Volume 10 Number 10 1982 Nucleic Acids Research

The nucleotide sequences of several tRNA genes from rat mitochondria: common features and relatedness to homologous species

P.Cantatore[1], C.De Benedetto[1], G.Gadaleta[1], R.Gallerani[2], A.M.Kroon[3], M.Holtrop[3], C.Lanave[1], G.Pepe[1], C.Quagliariello[2], C.Saccone[1] and E.Sbisa[1]

[1]Centro di Studio sui Mitocondri e Metabolismo Energetico, Istituto di 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 23 February 1982; Revised and Accepted 30 April 1982

ABSTRACT

We have determined the nucleotide sequences of thirteen rat mt tRNA genes. The features of the primary and secondary structures of these tRNAs show that those for Gln, Ser, and f-Met resemble, while those for Lys, Cys, and Trp depart strikingly from the universal type. The remainder are slightly abnormal. Among many mammalian mt DNA sequences, those of mt tRNA genes are highly conserved, thus suggesting for those genes an additional, perhaps regulatory, function. A simple evolutionary relationship between the tRNAs of animal mitochondria and those of eukaryotic cytoplasm, of lower eukaryotic mitochondria or of prokaryotes, is not evident owing to the extreme divergence of the tRNA sequences in the two groups. However, a slightly higher homology does exist between a few animal mt tRNAs and those from prokaryotes or from lower eukaryotic mitochondria.

INTRODUCTION

The complete nucleotide sequences of the mitochondrial (mt^+) genomes from human, bovine and mouse, recently determined (1-3) have confirmed the presence of 22 genes for tRNA in the mammalian mt genome. These genes are scattered along the circular molecule, fourteen being localized on the heavy strand and eight on the light strand. According to the punctuation model put forward by Attardi (4), the mRNA genes are adjacent to tRNA genes in such a way that the recognition of the tRNA cloverleaf structures by processing enzymes like RNAse P (5) should give rise to discrete RNA species from a single transcript. Besides this punctuation role, the animal mt tRNAs have structural anomalies to such an extent that they may be considered a separate class of tRNA molecules. Indeed, in most of them, many of the constant features of other tRNAs (6) are absent,

0305-1048/82/1010-3279$ 2.00/0

e.g., the sequence $G_{18}G$ in the D-loop (numbering according to Kim et al (7)), the sequence $T_{54}\Psi CPu$ in the T loop, and many of the usually invariant nucleotides. The most dramatic case is the $tRNA^{Ser}_{AGY}$ that completely lacks the D-loop and that, in addition, cannot form a cloverleaf structure (8-9).

In the course of our studies on the sequence of the mt genome of rat, we have completed the sequences of several genes for tRNA. The sequences of thirteen mt tRNAs are here reported and, together with the sequences of the rat mt tRNAs previously published by us (10) or by others (11)++, we have compared them to homologous sequences from prokaryotes, from eukaryotic cytoplasm and from lower eukaryotic mitochondria by using a suitable computer program.

These studies have permitted us to establish some features common to mt tRNAs from mammals. In addition, they provide the basis for studying the evolutionary relatedness of the animal mt tRNAs to homologous species.

MATERIALS AND METHODS

DNA sequencing: In order to determine the sequence of the rat mitochondrial genome, we have used DNA fragments Eco RI B, C, and D cloned in pSF2124 and Hind III B cloned in pBR 322 (13). The plasmid DNA, purified by CsCl-ethidium bromide centrifugation, was digested with Eco RI and Hind III. The mitochondrial inserts were gel purified and the DNA sequence was determined by using the method of Maxam and Gilbert (14). The DNA fragments were first digested with a restriction enzyme, then phosphorylated with $(\gamma^{32}P)$ ATP and polynucleotide kinase and finally digested with another restriction enzyme, or strand separated. In such a way, a population of fragments ready for the Maxam and Gilbert degradation reaction was obtained. The cleavage products were separated on 20%, 10%, and 8% polyacrylamide gels (40cm X 0,04cm) in the presence of 8 M urea. The tRNA genes have been positioned in the map either searching for cloverleaf structures or aligning the sequences with those of human, bovine and mouse.

