Volume 9 Number 16 1981 Nucleic Acids Research

The nucleotide sequence of the large ribosomal RNA gene and the adjacent tRNA genes from rat mitochondria

C.Saccone[1], P.Cantatore[1], G.Gadaleta[1], R.Gallerani[2], C.Lanave[1], G.Pepe[1] and A.M.Kroon[3]

[1]Centro di Studio sui Mitocondri e Metabolismo Energetico presso Istituto Chimica Biologica, Università di Bari, [2]Dipartimento di Biologia Cellulare, Università della Calabria, Cosenza, Italy, and [3]Laboratory of Physiological Chemistry, State University, Groningen, The Netherlands

Received 19 May 1981

ABSTRACT

We have sequenced the Eco R_1 fragment D from rat mitochondrial DNA. It contains one third of the $tRNA^{Val}$ gene (the remaining part has been sequenced from the 3' end of the Eco R_1 fragment A) the complete gene for the large mt 16S rRNA, the $tRNA^{Leu}$ gene and the 5' end of an unidentified reading frame.

The mt gene for the large rRNA from rat has been aligned with the homologous genes from mouse and human using graphic computer programs. Hypervariable regions at the center of the molecule and highly conserved regions toward the 3' end have been detected.

The mt gene for $tRNA^{Leu}$ is of the conventional type and its primary structure is highly conserved among mammals.

The mt gene for $tRNA^{Val}$ shows characteristics similar to those of other mt tRNA genes but the degree of homology is lower.

Comparative studies confirm that AGA and AGG are read as stop codons in mammalian mitochondria.

INTRODUCTION

The nucleotide sequence analysis of the mt genome represents the most direct way to look at the genetic organization of mtDNA and to estimate the upper limit of its genetic complexity. Such studies have confirmed that the organization of animal mtDNA is very compact. They have also provided very useful information on e.g. mt tRNAs whose structures are remarkably different from those of prokaryotic or eukaryotic type and they have revealed the peculiarity of the non-universal genetic code in mitochondria (1-6).

The physical map and the genetic organization of mtDNA from rat has been extensively studied in our laboratories. A detailed restriction map using several enzymes has been constructed and the location of rRNA, tRNA and presumptive mRNA genes, were deduced by hybridization experiments (7-9).

Non standard abbreviations: mt: mitochondrial; bp: basepair(s)

In order to carry out the nucleotide sequence analysis of the DNA we have used restriction fragments, recombined with the plasmid pSF2124 and cloned in E. coli W5445.

Preliminary results on the nucleotide sequence of some fragments have been already presented (9-11). In this paper we report the complete nucleotide sequence of the Eco R_1 fragment D containing part of $tRNA^{Val}$ gene, the gene for the large mt rRNA, the gene for the $tRNA^{Leu}$ and the 5' end of an unidentified reading frame. The termination of Eco R_1 fragment A containing the major part of the $tRNA^{Val}$ gene, which is on the junction with Eco R_1 fragment D, has also been sequenced.

MATERIALS AND METHODS

Fragment sequencing: the Eco R_1 fragment D was isolated from recombinant plasmid by polyacrylamide gel electrophoresis after single or double digestion with Eco R_1 or Eco R_1 plus Hind III restriction enzymes: in the latter case two subfragments previously designated EH_3 and EH_6 (7) were directly obtained. The DNA fragments separated on polyacrylamide gel were eluted by shaking overnight at 37°C in 0.5 M Ammonium acetate, 0.1% S.D.S., 0.1 mM EDTA, concentrated by Ethanol precipitation and digested with a series of endonucleases (Hinf I, Hpa II, Hae III, Hha I, Sau 3A). The entire population of fragments was first phosphorylated at the 3' or the 5' end, using Klenow DNA polymerase or T4 polynucleotide kinase, and then subjected to secondary digestion or strand separation.

