1992) and carp (Chang and Huang 1994)], and a bird [chicken (Desjardins and Morais 1990)].

Partial mtDNA sequences of the rainbow trout including a *Xba*I fragment 1,300 bp long (Davidson et al. 1988), its largely overlapping 2,200-bp *Hin*dIII fragment (Thomas and Beckenbach 1989), and the control region (Digby et al. 1992) have been previously reported. Furthermore, partial sequences of some other fish mtDNAs such as salmon (Thomas and Beckenbach 1989), cod (Johansen et al. 1990; Árnason and Rand 1992), and white sturgeon (Buroker et al. 1990) have been de-

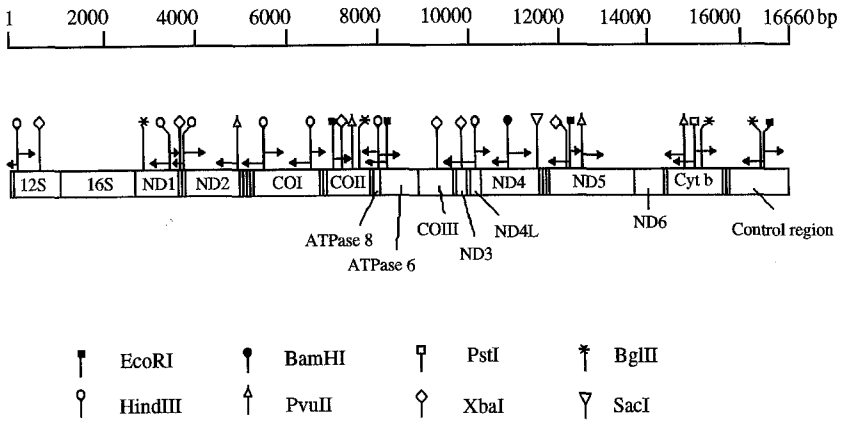

Fig. 1. Restriction map and sequencing strategy of the rainbow trout mitochondrial genome. The cleavage map shows representative restriction sites together with their relative position within rainbow trout mitochondrial genes. Direction of sequencing of the clones is denoted by *arrows.*

scribed. The sequence data of fish, however, are limited compared with those of mammals. In fact, the only two fish complete mitochondrial genomes reported correspond to Cyprinoidei species originally from Asia. Moreover, paucity of data among lower vertebrates has limited the understanding of the mechanisms involved in the evolution from the variable arrangements found in the invertebrate mitochondrial genomes (Hoffmann et al. 1992) to the more constrained organization shared by chordata.

In the present work, the complete nucleotide sequence of the rainbow trout (*Oncorhynchus mykiss*) mitochondrial genome has been obtained. This freshwater species originally from the Pacific coast of North America is widely spread to the rest of North America and Europe due to its particularly good adaptation to culture conditions. Furthermore, rainbow trout is considered as a fish model in nutritional and genetic research studies. Extensive comparative studies carried out with the three other lower vertebrate mtDNAs available—i.e., *Xenopus laevis, Crossostoma lacustre,* and *Cyprinus carpio*—reveal some new interesting features not found in other vertebrates yet sequenced.

Materials and Methods

Mitochondrial DNA Extraction. Rainbow trout (*Oncorhynchus mykiss*) fresh liver samples were minced with scissors and then homogenized in MSB buffer (200 mM mannitol, 70 mM sucrose, and 50 mM Tris-HCl, pH 7.5) by 3–4 strokes in a tight-fitting motor-driven glass/Teflon homogenizer. Intact nuclei and cellular debris were removed by a centrifugation at 1,000*g* for 5 min in a swinging bucket rotor. The recovered supernatant was subjected to a new round of centrifugation. The second low-speed supernatant was subsequently centrifuged at 10,000*g* for 20 min. The mitochondria-containing pellet was washed by performing four cycles of resuspension in 10–20 ml of MSB and centrifugation at 10,000*g* for 20 min. All previously described operations were carried at 4°C. In order to isolate mtDNA, the resulting mitochondrial pellet was subjected to a standard alkaline lysis procedure followed by a phenol/chloroform extraction (Palva and Palva 1985).

Cloning and Sequencing. mtDNA was cleaved with several restriction enzymes. *Eco*RI, *Hin*dIII, and *Pvu*II digestions yield the most suitable fragments in size for cloning. Two *Eco*RI fragments of 4 and 0.8 kb, six *Hin*dIII fragments of 3.5, 2.2, 1.8, 1.2, 1.1, and 0.25 kb, a *Eco*RI-*Bam*HI fragment of 1.2 kb, and two *Pvu*II fragments of 2.5 and 2.3 kb were cloned into pUC 18. Clones were obtained covering the whole rainbow trout mtDNA molecule except for a 700-bp fragment which was PCR amplified and subsequently cloned in a pGEM-T vector (Promega).

Sequencing was performed applying the dideoxy chain-termination technique (Sanger et al. 1977) on an automatic sequencer (A.L.F. from Pharmacia). Sequence was obtained from both strands using both universal and 31 specific oligonucleotide fluorescein-labeled primers according to the strategy depicted in Fig. 1. Sequence data were analyzed by use of the GCG program package (Devereux et al. 1984). The complete mtDNA sequence of the rainbow trout has been deposited at the EMBL/GenBank data libraries under the accession number L29771.

Results and Discussion

Genome Organization

A restriction map of the mtDNA of rainbow trout is given in Fig. 1. The restriction pattern obtained for *Hin*dIII and *Bgl*II cleavages revealed that the specimens used to clone mtDNA corresponded to the stock designated as "A" by Palva and Palva (1987).

