EVOLUTION OF TRANSFER RNA MOLECULES AS A REPETITIVE PROCESS

Thomas H. Jukes and Richard Holmquist

Spaces Sciences Laboratory
University of California, Berkeley, California 94720

Received August 7,1972

Summary: Thirty-six transfer RNAs were aligned on the basis of the "clover-leaf" model and compared for homology. All these tRNAs have identical residues at eleven loci. Additional homology may occur in pairs of tRNAs for different amino acids in the same organism. This type of homology indicates that tRNAs for 2 different amino acids may evolve from a common ancestor by gene duplication. Such an evolutionary separation must include changes in the anticodon and in the recognition site for the activating enzyme.

The transfer RNAs (tRNAs) are a group of molecules which can be represented in the form of "clover-leaf" secondary structures that contain four helical regions and 3 or 4 loops (1,2). Each tRNA contains at least 72 nucleotides plus a 3'-terminal trinucleotide $-CCA_{OH}$. The evolutionary relationship between the tRNAs is shown by their common secondary structure and by the presence of about eleven invariant bases in corresponding loci (2,3).

We have compared the sequences of tRNA molecules arranged as in the examples in Figure 1. Only 74 loci were used in the comparisons; the "extra" nucleotides, representing insertions that are present in some tRNAs, (mainly between positions 46 and 47), were omitted to make the comparisons uniform. Gaps were inserted in the portion of the sequence corresponding to loop I (residues 14-22 in Figure 1) to maximize the homology of this region, which varies in length (2). The 36 tRNAs compared were Phe Y, EC, wheat; Leu Y, EC; Ile TY, EC; Met EC; f-Met EC; Val Y, TY, EC; Ser Y, EC, rat; Ala Y; Tyr Y, TY, EC; His Salmonella; Lys Y; Asp Y; Glu EC; Trp EC; Arg Y, EC; Gly EC; and their alleles, totaling 630 comparisons. (Y = yeast; TY = torula yeast; EC = E. coli). The results when the tRNAs for different amino acids were compared gave average differences of 49% for yeast tRNAs, 48% for E. coli tRNAs, and 50% for yeast vs E. coli tRNAs. The average random difference between unrelated RNA sequences containing equal quantities of A, C, G and U is 75% because there is one chance in four that two bases will be identical by chance. The references for most of the sequences are in Nishimura (4).

Figure 1. Sequences of certain tRNAs showing evolutionary homology when compared as described in text.

	Yeast		E. coli					Yeast		E. coli					Yeast		E. coli			
	Lys H	Arg 3A(3B)	Met	Ile	Val 1	Val 2A(2B)		Lys H	Arg 3A(3B)			Val 1	Val 2A(2B)		Lys H	Arg 3A(3B)			Val 1	Val 2A(2B)
1	G	G	G	A	G	G		G	G	A	G	A	A	50	A	A	A	G	G	G
	C	C	G	G	G	C		U	U	C	C	C	C		G	U	C	G	G	G
	C	G	C	G	G	G		A	C	A	A	C	C		G	G	A	U	C	U
	U	C	U	C	U	U	30	U	U	U	C	U	A		G	G	G	G	G	G
	U	U	A	U	G	C(U)		G	G	C	C	C	C		G	G	G	G	G	G
	G	C(U)	C	U	A	C		A	A	A	C	C	C		U	U	U	U	U	U
	U	G	G	G	U	G(A)		C	C	C	C	C	U		U	U	U	U	U	U
	U	U	U	U	U	U		U	U	U	U	U	U		C	C	C	C	C	C
	G	G	A	A	A	A		C	U	C	G	U	G		G	G	G	A	G	G
10	G	G	G	G	G	G		U	C	A	A	A	A		A	A	A	A	A	A
	C	C	C	C	C	C		U	U	U	U	C	C	60	G	C	A	G	U	G
	G	G	U	U	U	U		A	A	A	A	A	A		C	C	U	U	C	U
	C	U	C	C	C	C		A	A	A	A	A	U		C	C	C	C	C	C
	A	A	A	A	A	A	40	U	U	U	G	G	G		C	C	C	C	C	C
	A	A	G	G	G	G		C	C	G	G	G	G		C	C	C	A	G	A
	U	U	U	G	C	U		A	A	A	G	A	U		C	A	G	C	U	C(A)
	C	-	U	U	U	U		U	G	U	U	G	G		U	U	U	U	C	U
	G	G	G	G	G	G		A	A	G	G	G	G		A	C	C	C	A	C(U)
	G	G	G	G	G	G		A	A	G	A	G	G		C	G	G	A	U	G
20	-	-	U	U	-	U		G	G	G	G	G	G		A	A	U	G	C	G(A)
	U	C	U	U	G	U		N	A	G	G	G	G	70	G	G	A	G	A	A
	A	A	A	A	A	A		U	U	N	N	U	N		G	U	G	C	C	C
	G	A	G	G	G	G		U	U	C	C	C	C		G	G	C	C	C	G
	C	C	A	A	A	A									C	C	C	U	C	C
	G	G	G	G	G	G									U	G(U)	A	A	A	A
	C	C	C	C	C	C														