The fragments generated in this way were finally ready for sequencing according to the Maxam and Gilbert method (12).
Computer analysis: the sequence analysis was carried out on a Digital Equipment PDP 11/70 computer; the computer program was written in the programming language Fortran IV-Plus.
tRNA program: it searches regions of sequences that could base-pair at a given distance (55 ↕ 75 bp) with a minimum number of complementary bases (in our case between 5-7). In the region in which the above criterium is satisfied the program looks for further complementarities (arms of tRNA with a length variable between 3-5 bp).
SEQTEK program: it allows to obtain a scatter-plot of data

stored on files by a program derived from the SEQFIT program of Staden (13). The TEKTRONIX terminal video 4006-1 was used with its own software package. The obtained graphic is printed on TEKTRONIX 4631 HARD COPY UNIT.

RESULTS AND DISCUSSION

Nucleotide sequence analysis: Fig. 1 gives the complete sequence of Eco R_1 fragment D. The sequence of the L strand of mtDNA, corresponding to the sequence of the RNA products of the genes, most of which are transcribed from the H strand in animal mitochondria (5,6), is reported. The restriction sites found in the sequence analysis are in agreement with those obtained by fine mapping of the original mtDNA. Eco R_1 fragment A, recovered from the plasmid after digestion with Eco R_1, was sequenced at the 3' end. The 44 terminal nucleotides belonging to the $tRNA^{Val}$ gene, are part of the structure given in Fig. 2.

Identification of genes: the 5' end of the 16S rRNA gene was identified by alignment with the sequence at the 5' end of the 16S rRNA which has been directly determined for HeLa cells (14) and hamster (15). The DNA sequence was also compared with that of the homologous region of the human and mouse genes (16,17). The 3' end of the gene was positioned just before the 5' end of the mt RNA gene for leucine. This was possible in view of the analogy with the situation in mouse and man, and taking into account the data of Dubin and coworkers (18) for the mt 16S rRNA from hamster. In agreement with the results of van Etten et al. (17), the experiments with S_1 nuclease (not shown) and electron microscopy analysis (19) indicated the absence of introns.

To detect typical cloverleaf structures in the nucleotide sequence we have used a tRNA computer program adapted for mt products. We found within Eco D only one stretch with the typical features of a tRNA gene whose position is indicated in Fig. 1. It corresponds to the tRNA gene for leucine. The $tRNA^{Val}$ gene was identified and localized for about two third in Eco A and one third in Eco D. No typical tRNA structures could be detected by the computer within the strand of Eco D complementary to that reported in Fig. 1 (H strand).

In order to identify possible reading frames the sequence of

Figure 1. The nucleotide sequence of Eco R_1 fragment D. Restriction sites and the positions of various genes are indicated.