The total length of the mitochondrial molecule was 16,660 bp and contained the two rRNA genes, 22 tRNA genes, and 13 peptide-encoding genes found so far in other vertebrate mtDNAs (Table 1). The major coding strand of rainbow trout mtDNA is the heavy strand and all the peptide-encoding genes are punctuated by one or more tRNA genes (Montoya et al. 1983). The relative position and the orientation of all genes and the control region are identical to that found in eutherian mammals, *X. laevis* (Roe et al. 1985), *C. lacustre* (Tzeng et al. 1992), and *C. carpio* (Chang and Huang 1994). The L-strand base composition is A, 27.9%; C, 28.9%, T, 26.2%; and G, 17.0%. Peptide-encoding genes were identified by comparison with *X. laevis* mtDNA (Roe et al. 1985) and by the presence of initiation and stop codons. Sequences encoding tRNA genes were recognized by the nucleotide sequence capability to fold into

Table 1. Localization of features in the mitochondrial genome of the rainbow trout, *Oncorhychus mykiss*

Feature	From	To	Size (bp)	Codon Start	Codon Stop
tRNA-Phe	1	68	68		
12S rRNA	69	1012	944		
tRNA-Val	1013	1084	70		
16S rRNA	1085	2764	1,680		
tRNA-Leu (UUR)	2765	2840	76		
NADH 1	2841	3812	972	ATG	TAG
tRNA-Ile	3816	3887	72		
tRNA-Gln	3955	3885	71 (L)		
tRNA-Met	3955	4023	69		
NADH 2	4024	5073	1,050	ATG	TAA
tRNA-Trp	5076	5144	69		
tRNA-Ala	5214	5146	69 (L)		
tRNA-Asn	5288	5216	73 (L)		
OriL	5289	5325	37		
tRNA-Cys	5392	5326	67 (L)		
tRNA-Tyr	5463	5393	71 (L)		
CO I	5465	7015	1,551	GTG	TAA
tRNA-Ser (UCN)	7086	7016	71 (L)		
tRNA-Asp	7091	7163	73		
CO II	7178	7868	691	ATG	T—
tRNA-Lys	7869	7942	74		
ATPase 8	7944	8111	168	ATG	TAA
ATPase 6	8102	8771	670	ATG	TA–
CO III	8773	9556	784	ATG	TA–
tRNA-Gly	9558	9627	70		
NADH 3	9628	9976	349	ATG	TA–
tRNA-Arg	9978	10,046	69		
NADH 4L	10,047	10,343	297	ATG	TAA
NADH 4	10,337	11,717	1,381	ATG	T—
tRNA-His	11,718	11,786	69		
tRNA-Ser (AGY)	11,787	11,855	69		
tRNA-Leu (CUN)	11,857	11,928	72		
NADH 5	11,929	13,767	1,839	ATG	TAA
NADH 6	14,303	13,764	540 (L)	ATG	TAG
tRNA-Glu	14,372	14,304	69 (L)		
Cyt b	14,376	15,516	1,141	ATG	T—
tRNA-Thr	15,517	15,588	72		
tRNA-Pro	15,657	15,588	70 (L)		
Control region	15,658	16,660	1,003		

[a] Gene nomenclature according to Attardi et al. (1986). L denotes light-strand sense

putative cloverleaf structures including their specific anticodons, by analogy with tRNA boundaries previously reported (Kumazawa and Nishida 1993) and by comparison with the *X. laevis* homologues.

Genetic Code and Codon Usage

All initiation codons in rainbow trout mtDNA protein-encoding genes are ATG except that of the COI gene, which is GTG (Table 1). Such initiation codon usage is shared by the two other fish mitochondrial genomes already completely sequenced (Tzeng et al. 1992; Chang and Huang 1994). Moreover, among vertebrates, chicken mtDNA (Desjardins and Morais 1990) also shares this condition.

Table 2. Percentage amino acid identity of four rainbow trout mitochondrial polypeptides against their homologues of representative species of vertebrates[a]

	COI	COII	ATPase 8	Cyt b
Cyprinus carpio	94.8	93.5	70.4	90.2
Crossostoma lacustre	93.8	90.9	65.4	90.5
Xenopus laevis	89.7	81.2	74.5	76.5
Gallus gallus	88.3	68.3	51.8	74.5
Homo sapiens	85.6	65.6	30.8	70.6

[a] Amino acid identities were calculated using the GAP program of the GCG package with a gap weight set at 3.0 and a gap length weight set at 0.1

Table 3. Pairwise comparison of mitochondrial rRNA sequences among fish species[a]

	Rainbow trout	Carp	Loach
Rainbow trout	—	167/69	161/70
Carp	333/137	—	97/35
Loach	368/155	240/96	—

[a] The observed nucleotide differences are given as total substitutions/transversions. Data corresponding 12S rRNA and 16S rRNA are shown above and below diagonal, respectively

On the other hand, five ORFs (ND2, COI, ATPase 8, ND4L, and ND5) end with TAA as stop codon while two (ND1 and ND6) end with TAG. The remainder of protein-encoding genes present incomplete stop codons, either T (COII, ND4, and Cyt b) or TA (ATPase 6, COIII, and ND3). By analogy with *X. laevis,* entire stop codons are probably generated by cleavage of the primary transcript and subsequent polyadenylation of the corresponding mRNAs (Roe et al. 1985). Neither AGR is used as a stop codon in rainbow trout, nor does either even appear internally in any of the rainbow trout protein-encoding genes. Apparently, the stop codon pattern of the rainbow trout is not shared by any other vertebrate.

The codon usage is similar to that inferred for cod (Johansen et al. 1990), *C. lacustre* (Tzeng et al. 1992), and carp (Chang and Huang 1994), with a clear bias to exclude guanosine in the third codon position. It is significant that ATG is mostly used as start codon whereas internal methionines are preferentially encoded by ATA. The amino acid tryptophan is mostly encoded by TGA instead of TGG, denoting the specificity of mitochondrial over nuclear genetic code.

