Volume 12 Number 5 1984 Nucleic Acids Research

A cluster of six tRNA genes in *Drosophila* mitochondrial DNA that includes a gene for an unusual $tRNA^{ser}_{AGY}$

Douglas O.Clary and David R.Wolstenholme

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

Received 22 November 1983; Accepted 1 February 1984

ABSTRACT

Genes for URF3, $tRNA^{ala}$, $tRNA^{arg}$, $tRNA^{asn}$, $tRNA^{ser}_{AGY}$, $tRNA^{glu}$, $tRNA^{phe}$, and the carboxyl terminal segment of the URF5 gene have been identified within a sequenced segment of the mtDNA molecule of Drosophila yakuba. The genes occur in the order given. The URF5 and $tRNA^{phe}$ genes are transcribed in the same direction as replication while the URF3 and remaining five tRNA genes are transcribed in the opposite direction. Considerable differences exist in the relative arrangement of these genes in D. yakuba and mammalian mtDNA molecules. In the $tRNA^{ser}_{AGY}$ gene an eleven nucleotide loop, within which secondary structure formation seems unlikely, replaces the dihydrouridine arm, and both the variable loop (six nucleotides) and the TψC loop (nine nucleotides) are larger than in any other D. yakuba tRNA gene. As available evidence is consistent with AGA codons specifying serine rather than arginine in the Drosophila mitochondrial genetic code, the possibility is considered that the 5'GCU anticodon of the D. yakuba $tRNA^{ser}_{AGY}$ gene can recognize AGA as well as AGY codons.

INTRODUCTION

Within nucleotide sequences of the mitochondrial DNA (mtDNA) molecule of Drosophila yakuba we have identified and reported the genes for two ribosomal RNAs, twelve tRNAs and eight polypeptides (cytochrome c oxidase (CO) subunits I, II and III, ATPase subunit 6, Unidentified Reading Frames (URF) 1, 2 and 3 and A6L) (1-4). Nucleotide sequences of sections of the D. melanogaster mtDNA molecule, corresponding to some of the sequenced sections of the D. yakuba mtDNA molecule have also been determined (1,2,5). While the genes found in Drosophila mtDNAs are the same as the genes found in a number of mammalian mtDNAs (6-8), their arrangement within the Drosophila and mammalian mtDNA molecules differs considerably (1-5). Also, evidence obtained from comparisons of sequences of Drosophila and mammalian polypeptide genes suggests that in the Drosophila mitochondrial genetic code the triplet AGA, which specifies arginine in the standard genetic code, and is used only as a rare termination codon in some mammalian mtDNAs (6-8), may specify serine (3-

5). Further, the COI gene of Drosophila mtDNA appears to be unique in that it may contain a four nucleotide initiation codon (4,5).

In this paper we report a sequence of the D. yakuba mtDNA molecule that contains a cluster of six tRNA genes ($tRNA^{ala}$, $tRNA^{arg}$, $tRNA^{asn}$, $tRNA^{ser}_{AGY}$, $tRNA^{glu}$ and $tRNA^{phe}$) bounded by the genes for URF3 and URF5. As shown previously, the URF3 gene of D. yakuba mtDNA is one of a series of polypeptide genes (URF2, COI, COII, URFA6L, ATPase6, COIII and URF3) which occur in the same relative order, and are transcribed in the same direction, in D. yakuba and mammalian mtDNAs (2-4). However, relative to the URF3 gene, all of the six presently described D. yakuba mt-tRNA genes, as well as the URF5 gene, are found in different locations in the mammalian mtDNA molecule. Further, the structure of the tRNA predicted from the $tRNA^{ser}_{AGY}$ gene contains some unusual features both in comparison to other Drosophila mt-tRNAs, and to its mammalian counterpart.

MATERIALS AND METHODS

Experimental details regarding isolation of mtDNA from D. yakuba (Stock 2371.6, Ivory Coast), preparation and identification of pBR322 and pBR325 clones of D. yakuba mtDNAs, restriction enzyme digestions, electrophoresis, cloning of fragments into M13mp8 or M13mp9 DNA, purification of single-stranded and double-stranded M13 DNAs and preparation of viral DNAs containing partial deletions of cloned restriction fragments of mtDNA are given or referred to in (1,3).

All DNA sequences were obtained from M13mp8- or M13mp9-cloned fragments by the extension-dideoxyribonucleotide termination procedure (9) using [α-^{32}P] dATP (600 Ci/mmol; New England Nuclear) as described (1). The sequencing strategies used are given in Fig. 1.

Computer analyses of DNA sequences were carried out using the programs referred to in (3).