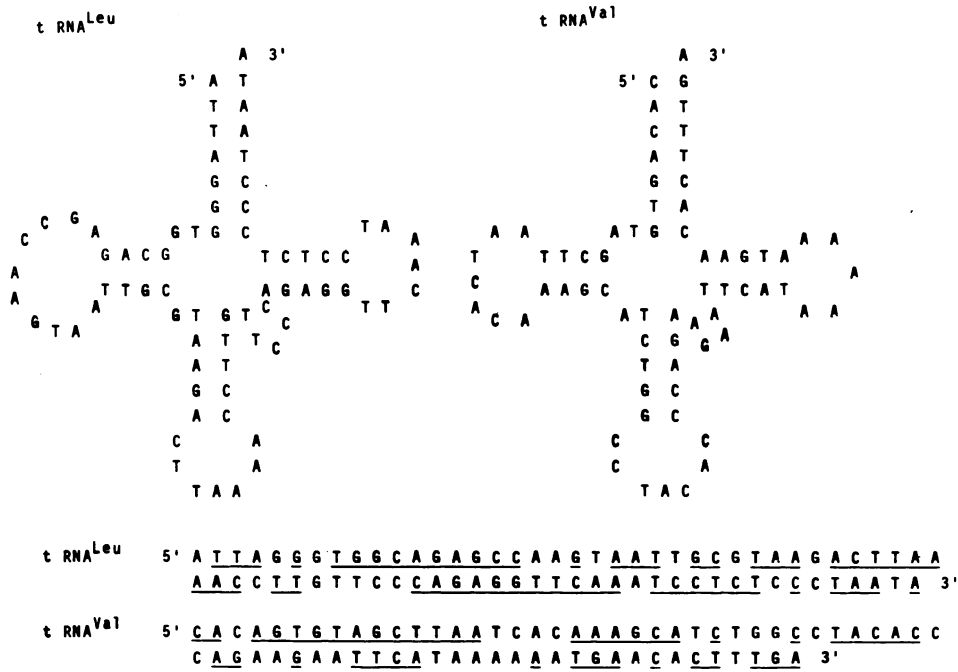

Figure 2. Primary and secondary structure of the $tRNA^{Leu}$ and $tRNA^{Val}$ genes of rat mtDNA. The nucleotides identical to mouse and human are underlined.

Eco R_1 fragment D has been analyzed with the computer program for translation (20) modified according to the codon recognition rules of the mitochondrial genetic system in animal cells (2). In the region of the ribosomal and tRNA genes the L strand contains only very short translatable sequences correctly initiated. On the other hand an open reading frame was identified after the $tRNA^{Leu}$ gene. We have analyzed by computer also the sequence of the complementary strand. Very short reading frames, probably meaningless, are present on this strand.

Homology with corresponding genes from mammals: while this work was in progress the sequences of mt DNA from human and mouse related to the same region were published (2,18). This allowed us to undertake a detailed comparison of the homologous genes from these three mammals.

Table I reports the content of the four nucleotides and the overall homology of the rat, mouse and human mt genes.

Fig. 2 shows the sequences related to mt tRNA genes for valine and leucine from rat. The regions conserved in the corresponding gene from mouse and human are underlined.

The tRNA genes for leucine and valine have a length of 75 and 69 basepairs respectively and contain no introns. Both display a typical cloverleaf structure but only the $tRNA^{Leu}$ gene contains the TTC sequence which may give rise to the TψC sequence in the actual tRNA. Other characteristics of mt tRNA are shared by both

TABLE I

Base composition and overall homology of the 16S rRNA, the $tRNA^{Val}$ and $tRNA^{Leu}$ genes of mitochondria of rat, mouse and human.

		Mouse	Mouse x Rat	Rat	Human x Rat	Human
16S rRNA						
base content (%)	A	35		38		38
	C	25		21		19
	T	22		24		26
	G	17		17		16
overall homology (%)			77		70	
$tRNA^{Val}$						
base content (%)	A	35		40		33
	C	25		23		25
	T	23		22		26
	G	16		14		16
overall homology (%)			85		71	
$tRNA^{Leu}$						
base content (%)	A	31		31		33
	C	25		27		28
	T	23		23		21
	G	21		19		19
overall homology (%)			97		77	

genes. However, some invariant bases are lacking especially in $tRNA^{Val}$ where the nucleotides G_{15} and U_{48} are absent. These bases are thought to interact according to the tertiary structural model available (21). The comparison of the nucleotide sequence of the two tRNA genes for leucine and valine among rat, mouse and human, reveals that the primary structure is more conserved in the $tRNA^{Leu}$ gene. This gene is furthermore of the conventional type similar to the tRNA genes also found in prokaryotes and eukaryotes, suggesting a possible regulatory role for the mt $tRNA^{Leu}$ which has been preserved during evolution.

