J Mol Evol (1989) 28:374–387

Journal of Molecular Evolution

Platyhelminth Mitochondrial DNA: Evidence for Early Evolutionary Origin of a tRNA^{ser}AGN that Contains a Dihydrouridine Arm Replacement Loop, and of Serine-Specifying AGA and AGG Codons

James R. Garey* and David R. Wolstenholme

Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

Summary. The nucleotide sequence of a segment of the mitochondrial DNA (mtDNA) molecule of the liver fluke *Fasciola hepatica* (phylum Platyhelminthes, class Trematoda) has been determined, within which have been identified the genes for tRNA^{ala}, tRNA^{asp}, respiratory chain NADH dehydrogenase subunit I (ND1), tRNA^{asn}, tRNA^{pro}, tRNA^{ile}, tRNA^{lys}, ND3, tRNA^{ser}AGN, tRNA^{trp}, and cytochrome *c* oxidase subunit I (COI). The 11 genes are arranged in the order given and are all transcribed from the same strand of the molecule. The overall order of the *F. hepatica* mitochondrial genes differs from what is found in other metazoan mtDNAs. All of the sequenced tRNA genes except the one for tRNA^{ser}AGN can be folded into a secondary structure with four arms resembling most other metazoan mitochondrial tRNAs, rather than the tRNAs that contain a TψC arm replacement loop, found in nematode mtDNAs. The *F. hepatica* mitochondrial tRNA^{ser}AGN gene contains a dihydrouridine arm replacement loop, as is the case in all other metazoan mtDNAs examined to date. AGA and AGG are found in the *F. hepatica* mitochondrial protein genes and both codons appear to specify serine. These findings concerning *F. hepatica* mtDNA indicate that both a dihydrouridine arm replacement loop-containing tRNA^{ser}AGN gene and the use of AGA and AGG codons to specify serine must first have occurred very early in, or before, the evolution of metazoa.

Key words: *Fasciola hepatica* — Platyhelminthes — Trematoda — Mitochondrial DNA — Mitochondrial genes — Nucleotide sequence — Genetic code — Gene arrangement — Transfer RNA genes — Evolution

Introduction

The mitochondrial (mt-)genome of metazoa is contained in a single circular molecule with a species-specific size that varies from 14 to 39 kb (Moritz et al. 1987; Snyder et al. 1987). The only known exceptions are the mt-genomes of species of the coelenterate genus *Hydra* that occur as two linear molecules, both of approximately 8.0 kb (Warrior and Gall 1985).

Complete nucleotide sequences and gene content have been determined for seven metazoan species to date: three mammals—human, mouse, and cow (Anderson et al. 1981, 1982b; Bibb et al. 1981), an amphibian—*Xenopus laevis* (Roe et al. 1985), an insect—*Drosophila yakuba* (Clary and Wolstenholme 1985), and two nematode worms—*Ascaris suum* and *Caenorhabditis elegans* (Wolstenholme et al. 1987; R. Okimoto, J.L. Macfarlane, D.O. Clary, J.A. Wahleithner, and D.R. Wolstenholme, unpublished data). Also, partial mtDNA sequences have been obtained from a number of other metazoans including some mammals (Grosskopf and Feldmann 1981; Pepe et al. 1983; Foran et al. 1988), echinoderms (Cantatore et al. 1987, 1988; Himeno et al. 1987), mosquito (HsuChen et al. 1984), and various *Drosophila* species (Clary et al. 1982; de

Offprint requests to: D.R. Wolstenholme
* *Present address:* Division of Hematology–Oncology, University of Utah, Salt Lake City, Utah 84132, USA

Bruijn 1983; Clary and Wolstenholme 1987; Garesse 1988).

Each metazoan mtDNA molecule that has been completely sequenced contains the genes for the structural RNA molecules of the mitochondrion's own protein-synthesizing machinery (two rRNAs and 22 tRNAs) and 12 or 13 proteins. These proteins are all components of the oxidative-phosphorylation system: cytochrome b (cyt b), subunits I–III of cytochrome *c* oxidase (COI, II, and III), subunit 6 and, in all except *A. suum* and *C. elegans,* subunit 8 of the ATPase complex (ATPase6 and ATPase8), and subunits 1–6 and 4L of the respiratory chain NADH dehydrogenase (ND1–6 and 4L) (Chomyn and Attardi 1987). Also, there is a region of variable length (<1 kb to 5 kb) that contains the replication origin and control signals for transcription. Metazoan mt-genomes are extremely compact, none of the genes contain introns, and there are very few nucleotides between genes. However, segments of various sizes of some mtDNA molecules are duplicated (Moritz and Brown 1987; Snyder et al. 1987).

Protein genes of metazoan mtDNAs use genetic codes that are modified (Barrell et al. 1979, 1980; Bibb et al. 1981; Clary and Wolstenholme 1985). Rather than being a stop codon, TGA specifies tryptophan in all metazoan mtDNAs examined. In both mammalian and *D. yakuba* mtDNAs, and possibly all other metazoan mtDNAs, ATA specifies methionine rather than isoleucine. AGA and AGG that specify arginine in the standard genetic code either are absent or are used as rare stop codons in vertebrate mtDNAs. However, in *D. yakuba* mtDNA, AGA (but not AGG) specifies serine, and in nematode and echinoderm mtDNAs both AGA and AGG specify serine (Clary and Wolstenholme 1985; Himeno et al. 1987; Wolstenholme et al. 1987; Cantatore et al. 1988).

Among mt-protein genes from different metazoa there is frequent use of unorthodox translation initiation codons that include all ATN codons, GTG, and TTG (Anderson et al. 1981, 1982b; Bibb et al. 1981; Clary and Wolstenholme 1985; R. Okimoto, J.L. Macfarlane, and D.R. Wolstenholme, unpublished data). Some vertebrate and *D. yakuba* mt-protein genes end in a T rather than a termination codon (TAA or TAG), and this T is immediately adjacent to the 5′ terminal nucleotide of the sense sequence of a tRNA gene. The corresponding U in the mt-protein gene transcript is converted to a complete termination codon by polyadenylation, following precise cleavage of a primary multicistronic transcript, between the mt-protein and tRNA sequences (Anderson et al. 1981, 1982b; Bibb et al. 1981; Ojala et al. 1981; Roe et al. 1985; Clary and Wolstenholme 1985).

Transfer RNA genes of mammalian and *D. yakuba* mtDNAs resemble prokaryotic tRNAs and eukaryotic nuclear-encoded tRNAs with regard to both primary and secondary structure, except that in these mt-tRNAs there is considerable variation in both the size and the sequence of the dihydrouridine and TψC loops (Barrell et al. 1979; Crews and Attardi 1980; Van Etten et al. 1980; Anderson et al. 1982a; Roe et al. 1982; Clary and Wolstenholme 1985). Also, in all metazoan mtDNAs so far sequenced, the dihydrouridine arm of the tRNA that recognizes serine-specifying AGN codons is replaced by a loop of between 5 and 11 nucleotides (nt) (Arcari and Brownlee 1980; de Bruijn et al. 1980; Clary and Wolstenholme 1984; Dubin et al. 1984; Himeno et al. 1987; Wolstenholme et al. 1987; Cantatore et al. 1988). However, all of the mt-tRNAs of the nematodes *A. suum* and *C. elegans* are unusual in that each contains only three arms: in 20 of the 22 mt-tRNAs of each species, the TψC arm and variable loop are replaced by a loop of 4–12 nt. In each of the two serine-specifying tRNAs (AGN and UCN) the dihydrouridine loop is replaced by a loop of 5–8 nt (Wolstenholme et al. 1987; R. Okimoto and D.R. Wolstenholme, unpublished data).

In order to gain further information on the evolution of the various genetic peculiarities found in metazoan mtDNA we have determined the nucleotide sequence of a segment of the mtDNA molecule from the liver fluke *Fasciola hepatica* (phylum Platyhelminthes, class Trematoda). The results of our analysis of this sequence, which are the subject of this paper, indicate that the use of AGA and AGG codons to specify serine and the loss of a standard dihydrouridine arm in the tRNA that recognizes mitochondrial AGN codons must have occurred very early in, or before, the evolution of metazoa.

Materials and Methods

Adult *Fasciola hepatica* were obtained from the bile ducts of fresh bovine livers at a local slaughterhouse and transported to the laboratory in isotonic saline at 4°C. Mitochondria were isolated from whole *F. hepatica* as previously described (Wolstenholme and Fauron 1976) except that in all solutions mannitol was used in place of sucrose. Mitochondrial DNA was obtained from sarkosyl lysates of mitochondria by cesium chloride–ethidium bromide centrifugation (Wolstenholme and Fauron 1976).