Computer programs: We have used the tRNA program of Staden (15) which we have adapted to the peculiarities of the structure of mt

tRNA considering the following features common to all mt tRNAs: amino acid stem of 7 bp, anticodon stem of 5 bp, anticodon loop of 7 bases, D stem of 3-4 bp, T stem of 5 bp, two bases between the amino acid stem and the D stem, and a single base between the D arm and the anticodon arm, variable D and T loops, and variable extra-arm. The program has been further modified as follows: The anticodon and its position was fixed for each tRNA: once fixed, the program generates a limited number (between 1 and 4) of secondary structures. From these structures, the operator selects the most probable one (see Results for criteria). This is then stored in linear form according to the tRNA representation of Sprinzl et al (16). The program was implemented on VAX/780 (Digital) under the VAX-VMS operating system.

RESULTS AND DISCUSSION

In Fig. 1, the secondary structures of the thirteen rat mt tRNAs sequenced in our laboratories are shown. Mt tRNAs have many mismatches in the cloverleaf structure. The most common type of mismatch is the GT bp, but AC (Gln, Gly, Ser_{UCN}), RR (Ser_{UCN}, Ile, Fmet, Asn, Tyr) and Y-Y mismatches occur (Ser_{UCN}). Owing to the high degree of mismatching in the stems in the case of some tRNAs, several possible secondary structures can be elaborated by the program. On the basis of the highest number of matchings in the stems, the most probable ones have been compared with the secondary structures of the homologous mt tRNA species from mouse, bovine and human cells, which have been analyzed with the aid of the same computer program.The structures reported in Fig. 1 represent the more stable structures displaying the highest degree of homology. As described in Material and Methods, these structures are automatically converted by the computer to linear forms containing the different regions of the tRNAs separated and organized as described by Sprinzl et al (16). This type of storage produces an alignment of the different regions of the tRNA molecules allowing an easy and precise comparison, and offering at the same time the possibility of checking the sequences already published. In the case of the gene for $tRNA^{Ser}_{UCN}$, for example, it has been found that a better correlation can be established if the gene in the mouse starts