RESULTS AND DISCUSSION

The sequence of a continuous 1629 nucleotide section of the D. yakuba mtDNA molecule, determined using the strategies indicated in Fig. 1, is shown in Fig. 2. Within the sequence are five regions each of which can fold into the characteristic secondary structure of a tRNA, with anticodons indicating them to be the genes for $tRNA^{ala}$, $tRNA^{arg}$, $tRNA^{asn}$, $tRNA^{glu}$ and $tRNA^{phe}$ (Fig. 3). A sixth region, which lies between the $tRNA^{asn}$ and $tRNA^{glu}$ genes, can also fold into a structure which resembles a tRNA, but this structure lacks a

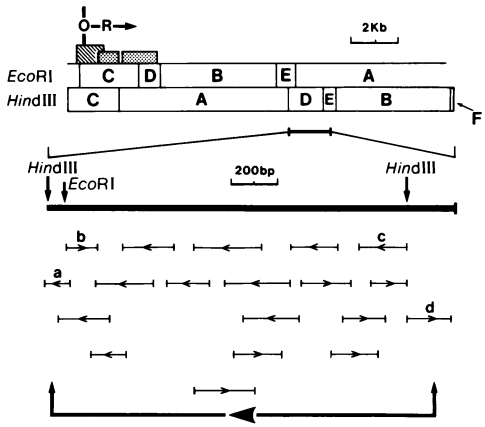
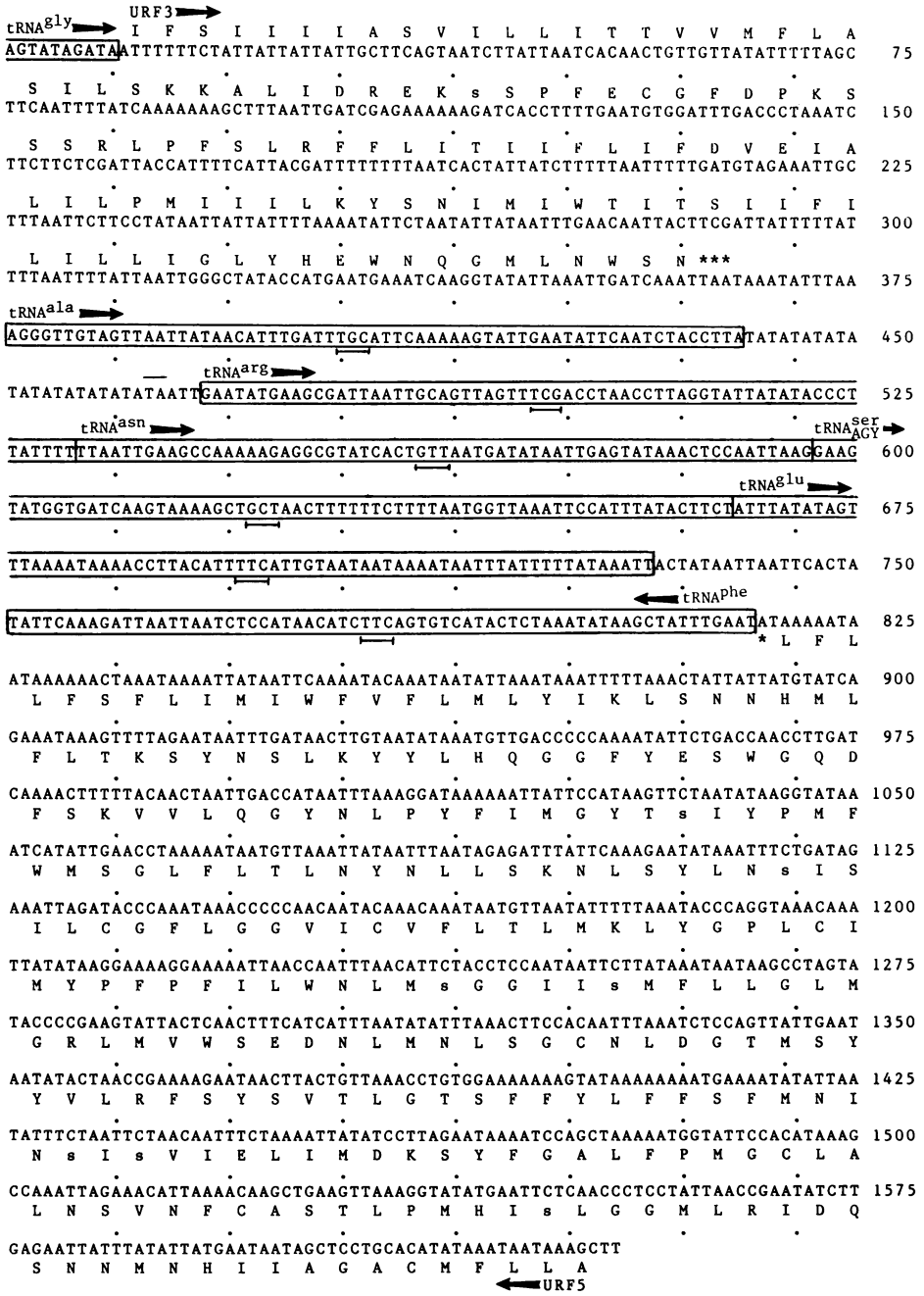


Figure 1. A map of the D. yakuba mtDNA molecule showing the relative locations of the A+T-rich region (crosshatched) the two rRNA genes (dotted), the origin (O) and direction (R) of replication, and EcoRI and HindIII sites and fragments (A-E and A-F respectively), (see (1) for details and references). The bar under the map indicates the segment sequenced. This segment is expanded below and the restriction sites and strategies employed to obtain the entire nucleotide sequence are shown. The origin of each sequence is as follows: a, the EcoRI-E fragment. b and c, the two ends of the 1.5 kb HindIII-EcoRI subfragment of the pBR325-cloned EcoRI-A fragment. d, the HindIII-E fragment which was subcloned from the cloned EcoRI-A fragment. Unlabelled sequences with arrows pointing to the right are DNaseI-EcoRI deletions (cloned in M13mp9) of the 1.5 kb HindIII-EcoRI subfragment of the EcoRI-A fragment. Unlabelled sequences with arrows pointing to the left are DNaseI-BamHI deletions (cloned in M13mp9) of the HindIII-D fragment. The small vertical terminal arrows on the lower bar indicate the extent of the sequence shown in Fig. 2, and the solid arrowhead shows the 5'-3' direction of the sequence.

standard dihydrouridine arm. Because 5'GCT is found in the anticodon position of this structure, and for other reasons discussed below, this sequence is interpreted as being a gene for $tRNA^{ser}_{AGY}$. The $tRNA^{phe}$ gene is transcribed in the same direction as that in which replication proceeds around the molecule, while the remaining tRNA genes are transcribed in the opposite direction (Fig. 4).