Fasciola hepatica mtDNA was digested with the restriction enzymes, *Bam*HI, *Eco*RI, and *Hin*dIII, singly and sequentially, under conditions specified by the manufacturers, and the products were analyzed in 1% agarose gels. *Hin*dIII–*Bam*HI fragments of *F. hepatica* mtDNA were cloned into pUC8 DNA and amplified in *Escherichia coli* JM101. A pUC8-cloned 2335-ntp *Hin*dIII–*Bam*HI fragment was recloned into the DNAs of bacteriophages M13mp8 and M13mp9 (Messing and Vieira 1982). *Eco*RI fragments of *F. hepatica* mtDNA were cloned directly into M13mp8. Two of these *Eco*RI fragments (1344 ntp and 2116 ntp) were shown by plaque hybridization (Maniatis et al. 1982) to overlap the 2335-ntp *Hin*dIII–*Bam*HI fragment. Using dou-

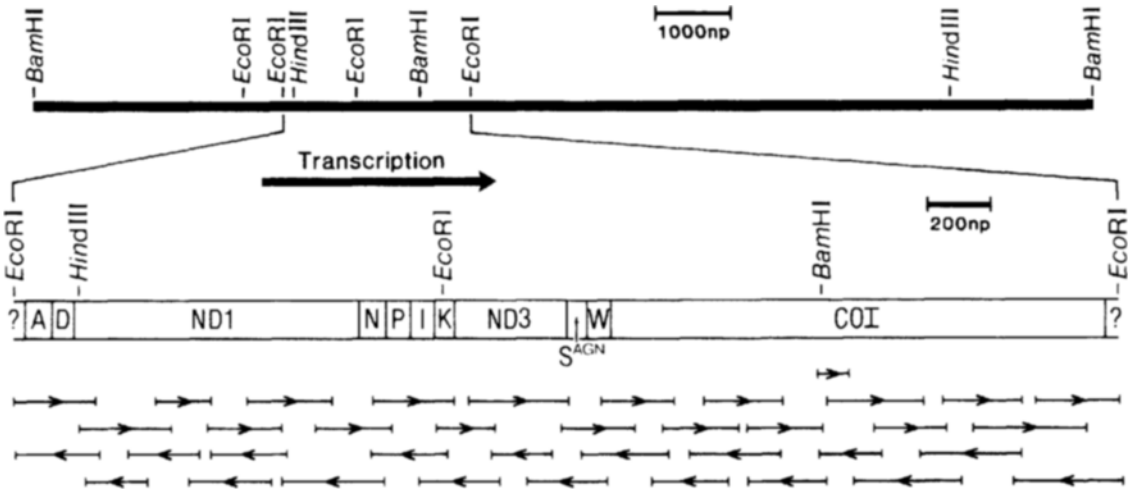

Fig. 1. A restriction enzyme cleavage site map of the *Fasciola hepatica* mtDNA molecule. Beneath this map is shown the arrangement of genes within the 3466-ntp segment sequenced. The eight tRNA genes are identified by the one-letter amino acid code. The three protein genes are subunit I of cytochrome *c* oxidase (COI) and subunits 1 and 3 of the respiratory chain NADH dehydrogenase (ND1 and ND3). All genes are transcribed in the same direction, shown by the arrow. The question marks indicate unidentified terminal sequences (Fig. 2). The arrows beneath the gene map show the extent and direction of individual sequences obtained, either from the ends of various restriction fragments or from deletion subfragments of restriction fragments. The 5′–3′ direction of the sequence of this mtDNA molecule segment given in Fig. 2 is from left to right.

ble-stranded and single-stranded forms of the three M13–mtDNA hybrid molecules, partial deletions of each of the three-component *F. hepatica* mtDNA fragment inserts were generated (Hong 1982; Dale et al. 1985). All DNA sequences were obtained from M13mp8- or M13mp9-cloned fragments by the extension-dideoxyribonucleotide termination procedure of Sanger et al. (1977), using [α^{32}P]dATP. Some sequencing reactions utilized oligonucleotide primers prepared in an Applied Biosystems Synthesizer. When a band compression was observed in a sequencing gel, an additional sequence determination of the fragment under study was made using dITP in place of dGTP (Barnes et al. 1983). The sequencing strategies used are given in Fig. 1.

Nucleotide sequences were assembled and analyzed using the computer compiling program of Staden (1982). Mitochondrial protein genes were identified and analyzed using the FASTP and FASTN programs (Lipman and Pearson 1985). Transfer RNA genes were identified by eye, within sequences lying between or at the end of protein genes, from the ability of these sequences to fold into secondary structures characteristic of tRNAs, and from the trinucleotide in the anticodon position of the folded sequence.

Results and Discussion

Gene Content and Arrangement

From analyses of single and double digests of mtDNA of *F. hepatica* with *Bam*HI, *Eco*RI, and *Hin*dIII it was determined that the native mtDNA molecule is in the form of a circle of approximately 14.5 kb (Fig. 1). The complete nucleotide sequence of a 3466-nucleotide pair (ntp) segment of the *F. hepatica* mtDNA molecule was compiled from the sequences of the 2335-ntp *Hin*dIII–*Bam*HI fragment and two overlapping 1344- and 2116-ntp *Eco*RI fragments (Fig. 1). This sequence (Fig. 2) contains three open reading frames that, from comparisons of amino acid sequences to the amino acid sequences of previously identified mt-genes of mouse, *Drosophila yakuba,* and *Ascaris suum* mtDNAs, were identified as the genes for ND1, ND3, and COI. Segments of sequences lying upstream from the ND3 gene, between the ND3 and ND1 genes, and between the ND1 and COI genes can be folded into the characteristic secondary structures of mt-tRNA genes that contain anticodons expected for $tRNA^{ala}$, $tRNA^{asp}$, $tRNA^{asn}$, $tRNA^{pro}$, $tRNA^{ile}$, $tRNA^{lys}$, and $tRNA^{trp}$ genes (Fig. 4). An eighth sequence, that lies between the ND3 gene and the $tRNA^{trp}$ gene, can also be folded into a structure that resembles a tRNA gene, but this structure lacks a normal dihydrouridine arm (Fig. 5). Because a 5′ GCT is located at the anticodon position of this structure, and for other reasons discussed below, it is interpreted as a $tRNA^{ser}$AGN gene. The 5′ terminal 58 ntp of the 3466-ntp sequence (Fig. 2) have not been identified with certainty, but a portion of this sequence can be folded into a secondary structure that resembles the anticodon and TψC arm of a mt-tRNA gene with the anticodon (5′ TAC) expected for $tRNA^{val}$. The 3′-terminal 46 nucleotides of the sequence also have not been identified.

All of the three protein genes and eight tRNA genes would be transcribed in the same direction (Figs. 1 and 2). The overall order of the 11 genes is different from what has been reported for other metazoan mtDNAs (see Bibb et al. 1981; Clary and Wolstenholme 1985; Wolstenholme et al. 1987; Cantatore et al. 1988). However, in both *F. hepatica*