In order to compare the mt 16S rRNA genes of the three mammals we have searched for regions of 20 nucleotides length having a 75% or greater homology. The data were plotted using the SEQTEK program. The degree of homology along the 16S rRNA genes between rat and human and between rat and mouse are shown in Figs. 3a and 3b. Figs. 3c and 3d depict similar results if 20 nucleotide blocks with 100% homology are considered. The results clearly indicate regions of hypervariability, the largest one at about the center of the molecule between the positions 550 and 750. An extended region, highly conserved in the three genes can be localized at the 3' end. This region, which contains the U^mG^m loop, is probably involved in the interaction between the two ribosomal subunits (22). It is also homologous to the corresponding region of E. coli 23S rRNA gene and ribosomal genes from yeast mitochondria and chloroplasts (23). It is now generally accepted that the region at the 3' end, whose secondary structure is highly maintained also in cases of divergence in the primary structure, is the most conserved one in the large rRNA gene. The bases involved in chloramphenicol sensitivity, which are changed in mouse and human mutant cells (24) are the same in our sequence as in the wild types (positions 1302 and 1353 in Fig. 1).

According to the data reported in the literature, the gene coding for $tRNA^{Leu}$ should be followed by a reading frame identified in HeLa cells as the poly A-RNA 13 (6). In our sequence a possible reading frame could start at the position 1657 within the 3' end of the $tRNA^{Leu}$ gene with a AUA codon or at the position 1669 with the AUU codon and it extends to Eco R_1

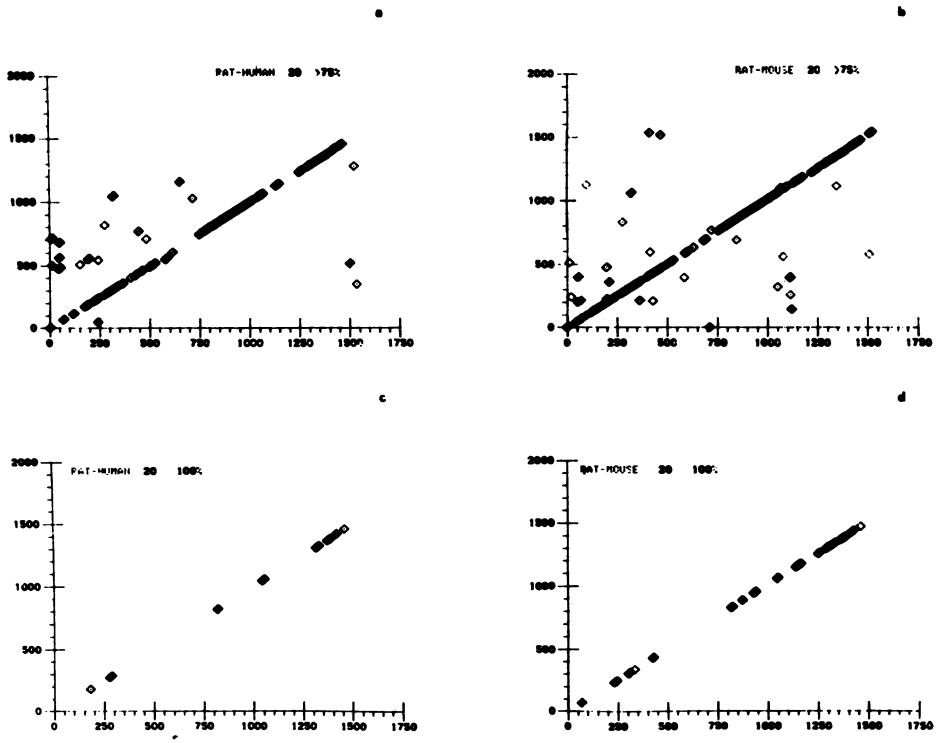

Figure 3. Graphic representation of the degrees of homology between the 16S rRNA genes of mtDNA from rat, mouse and human. For details see text.

fragment E. If we compare our sequence with that reported by van Etten et al. (18) we can observe a similar reading frame starting at corresponding positions, but only if we assign a stop codon function to the codons AGG and AGA, which universally code for arginine. These results confirm (2-5) that AGA and AGG are indeed read as stop codons within the mitochondrion. The aminoacid sequence coded for in this reading frame shows 63% homology to the corresponding sequence in human mitochondria (2) and 92% to that of mouse (18). The degree of divergence between rat and man is roughly similar to that reported between ox and man (2).