The sequence to the left of the $tRNA^{ala}$ gene contains an open reading frame of 351 nucleotides which would be transcribed in the same direction as the $tRNA^{ala}$ gene and which has been identified as the URF3 gene; this nucleotide sequence and the amino acid sequence predicted from it are 52.4% and 41.9% homologous, respectively, to the corresponding sequences of the mouse URF3 gene (8). The 5' end of the D. yakuba URF3 gene is located


URF3
tRNAgly
I F S I I I I A S V I L L I T T V V M F L A
AGTATAGATAATTTTTTCTATTATTATTATTGCTTCAGTAATCTTATTAATCACAACTGTTGTTATATTTTTAGC 75
S I L S K K A L I D R E K s S P F E C G F D P K S
TTCAATTTTATCAAAAAAAGCTTTAATTGATCGAGAAAAAAGATCACCTTTTGAATGTGGATTTGACCCTAAATC 150
S S R L P F S L R F F L I T I I F L I F D V E I A
TTCTTCTCGATTACCATTTTCATTACGATTTTTTTTAATCACTATTATCTTTTTAATTTTTGATGTAGAAATTGC 225
L I L P M I I I L K Y S N I M I W T I T S I I F I
TTTAATTCTTCCTATAATTATTATTTTAAAATATTCTAATATTATAATTTGAACAATTACTTCGATTATTTTTAT 300
L I L L I G L Y H E W N Q G M L N W S N ***
TTTAATTTTATTAATTGGGCTATACCATGAATGAAATCAAGGTATATTAAATTGATCAAATTAATAAATATTTAA 375
tRNAala
AGGGTTGTAGTTAATTATAACATTTGATTTGCATTCAAAAAGTATTGAATATTCAATCTACCTTATATATATATA 450
tRNAarg
TATATATATATATAATTGAATATGAAGCGATTAATTGCAGTTAGTTTCGACCTAACCTTAGGTATTATATACCCT 525
tRNAasn
tRNAserAGY
TATTTTTTAATTGAAGCCAAAAAAGAGGCGTATCACTGTTAATGATATAATTGAGTATAAACTCCAATTAAGGAAG 600
tRNAglu
TATGGTGATCAAGTAAAAGCTGCTAACTTTTTTCTTTTAATGGTTAAATTCCATTTATACTTCTATTTATATAGT 675
TTAAAATAAAACCTTACATTTTCATTGTAATAATAAAATAATTTATTTTTATAAATTACTATAATTAATTCACTA 750
tRNAphe
TATTCAAAGATTAATTAATCTCCATAACATCTTCAGTGTCATACTCTAAATATAAGCTATTTGAATATAAAAAATA 825
* L F L
ATAAAAAACTAAATAAAATTATAATTCAAAATACAAATAATATTAAATAAATTTTTAAACTATTATTATGTATCA 900
L F S F L I M I W F V F L M L Y I K L S N N H M L
GAAATAAAGTTTTAGAATAATTTGATAACTTGTAATATAAATGTTGACCCCCAAAATATTCTGACCAACCTTGAT 975
F L T K S Y N S L K Y Y L H Q G G F Y E S W G Q D
CAAAACTTTTTACAACTAATTGACCATAATTTAAAGGATAAAAAATTATTCCATAAGTTCTAATATAAGGTATAA 1050
F S K V V L Q G Y N L P Y F I M G Y T s I Y P M F
ATCATATTGAACCTAAAAATAATGTTAAATTATAATTTAATAGAGATTTATTCAAAGAATATAAATTTCTGATAG 1125
W M S G L F L T L N Y N L L S K N L S Y L N s I S
AAATTAGATACCCAAATAAACCCCCAACAATACAAACAAATAATGTTAATATTTTTAAATACCCAGGTAAACAAA 1200
I L C G F L G G V I C V F L T L M K L Y G P L C I
TTATATAAGGAAAAGGAAAAATTAACCAATTTAACATTCTACCTCCAATAATTCTTATAAATAATAAGCCTAGTA 1275
M Y P F P F I L W N L M s G G I I s M F L L G L M
TACCCCGAAGTATTACTCAACTTTCATCATTTAATATATTTAAACTTCCACAATTTAAATCTCCAGTTATTGAAT 1350
G R L M V W S E D N L M N L S G C N L D G T M S Y
AATATACTAACCGAAAAGAATAACTTACTGTTAAACCTGTGGAAAAAAAGTATAAAAAAAATGAAAATATATTAA 1425
Y V L R F S Y S V T L G T S F F Y L F F S F M N I
TATTTCTAATTCTAACAATTTCTAAAATTATATCCTTAGAATAAAATCCAGCTAAAAATGGTATTCCACATAAAG 1500
N s I s V I E L I M D K S Y F G A L F P M G C L A
CCAAATTAGAAACATTAAAACAAGCTGAAGTTAAAGGTATATGAATTCTCAACCCTCCTATTAACCGAATATCTT 1575
L N S V N F C A S T L P M H I s L G G M L R I D Q
GAGAATTATTTATATTATGAATAATAGCTCCTGCACATATAAATAATAAAGCTT
S N N M N H I I A G A C M F L L A
URF5