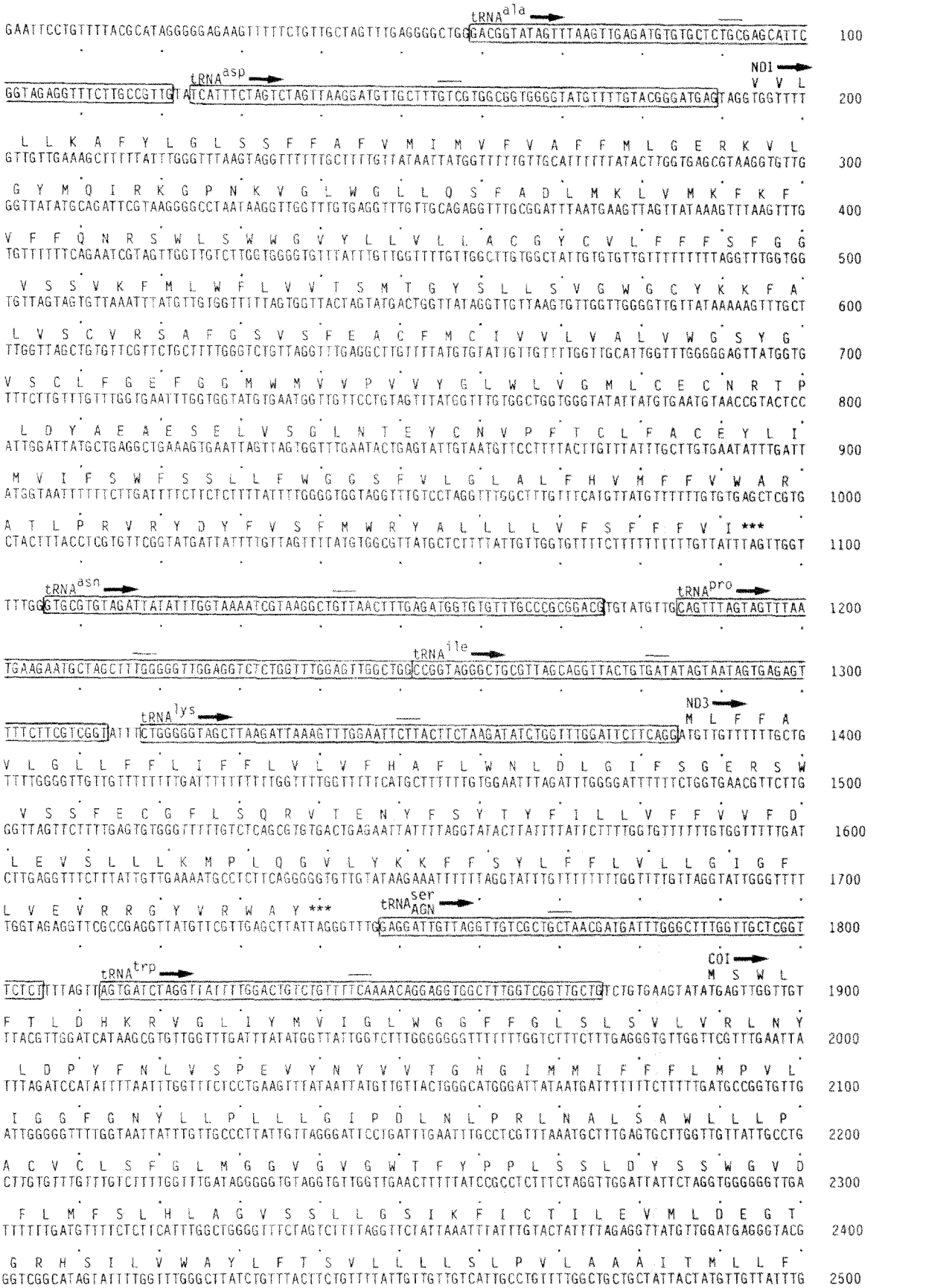

Fig. 2. Continued on next page

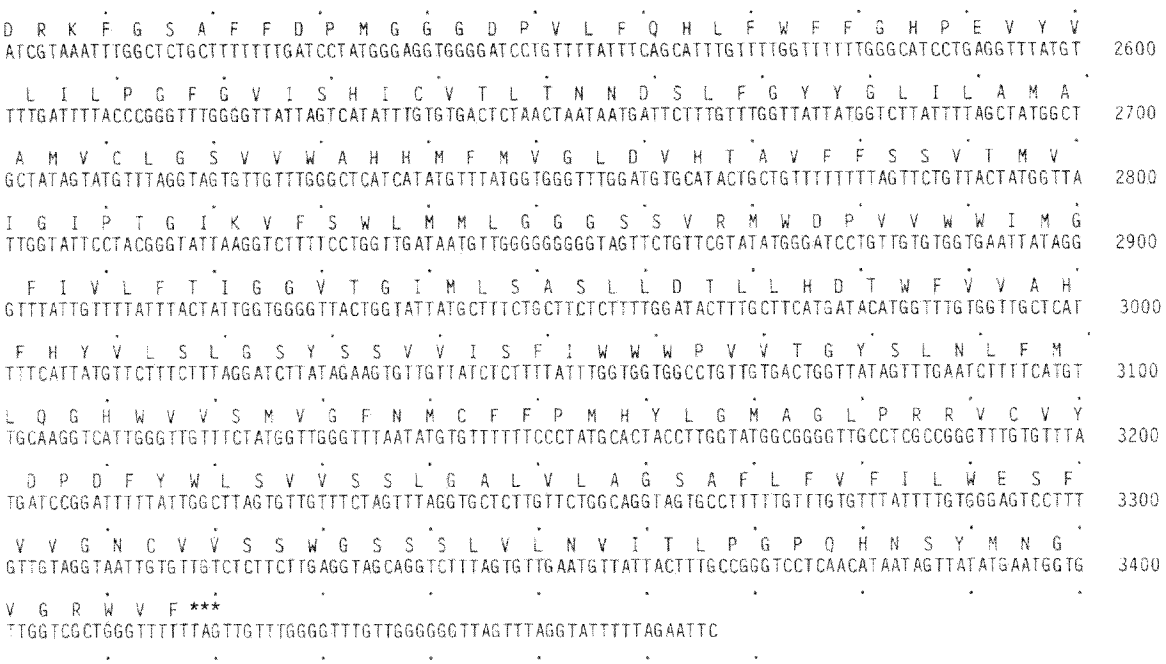

Fig. 2. Continued from previous page. Nucleotide sequence of the 3466-ntp segment of the *Fasciola hepatica* mtDNA molecule shown in Fig. 1. From considerations of amino acid sequence similarities to mouse, *Drosophila yakuba*, and *Ascaris suum* mtDNAs (Bibb et al. 1981; Clary and Wolstenholme 1985; R. Okimoto, J.L. Macfarlane, D.O. Clary, J.A. Wahleithner, and D.R. Wolstenholme, unpublished data), this sequence contains the genes for cytochrome *c* oxidase, subunit I (COI), and the respiratory chain NADH dehydrogenase subunits 1 and 3 (ND1 and ND3). Sequences of eight tRNA genes are shown boxed and the anticodon in each is overlined. The nucleotide sequence shown is the (5′–3′) sense strand for the 11 genes; all of these genes are transcribed (arrows) in the same direction. Asterisks indicate termination codons.

and mammalian mtDNAs the ND1 and COI genes are separated by a single protein gene: ND2 in mammals and ND3 in *F. hepatica*. Also, in *F. hepatica* mtDNA, as in mammalian mtDNAs, the $tRNA^{ile}$ gene lies between the ND1 gene and the downstream protein gene, and the $tRNA^{trp}$ gene lies between the COI gene and the upstream protein gene. However, other tRNA genes found in these locations differ in mammalian and *F. hepatica* mtDNAs.

Protein Genes

Similarities of the predicted amino acid sequences of the three *F. hepatica* mt-protein genes and the corresponding genes of *A. suum, D. yakuba,* and mouse are summarized in Table 1. As has been found for all other metazoan mtDNAs, greatest identity is found in all species comparisons for the COI genes (45.2% to 47.6%). Alignments of the amino acid sequences of the *F. hepatica* COI, ND1, and ND3 genes with the *D. yakuba* and mouse COI, ND1, and ND3 genes are shown in Fig. 3. The *F. hepatica* COI protein (510 amino acids) is predicted to be two and four amino acids shorter than the *D. yakuba* (512 amino acids) and mouse (514 amino acids) COI proteins. There are internal codon insertions/deletions as well as codon differences at the beginnings and ends of the COI genes of the three species. The *F. hepatica* ND1 protein (301 amino acids) is predicted again to be smaller than the *D. yakuba* and mouse ND1 proteins (324 and 315 amino acids, respectively). There are variations in codon numbers at the beginnings and ends of the three ND1 genes, and there appear to be substantial codon insertions/deletions toward the 3′ end of these genes. A greater size (118 amino acids) is predicted for the *F. hepatica* ND3 protein than for the *D. yakuba* (117 amino acids) and mouse (114 amino acids) ND3 proteins.

As in other metazoan mtDNAs, there are few or

Table 1. Amino acid sequence similarities (percentage identity) of the COI, ND1, and ND3 genes of *Fasciola hepatica* with the corresponding genes of *Ascaris suum, Drosophila yakuba,* and mouse[a]

Species	Gene		
	COI	ND1	ND3
A. suum	45.2	33.0	35.1
D. yakuba	47.2	33.5	25.8
Mouse	47.6	29.6	24.2

[a] Data for *A. suum, D. yakuba,* and mouse are taken from J.L. Macfarlane, D.O. Clary, J.A. Wahleithner, R. Okimoto, and D.R. Wolstenholme, unpublished; Clary and Wolstenholme (1985); and Bibb et al. (1981)

COI

```
F. hepatica           MSWLFTLDHK RVGLIYMVIG LWGGFFGLSL SVLVRLNYLD PYFNLVSPEV YNYVVTGHGI MMIFFFLMPV LIGGFGNYLL  80
D. yakuba          MS RQ...STN.. DI.TL.FIF. A.A.MV.T.. .I.I.AELGH .GALIGDDQI ..VI..A.AF I....MV..I M......W.V
mouse              MFI NR...STN.. DI.TL.LLF. A.AGMV.TA. .I.I.AELGQ .GAL.GDDQI ..VI..A.AF V....MV..M M......W.V

PLLLGIPDLN LPRLNALSAW LLLPAC-VCL SFGLM-GGVG VGWTFYPPLS SLD-YSSWGV DFLMFSLHLA GVSSLLGSIK FICTILEVML DEGTG-RHSI 180
..M..A..MA F..M.NM.F. ..P..LSLL. VSSMVEN.A. T...V..... .GIAHGGAS. .LAI..... .I..I..AVN ..T.VINMRS TGI.LD.MPL
..MI.A..MA F..M.NM.F. ..P.SFLLL. ASSMVEA.A. T...V....A GNPVHAGAS. .LTI..... ....I..A.N ..T..INMKP PAM.QYQTPL

LVWAYLFTSV LLLLSLPVLA AAITMLLFDR KFGSAFFDPM GGGDPVLFQH LFWFFGHPEV YVLILPGFGV ISHICVTLTN NDSLFGYYGL ILAMAAMVCL 280
F..SVVI.AL .......... G......T.. NLNTS....A .....I.Y.. .......... .I.......M ....ISQESG KKET..SL.M .Y..L.IGL.
F..SV.I.A. .......... .G.....T.. NLNTT....A .....I.Y.. .......... .I.......I ...VVTYYSG KKEP...M.M VW..MSIGF.

GSVVWAHHMF MVGLDVHTAV FFSSVTMVIG IPTGIKVFSW LMMLGGGSSV RMWDPVVWWI MGFIVLFTIG GVTGIMLSAS LLDTLLHDTW FVVAHFHYVL 380
.FI....... T..M..D.RA Y.T.A..I.A V.....I... .AT.H.TQL- -SYS.AIL.A L..VF...V. .L..VV.AN. SV.II....Y Y.........
.FI....... T.....D.RA C.T.A..I.A ....V..... .AT.H..NI- -K.S.AML.A L...F...V. .L...V..N. S..IV....Y Y.........

SLGSYSSVVI SFIWWWPVVT GYSLNLFMLQ GHWVVSMVGF NMCFFPMHYL GMAGLPRRVC VYDPDFYWLS -VVSSLGALV LAGSAFLFVF ILWESFVVGN 480
.M.AVFAIMA G..H.Y.LF. .LT..NKW.K SQFIIMFI.V .LT...Q.F. .L..M...YS D.-...A.TTW N...TI.STI SLLGILF.FY .I...L.SQR
.M.AVFAIMA G.VH.F.LFS .FT.DDTWAK A.FAIMF..V ..T...Q.F. .LS.M...YS D.-...A.TTW NT...M.SFI SLTAVLIMI. MI..A.ASKR

CVVS-SWGSS SLVLNVITLP GPQHNSYMNG VGRWVF*
Q.IYPIQLN. .IEWYQN.P. AEHSY.ELPL LTN*
E.M.V.YA.T N.EWLHGCP. PYHTFEEPTY .KVK*
```