ACKNOWLEDGEMENTS

The Authors are indebted to Dr. J. Doly and Prof. G. Bernardi for the collaboration in recombinant DNA experiments. The assistance of Miss M. Holtrop, Mr G. Rainaldi and Miss E. Sbisà in some of the experiments reported and of Mrs M. Badini for the preparation of the manuscript is also gratefully acknowledged. The cooperation between the laboratories was facilitated by NATO Research Grant No 1484. Computerized data analysis was performed using equipments of the GNCB - Unità di Bari of CNR with their technical assistance.

REFERENCES

1 Kroon, A.M. and Saccone, C. (eds) (1980) The Organization and Expression of the Mitochondrial Genome, North-Holland Amsterdam.
2 Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H. L., Coulson, A.R., Drouin, J., Eperon, J.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, J.G. (1981) Nature 290, 457-465.
3 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. U.S.A. 77, 3167-3170.
4 Heckman, J.E., Sarnoff, J., Alzner-De Weerd, B., Yin, S. and RajBhandary, U.L. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 3159-3163.
5 Montoya, J., Ojala, D. and Attardi, G. (1981) Nature 290, 465-470.
6 Ojala, D., Montoya, J. and Attardi, G. (1981) Nature 290, 470-474.
7 Kroon, A.M., Pepe, G., Bakker, H., Holtrop, M., Bollen, J.E., van Bruggen, E.F.J., Cantatore, P., Terpstra, P. and Saccone, C. (1977) Biochim. Biophys. Acta 478, 128-145.
8 Greco, M., Pepe, G., Bakker, H., Kroon, A.M. and Saccone, C. (1979) Biochem. Biophys. Res. Commun. 88, 199-207.
9 De Vos, W.M., Bakker, H., Saccone, C. and Kroon, A.M. (1980) Biochim. Biophys. Acta 607, 1-9.
10 Saccone, C., Cantatore, P., Pepe, G., Holtrop, M., Gallerani, R., Quagliariello, C., Gadaleta, G. and Kroon, A.M. (1980) In: The Organization and Expression of the Mitochondrial Genome (Kroon, A.M. and Saccone, C. eds) pp. 211-219. North-Holland, Amsterdam.
11 Saccone, C., Cantatore, P., Gadaleta, G., Gallerani, R., Lanave, C., Pepe, G. and Kroon, A.M. (1981) Boll. Soc. It. Biol. Sper. (in press).
12 Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560-564.
13 Staden, R. (1978) Nucl. Ac. Res. 4, 4037-4051.
14 Crews, S. and Attardi, G. (1980) Cell 19, 775-784.
15 Dubin, D.T. and Baer, R.J. (1980) In: The Organization and Expression of the Mitochondrial Genome (Kroon, A.M. and Saccone, C. eds) pp. 231-240. North-Holland. Amsterdam.
16 Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980) Nature 286, 460-467.
17 Van Etten, R.A., Walberg, M.W. and Clayton, D.A. (1980) Cell 22, 157-170.

18 Dubin, D.T., Timko, K.D. and Baer, R.J. (1981) Cell 23, 271-278.
19 Cantatore, P., Quagliariello, C., Gadaleta, G., Gallerani, R., Pepe, G., Bleve, T., Zacheo, G., Kroon, A.M. and Saccone, C. (1980) Eur. J. Cell Biol. 22, 125.
20 Staden, R. (1978) Nucl. Ac. Res. 5, 1013-1015.
21 Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.S., Wang, A.H.J., Seeman, N.C. and Rich, A. (1974) Science 185, 435-440.
22 Baer, R. and Dubin, D.T. (1981) Nucl. Ac. Res. 21,4927-4941.
23 Gallerani, R., Cantatore, P., Gadaleta, G., Pepe, G., Quagliariello, C., Lanave, C., Kroon, A.M. and Saccone, C. (198) In: Proceedings of the 2nd Soviet-Italian Symp., Pushino on Oka, U.S.S.R., in press.
24 Kearsey, S.E. and Craig, I.W. (1981) Nature 290, 607-608.