adjacent to the $tRNA^{gly}$ gene, as shown previously (Fig. 2; (2)). Alignment of the nucleotide and amino acid sequences of the URF3 genes of D. yakuba and mouse indicate that three codons (nine nucleotides; 68-70; 104-106; 113-115, Fig. 2) of the D. yakuba gene are absent from the mouse gene. The URF3 genes of both D. yakuba and mouse have a TAA termination codon. In mouse this codon is separated by a single A from the 5' terminus of the $tRNA^{arg}$ gene, but in D. yakuba it is separated by eleven nucleotides from the 5' terminus of the $tRNA^{ala}$ gene.

The sequence of 811 nucleotides to the right of the $tRNA^{phe}$ gene contains only one open reading frame which proceeds in the same direction as the sense strand of the $tRNA^{phe}$ gene. The 270 amino acid sequence predicted from this open reading frame is 25.7% homologous to the 271 amino acid sequence which precedes the carboxyl terminus of the mouse URF5 gene (8). This alignment (not shown) includes insertion/deletions of only three internal codons between the Drosophila and mouse genes, all within 108 nucleotides of the carboxyl terminus. Interpretation of this D. yakuba sequence as the URF5 gene is strengthened by the finding that the amino acid sequences corresponding to the first 67 codons (nucleotides 1427-1627, Fig. 2) of the D. yakuba and mouse sequences have a homology of 55.2%. In mouse mtDNA the URF5 and URF6 genes are transcribed from opposite strands of the molecule, the carboxyl termini of the two genes overlap and each gene terminates with a standard TAA codon (8). In D. yakuba mtDNA the last codon of the URF5 gene is separated by a single T from the 5' terminus of the $tRNA^{phe}$ gene. Presumably, as has been argued for other genes with similar arrangements in both mammalian (6,10) and Drosophila (2-5) mtDNAs, a translation termination codon is created at the end of the D. yakuba URF5 transcript by polyadenylation when this transcript is

Figure 2. Nucleotide sequence of the segment of the D. yakuba mtDNA molecule identified in Fig. 1. The sequence contains a cluster of six tRNA genes (boxed) the anticodons of which are underlined. From considerations of nucleotide and predicted amino acid sequence homologies to mouse mtDNA (8) the sequence to the left of the tRNA genes is the entire URF3 gene, and the sequence to the right of the tRNA genes is the carboxyl terminal segment of the URF5 gene. The nucleotide sequence shown is the (5'-3') sense strand of all of the genes except $tRNA^{phe}$ and URF5, and the arrows indicate the direction of transcription of each gene. Asterisks indicate partial or complete termination codons. In the amino acid sequences a lower case letter s indicates the tentative assignment of serine to an AGA codon (see text). The line above nucleotides 463 and 464 indicates the lack of a TA dinucleotide in one of the cloned fragments of this region of the sequence. The sequence containing the amino terminal 88 nucleotides of the URF3 gene and the entire $tRNA^{gly}$ gene are published in (2).