Protein-Encoding Genes

The rainbow trout mitochondrial genome contains 13 large open reading frames and, as in other vertebrate mtDNAs, there are two cases of reading frame overlap. (ATPases 8 and 6 overlap in ten nucleotides; ND4L and ND4 share seven nucleotides.) These gene pairs are translated in two different reading frames shifted by one

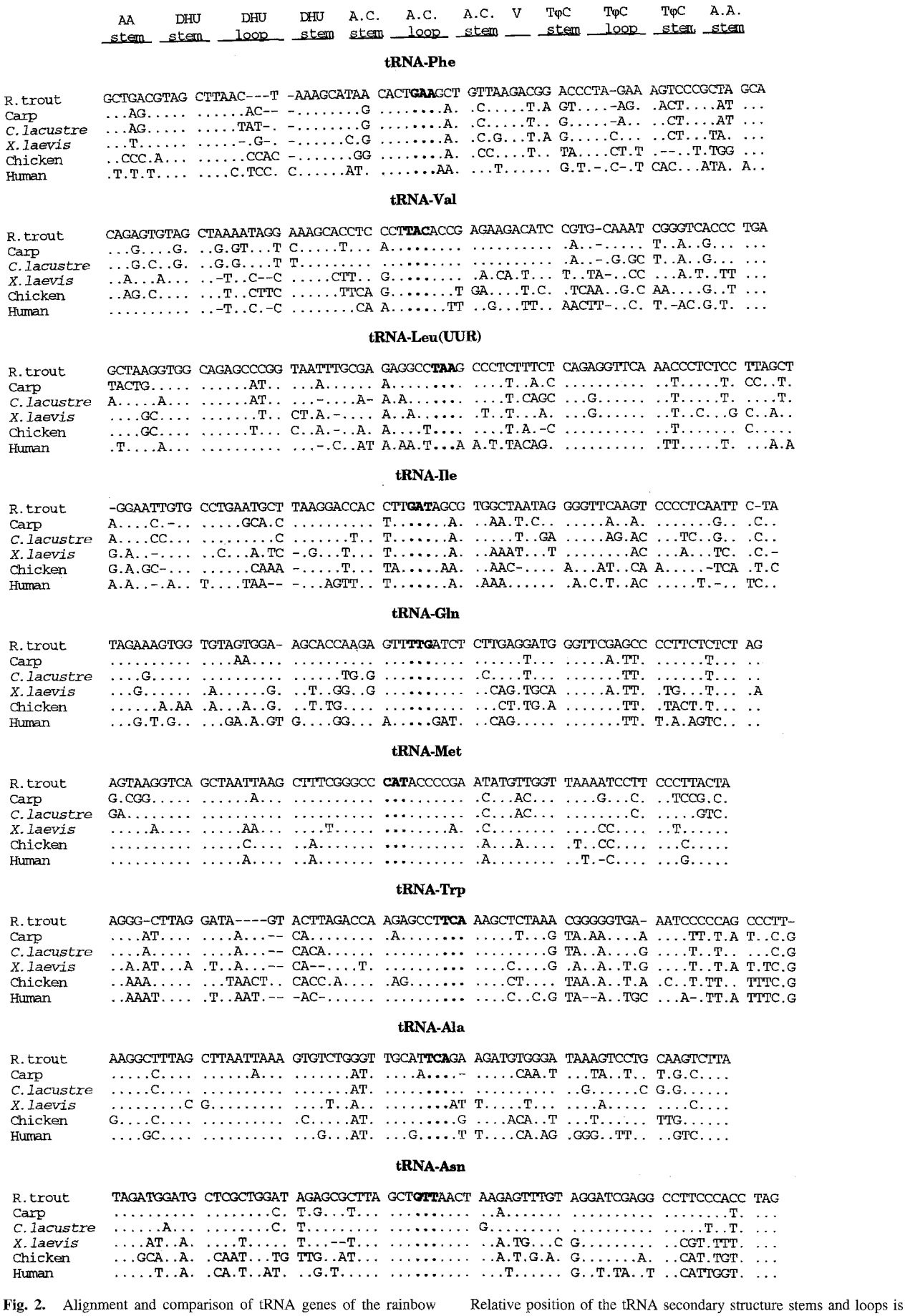

Fig. 2. Alignment and comparison of tRNA genes of the rainbow trout and those of other representative vertebrates. The anticodons of each tRNA are shown in *bold*. *Dashes* indicate gaps introduced to optimize the sequence homology. Base matching is indicated by *dots*. Relative position of the tRNA secondary structure stems and loops is also depicted. Abbreviations: *AA*, amino acid; *DHU*, dihydrouridine; *AC*, anticodon; *V*, variable; *TψC*, ribothymidine-pseudouridine-cytosine.

	AA stem	DHU stem	DHU loop	DHU stem	A.C. stem	A.C. loop	A.C. stem	V	TφC stem	TφC loop	TφC stem	A.A. stem

tRNA-Cys

```
R.trout      AAGCCCTGTG GTGTACCACA CGTTAGATTG CAAATCTGAA GAAGTAGGCT AATAGCCTGC CGGGGCTTT
Carp         C...T.C... ...A.GA... .A..G..... ......C... ..C.C..AT. ....CT.... .........
C.lacustre   G....T.... ..AG.GG.T. .A........ .......A.. ..C.C..AT. ..C.-T.... .........
X.laevis     ....-...C. ....TTG... T.CC...... ........CG. ....C.AA.G ..-G.TT... ....-...C
Chicken      -GA.T....A ...A.GTT.. TAA.GAG... ....CTC.TT ..T...CA.. ..-..TG... .....-.C.
Human        -...T.C.A. ...ATTTT.. TA..GA.... .....TC... ....C..CT. C.-.A..... ........-
```