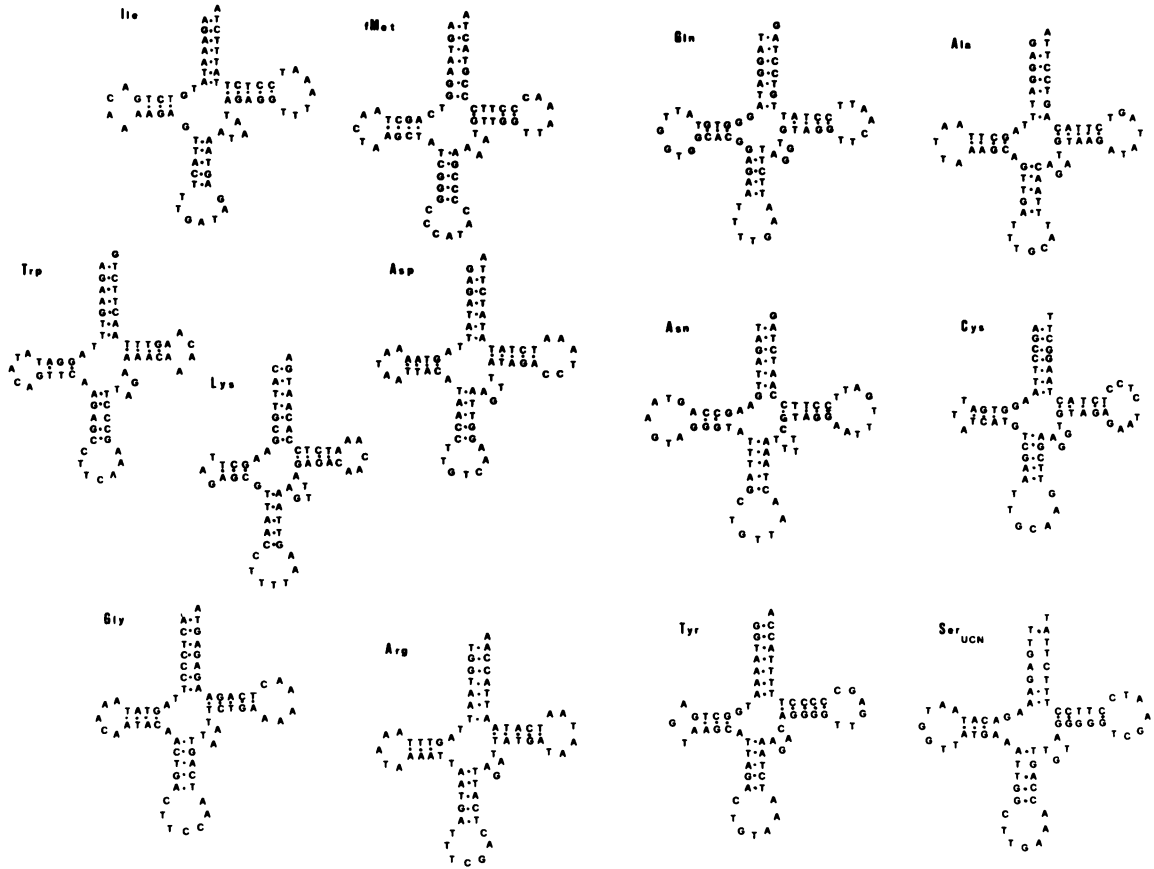

Figure 1 Secondary structure of several mt tRNAs from rat as deduced from the DNA sequences by our computer program (see text).

one nucleotide ahead (nucleotide 6940 in ref. 3), although in this case a very poor cloverleaf structure is obtained.
We have distinguished the tRNAs, according their degree of similarity to all other not organelles tRNAs (16).

From Fig. 1 it appears that the mt tRNAs that have features common to the other tRNAs are $tRNA^{Gln}$ and $tRNA^{Ser}_{UCN}$. They contain most of the invariant nucleotides (the $tRNA^{Gln}$ does not contain T_8, R_{15} and A_{21} and the $tRNA^{Ser}_{UCN}$ lacks only of T_8, A_{21}, Y_{48} and T_{55}) and have rather regulars T and D loops (made by seven bases in all the cases). We put the $tRNA^{Fmet}$ also in this category even if it lacks 6 invariant nucleotides. In fact,as shown in Tab. 1 and 2 by the comparison with either mitochondrial or not organelles tRNAs,it results the most conserved. The structures that depart strikingly from the usual ones are $tRNA^{Lys}$, $tRNA^{Cys}$, and $tRNA^{Trp}$ which have either a small D loop or lack the TTC sequence and many of the invariant nucleotides. The other tRNAs can be considered somewhat abnormal; they have a rather regular secondary structure but depart from the other tRNAs in that they lack many invariant nucleotides and, in some cases, also the TTC sequence.

It is well known that the tertiary structure envisaged for the non-mt tRNAs involves interactions among bases that are in the polynucleotide sequences of all tRNAs (7). Since mt tRNAs often contain sequences that depart somewhat from the sequences common to other tRNAs, particularly in the D and T loops, it is quite likely that these differences are associated with a tertiary structure slightly different from that of their non-mt counterparts. However, very little is known about the minor bases of mt RNAs and therefore any prediction about the possible base-base interactions in these molecules is premature.

The comparison of the secondary structures of homologous mt tRNA species from mammals reveals a rather high variability in the D-loop, T loop and extra arm. For the other parts of the molecule, by considering the nucleotide sequence homology, we have found that the order of conserved regions (in order of decreasing conservation), is the following: anticodon arm, D-stem, T stem. The amino acid stem is the region least conserved with a great deal of mismatching. In the case of $tRNA^{Lys}$, the

TABLE I

Percentage of homology between rat mt tRNAs and homologous mammalian genes.

tRNAs	Mouse	Bovine	Human
Leu_{uur}	97 (100)	80 (78)	76 (78)
F-met	94 (100)	94 (100)	97 (100)
Arg	93 (100)	90 (96)	85 (90)
Asn	92 (96)	77 (86)	76 (82)
Ile	91 (98)	97 (100)	91 (94)
His	91 (90)	84 (90)	75 (82)
Ser_{ucn}	90 (96)	88 (92)	78 (84)
Gln	90 (94)	79 (80)	83 (84)
Tyr	89 (96)	88 (96)	92 (96)
Trp	89 (97)	71 (74)	77 (76)
Ala	88 (88)	87 (88)	71 (69)
Asp	88 (92)	69 (74)	81 (86)
Gly	88 (94)	68 (71)	74 (73)
Val	87 (96)	73 (76)	71 (71)
Cys	87 (92)	63 (69)	66 (71)
Phe	86 (94)	74 (81)	74 (81)
Lys	84 (86)	67 (71)	72 (78)

The mt tRNA genes from rat and other mammals were aligned and the homology was determined considering either the entire sequence or the invariable parts only (values in parentheses). The percent of homology has been determined dividing the number of bases identical by the length of the rat tRNA gene. For rat mt tRNAs which are not presented in this paper, see ref.10-11.

D-stem contains only 3 bp.