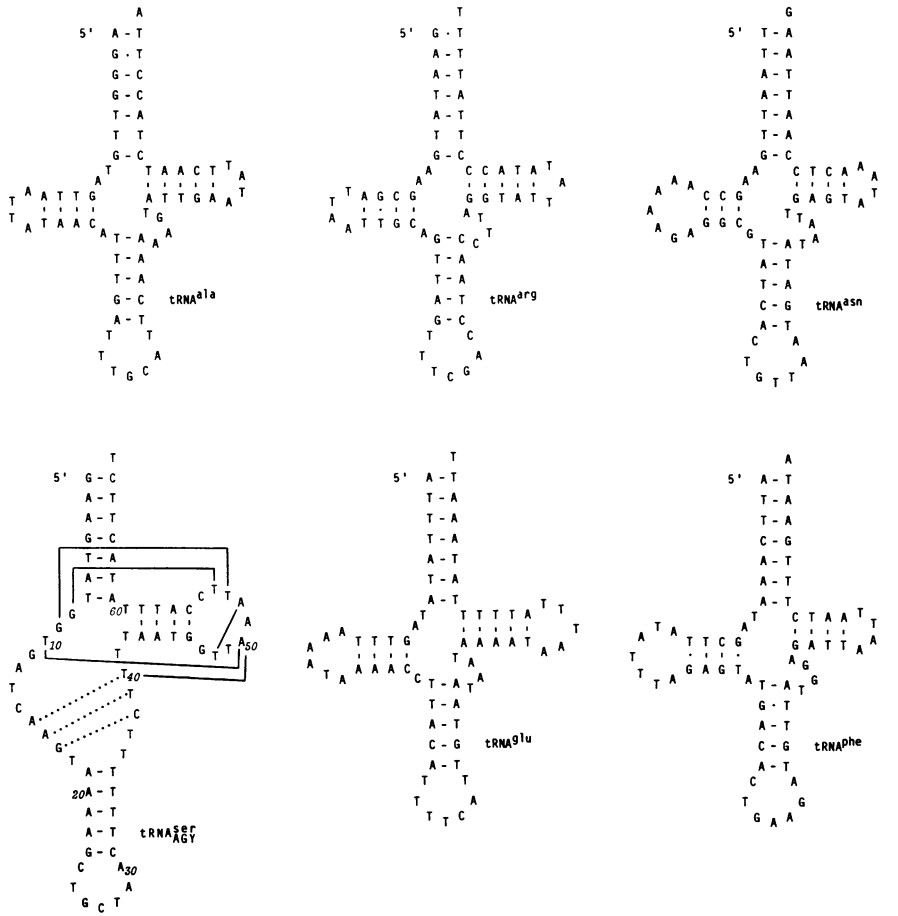


Figure 3. Six tRNA genes of D. yakuba mtDNA shown in the presumed secondary structure of the corresponding tRNAs. Also indicated in the diagram of the $tRNA^{ser}_{AGY}$ gene are possible secondary interactions between nucleotides in the dihydrouridine replacement loop and the variable loop (dotted lines), and other possible tertiary interactions (solid lines) proposed for $tRNA^{ser}_{AGY}$ from bovine and other mammalian mitochondria (23; see text). The numbering shown is specific for this tRNA gene.

excised from a multicistronic primary transcript.

Intervening nucleotides are not found between the central four tRNA genes (arg, asn, ser(AGY) and glu) of the six tRNA gene cluster (Fig. 2), nor do any of these genes overlap. However, between the $tRNA^{ala}$ and $tRNA^{asn}$ genes are 27 nucleotides which include a twelve unit TA repeat (Fig. 2). In a second sequence of this region which was derived from a different mtDNA molecule, the

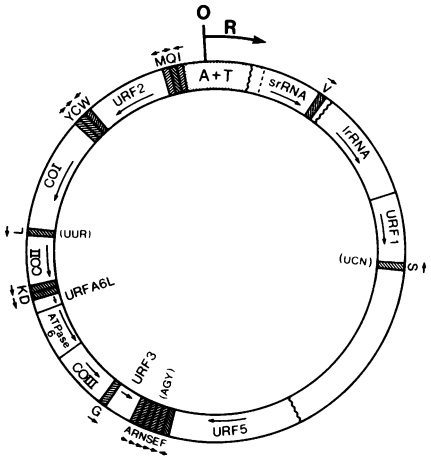


Figure 4. The arrangement of genes determined to date in the circular mtDNA molecule of D. yakuba, from the results of various studies (for references see 1-4) and the present data. Each tRNA gene is identified by the one letter amino acid code. Arrows within and outside the molecule indicate the direction of transcription of each gene. Wavy lines indicate uncertain gene boundaries. O and R indicate the origin and direction of replication respectively.

sequence between $tRNA^{ala}$ and $tRNA^{asn}$ genes included only an eleven unit TA repeat. Whether this difference reflects heterogeneity in the size of the intervening sequence between the $tRNA^{ala}$ and $tRNA^{asn}$ genes in different mtDNA molecules, either within or between flies, or is the product of a replication error or of recombination during clone amplification, remains unknown. Between $tRNA^{glu}$ and $tRNA^{phe}$ there are 18 nucleotides, the only notable feature of which is a direct tetranucleotide repeat at each end (Fig. 2).

We have shown previously that the genes for URF2, COI, COII, URFA6L, ATPase6, COIII and URF3 are found in the same relative order in D. yakuba and mammalian mtDNAs, but that the tRNA genes which separate some of these polypeptide genes differ between the two mtDNAs. All of the six tRNA genes reported here as well as URF5 are in different relative locations in the D. yakuba and mammalian mtDNA molecules. In mammalian mtDNA the $tRNA^{arg}$ gene is adjacent to the URF3 gene; the $tRNA^{ala}$ and $tRNA^{asn}$ genes are adjacent to each other and located between the $tRNA^{trp}$ and $tRNA^{cys}$ genes; the $tRNA^{ser}_{AGY}$ gene lies between the $tRNA^{his}$ and $tRNA^{leu}_{CUN}$ genes (which are located between the URF4 and URF5 genes); the $tRNA^{phe}$ gene is located between the replication origin-containing region and the large rRNA gene; and the URF5 gene is located between the $tRNA^{leu}_{CUN}$ and URF6 genes. Also, of these six tRNA genes and the URF5 gene, only the $tRNA^{arg}$ and $tRNA^{ser}_{AGY}$ genes are transcribed in the same direction in D. yakuba and mammalian mtDNAs, relative to both the direction of replication and to the seven polypeptide genes listed above.

The relative locations, and directions of transcription of all of the D. yakuba mitochondrial genes which have been determined to date are shown in

Fig. 4. The directions of transcription of the newly mapped six tRNA genes and of the URF5 gene add to the trend which we have noted earlier (3); genes in the approximate half of the D. yakuba mtDNA molecule which lies to the left of the A+T-rich region (Fig. 4) are transcribed from one strand of the molecule, while genes in the other half of the molecule are transcribed from the other strand (Fig. 4). The only exceptions to this generalization so far are three tRNA genes (gln, cys, tyr) to the left of the A+T-rich region, and one tRNA gene (ser, UCN) to the right. This contrasts with what is found in mammalian mtDNA, where all genes except the URF6 gene and eight tRNA genes are transcribed from one strand (the H strand) of the molecule (6-8).