tRNA-Tyr

```
R.trout      GGTAGGGTGG CTGAGAGCTT AAGCGGTGGA TTGTAGCTCC ATAAACAGAG GTTTGACTCC TCTCCTTATC A
Carp         ........A. .C...T.... .G.......G ........C.- GA.G...... ......G... ...T.C.... .
C.lacustre   ......A... .C..A.---. .G.T.....G ........C.. GA.G.T.... ....A.G... ...T.C.... .
X.laevis     ....A..... .C...TAA.- .G....C... .......... G.GT...... ...CA.G... ...T...... .
Chicken      ....A.A... .....T.T.G .....T.A.G C.....TC.T T.TT...... ...CA.T... ...T...... -
Human        ....AAA... .....T.--- ....AT.... C....AA..T .A.G....G. ...--.GG.. ...TT...C. .
```

tRNA-Ser (UCN)

```
R.trout      AAGAAAGTGG CAG--AGTGG TT--ATGCGG TTGGCTTGAA ACCAGCACAT GGGGGTTCAA TTCCTCCCTT TCTCG
Carp         ...G.....A ...--..... ..--...T.. C.......... .......T.. .......... .......... .....
C.lacustre   -..G..... ...--..... ..--...T.A C.......... .....T.T.. .......... .......... C...-
X.laevis     ......A... ...--..... .G--....AA C..A..... .T...AGT.. ........G. ....CT.T.. .....
Chicken      .......AA. ..TTA..... ..TG..... .......... ....A..TG. .A......G. .....T.... ...T.
Human        T.....AA.T ..T--G.A.. CC--...G.. .......... ......TT-. ........G. .......T.. .T.GT
```

tRNA-Asp

```
R.trout      AAGACACTAG TAAAACTAGT CTATTACACT GCCTGGTCAA GGCAAAATTG TGGGTTAAAC CCC-GCGTGT CTTA
Carp         ......T... ......--.. AA......TC A...T..... ..TG....C. .A.......T ...T..A... ....
C.lacustre   ......T... ....CT.-.. AA......TC A.T.T..... ..TG....A. .A........ T..T..A..A ...G
X.laevis     G...TGT... ..........- -A....GCAC ....T..... ...G....A. CT.....G.. T..G..ACA. ..C.
Chicken      G....GT... ....C.---- -A......TA .A.CT..... .A.T....CA CA...GC..A ...T.TACA. ..C.
Human        ...GT.T... A.....C.-- ---..T..TA A.T.T..... A.TT.....A .A..C....T -..TATA.A. ....
```

tRNA-Lys

```
R.trout      CACTAAGAAG CTAAA--TCG GGAAT-AGCG TTAGCCTTTT AAGCTAAAGA TTGGTGGCCC CCA-ACCACC CC-TAGTGA
Carp         .G...G.... .....TA.T. .AC.A..... ..G..... ....C..... ......ATT. ..G..... T......--
C.lacustre   .....G.... .....C..A. ..T.AC...A ..G..... ....C..... ......C.T. .......... T........
X.laevis     .......... ........A. ..C..T.... AC........ .....GT... ......A.T. .......... .T...A...
Chicken      ..T....... ...TGCAC-- -----C...A C.....G... ......G... GA..G.A.A. ..-.-T.C.. .T...A...
Human        ....GTA... ....CT.--- -----T...A ...A..... ...T..... ..AAGA.AA. .A.C...T.T TTAC.....
```

tRNA-Gly

```
R.trout      ATCTTTCTAG TATTAATAC- -GTATAAGTG ACTTCCAATC ACCCGGTCTT GGTTAAAATC CAAGGAAAGA TA
Carp         .......... ...C..AGTT A...C..... .......... ..A.A..... .......CC. .......... ..
C.lacustre   .......... .......G.T A...CG.... .......... ....C..... .......CC. .......... ..
X.laevis     .CT..CT... ......--.C A...C.C... .......... ..AAA..... A....G.... T...AG..AG ..
Chicken      GCTC...... ...--.C-TC AT..C..C.. .......... TTTAAAATC. ...ATC..C. ..GA...GAG C.
Human        .CTC..T... ...A..---T A...CCGT.A .........T .A.TA..T.. .AC-..C..T ...AA..GAG ..
```

tRNA-Arg

```
R. trout     CGG-AGTTAG TCCAAAACAA GACCCTTGAT TTCGGCTCAA AAGACCATGG TTTAAGTCCA TGACCGCCTT
Carp         A..G...... .......... ....TC.... .........G ..A.T..... .....T.... .....C....
C.lacustre   A..G.A.... ..T....T.. ....TC.... .........G .GA.T.G... ..A....... C..T.C....
X.laevis     GA.TT..... ..T...-... ...AG..... .......... C.A.TT.... ..A..CC... .A.TAA.TC.
Chicken      A.AA...... ..T..C-T.. ...AGC..G. ....A.C..G C.A.TT..A. ACCC.CCT-. .A..TTT...
Human        T..T.TA... .TT...-... A..GAA.... ....A....T T.A.TT...A .AA----T.. .ATTTA..AA
```

tRNA-His

```
R.trout      GTAGATATAG TTTAACC-AA GACATTAGAT TGTGATTCTA AAAATAGAGG TTAAAATCCT CTTATCCACC
Carp         ...AG..... ......T... A.......... .......... ....C..G.. .....G...C ....CT....
C.lacustre   ...AGC.... .....TTT.. A.......... .......... .....G.... .......... .C..CT....
X.laevis     .......... ......--T.. A...C..... .......... G.GTC..... .....CC... ......A...
Chicken      .C.A.C.... ........C. A.......... .......... .......GA. ..T..CC.TC ...G.T.G..
Human        ...A...... .......... A....C.... .....A...G .C..C..... C.T.CGA..C .....TT...
```

tRNA-Ser (AGY)

```
R.trout      GAGAGAAATC TGTTGATAAC AGAGACTGCT AATCTTCTGC CCCCTCAGTT AAATTCTGTG GTTCACTCG
Carp         A..GA.G.A. A.AAA...GT .AGC...... ....C.TACA TT..ATG... ...A..C... .C.TC..T.
C.lacustre   CGAGA..GG. CCCA.GC..T .AGC...... ...AC.TATA A...A.G... ...C..C... .C..T....
X.laevis     ..ACTTG.-. ..-G..CCCT .AGA...... ...TACT-TA .G.TGT-... C.....CAC. .C.TGT...
Chicken      AG.G..GGC. CA-A.CC.G. .AGA...... ...TCCTGCA T.TGAGCT.. ...CCTCA-- ..C.C..T-
Human        ..----G.-- -A-A.C.C.. .AGA...... ..CTCAT-.. ....AT-..C T..CAACA.. .C.TT...A
```

Fig. 2. Continued.