ND1

```
F. hepatica           VVLLLKAFYL GL-SSFFAFV MIMVFVAFFM LGERKVLGYM QIRKGPNKVG LWGLLQSFAD LMKLVMKFKF VFFQNRSWLS  80
D. yakuba                  ME.I. S.IG.LLLII CVL.S...LT .L.......I .......... .M.IP.P.C. AI..FT.EQ- TYPLLSNY..
mouse                        INILTLL. P.LIAM..LT .V...I.... .L.....I.. PY.I..P... A...F..EPM RPLTTSMS.F

WWGVYLLVLL ACGYCVLFFF SFGGVSSVKF -MLWFLVVTS MTGYSLLSVG WGCYKKFALV SCVRSAFGSV SFEACFMCIV VLVALVWGSY GVSCLFGEFG 180
YYISPIFS.F LSLFVWMCMP F.VKLY.FNL GG.F..CC.. LGV.TVMVA. .SSNSNY..L GGL.AVAQTI .Y.VSLAL.M LSFIFLI... NMIYF.YYQI
IIAPT.SLT. .LSLW.PLPM PHPLINLNLG -I.FI.ATS. LSV..I.WS. .ASNS.YS.F GAL.AVAQTI .Y.VTMAI.L LS.L.MN... SLQT.ITTQE

GMWMVVPV-- VYGLWLVGML CECNRTPLDY AEAESELVSG LNTEYCNVPF TCLFACEYLI MVIFSWFSSL LFWGGSFVLG L--------- --ALFHVMFF 280
Y..FLIILFP MSLV..TIS. A.T....F.F ..G..... F.V..SSGG. ALI.MA..AS ILFM.MLFCV I.L.CDVFNL LFYVKLT--- ----.ISFV.
H..LLL.AWP MAMM.FIST. A.T..A.F.L T.G..... F.V..AAG.. ALF.MA..TN IILMNALTTI I.L.PLYYIN .PELYSTNFM ME..LLSST.

VWARATLPRV RYD---Y--- --FVSFMWRY ALLL-----L VFSFFFVI*
I...G....F ...KLM.LAW KC.L..SLN. L.FFIGFKI. L...LLW.FF SKKLMEN*
L.I..SY..F ...QLMHLLW KN.LPLTLAL CMWHISLPIF TAGVPPYM*
```

ND3

```
F. hepatica           MLFFAVLGLL FFLIFF-LVL VFHAFLWNLD LGIFSGERSW VSSFECGFLS QRVTENYFSY TYFILLVFFV VFDLEVSLLL  80
D. yakuba             IFSIIIIASV IL..TT-V.M FLASI.SKKA .----ID.EK S.P.....DP KSSSRLP..L RF.LITII.L I..V.IA.I.
mouse                    INLYTVI .IN.LLS.T. ILV..-.LPQ M----NLY.E ANPY....DP TSSARLP..M KF.LVAIT.L L....IA...

KMP----LQG VLYKKFFSYL FFLVLL-GIG FLVEVRRGYV RWAY*
P.IIILKYSN IMIWTIT.II .I.I..--.. LYH.WNQ.ML N.SN*
PL.--WAI.T IKTSTMMIMA .IL.TILSL. LAY.WTQKGL E.TE*
```

Fig. 3. Comparisons of the amino acid sequences of COI, ND1, and ND3 predicted from the nucleotide sequences of the respective genes of mtDNA of *Fasciola hepatica* with the corresponding amino acid sequences of *Drosophila yakuba* (Clary and Wolstenholme 1985) and mouse (Bibb et al. 1981). A dot indicates an amino acid that is conserved relative to *F. hepatica*. A dash indicates an amino acid that is absent. An asterisk indicates a termination codon or, for mouse ND1, a terminal T.

no nucleotides between the various *F. hepatica* mt-genes. All three of the *F. hepatica* mt-protein genes end with TAG (Fig. 2). Both the *F. hepatica* COI and ND3 genes begin with an ATG codon. However, in the ND1 gene the first ATG is located 67 nt downstream from the preceding $tRNA^{asp}$ gene (Fig. 2).

In mammalian mtDNA, all ATN codons, and in *D. yakuba* mtDNA ATT, ATA, and ATG are recognized as translation initiation codons (Anderson et al. 1981, 1982b; Bibb et al. 1981; Clary and Wolstenholme 1985). Arguments for the occasional use of GTG (valine; including the mouse ND1 gene) and in one case ATAA (the COI gene of *Drosophila*) as translation initiation codons in mt-protein genes have also been presented (see Clary and Wolstenholme 1983b, 1985 for discussion). Further, there is evidence that in nematode mtDNAs, TTG (leucine) is used as the translation initiation codon by 25–50% of the protein genes (R. Okimoto and D.R. Wolstenholme, unpublished data). The first ATN codon in the *F. hepatica* ND1 gene is located 61 nt downstream from the 3′ end of the $tRNA^{asp}$ gene, and this codon (ATA) corresponds to codon 16 (TGT, cysteine) in the *D. yakuba* ND1 gene and codon 9 (CCC, proline) in the mouse ND1 gene. However, beginning 4 nt downstream from the *F. hepatica* $tRNA^{asp}$ gene (and following a TAG) is an in-frame GTG codon, and separated from this GTG by one codon is a series of three TTG codons. From

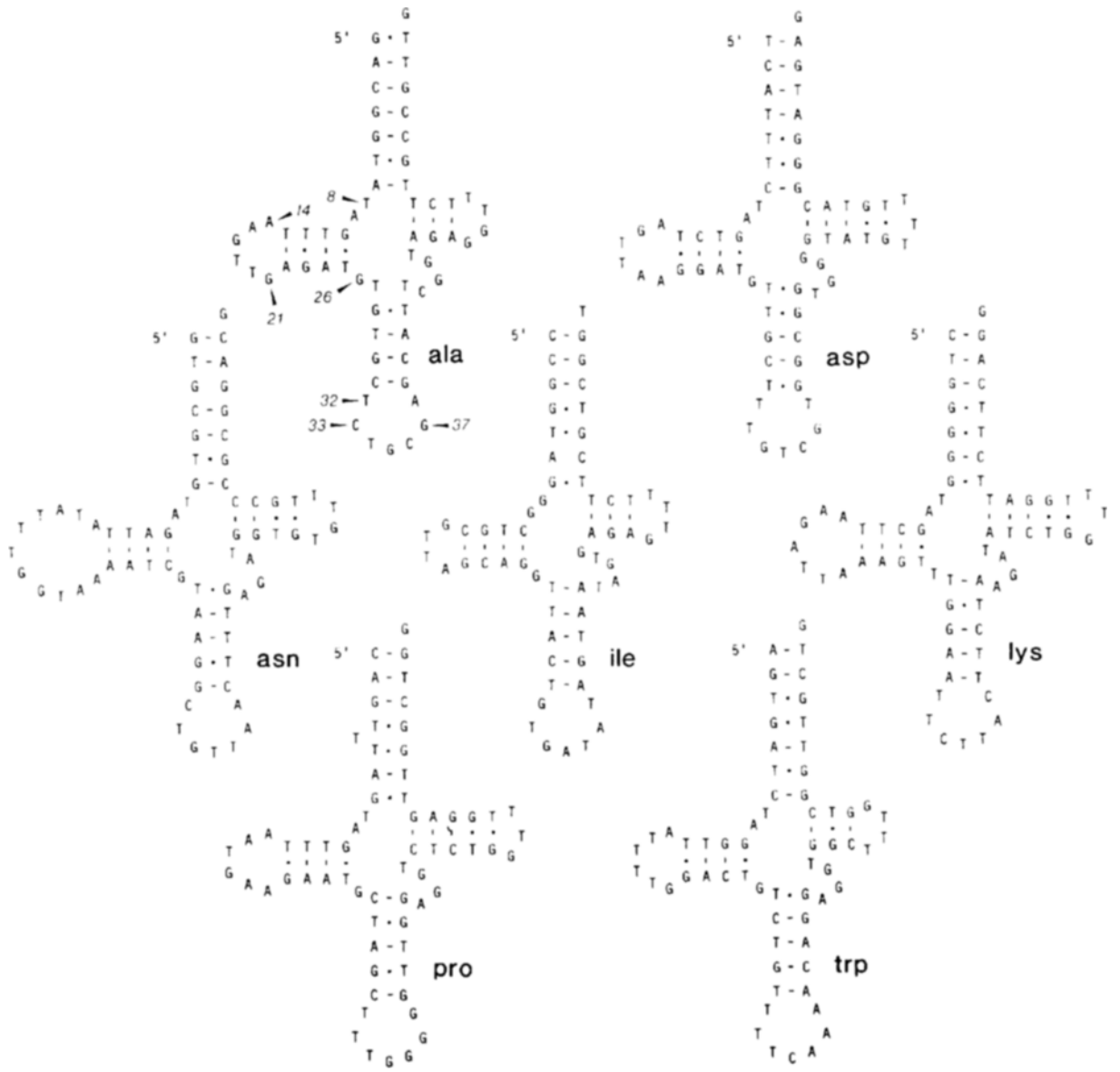

Fig. 4. Seven tRNA genes of *Fasciola hepatica* mtDNA shown in the presumed secondary structures of the corresponding tRNAs. The numbers shown on $tRNA^{ala}$ correspond to those used in Table 2 and follow the numbering system for yeast $tRNA^{phe}$ (Sprinzl et al. 1987).

the above considerations it seems plausible that either GTG or TTG could be the translation initiation codon of the *F. hepatica* ND1 gene.