Transfer RNA genes. The $tRNA^{ala}$, $tRNA^{arg}$, $tRNA^{asn}$, $tRNA^{glu}$ and $tRNA^{phe}$ genes of D. yakuba mtDNA are 62%, 32%, 56%, 54% and 56% homologous, respectively, to the corresponding tRNA genes of mouse mtDNA (8). The value for $tRNA^{arg}$ is the lowest for any of the 18 D. yakuba-mouse tRNA gene homology comparisons made so far. However, the D. yakuba mt-$tRNA^{arg}$ gene is 84% homologous to the recently published sequence of $tRNA^{arg}$ from mitochondria of mosquito, Aedes albopictus (11). The presently reported five D. yakuba mt-tRNA genes are similar to the other twelve D. yakuba mt-tRNA genes we have described previously (1-4) in regard to their major secondary structural characteristics (Fig. 3). Also, each of these tRNA genes lacks some of the nucleotides which are constant in prokaryotic and non-organelle eukaryotic tRNAs. However, in each of the $tRNA^{ala}$, $tRNA^{arg}$, $tRNA^{asn}$ and $tRNA^{phe}$ genes of D. yakuba mtDNA, as in all other D. yakuba mt-tRNA genes, the constant Pu_{26}, T_{33} and Pu_{37} nucleotides are maintained (numbering system in 12). In the $tRNA^{glu}$ gene T_{33} and Pu_{37} are also found, but nucleotide 26 is a pyrimidine (C) rather than a purine. The $tRNA^{glu}$ gene is again unusual in that its G+C content is only 9%, compared to a range of 24% to 31% for the remaining five newly described mt-tRNA genes. Among the other D. yakuba tRNA genes sequenced, only the $tRNA^{asp}$ gene has a similar low (9%) G+C content. In both the $tRNA^{glu}$ and $tRNA^{asp}$ genes both the amino-acyl and the TΨC stems lack G-C pairs (3). Among stems of other D. yakuba mt-tRNA genes G-C pairs are absent only in the amino-acyl stem of $tRNA^{gln}$. Within the $tRNA^{glu}$ gene stability of secondary structure may be aided by the absence of unusual nucleotide pairs within stems (2), and by the stacking effect expected from runs of A's and T's (Fig. 3), as has been noted for the amino-acyl stems of the $tRNA^{asp}$ and $tRNA^{gln}$ genes (3).

As mentioned above, we have interpreted the sequence between the $tRNA^{asn}$ and $tRNA^{glu}$ genes as a $tRNA^{ser}_{AGY}$ gene. This sequence can be folded into a

structure (Fig. 3) which has an amino-acyl stem and an anticodon stem with the same number of nucleotide pairs (seven and five, respectively), and an anticodon loop with the same number of nucleotides (seven) as are found in all other D. yakuba mt-tRNA genes, and almost all other known tRNAs (12,13). The TψC arm could comprise either a stem of five nucleotide pairs (other D. yakuba mt-tRNA genes contain four or five) and a loop of nine nucleotides (other D. yakuba mt-tRNA genes contain between three and eight), or a six nucleotide pair stem and a seven nucleotide loop. The variable loop of six nucleotides is larger than the variable loops (four or five nucleotides) of other D. yakuba, other insect, and mammalian mt-tRNAs (2,5,11-14), but is quite small compared to the variable loops of some fungal mt-tRNAs (15,16), prokaryotic and non-organelle eukaryotic tRNAs (12,13). Eleven nucleotides occur between the last nucleotide of the amino-acyl stem (nucleotide 7, numbering system shown in Fig. 3) and the first nucleotide of the anticodon stem (nucleotide 19), but these nucleotides cannot be arranged to form a structure resembling the usual dihydrouridine arm of other tRNAs. Nucleotides T_{10}, G_{11} and A_{12} are complementary to T_{13}, C_{14} and A_{15}, as might be expected if they were the conserved dihydrouridine stem; however, if this is the case then two of the A_{16}, G_{17} and T_{18} nucleotides must have been added during the evolution of this gene from a gene coding a tRNA with a standard cloverleaf configuration. The three nucleotides in the anticodon position, 5'GCT, would be expected to recognize the triplets AGU and AGC, which in all genetic codes known specify serine. Recently, the sequence of an mt-tRNA isolated from mosquito tissue culture cells has been obtained (D. T. Dubin, personal communcation) which is 91% homologous to the D. yakuba $tRNA^{ser}_{AGY}$ gene sequence, and can be folded into a structure similar to that shown in Fig. 3 for the D. yakuba $tRNA^{ser}_{AGY}$ gene sequence.