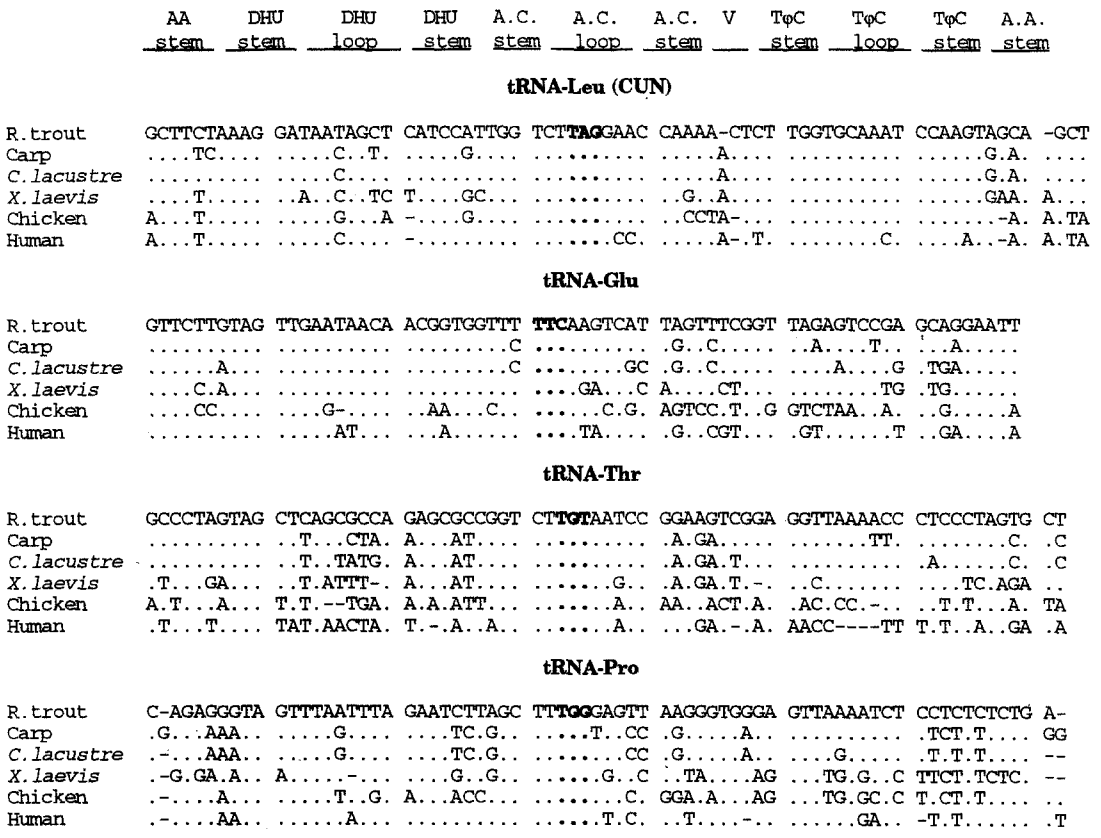

Fig. 2. Continued.

nucleotide, but only one transcript is detected in each case (R. Zardoya, unpublished results). The length of the ND4L/ND4 overlap is conserved in all vertebrate mitochondrial genomes already sequenced; however, the ATPase 8/6 overlap presents more variation in its length: from ten nucleotides in fish, *X. laevis,* and chicken to about 40 nucleotides in mammals. This increased length overlapping in mammals is due to a single base change that deletes the stop codon of ATPase 8 present in fish, *X. laevis,* and chicken (Roe et al. 1985).

The alignment of the deduced amino acid sequences from the 13 mitochondrial polypeptides encoded by the rainbow trout mtDNA with their homologues of other vertebrates reveals that cytochrome oxidase subunits and cytochrome b are highly conserved whereas NADH dehydrogenase and ATPase subunits present more variation. This fact can be explained by taking into account that cytochrome oxidase and cytochrome b mitochondrial genes encode the redox active centers of the corresponding holoenzymes (complex IV and complex III, respectively), whereas NADH dehydrogenase and ATPase mitochondrial genes encode for the regulatory subunits of complex I and complex V, respectively. In all cases, the carboxyl end of the polypeptides is highly variable in length and sequence. The COI polypeptide sequence is the most conserved when compared to that of carp, *C. lacustre, X. laevis,* chicken, and human, whereas ATPase 8 is the more divergent (Table 2).

Ribosomal RNA Genes

The 12S and 16S rRNA genes in rainbow trout mitochondria are 944 and 1,680 nucleotides long, respectively. The inferred secondary structure of the 12S rRNA gene was found to be essentially equivalent to that of *X. laevis* (Roe et al. 1985). The sequence similarity between rainbow trout rRNA genes and their homologues of the two fish species already sequenced—i.e., *C. lacustre* and *C. carpio*—was estimated. We restricted the analysis to those positions at which substitutions could be unambiguously determined based on a correct alignment. This fact limited the analysis of the 12S rRNA gene to 97–167 positions, and that of the 16S rRNA gene to 240–368 substitutions (Table 3). Our data show a high transition bias among fish species, which is in full agreement with the pronounced predominance of transitions over transversions reported in mtDNA (Meyer 1994) and probably reflects differences in mutation rates specific to mitochondria. Moreover, the transition/transversion ratios inferred from the rRNA sequences are consistent with those reported for rainbow trout ATPase6, COIII, ND3, and ND4L genes (Thomas and Beckenbach 1989).