Transfer RNA Genes

Each of the $tRNA^{ala}$, $tRNA^{asn}$, $tRNA^{asp}$, $tRNA^{ile}$, $tRNA^{lys}$, $tRNA^{pro}$, and $tRNA^{trp}$ genes identified in the *F. hepatica* mtDNA sequence can be folded into the four-arm cloverleaf structure characteristic of almost all known tRNAs (Sprinzl et al. 1987). The major structural features of these seven *F. hepatica* tRNA genes resemble those of tRNA genes from mammalian and *D. yakuba* mtDNAs (Fig. 4). The *F. hepatica* mt-tRNA genes range in size from 62 nt ($tRNA^{ile}$) to 70 nt ($tRNA^{asn}$) and vary in their sequence similarity to the corresponding mt-genes of *D. yakuba* from 35.3% (asp) to 44.8% (ala), and to the corresponding mt-tRNA genes of mouse from 36.6% (asp and lys) to 59.4% (pro). There is strict conservation of the size of the aminoacyl stem (7 ntp, with the possible exception of $tRNA^{pro}$—see below), the anticodon stem (5 ntp) and the anticodon loop (7 nt), and the dihydrouridine stem (4 ntp). The dihydrouridine loop ranges from 4 to 11 nt and the variable loop (between the anticodon stem and the TψC stem) is 5 nt in the $tRNA^{ile}$ gene and 4 nt in each of the remaining six tRNA genes. The TψC stem ranges from 3 to 5 ntp and the TψC loop is either 3 or 4 nt. However, the numbers of nucleotides in the TψC stem and loop are limited to three combinations: a stem of 3 ntp and a loop of 4 nt

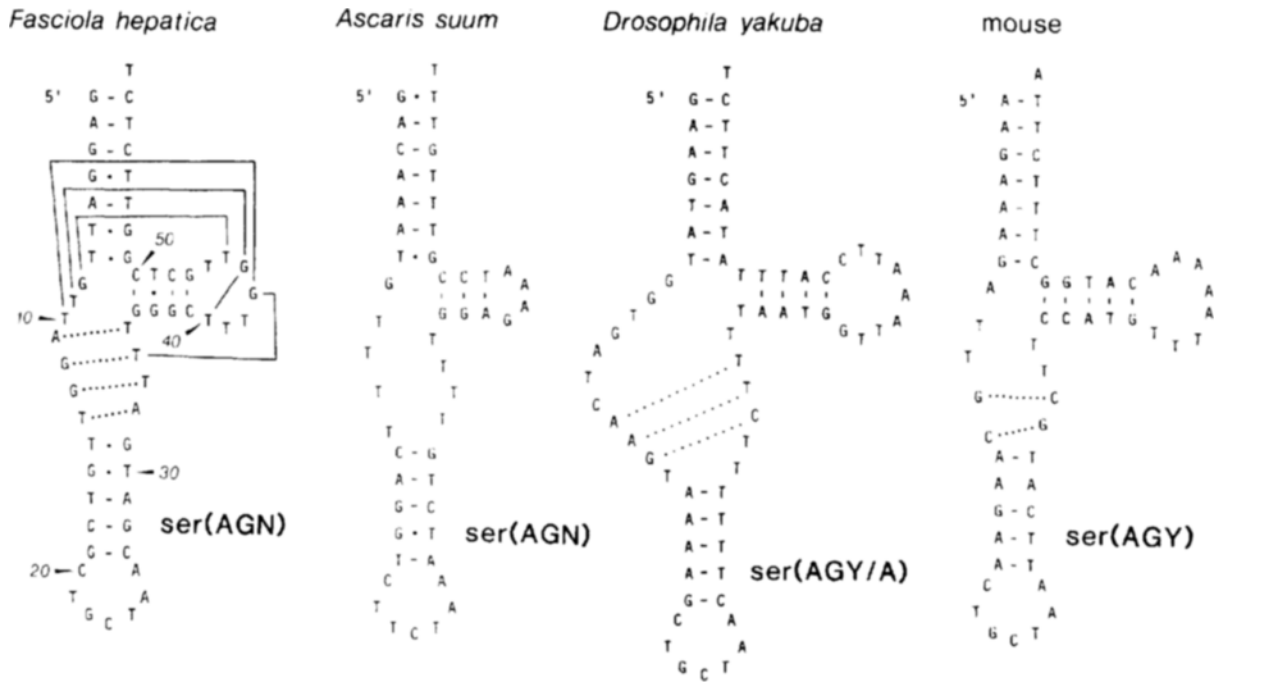

Fig. 5. Comparisons of the genes for the $tRNA^{ser}$ that recognizes AGN, AGY/A, or AGY codons in mtDNAs of *Fasciola hepatica*, *Ascaris suum*, *Drosophila yakuba*, and mouse. The latter three sequences are taken from Wolstenholme et al. (1987), Clary and Wolstenholme (1984), and Van Etten et al. (1982). In the diagram of the *F. hepatica* $tRNA^{ser}AGN$ gene, possible secondary interactions between nucleotides in the dihydrouridine arm replacement loop and the variable loop (dotted lines) and possible tertiary interactions (solid lines) similar to those proposed for $tRNA^{ser}AGY$ of bovine and other mammalian mitochondria (de Bruijn and Klug 1983) are shown. The numbering of *F. hepatica* $tRNA^{ser}AGN$ is specific for this gene.

($tRNA^{ala}$, $tRNA^{ile}$, and $tRNA^{trp}$); a stem of 5 ntp and a loop of 3 nt ($tRNA^{asp}$, $tRNA^{lys}$, and $tRNA^{pro}$); a stem of 4 ntp and a loop of 4 nt ($tRNA^{asn}$). Also, only G and T nucleotides are found in the TψC loops. In each of the four tRNA genes that contain 4 or 5 ntp in the TψC stem, two G·T pairs close the stem. Among the 28 stems of the seven *F. hepatica* tRNA genes, G·T pairs are absent from only five (TψC stem, $tRNA^{ala}$; dihydrouridine stem, $tRNA^{asn}$; dihydrouridine, anticodon, and TψC stems, $tRNA^{ile}$). Although G·T pairs are found in some of the tRNA genes of other metazoan mtDNAs, in no case is the frequency of distribution among stems as great as in the seven *F. hepatica* mt-tRNAs. Among the 28 stems of the corresponding tRNA genes, G·T pairs are absent from 21 in mouse mtDNA and 24 in *D. yakuba* mtDNA (Bibb et al. 1981; Clary et al. 1982, 1984; Clary and Wolstenholme 1983a,b, 1984). In the corresponding mt-tRNA genes of *A. suum* that lack the TψC arm, G·T pairs are found in 12 of the 21 stems (Wolstenholme et al. 1987). The wider distribution of G·T pairs among tRNA stems in both *F. hepatica* and *A. suum* mtDNAs may be a function of the higher frequency of G and lower frequency of A nucleotides in the sense strand of these mtDNAs, than is found in the sense strands of mouse and *D. yakuba* mtDNAs (Table 4).

As there is an unusually high frequency of G·T pairs in *F. hepatica* mt-tRNA gene stems, it is relevant to note that Freier et al. (1987) recently have shown significant negative free energy changes to be associated with G·U pairs in RNA secondary structures. It is likely, therefore, that the contribution of G·U pairs to tRNA stability is greater than expected from previously reported free energy increments (Borer et al. 1973; Bloomfield et al. 1974). Also, with regard to stem stability, it is noted that although there is a low frequency of C nucleotides in the sense strand of the *F. hepatica* mt-tRNA genes (12.2% versus 9.2% for protein genes, Table 4) there is at least one G–C pair in all of the 28 stems of the seven *F. hepatica* mt-tRNA genes, except two, the dihydrouridine stems of $tRNA^{ala}$ and $tRNA^{pro}$. Further, only three mismatched nucleotide pairs are found among the stems of the seven tRNA genes: a T T pair at the top of the anticodon stem of the $tRNA^{ala}$ gene; a G G pair and an A C pair in the aminoacyl stems of $tRNA^{asn}$ and $tRNA^{ile}$, respectively.

In order to fold the $tRNA^{pro}$ into a cloverleaf structure with a fully paired aminoacyl stem it is necessary to permit either nucleotide T_4 or nucleotide T_5 to loop out (Fig. 4). The occurrence of a similar loopout (of nucleotide T_5) has been suggested for the aminoacyl stem of $tRNA^{pro}$ of *D. yakuba* (Clary et al. 1984).

As has been found for other metazoan mt-tRNA genes, the trinucleotide CCA that occurs at the 3′ end of eukaryotic nuclear-encoded tRNA genes and some prokaryotic tRNA genes is absent from this position in *F. hepatica* mt-tRNA genes.

There is variation among vertebrate and *D. yakuba* mt-tRNA genes regarding the presence of nucleotides that are either invariant or semi-invariant among prokaryotic tRNAs and eukaryotic nuclear-encoded tRNAs (Barrell et al. 1979; Crews and Attardi 1980; Van Etten et al. 1980; Anderson et al. 1982a; Clary and Wolstenholme 1985; Roe et al. 1985). The least conserved of these nucleotides are those that are normally found in the dihydrouridine loop and TψC loop; among metazoan mt-tRNA genes these loops have considerable differences in size as well as sequence. Among the seven *F. hepatica* mt-tRNA genes many of the invariant and semi-invariant nucleotides of prokaryotic tRNAs and eukaryotic nuclear-encoded tRNAs that lie mainly outside of the dihydrouridine and TψC loops occur with high frequency (Table 2). These are the nucleotides that show the greatest degree of conservation in mouse, *D. yakuba,* and *A. suum* mt-tRNA genes (Bibb et al. 1981; Clary and Wolstenholme 1985; Wolstenholme et al. 1987). The G_{11}–C_{24} pair (rather than the standard pyrimidine$_{11}$–purine$_{24}$ pair) in the *F. hepatica* mt-tRNA^{trp} gene is also found in the mt-tRNA^{trp} genes of *A. suum, D. yakuba,* and various vertebrate mtDNAs (Wolstenholme et al. 1987). The standard T_{33} nucleotide in *F. hepatica* mt-tRNA^{ala} is a C as is the case in mammalian mt-tRNA$^{f\text{-}met}$ genes (Bibb et al. 1981; Anderson et al. 1982a).