In each of the mt-$tRNA^{ser}_{AGY}$ genes of nine different mammals (hamster (17), bovine (18,19), human (19), mouse (8,20), rat (21), gorilla, chimpanzee, orangutan and gibbon (22)) the dihydrouridine arm is replaced by a loop of only five nucleotides. The variable loop has a maximum of five nucleotides, the TψC stem comprises five nucleotide pairs and the TψC loop is always eight or nine nucleotides. Based on information concerning tertiary structure obtained from experiments involving chemical probing, de Bruijn and Klug (23) have derived a structural model for bovine $tRNA^{ser}_{AGY}$. This model is similar in overall shape to the crystal structure of yeast $tRNA^{phe}$ (24-26) but is smaller, and involves a unique set of tertiary interactions. The latter include interactions of all five nucleotides in the dihydrouridine arm

replacement loop, two with nucleotides in the variable loop which effectively extends the anticodon stem, and three with nucleotides in the TψC loop. The structural model proposed for bovine $tRNA^{ser}_{AGY}$ is also compatible with the tRNAs predicted from the other eight described mammalian $tRNA^{ser}_{AGY}$ genes, if some A-C base pairing is permitted (23). In the D. yakuba tRNA predicted from the $tRNA^{ser}_{AGY}$ gene (Fig. 3), A_{15}, A_{16}, and G_{17} in the dihydrouridine arm replacement loop could pair with C_{38}, T_{39} and T_{40} in the variable loop. Also, the three tertiary interactions between nucleotides in the dihydrouridine arm replacement loop and the TψC loop thought to occur in the mammalian $tRNA^{ser}_{AGY}$ (23) might be maintained (G_8-T_{54}; G_9-T_{53}; T_{10}-A_{50}, Fig. 3) as well as interaction between T_{40} and A_{50}. The structural model presented for mt-$tRNA^{ser}_{AGY}$ from bovine and other mammals includes an intraloop U-A pair (also found in yeast $tRNA^{phe}$ (24,25)) formed between the first and fifth nucleotides of the TψC loop. Such an interaction in the D. yakuba $tRNA^{ser}_{AGY}$ gene (Fig. 3) seems more likely between the second and sixth nucleotides (T_{48}-A_{52}) of the TψC loop than between G_{47} and A_{51}.

Codon usage and the genetic code. The frequency of codons ending in A or T in the URF3 gene (93.2%) and in the carboxyl terminal segment of the URF5 gene (91.9%) are within the range of values (91.6% to 98.1%) found for the other six polypeptide genes of D. yakuba mtDNA which have been analysed (3,4). The occurrence of the leucine-specifying codon CTG (nucleotides 899-901, Fig. 2) in the URF5 gene leaves only four expected sense codons (CAG, GAG, CGC, and AGG) unrecorded among the seven completed polypeptide gene sequences and the partial URF1 and URF5 sequences (1-4).

The G+C content of the URF3 gene (20.6%) and the partial URF5 gene (21.9%) are slightly higher than the G+C contents of the URF2 and URFA6L genes (18.6% and 17.2%, respectively), but still distinctly lower than the range of G+C contents (24.2% to 30.1%) found for the D. yakuba COI, COII, COIII and ATPase6 genes (3,4).

As in all other D. yakuba genes for which sequences have been obtained, internal TGA and internal AGA codons are found in both the URF3 gene and the partial URF5 gene (Fig. 2). It appears that, as in fungal and mammalian mtDNAs (8,27-29), TGA codons specify tryptophan in the Drosophila mitochondrial genetic code (1,3-5). While two out of three of the TGAs of the D. yakuba URF3 sequence correspond to tryptophan-specifying codons (TGAs) in the mouse URF3 sequence, none of the three TGA codons of the D. yakuba URF5 sequence correspond to tryptophan-specifying codons in the mouse URF5 sequence. The difference in conservation of TGA codons between the URF3 and

URF5 genes of D. yakuba and mouse correlates with differences in overall nucleotide sequence homology between these genes (52.4% and 39.7%, respectively), as noted previously for other D. yakuba-mouse mitochondrial gene comparisons (4).

It has been argued from comparisons of mtDNA sequences of Drosophila and mammals (3-5), and from considerations of the characteristics of mitochondrial genetic codes (3), that AGA may specify serine rather than arginine in the Drosophila mitochondrial genetic code. The D. yakuba URF3 and URF5 genes contain one and seven AGA codons respectively, which correspond in position to codons specifying serine, asparagine (two), phenylalanine, alanine, threonine, histidine and cysteine in the mouse URF3 and URF5 genes. A similar low correspondence of serine-specifying codons to AGA codons has been found previously for other mammal-Drosophila comparisons involving mitochondrial genes with relatively low overall nucleotide sequence homologies (3-5).

As D. yakuba mtDNA contains a gene for a tRNA with a 5'GCU anticodon which is expected to recognize only AGY codons, it remains unclear as to how AGA codons are decoded. Two possibilities seem worth considering: The D. yakuba mtDNA molecule may contain one extra tRNA gene with a TCT anticodon (1) which is specific for AGA (and possibly AGG). If this is the case then D. yakuba mtDNA would contain three tRNAs to decode serine, a situation which is inconsistent with the otherwise overall economy of structure and organization of this mtDNA molecule. The alternative possibility is that there is stable pairing between the GCU anticodon and each of the three AGC, AGU and AGA codons. This would necessitate either selective two-out-of-three nucleotide pair recognition (30,31), or that the G in the wobble position of the anticodon can effectively pair with C, U or A. As AGG codons have not been found in Drosophila gene sequences so far, there is no need to include the prediction of G-G pairing. Other cases of unusual base pairing of the nucleotide in the wobble position of the anticodon have been indicated for mtDNAs. In mammalian mtDNAs the 5'CAU anticodon of $tRNA^{f\text{-}met}$ appears to be able to recognize as methionine-specifying codons all four AUN codons when they occur as initiator codons, and AUG and AUA when they occur internally. A similar situation is indicated in Drosophila mtDNA, except that evidence for only AUG and AUU initiation codons (and one possible special case of an AUA initiation codon (4,5)) has been found so far (1-5). Also, in Drosophila mtDNA, while genes contain both AAA and AAG codons, only a $tRNA^{lys}$ gene with a 5'CTT anticodon has been located (3,5). It seems reasonable, therefore, to give serious consideration to the view that in Drosophila mitochondria, the

GCU anticodon can recognize AGA as well as AGY codons.