Transfer RNA Genes

The rainbow trout mitochondrial genome contains 22 tRNA genes interspersed between ribosomal RNA and

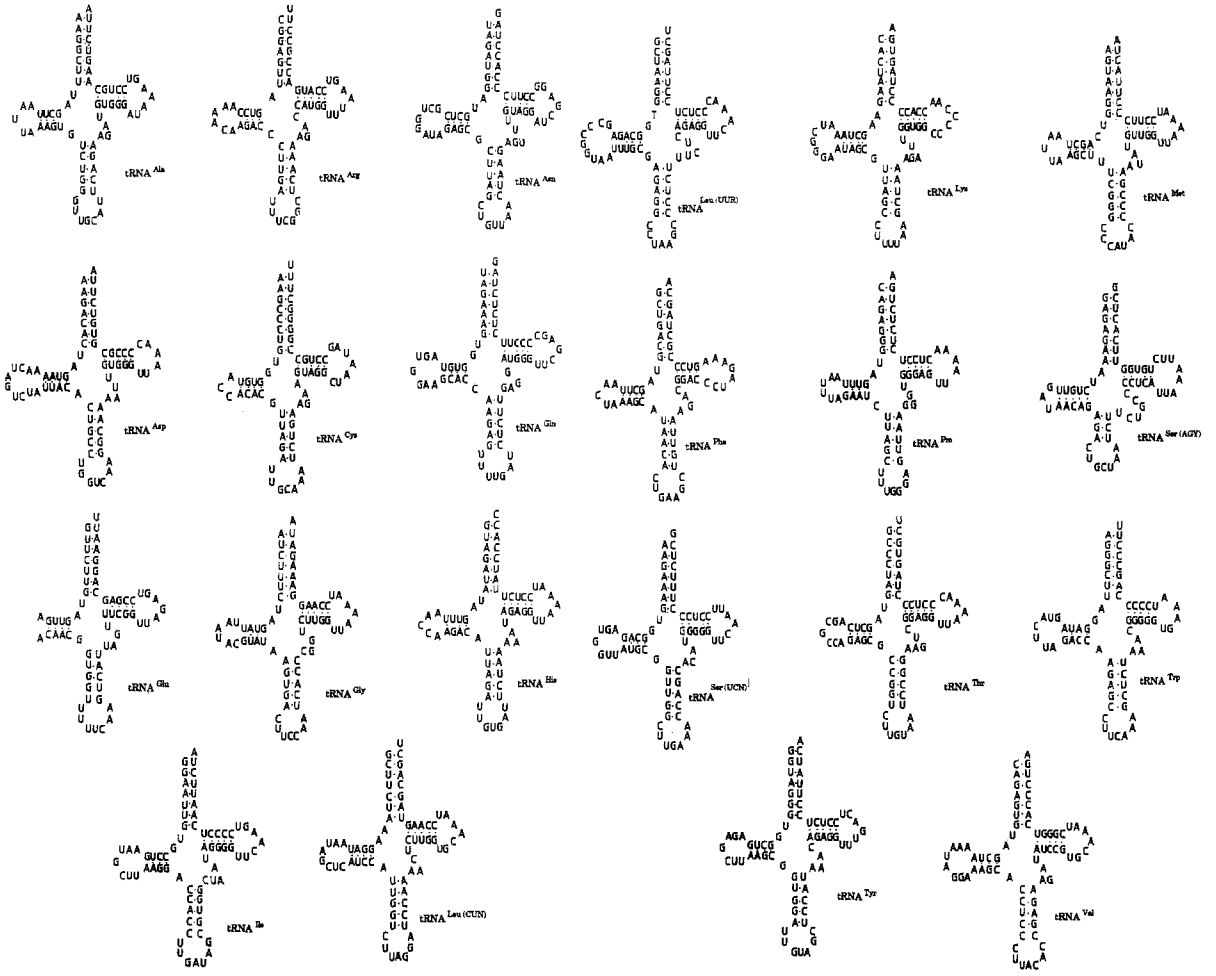

Fig. 3. Sequences of the rainbow trout 22 tRNAs folded into their proposed cloverleaf secondary structures.

protein coding regions. The alignment and comparison of the sequences of the rainbow trout tRNA genes with their counterparts in other vertebrates is shown in Fig. 2. $tRNA^{Met}$ gene is the most conserved among all vertebrates whereas $tRNA^{Ser(AGY)}$ is the most divergent. In general, the anticodon and amino acid acceptor arms are more conserved than DHU and TψC arms, which show less similarity. However, it is striking that the DHU loop of $tRNA^{Leu(UUR)}$ is highly conserved among vertebrates and even identical between rainbow trout and human. The sequence that forms this loop has been reported to be necessary for the correct ending of the transcription event that includes the two ribosomal RNA genes in human (Christianson and Clayton 1988). It is likely that the conservation of this sequence is required in all vertebrates for this purpose, and therefore this functional constraint would explain the unusual preservation of $tRNA^{Leu(UUR)}$ DHU loop sequence.

All the rainbow trout tRNAs can be folded into a cloverleaf secondary structure which is depicted in Fig. 3. As in other vertebrates, folding of the tRNA sequences specifically requires the formation of G + U and other atypical pairings. The rainbow trout $tRNA^{Ser(AGY)}$ comprises a complete DHU arm that is missing in the rest of the homologous tRNAs already reported, being therefore an exception to the universal rule proposed so far (Wolstenholme 1992). The putative secondary structure of this tRNA included a DHU arm with a three-nucleotide loop, a 5-bp stem, and an exceptionally short anticodon stem. The unconventional secondary structure proposed

	TAS	Position
1.	5′-ACIGTAAATGTTAT-3′	20–33
2.	5′-ACATCT-ATG-TAT-3′	58–69
3.	5′-ACATATTATG-TAT-3′	76–88
4.	5′-CCATAT-ATA--AT-3′	93–103
5.	5′-ACATTATATG-TAT-3′	120–132
6.	5′-ACATACGGTG--AT-3′	138–149
7.	5′-ACATCAGCACAAAT-3′	164–177
8.	5′-ACATTAAGCCAAAC-3′	187–200
Mouse	5′-ACATTAAAYY-AAT-3′	

Fig. 4. Comparative alignment of the termination-associated sequences (TASs) identified in the rainbow trout D-loop-containing region. *Dashes* indicate gaps introduced to maximize sequence identity. Position of each TAS within the complete sequence of the rainbow trout mtDNA is indicated. Mouse consensus TAS (Doda et al. 1981) is also shown.