The anticodon of the tRNA transcribed from the *F. hepatica* tRNA^{lys} gene would be CUU (Fig. 4), rather than UUU expected for a tRNA that recognizes both AAA and AAG codons. Although the tRNA^{lys} genes of vertebrate and *A. suum* mtDNAs have a TTT anticodon, the tRNA^{lys} gene of *D. yakuba* and mosquito (*Aedes albopictus*) mtDNAs also contain a CTT anticodon (Bibb et al. 1981; Clary and Wolstenholme 1983a; HsuChen et al. 1983; Wolstenholme et al. 1987). As in *F. hepatica* mtDNA, both AAA and AAG codons are used in *D. yakuba* mt-protein genes. The only other case of a C in the wobble position of the anticodon among metazoan mt-tRNA genes is the 5′ CAT anticodon of all known tRNA$^{f\text{-}met}$ genes. In this case, the C appears to recognize ATA and ATG codons and also ATT and ATC codons when they occur as translation initiation codons (Bibb et al. 1981).

The *F. hepatica* sequence that we have interpreted as a mt-tRNA^{ser}AGN gene has many of the features found in the corresponding mt-gene of vertebrates, *D. yakuba,* and *A. suum* (Fig. 5). The segment of nucleotides that lies between the aminoacyl and anticodon stems, that in other tRNA genes can be folded into the dihydrouridine stem and loop, comprises a loop (the dihydrouridine arm replacement loop) of only 7 nt in the *F. hepatica* mt-tRNA^{ser}AGN gene. The remaining secondary structural features of this tRNA^{ser}AGN gene are consistent with those of other mt-tRNA genes: an aminoacyl stem of 7 ntp, an anticodon arm with a stem of 5 ntp and a loop of 7 nt, a variable loop of 4 nt, and a TψC arm with a stem of 4 ntp and a loop of 7 nt. Also, the anticodon loop contains the highly conserved Y_{32}, T_{33}, and R_{37} nucleotides (Table 2). In the *F. hepatica* tRNA predicted from the tRNA^{ser}AGN gene (Fig. 5), the last 4 nt of the dihydrouridine arm replacement loop (A_{11}, G_{12}, C_{13}, T_{14}) could pair with the 4 nt of the variable loop (A_{32}, T_{33}, T_{34}, T_{35}), thus greatly extending the anticodon stem. A similar extension would be plausible in the tRNA^{ser}AGY of mammals and the tRNA^{ser}AGY/A of *Drosophila,* but not in the tRNA^{ser}AGN gene of *A. suum* (Fig. 5). A model for the tertiary structure of bovine tRNA^{ser}AGY, which resembles that of yeast tRNA^{phe} (Kim 1979) but is smaller and involves a unique set of interactions, has been derived by de Bruijn and Klug (1983). Both secondary and tertiary interactions similar to those proposed for the mammalian tRNA^{ser}AGY would

Table 2. The occurrence among eight mt-tRNA genes of *Fasciola hepatica* of some nucleotide or nucleotide combinations that are considered invariant (marked with an asterisk) or semi-invariant in prokaryotic and eukaryotic nuclear-encoded tRNAs (Dirheimer et al. 1979; Singhal and Fallis 1979)[a]

Nucleotide or nucleotide combination[b]	*F. hepatica* tRNA gene							
	ala	asn	asp	ile	lys	pro	trp	ser-(AGN)
*T8	+	+	+	G	+	+	+	o
A9	+	+	+	G	+	+	+	o
R10–Y25	+	+	+	CG	+	+	+	o
*Y11–R24	+	AT	+	+	+	+	GC	o
Y13–R22	+	+	+	+	+	+	+	o
*A14	+	+	+	G	+	+	+	o
*R21	+	+	+	+	+	+	+	o
R26	+	+	+	+	T	+	+	o
Y27–R43	TT	+	+	+	+	+	+	+
*Y32	+	+	+	G	+	+	+	+
*T33	C	+	+	+	+	+	+	+
*R37	+	+	+	+	+	+	+	+

[a] The numbering system used is that given for yeast tRNA^{phe} (Sprinzl et al. 1987) and the numbers correspond to those shown on tRNA^{ala} in Fig. 4. For each *F. hepatica* tRNA gene, + indicates the presence of the invariant or semi-invariant nucleotide or nucleotide combination. Open circles in the tRNA^{ser}AGN gene column indicate a lack of corresponding nucleotides due to the peculiar structure of this tRNA gene

[b] A, adenine; T, thymine; R, adenine or guanine; Y, cytosine or thymine

Table 3. Codon usage in *Fasciola hepatica* mtDNA

Phe	TTT	100	Ser	TCT	36	Tyr	TAT	39	Cys	TGT	21
	TTC	3		TCC	2		TAC	1		TGC	0
Leu	TTA	31		TCA	1	Ter	TAA	0	Trp	TGA	8
	TTG	91		TCG	0		TAG	3		TGG	29
Leu	CTT	20	Pro	CCT	20	His	CAT	17	Arg	CGT	16
	CTC	0		CCC	2		CAC	1		CGC	3
	CTA	2		CCA	2	Gln	CAA	2		CGA	1
	CTG	3		CCG	4		CAG	6		CGG	3
Ile	ATT	35	Thr	ACT	22	Asn	AAT	20	Ser	AGT	25
	ATC	1		ACC	0		AAC	1		AGC	2
Met	ATA	10		ACA	1	Lys	AAA	7		AGA	1
	ATG	33		ACG	3		AAG	10		AGG	12
Val	GTT	83	Ala	GCT	33	Asp	GAT	21	Gly	GGT	57
	GTC	3		GCC	1		GAC	0		GGC	2
	GTA	6		GCA	3	Glu	GAA	7		GGA	2
	GTG	21		GCG	2		GAG	12		GGG	26

also be plausible in the *F. hepatica* tRNA^{ser}AGN gene: the aforementioned pairing between nucleotides in the dihydrouridine arm replacement loop and the variable loop; three tertiary interactions between nucleotides in the dihydrouridine arm replacement loop and the TψC loop (G_8–T_{45}, T_9–G_{44}, T_{10}–G_{43}); T_{34} in the variable loop and G_{43} in the TψC loop (Fig. 5). Also, the intraloop interaction between the first and fifth nucleotides of the TψC loop that is a U–A pair in both bovine mt-tRNA^{ser}AGY and yeast tRNA^{phe} might occur but in the form of a U·G pair (T_{40}–G_{44}). Similar secondary and tertiary interactions are possible in the *D. yakuba* tRNA^{ser}AGY/A (Clary and Wolstenholme 1984), but most seem unlikely to occur in the *A. suum* tRNA^{ser}AGN (Fig. 5).

Codon Usage, Nucleotide Composition, and the Genetic Code

Codon usage among the three *F. hepatica* mt-protein genes is summarized in Table 3. The frequency of codons ending in G (27.8%) or T (60.8%) is remarkably high (88.6%) compared to *D. yakuba* and mouse mtDNAs and reflects the overall nucleotide composition of the sense strand of the three protein genes (Table 4). A high frequency of codons ending in G or T has also been found in *A. suum* mt-protein genes (Table 4; Wolstenholme et al. 1987). Codons ending in C are rare in *F. hepatica* mtDNA as is the case in both *D. yakuba* and *A. suum* mtDNAs, but not in mammalian mtDNAs. Among *D. yakuba* and mammalian mtDNAs the ratio of TTR to CTN triplets used as leucine-specifying codons is positively correlated with the differential use of T and C nucleotides in the third position of codons (Clary and Wolstenholme 1985). Data from the *F. hepatica* mt-protein genes support the validity of this correlation: the ratio of TTR:CTN codons is 4.8:1 and the ratio of T:C nucleotides in the third position of codons is 25.3:1 (Tables 3 and 4).

All codons expected to specify an amino acid are found among the three *F. hepatica* mt-protein genes except CTC (leucine), TCG (serine), ACC (threonine), GAC (aspartic acid), and TGC (cysteine), all of which contain at least one C, the least frequently used nucleotide in the sense strand of these genes (Table 3).

Internal TGA codons occur in each of the three *F. hepatica* mt-protein genes (Fig. 2 and Table 3). Of a total of eight TGA codons, five and four correspond in position to tryptophan-specifying codons (TGA or TGG) in the corresponding *D. yakuba* and mouse mt-protein genes, respectively (Figs. 2 and 3). This is consistent with the interpretation that in *F. hepatica* mtDNA, as in other metazoan mtDNAs (and fungal mtDNAs), TGA specifies tryptophan (Barrell et al. 1979, 1980; Fox 1979; Bonitz et al. 1980; Heckmann et al. 1980; Clary and Wolstenholme 1985). The much higher ratio of TGG:TGA codons (29:8) in *F. hepatica* than in *D. yakuba* mtDNA (6:96) and mouse mtDNA (7:97) reflects different proportions of G and A nucleotides in the sense strands of these three respective mtDNAs (Table 4).