In Table I, the homology of the rat mt tRNAs to the corresponding tRNAs from the mouse, bovine and human cell is reported. The comparison has been done both on the entire sequence and with the hypervariable regions (D loop, T loop and extra arm) excluded; the three anticodon nucleotides, which are the same in corresponding tRNAs of different mammals, have been also excluded. This latter criterion used also by others (17) allows to compare the molecules independently of their length excluding also a part which is for all identical. With this

TABLE II

Percentage of homology of some rat mitochondrial tRNAs with homologous species from several organisms.

	E. coli	A.Nidulans mt	Yeast mt	Yeast cyt.	Mammalian cyt.
F-met	50	34	41	35	46
Gly	48	34	41	34	n.d.
Gln	46	48	42	n.d.	n.d.
His	46	n.d.	50	n.d.	n.d.
Asn	43	n.d.	n.d.	n.d.	46
Ile	43	38	n.d.	n.d.	n.d.
Ala	39	48	48	33	n.d.
Asp	38	41	40	32	n.d.
Val	36	36	43	29	n.d.
Leu_{uur}	35	33	33	33	n.d.
Tyr	36	n.d.	n.d.	42	n.d.
Lys	34	20	23	33	33
Phe	32	42	36	33	30
Ser_{ucn}	30	51	n.d.	n.d.	n.d.
Arg	29	n.d.	n.d.	26	n.d.
Cys	28	n.d.	25	n.d.	n.d.
Trp	n.d.	30	32	n.d.	n.d.

The rat mt tRNA gene sequences have been aligned with those of different organisms here reported (taken mainly by Sprinzl et al (16). The percent of homology has been calculated subtracting the bases not identical (as a result of either a replacement or an insertions/deletions) from the length of the rat mt genes and dividing the result by the same number.

exclusion, it can be seen that the percentage of homology is generally higher, confirming that the above mentioned regions are the most variable. The results reported in Table I clearly demonstrate a very high homology among mt tRNAs from mammals, the extent of which is generally not lower than 70% (63% only in the case of the tRNA for cysteine). As expected, the greatest homology is between rat/mouse (90%): it is, however, difficult to establish a general priority between rat/bovine (79%) and

rat/human (79%) (This values represent the average homology of the tRNAs reported in Table).

Some of the mt tRNA genes like that for F-met, Arg, Ser, Tyr and Ala appear to be highly conserved particularly between rat/mouse, and rat/bovine; for the others the degree of divergence is higher. This indicates that some mt tRNA genes have higher functional constraints than others and suggests that some tRNAs could have an additional, probably regulatory, role. In accord with the model of Attardi (4), these species could have a peculiar role during the processing of the primary transcript.

When the mt tRNAs from rat are compared with the sequences of other non-mt tRNA species and also with the mt tRNAs from lower eukaryotes (Table II), we found that the pattern of homology change depending upon the type of tRNA. Those mt tRNAs that have a structure close to that of the universal tRNAs are more homologous to E. coli tRNA. However, the degree of homology of mammalian mt tRNAs with non-mt species is generally very low (between 30% and 50%). The divergence in base pairs greatly exceeds that among non-mitochondrial counterparts (eukaryotic-eukaryotic, prokaryotic-prokaryotic or eukaryotic-prokaryotic) making it virtually impossible for us or others (17) to integrate the mitochondrial sequences into existing tRNA evolutionary trees or to construct new trees. Owing to this extreme divergence of primary structures, the mt tRNA species were assigned to a separated branch whose evolutionary relationship with other species cannot be easily established at this time. If the degree of homology is calculated excluding the hypervariable regions, as in Table I, the results (not shown) do not change significantly, suggesting a random evolution of the different regions of the molecules. Taking all the results together, a greater homology is observed among prokaryotic or mt tRNA species than among cytoplasmic ones.

The position of mitochondria in the evolutionary tree is, at present, a matter of speculation and some controversy. According to the endosymbiotic theory of mitochondrial biogenesis, mitochondria are derived from bacteria which became associated with primitive eukaryotic cells by endosymbiotic events. On the other hand, the recent discoveries of the very peculiar

properties and organization of the mt genetic system argues, according to some authors (18), against such a theory, rather suggesting an independent origin and evolution for the mitochondria. Schwartz and Dayhoff (19), constructing a composite evolutionary tree, have presented evidence in favour of the endosymbiotic theory. More recently, Küntzel and Kochel (20) have drawn a tree for the small subunit rRNA which also support the bacterial origin for fungal mitochondria and they further suggest an independent bacterial origin for animal mitochondria. Our data are in agreement with the conclusions reached by Küntzel and are supported by our earlier studies on the comparison of the gene for the large ribosomal rRNA from mammals. These studies made with the aid of a graphic computer program demonstrate a higher homology of mt rRNAs with prokaryotic rRNAs (21).