ACKNOWLEDGEMENTS

We wish to thank Donald T. Dubin for sending us data on mosquito tRNAs prior to publication, and Felix Machuca for excellent technical assistance. This work was supported by National Institutes of Health Grant No. GM 18375.

REFERENCES

1. Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M.-R. and Wolstenholme, D.R. (1982) Nucl. Acids Res. 10, 6619-6637.
2. Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R. (1983) Nucl. Acids Res. 11, 2411-2425.
3. Clary,D.O. and Wolstenholme,D.R. (1983a) Nucl. Acids Res. 11, 4211-4227.
4. Clary,D.O. and Wolstenholme,D.R. (1983b) Nucl. Acids Res. 11, 6859-6872.
5. de Bruijn,M.H.L. (1983) Nature 304, 234-241.
6. Anderson,S., Bankier,A.T., Barrell,B.G., de Bruijn,M.H.L., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A., Sanger,F., Schreier, P.H., Smith,A.J.H., Staden,R. and Young,I.G. (1981) Nature 290, 457-465.
7. Anderson,S., de Bruijn,M.H.L., Coulson,A.R., Eperon,I.C., Sanger,F. and Young,I.G. (1982) J. Mol. Biol. 156, 683-717.
8. Bibb,M.J., Van Etten,R.A., Wright,C.T., Walberg,M.W. and Clayton,D.A. (1981) Cell 26, 167-180.
9. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
10. Ojala,D., Montoya,J. and Attardi,G. (1981) Nature 290, 470-474.
11. HsuChen,C.C., Cleaves,G.R. and Dubin,D.T. (1983) Plasmid 10, 55-65.
12. Gauss,D.H. and Sprinzl,M. (1983a) Nucl. Acids Res. 11, r1-r53.
13. Gauss,D.H. and Sprinzl,M. (1983b) Nucl. Acids Res. 11, r55-r103.
14. HsuChen,C.C., Cleaves,G.R. and Dubin,D.T. (1984) Nucl. Acids. Res. in press.
15. Kochel,H.G., Lazarus,C.M., Basak,N. and Kuntzel,H. (1981) Cell 23, 625-633.
16. Netzker,R., Kochel,H.G., Basak,N. and Kuntzel,H. (1982) Nucl. Acids Res. 10, 4783-4794.
17. Baer,R.J. and Dubin,D.T. (1980) Nucl. Acids Res. 8, 3603-3610.
18. Arcari,P. and Brownlee,G.G. (1980) Nucl. Acids Res. 8, 5207-5212.
19. de Bruijn,M.H.L., Schreier,P.H., Eperon,I.C., Barrell,B.G., Chen,E.Y., Armstrong,P.W., Wong,J.F.H. and Roe,B.A. (1980) Nucl. Acids Res. 8, 5213-5222.
20. Van Etten,R.A., Michael,N.L., Bibb,M.J., Brennicke,A. and Clayton,D.A. (1982) in Mitochondrial Genes, Slonimski,P., Borst,P. and Attardi,G., Eds., pp. 73-88, Cold Spring Harbor Laboratory.
21. Grosskopf,R. and Feldmann,H. (1981) Curr. Gen. 4, 191-196.
22. Brown,W.M., Prager,E.M., Wang,A. and Wilson,A.C. (1982) J. Mol. Evol. 18, 225-239.
23. de Bruijn,M.H.L. and Klug,A. (1983) EMBO J. 2, 1309-1321.
24. Kim,S.-H., Suddath,F.L., Quigley,G.J., McPherson,A., Sussman,J.C., Wang,A., Seeman,N.C. and Rich,A. (1974) Science 185, 435-440.
25. Robertus,J.D., Ladner,J.E., Finch,J.T., Rhodes,D., Brown,R.S., Clark,B.F.C. and Klug,A. (1974) Nature 250, 546-551.
26. Ladner,J.E., Jack,A., Robertus,J.D., Brown,R.S., Rhodes,D., Clark,B.F.C. and Klug,A. (1975) Proc. Natl. Acad. Sci. USA 72, 4414-4418.

27. Heckman,J.E., Sarnoff,J., Alzner-DeWeerd,B., Yin,S. and RajBhandary,U.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3159-3163.
28. Bonitz,S.G., Berlani,R., Coruzzi,G., Li,M., Macino,G., Nobrega,F.G., Nobrega,M.P., Thalenfeld,B.E. and Tzagoloff,A. (1980) Proc. Natl. Acad. Sci. USA 77, 3167-3170.
29. Barrell,B.G., Bankier,A.T., de Bruijn,M.H.L., Chen,E., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A., Sanger,F., Schreier, P.H., Smith,A.J.H., Staden,R. and Young,I.G. (1980) Proc. Natl. Acad. Sci. USA 77, 3164-3166.
30. Lagervist, U. (1978) Proc. Natl. Acad Sci. USA 75, 1759-1762.
31. Lagervist, U. (1981) Cell 23, 305-306.