Twelve AGG and AGA triplets occur as internal codons among the *F. hepatica* COI, ND1, and ND3 genes (Fig. 2 and Table 3). None of these 13 codons correspond in position to codons (CGN) specifying arginine in the *D. yakuba* and mouse COI, ND1, and ND3 genes (Figs. 2 and 3). In the three *F. hepatica* mt-protein genes, the 13 AGG and AGA codons correspond in position to codons that specify a total of nine different amino acids in the corresponding *D. yakuba* and mouse genes [for *D. yakuba*, serine (7), proline (1), phenylalanine (1), threonine (1), glycine (1), valine (1), aspartic acid (1); for mouse, serine (4), isoleucine (2), proline (2), glycine (2), leu-

Table 4. Nucleotide comparison data for the protein genes of various metazoan mtDNAs[a]

	Percentage nucleotide composition of the sense strand						Percentage of codons ending in:					
	T	C	A	G	G+C	G+T	T	C	A	G	G+C	G+T
F. hepatica	49.3	9.2	14.8	26.7	35.9	76.0	60.8	2.4	9.9	27.8	30.2	88.6
A. suum[b]	51.8	8.0	18.7	21.5	29.5	73.3	62.7	2.4	11.5	23.4	25.8	86.1
D. yakuba[c]	44.4	11.1	32.3	12.2	23.3	56.6	48.4	3.3	45.4	2.9	6.2	51.3
Mouse[d]	29.4	26.0	33.2	11.4	37.4	40.8	22.9	27.4	46.6	3.2	30.6	26.1

[a] Data for mouse mtDNA are derived from all protein genes except the ND6 gene as the frequency of nucleotides in the third positions of codons of mammalian ND6 genes are complementary to those of the other 12 genes (see Clary and Wolstenholme 1985)
[b] Unpublished data from J.L. Macfarlane, D.O. Clary, J.A. Wahleithner, R. Okimoto, and D.R. Wolstenholme
[c] Data from Clary and Wolstenholme (1985)
[d] Data from Bibb et al. (1981)

cine (1), valine (1), threonine (1); Figs. 2 and 3]. However, the AGR codons in the *F. hepatica* mt-protein genes correspond in position to more than twice as many serine-specifying codons than to codons specifying any other amino acid in the equivalent *D. yakuba* and mouse protein genes. In view of this and the findings that AGA and AGG specify serine in other invertebrate mt-genetic codes (Wolstenholme and Clary 1985; Himeno et al. 1987; Wolstenholme et al. 1987; R. Okimoto, J.L. Macfarlane, and D.R. Wolstenholme, unpublished data), it seems reasonable to conclude that AGA and AGG also specify serine in the *F. hepatica* mt-genetic code.

In *A. suum* mtDNA there is a gene for a $tRNA^{ser}$ with a UCU anticodon that is available to decode all four AGN codons (Wolstenholme et al. 1987). This is consistent with the finding that in mammalian mitochondria, a single tRNA with a U in the wobble position of the anticodon can apparently decode all four codons of a four-codon family (Barrell et al. 1980). In mammalian mtDNA, in which AGT and AGC (but not AGA and AGG) are used to specify serine, the anticodon of the tRNA that decodes these codons is GCU (Anderson et al. 1981; Van Etten et al. 1982). A GCU anticodon is also found in the single tRNA that recognizes the AGT, AGC, and AGA serine-specifying codons in *D. yakuba* mtDNA (Clary and Wolstenholme 1984), implying either that the G in the wobble position can pair with A or that the GCU–AGA anticodon–codon interaction is dependent upon two out of three pairing (Lagerkvist 1981). The anticodon of the *F. hepatica* mt-tRNA that would be expected to recognize AGT and AGC codons is also GCU. If, in accordance with what has been found in *D. yakuba* and *A. suum* mtDNAs, there is only one tRNA to decode all four AGN codons in *F. hepatica* mtDNA, then either the G in the wobble position of the GCU anticodon would have to pair with both A and G in AGA and AGG codons, or the two remaining base pairs (G–C and A–U) must be sufficient to stabilize the GCU–AGA and GCU–AGG interactions. A similar situation seems likely in echinoderms: only one $mt\text{-}tRNA^{ser}AGN$ gene per mt-genome has been found, and all AGN codons are used to specify serine in mt-protein genes (Himeno et al. 1987; Cantatore et al. 1988).

As ATA specifies methionine in the mt-genetic code of all other metazoans examined (and yeast) (Barrell et al. 1979, 1980; Hudspeth et al. 1982; Wolstenholme and Clary 1985), we have tentatively assigned methionine to ATA in the predicted amino acid sequences of the three *F. hepatica* mt-protein genes (Figs. 2 and 3). However, it is noted that comparisons of amino acid sequences of the *F. hepatica* mt-protein genes and the equivalent *D. yakuba* and mouse genes (Fig. 3) provide insufficient evidence to confirm this assignment.

Conclusions

It is generally agreed that the ancestral line of modern Platyhelminthes diverged from the vertebrate ancestral line before divergence from the vertebrate line of the lines that led to nematodes and insects (see for example Wilson et al. 1978). Our finding that AGA and AGG specify serine in the *F. hepatica* mt-genetic code indicates, therefore, that the switch in specificity of these codons from arginine to serine must have occurred very early in, or before, the evolution of metazoa.

Data from the present and other sequencing studies (Wolstenholme and Clary 1985; Himeno et al. 1987; Wolstenholme et al. 1987; Cantatore et al. 1988) suggest that the serine-specifying capacity of AGA and AGG codons may have been retained in all invertebrates. The apparent absence of AGG codons from *Drosophila* mtDNAs may be a function of the extremely low use of G in the third position of codons in these mtDNAs (Clary and Wolstenholme 1985). In amphibian and mammalian mt-protein genes AGA and AGG are either not used, or they signal translation termination (see Bibb et al. 1981; Roe et al. 1985). This further switch in specificity could have taken place either just before

or following the development of vertebrates. If the AGA and AGG codons that occur in the 3′ region of some vertebrate mt-protein genes do, in fact, signal termination, then this could be accomplished either by recognition of AGA and AGG by a protein synthesis termination factor, or simply by failure of the anticodon of any tRNA to interact with an AGA or AGG codon. However, in the latter case, lack of recognition of AGA and AGG by the tRNA that decodes AGT and AGC (serine) codons must involve more than specificity of the tRNA anticodon. This is because the anticodon of the vertebrate $tRNA^{ser}AGY$ is GCU, the same anticodon that in the *Drosophila* $tRNA^{ser}AGY/A$ recognizes AGA as well as AGT and AGC, and that in the *F. hepatica* $tRNA^{ser}AGN$ and the echinoderm $tRNA^{ser}AGN$ may recognize all four AGN codons.

Our finding that *F. hepatica* mtDNA, like all other metazoan mtDNAs examined to date, contains a gene for a $tRNA^{ser}AGN$ that has a dihydrouridine arm replacement loop, indicates that loss of the dihydrouridine arm in this tRNA gene again occurred at a very early time with regard to metazoan evolution.

The relative arrangement of tRNA and protein genes in the sequenced segment of the *F. hepatica* mtDNA molecule suggests that there is a considerable difference in relative gene order between the mtDNAs of *F. hepatica,* nematodes, *Drosophila,* echinoderms, and vertebrates. Such a difference is consistent with the accepted evolutionary relationships of Platyhelminthes to the other metazoans whose mtDNAs have been analyzed: within vertebrates (amphibia and mammals), mtDNA gene order is identical; between vertebrate and *Drosophila* mtDNAs there is a difference in gene order for only 11 tRNA genes, the 2 rRNA genes, and 5 protein genes, but between mtDNAs of nematodes and *Drosophila,* and nematodes and mammals extensive rearrangements involving almost all tRNA, rRNA, and protein genes have occurred (Bibb et al. 1981; Clary and Wolstenholme 1985; Roe et al. 1985; Wolstenholme et al. 1987).

Examination of metazoan mt-tRNA gene sequences indicates that the nucleotides that are highly conserved in the dihydrouridine loop and TψC stem and loop of prokaryotic tRNAs and eukaryotic nuclear-encoded tRNAs have been substituted or lost in corresponding mt-tRNA genes at different times in the various evolutionary lines. The finding that tRNA genes of *F. hepatica* mtDNA, other than the $tRNA^{ser}AGN$ gene, can be folded into four-armed secondary structures suggests that mt-tRNAs containing a TψC arm replacement loop are limited to either nematodes or to nematodes and members of closely related phyla. This supports the view that loss of the TψC arm in mt-tRNAs occurred only once, sometime after the evolutionary lines that led to nematodes and to higher invertebrates and vertebrates diverged from each other (Wolstenholme et al. 1987).