The different organization of mitochondrial genomes in different kingdoms (low eukaryotes, animal cells and probably also plants) could be explained by an independent invasion of different bacteria into already diverged prokaryotic precursors of animal and fungal cells (19-20). However it might also be speculated that an important but unknown evolutionary event took place at the branching point of animals, generating conditions (e.g. a lower dependence upon the nucleus) favourable to an extreme divergence and a faster evolution of the mt genome. Such a genome might well evolve along its own special path, independent and different from the pathway of other mitochondria. The large extent of such divergence would serve to obscure the primitive origin of the animal mitochondrial genome.

ACKNOWLEDGEMENTS

The Authors are indebted to Prof. M. Simpson for the helpful discussion and criticism on the manuscript. Thanks are due to Mr D. Munno and Mr F. Fanizzi for their assistance in some of the experiments reported and to Mrs M. Badini for the preparation of the manuscript. The cooperation between the laboratories was facilitated by NATO Research Grant No 1484, and by ZWO fellowship to Miss M. Holtrop. Computerized data analysis was performed using equipments and technical support

of the GNCB - Unità di Bari of CNR.

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

++ After the completion of this manuscript, the paper of Grosskopf and Feldmann (12) came to our attention in which the sequence of five tRNA gene (for Lys, Asp, Ser_{UCN}, Gln, Arg) is reported. We note only a very few minor differences which do not alter our conclusion in any way.

REFERENCES

1 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., Schreirer, P.H., Smith, A.J.H., Staden, R., Young, I.G. (1981) Nature 209, 457-465.
2 Anderson, A., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. Personal communication.
3 Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167-180.
4 Attardi, G. (1981) Trends Biochem. Sci. 6, 100-103.
5 Apirion, D. and Gegenheimer, P. (1981) FEBS Lett. 125, 1-9.
6 Gauss, D.H., Grüter, F., Sprinzl, M. (1979) in: Cold Spring Harbor Symposia on Quantitative Biology (Abelson, J.,Schimmel, P.R. and Söll, D. eds) pp. 518-519, Cold Spring Harbor, New York.
7 Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang, A.H.J., Seeman, N.C. and Rich, A. (1974) Science 185, 435-440.
8 Arcari, P. and Brownlee, G. (1980) Nucl. Acid Res. 8, 5207-5212.
9 de Bruijn, M.H.L. (1980) Nucl. Acid Res. 8, 5213
10 Saccone, C., Cantatore, P., Gadaleta, G., Gallerani, R., Lanave, C., Pepe, G. and Kroon, A.M. (1981) Nucl. Acid Res. 9, 4139-4148.
11 Kobayhashi, M., Yaginuma, K., Seki, T. and Koike, K. (1980) in: The Organization and the Expression of Mitochondrial Genome (Kroon, A.M. and Saccone, C. eds) pp. 211-219, North-Holland, Amsterdam.
12 Grosskopf, R. and Feldmann, H. (1981) Curr. Genet. 4,151-158
13 Kroon, A.M., Pepe, G., Bakker, H., Holtrop, M., Bolten,J.E., van Bruggen, E.F.J., Cantatore, P., Terpstra, P. and Saccone, C. (1977) Biochim. Biophys. Acta 478, 128-145.
14 Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560-564.
15 Staden, R. (1980) Nucl. Acid Res. 8, 819-823.
16 Sprinzl, M., Grueter, F., Spelzhaus, A. and Gauss, D.H. (1980) Nucl. Acid Res. 8, r1-r21.
17 Cedergren, R.J., La Rue, B., Sankoff, D., La Palme, G. and Grosjan, H. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2791-2795.
18 Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980) Nature 286, 460-467.
19 Schwartz, R.M. and Dayhoff, M.O. (1978) Scienze 199,359-403.

20 Küntzel, H. and Köchel, H.G. (1981) Nature 293, 751-755.
21 Saccone, C., Cantatore, P., De Benedetto, C., Gadaleta, G., Gallerani, R., Holtrop, M., Kroon, A.M., Lanave, C., Pepe, G., Quagliariello, C., Rainaldi, G. and Sbisà, E. (1982) in: Mitochondrial Genes (Attardi, G., Borst, P. and Slonimski, P. eds) Cold Spring Harbor, New York. In press.