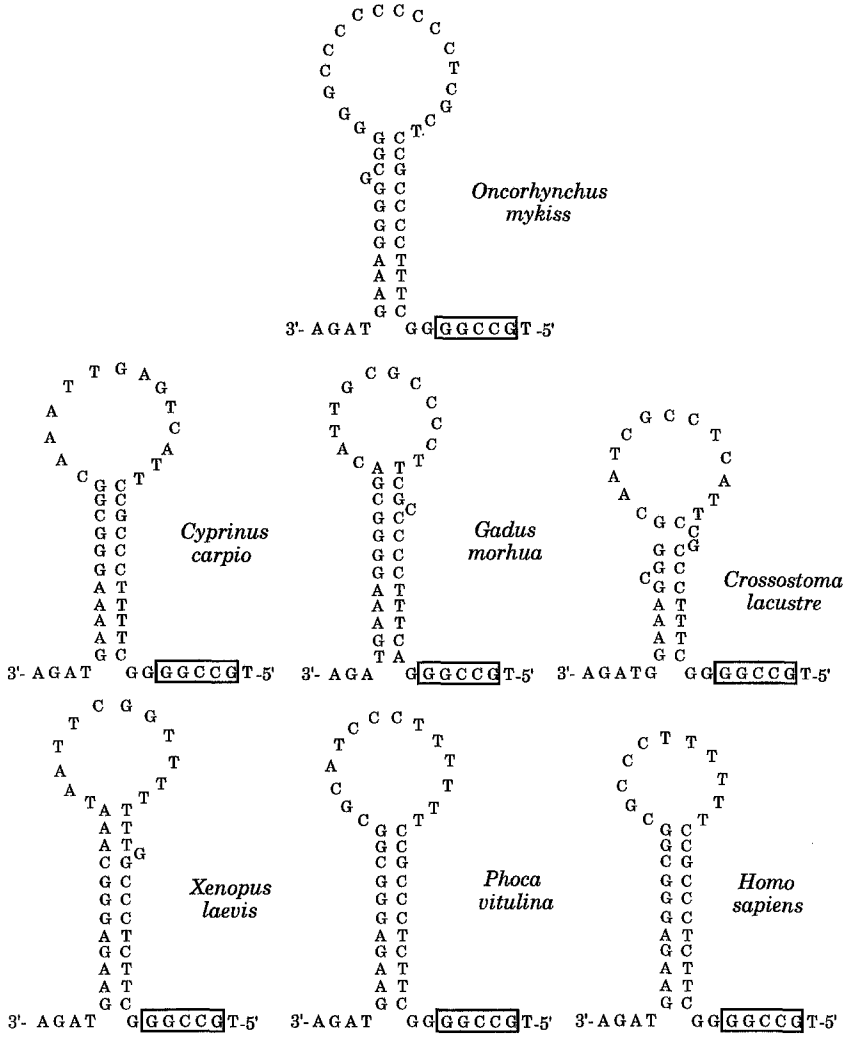

Fig. 5. Proposed stem-loop structures for the L-strand origin of replication of rainbow trout mtDNA and other representative vertebrates. Nucleotide sequences of the H-strand template are shown. The sequence associated to the transition of RNA synthesis to DNA synthesis is indicated by a box.

might be explained in terms of a structural compensation among tRNA arms as described by Steinberg and Cedergren (1994). A less-truncated $tRNA^{Ser(AGY)}$ DHU arm has also been proposed in cod (Johansen et al. 1990), but in this case, the anticodon stem inferred is unusually large. In any case, it could be suggested that the DHU arm of this tRNA, lost in the majority of species, is conserved in fish.

Noncoding Sequences

The major noncoding region found in rainbow trout mtDNA is a 1,003-bp sequence localized between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes which corresponds to the control region. The region includes the origin of H-strand replication and the sites of initiation of both H- and L-strand transcription. The control region now presented is identical to that reported by Digby et al. (1992). In addition to the features described by Digby et al. (1992) we have found a putative L-strand-encoded open reading frame between bp 820 and bp 985. Its localization and coding strand resemble those of the ORF reported in the human mitochondrial control region (Ojala et al. 1981). However, unlike human, no transcript for this ORF has been found (data not shown). The amino acid sequence predicted for this ORF shows limited similarity to a yeast nuclear-encoded peptide that activates the translation of COIII mitochondrial mRNA through a site in the 5′-untranslated mRNA leader (Haffter and Fox 1992). Analysis of the control region sequence permitted the identification of a total of eight termination-associated sequences (TASs) in the left domain [five more than the previously identified (Digby et al. 1992)] in comparison with the consensus sequence described by Doda et al. (1981) (Fig. 4).