Acknowledgments. This work was supported by a grant from the University of Utah Research Committee and by National Institutes of Health Grant no. GM-18375.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1982a) Comparison of the human and bovine mitochondrial genomes. In: Slonimski P, Borst P, Attardi G (eds) Mitochondrial genes. Cold Spring Harbor Laboratory, Cold Spring Harbor NY, pp 5–43

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

Arcari P, Brownlee GG (1980) The nucleotide sequence of a small (3S) seryl-tRNA (anticodon GCU) from beef heart mitochondria. Nucleic Acids Res 8:5207–5212

Barnes WM, Bevan M, Son PH (1983) Kilo-sequencing: creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. Methods Enzymol 101: 98–122

Barrell BG, Bankier AT, Drouin J (1979) A different genetic code in human mitochondria. Nature 282:189–194

Barrell BG, Anderson S, Bankier AT, de Bruijn MHL, Chen E, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1980) Different pattern of codon recognition by mammalian mitochondrial tRNAs. Proc Natl Acad Sci USA 77:3164–3166

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

Bloomfield VA, Crothers DM, Tinoco I Jr (1974) Physical chemistry of nucleic acids. Harper and Row, New York

Bonitz SG, Berlani R, Coruzzi G, Li M, Macino G, Nobrega FG, Nobrega MP, Thalenfeld BE, Tzagoloff A (1980) Codon recognition rules in yeast mitochondria. Proc Natl Acad Sci USA 77:3167–3170

Borer PN, Dengler B, Tinoco I Jr (1973) Stability of ribonucleic acid double stranded helices. J Mol Biol 86:843–853

Cantatore P, Roberti M, Morisco P, Rainaldi G, Gadaleta MN, Saccone C (1987) A novel gene order in the *Paracentrotus lividus* mitochondrial genome. Gene 53:41–54

Cantatore P, Roberti M, Rainaldi G, Saccone C, Gadaleta MN (1988) Clustering of tRNA genes in *Paracentrotus lividus* mitochondrial DNA. Curr Genet 13:91–96

Chomyn A, Attardi G (1987) Mitochondrial gene products. Curr Top Bioenerg 15:295–329

Clary DO, Wolstenholme DR (1983a) Nucleotide sequence of a segment of *Drosophila* mitochondrial DNA that contains the genes for cytochrome *c* oxidase subunits II and III and ATPase subunit 6. Nucleic Acids Res 11:4211–4227

Clary DO, Wolstenholme DR (1983b) Genes for cytochrome *c* oxidase subunit I, URF2 and three tRNAs in *Drosophila* mitochondrial DNA. Nucleic Acids Res 11:6859–6872

Clary DO, Wolstenholme DR (1984) A cluster of six tRNA genes in *Drosophila* mitochondrial DNA that includes a gene for an unusual $tRNA^{ser}AGY$. Nucleic Acids Res 12:2367–2379

Clary DO, Wolstenholme DR (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J Mol Evol 22:252–271

Clary DO, Wolstenholme DR (1987) *Drosophila* mitochondrial DNA: conserved sequences in the A+T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. J Mol Evol 25:116–125

Clary DO, Goddard JM, Martin SC, Fauron CMR, Wolstenholme DR (1982) *Drosophila* mitochondrial DNA: a novel gene order. Nucleic Acids Res 10:6619–6637

Clary DO, Wahleithner JA, Wolstenholme DR (1984) Sequence and arrangement of the genes for cytochrome b, URF1, URF4L, URF4, URF5, URF6 and five tRNAs in *Drosophila* mitochondrial DNA. Nucleic Acids Res 12:3747–3762

Crews S, Attardi G (1980) The sequences of the small ribosomal RNA gene and the phenylalanine tRNA gene are joined end to end in human mitochondrial DNA. Cell 19:775–784

Dale RMK, McClure BA, Houchins JP (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid 13:31–40

de Bruijn MHL (1983) *Drosophila melanogaster* mitochondrial DNA: a novel organization and genetic code. Nature 304: 234–241

de Bruijn MHL, Klug A (1983) A model for the tertiary structure of mammalian mitochondrial transfer RNAs lacking the entire dihydrouridine loop and stem. EMBO J 2:1309–1321

de Bruijn MHL, Schreier P, Eperon IC, Barrell BG, Chen EY, Armstrong PW, Wong JFH, Roe BA (1980) A mammalian mitochondrial serine transfer RNA lacking the "dihydrouridine" loop and stem. Nucleic Acids Res 8:5213–5222

Dirheimer G, Keith G, Sibler A-P, Martin RP (1979) The primary structure of tRNAs and their rare nucleosides. In: Schimmel PR, Soll D, Abelson JN (eds) Transfer RNA: structure, properties and recognition. Cold Spring Harbor Laboratory, Cold Spring Harbor NY, pp 19–41

Dubin DT, HsuChen CC, Cleaves GR, Timko KD (1984) Sequence and structure of a serine transfer RNA with GCU anticodon from mosquito mitochondria. J Mol Biol 176:251–260

Foran DR, Hixson JE, Brown WM (1988) Comparisons of ape and human sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. Nucleic Acids Res 16: 5841–5861

Fox TD (1979) Five TGA "stop" codons occur within the translated sequence of the yeast mitochondrial gene for cytochrome *c* oxidase subunit II. Proc Natl Acad Sci USA 76: 6534–6538

Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Nielson T, Turner DH (1986) Improved free-energy parameters for prediction of RNA duplex stability. Proc Natl Acad Sci USA 83:9373–9377

Garesse R (1988) *Drosophila melanogaster* mitochondrial DNA: gene organization and evolutionary considerations. Genetics 118:649–663

Grosskopf R, Feldmann H (1981) Analysis of a DNA segment from rat liver mitochondria containing the genes for cytochrome oxidase subunits I, II and III, ATPase6 and several tRNA genes. Curr Genet 4:151–158

Heckman JE, Sarnoff J, Alzner-de Weerd B, Yin S, RajBhandary UL (1980) Novel features in the genetic code and codon reading patterns in *Neurospora crassa* mitochondria based on sequences of six mitochondrial tRNAs. Proc Natl Acad Sci USA 77:3159–3163

Himeno H, Masaki H, Kawai T, Ohta T, Kumagi I, Miura I, Watanabe K (1987) Unusual genetic codes and a novel gene structure for $tRNA^{ser}AGY$ in starfish mitochondrial DNA. Gene 56:219–230

Hong GF (1982) A systematic DNA sequencing strategy. J Mol Biol 158:539–549

HsuChen CC, Cleaves GR, Dubin DT (1983) A major lysine tRNA with a CUU anticodon in insect mitochondria. Nucleic Acids Res 11:8659–8662

HsuChen CC, Kotin RM, Dubin DT (1984) Sequence of the coding and flanking regions of the large ribosomal subunit RNA gene of mosquito mitochondria. Nucleic Acids Res 12: 7771–7785

Hudspeth MES, Ainley WM, Shumard DS, Butow RA, Grossman LI (1982) Location and structure of the var1 gene of yeast mitochondrial DNA: nucleotide sequence of the 40.0 allele. Cell 30:617–626

Kim SH (1979) Crystal structure of yeast $tRNA^{phe}$ and general structural features of other tRNAs. In: Schimmel PR, Soll D, Abelson JN (eds) Transfer RNA: structure, properties and recognition. Cold Spring Harbor Laboratory, Cold Spring Harbor NY, pp 83–100

Lagerkvist U (1981) Unorthodox codon reading and the evolution of the genetic code. Cell 23:305–306

Lipman DJ, Pearson WR (1985) Rapid and sensitive protein similarity searches. Science 227:1435–1441

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor NY

Messing J, Vieira J (1982) A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. Gene 19:269–276

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

Moritz C, Dowling T, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu Rev Ecol Syst 18:269–292

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470–474

Pepe G, Holtrop M, Gadaleta G, Kroon AM, Cantatore P, Gallerani R, De Benedetto C, Quagliariello E, Sbisa E, Saccone C (1983) Non random patterns of nucleotide substitutions and codon strategy in the mammalian mitochondrial genes coding for identified and unidentified reading frames. Biochem Int 6:553–563

Roe BA, Wong JFH, Chen EY, Armstrong PW, Stankiewicz A, Ma DP, McDonough J (1982) Mammalian mitochondrial tRNAs: a modified nucleotide 3′ to the anticodon may modulate their codon response. In: Slonimski P, Borst P, Attardi G (eds) Mitochondrial genes. Cold Spring Harbor Laboratory, Cold Spring Harbor NY, pp 45–49

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

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

Singhal RP, Fallis PAM (1979) Structure, function, and evolution of transfer RNAs. Prog Nucleic Acid Res Mol Biol 23: 227–290

Snyder M, Fraser AR, LaRoche J, Gartner-Kepkay KE, Zouros E (1987) A typical mitochondrial DNA from the deep-sea

scallop *Placopecten magellanicus*. Proc Natl Acad Sci USA 84:7595–7599

Sprinzl M, Hartmann T, Meissner F, Moll J, Vorderwülbecke T (1987) Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res 15:r53–r188

Staden R (1982) Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res 10:4731–4751

Van Etten RA, Walberg MW, Clayton DA (1980) Precise localization and nucleotide sequence of the two mouse mitochondrial rRNA genes and three immediately adjaent novel tRNA genes. Cell 22:157–170

Van Etten RA, Michael NL, Bibb MJ, Brennicke A, Clayton DA (1982) Expression of the mouse mitochondrial DNA genome. In: Slonimski P, Borst P, Attardi G (eds) Mitochondrial genes. Cold Spring Harbor Laboratory, Cold Spring Harbor NY, pp 73–88

Warrior R, Gall J (1985) The mitochondrial DNA of *Hydra attenuata* and *Hydra littoralis* consists of two linear molecules. Arch Sci (Geneva) 38:439–445

Wilson EO, Eisner T, Briggs WR, Dickerson RE, Metzenberg RL, O'Brien RD, Susman M, Boggs WE (1978) Life on earth. Sinauer Assoc Inc, Sunderland MA

Wolstenholme DR, Clary DO (1985) Sequence evolution of *Drosophila* mitochondrial DNA. Genetics 109:725–744

Wolstenholme DR, Fauron CMR (1976) A partial map of the circular mitochondrial genome of *Drosophila melanogaster*: location of *Eco*RI-sensitive sites and the adenine–thymine-rich region. J Cell Biol 71:434–448

Wolstenholme DR, Macfarlane JL, Okimoto R, Clary DO, Wahleithner JA (1987) Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. Proc Natl Acad Sci USA 84:1324–1328

Received August 22, 1988/Accepted November 7, 1988