The origin of light-strand replication (O_L) is located in a cluster of five tRNA genes and comprises 52 nucleotides in length. This region has the potential to fold in a stem-loop secondary structure with a stem formed by 11 paired nucleotides and a loop of 17 nucleotides. The 5′-GCCGG-3′ motif that in human mtDNA is involved in the transition from RNA synthesis to DNA synthesis (Hixson et al. 1986) is entirely conserved in the rainbow trout mtDNA, both in sequence and in its location in the

base of the stem, within the $tRNA^{Cys}$ gene. The stem sequence of O_L is highly conserved among vertebrate mtDNAs whereas the loop sequence is more variable (Fig. 5). Striking features of the rainbow trout O_L loop are its great size and the presence of a stretch of nine cytosine residues. O_L of human mtDNA contains a T-rich sequence involved in the initiation of the L-strand replication by a RNA primase (Wong and Clayton 1985). All mammalian mtDNAs already sequenced share this T-rich sequence; however, it is not found in fish mitochondrial genomes (Fig. 5). Therefore, in contrast to mammals, it seems that in rainbow trout O_L loop, RNA primer synthesis is most probably initiated in the C-rich sequence, and consequently, the initiation in vertebrates is not restricted to a stretch of thymines as previously suggested (Wong and Clayton 1985) but to a polypyrimidine tract. In this sense, the O_L loop of cod (Johansen et al. 1990) containing four cytosine residues recalls that of the rainbow trout.

A minor noncoding sequence (14 nucleotides) is localized in rainbow trout mtDNA between the $tRNA^{Asp}$ gene and the COII gene. A similar 14-nt region is found in *C. lacustre* (Tzeng et al. 1992) and carp (Chang and Huang 1994), but not in any other vertebrate. Loss of this sequence during evolution of vertebrates emphasizes the tendency of mtDNA to economy.

Acknowledgments. Research was partially supported by a grant from the Spanish CICYT (ANT93-1005). R.Z. was recipient of a UCM predoctoral fellowship. We thank Manuel Mazin for excellent technical assistance and M.I. Garcia-Saez for assistance in the automatic sequencing at the UCM sequencing service.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–464

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

Árnason E, Rand DM (1992) Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua.* Genetics 132: 211–220

Árnason Ú, Gullberg A, Widegren B (1991) The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus.* J Mol Evol 33:556–568

Árnason Ú, Johnsson E (1992) The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina.* J Mol Evol 34:493–505

Árnason Ú, Gullberg A (1993) Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. J Mol Evol 37:312–322

Árnason Ú, Gullberg A, Johnsson E, Ledje C (1993) The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, *Halichoerus grypus,* and a comparison with mitochondrial sequences of other true seals. J Mol Evol 37:323–330

Attardi G (1985) Animal mitochondrial DNA: an extreme example of genetic economy. Int Rev Cytol 93:93–145

Attardi G, Chomyn A, Doolittle RF, Mariottini P, Ragan CI (1986) Seven unidentified reading frames of human mitochondrial DNA encode subunits of the respiratory chain NADH dehydrogenase. Cold spring Harbor Symp Quant Biol, pp 103–114.

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

Brown WM, George Jr M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76:1967–1971

Buroker NE, Brown JR, Gilbert TA, O'Hara PJ, Beckenbach AT, Thomas WK, Smith MJ (1990) Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics 124:157–163

Chang YS, Huang FL (1994) The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. J Mol Evol 38:138–155

Chomyn A, Mariottini P, Cleeter WJ, Ragan CI, Matsuno-Yagi A, Hatefi Y, Doolittle RF, Attardi G (1985) Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature 314:592–597

Christianson TW, Clayton DA (1988) A tridecamer DNA sequence supports human mitochondrial RNA 3′-end formation in vitro. Mol Cell Biol 8:4502–4509

Davidson WS, Bartlett SE, Birt TP, Green JM (1988) Cloning and sequence analysis of an XbaI fragment of rainbow trout mitochondrial DNA. Curr Genet 14:483–486

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

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395

Digby TJ, Gray MW, Lazier CB (1992) Rainbow trout mitochondrial DNA: sequence and structural characteristics of the non-coding control region and flanking tRNA genes. Gene 113:197–204

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

Gadaleta G, Pepe G, DeCandia G, Quagliariello C, Sbisá E, Saccone C (1989) The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J Mol Evol 28:497–516

Haffter P, Fox TD (1992) Suppression of carboxy-terminal truncations of the yeast mitochondrial mRNA-specific translational activator PET122 by mutations in two genes, MRP17 and PET127. Mol Gen Genet 235:64–73

Hixson JE, Wong TW, Clayton DA (1986) Both the conserved stem-loop and divergent 5′-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA replication. J Biol Chem 261:2384–2390

Hoffmann RJ, Boore JL, Brown WM (1992) A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis.* Genetics 131:397–412

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

Johansen S, Guddal PH, Johansen T (1990) Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua.* Nucleic Acids Res 18:411–419

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

Meyer A (1994) DNA technology and phylogeny of fish: molecular

phylogenetic studies of fish. In: Beaumont AR (ed) Genetics and evolution of aquatic organisms. Chapman and Hall, pp 219–249

Montoya J, Gaines G, Attardi G (1983) The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. Cell 34:157–159

Ojala D, Crews S, Montoya J, Geldfand R, Attardi G (1981) A small polyadenylated RNA, containing a putative ribosome attachement site, maps near the origin of human mitochondrial DNA replication. J Mol Biol 290:470–474

Palva TK, Palva ET (1985) Rapid isolation of animal mitochondrial DNA by alkaline extraction. FEBS Lett 192:267–270

Palva TK, Palva ET (1987) Restriction site polymorphism in mitochondrial DNA of rainbow trout, *Salmo gairdneri* Richardson stocks in Finland. Aquaculture 67:283–289

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

Sanger F, Nicklen FS, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Steinberg S, Cedergln R (1994) Structural compensation in atypical mitochondrial tRNAs Nature struct Biol 1: 507–511

Thomas WK, Beckenbach AT (1989) Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. J Mol Evol 29:233–245.

Tzeng CS, Hui CF, Shen SC, Huang PC (1992) The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variations among vertebrates. Nucleic Acids Res 20:4853–4858

Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution. Int Rev Cytol 141:173–216

Wong TW, Clayton DA (1985) In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. Cell 42:951–958

Xu X, Árnason Ú (1994) The complete mitochondrial DNA sequence of the horse, *Equus caballus:* extensive heteroplasmy of the control region. Gene 148:357–362