The horizontal lines mark the regions of the cloverleaf tRNA model (2) as follows: residues 1-7 and 67-73, helix a; 10-13 and 23-26, helix b; 14-22, loop I; 28-32 and 40-44, helix c; 33-39, loop II (anticodon bases underlined); 45-49, loop III; 50-54 and 62-66, helix e. The terminal -CCA is omitted. The sequences are taken from references 6-11. All bases are shown in their unmodified forms.

Additional homology in pairs of tRNAs occurs as follows:

Class (i). TRNAs for the same amino acid in the same organism may differ in only 1 to 6 corresponding loci (1 to 8 percent). Examples are yeast Ser 1

and 2 (3 base differences); E. coli Tyr 1 and 2 (6 differences); yeast Arg 3A and 3B (2 differences); and E. coli Val 2A and 2B (5 differences). Other examples of type (i) homology are found in pairs which have undergone more pronounced evolutionary separation but are still evidently homologous. Such sequences may differ in 14 to 21 loci (20 to 30 percent). Examples are E. coli Arg 1 and 2; two yeast Lys tRNAs; E. coli Val 1 and 2A (or 2B). We conclude that continuing evolutionary divergence may proceed in such pairs, sometimes following phylogenetic separation of two species of organisms from a common ancestor, until eventually the homology of such a pair of tRNAs is not significantly greater than that represented by the average difference (50%, range 38% to 69%) of tRNAs for different amino acids, excluding pairs in Class (iii) below.

Class (ii) consists of tRNAs for the same amino acid in two different organisms. Examples are yeast and wheat Phe tRNAs (23% difference); and yeast and rat Ser tRNAs (19% difference).

Class (iii), the third type of "higher-than-average" homology, occurs between tRNAs for different amino acids in the same organism. This was first reported by Squires and Carbon (5), who found a relationship between the primary structures of E. coli Gly 3 and Val 2 A (26%); and Gly 3 and Val 2B (24%). We have found other examples of this third type of homology in the following tRNA pairs: yeast Lys and Arg 3A and 3B (25 and 24% differences), and a strong suggestion of such "higher-than-average" homology in the group: E. coli Met, Ile and Val 1, 2A and 2B (Met: Ile, 35% difference; Met: Val 1, 30%; Met: Val 2A or 2B, 33%; Ile: Val 1, 38%; Ile: Val 2A, 28%; Ile: Val 2B, 32%). The comparisons are in Figure 1.

These findings indicate that tRNAs for 2 different amino acids may evolve from a common ancestor by gene duplication and subsequent incorporation of mutational nucleotide changes. Such an evolutionary separation must include changes both in the anticodon and in the site that recognizes the specific aminoacyl-tRNA ligase (the "charging" enzyme). How could either one of these

changes take place without lethal results? A change to an anticodon corresponding to a different amino acid, without a change in the amino acid carried by the tRNA, would place the "wrong" amino acid in the sequences of numerous proteins. The same effect would be produced if the anticodon remained constant and the recognition site of the charging enzyme was altered to receive the "wrong" amino acid. It would therefore be necessary that the tRNA be "withdrawn from circulation" while these changes are taking place. This could conceivably be accomplished by one or more of the following procedures following gene duplication: (i) The DNA recognition site for transcription of one of the two tRNA cistrons would mutate so that it would not be transcribed; the mutated cistron would then remain in a dormant condition without producing tRNA. During this stage, it could accumulate base replacements by mutations followed by genetic drift. Eventually, a back mutation could restore the transcription recognition site to re-functioning. (ii) a mutation could take place which would prevent the tRNA from becoming bound to ribosomes, so that it could not participate in protein synthesis; conceivably, the sequence in loop IV (residues 55-61, Figure 1) is concerned with ribosomal binding. Also, it is likely that the tertiary structure of tRNA must "fit" the ribosome. If a mutation made a critical change in the tertiary structure, the tRNA would perhaps not be received by the ribosomal site (12). (It may be recalled that a single amino acid change can disrupt the tertiary structure of a hemoglobin molecule, e.g. hemoglobin S, to impair its function drastically.) (iii) the recognition site for the charging enzyme could be inactivated by a mutation, so that the tRNA could not combine with an amino acid until another, correcting, mutation subsequently took place, perhaps to change the recognition site to correspond to another aminoacyl tRNA ligase.

It is conceivable that changes in the specifity of tRNAs are a dynamic process that operates continually during evolution so that functional tRNA cistrons undergo disappearances from and reappearances in the transcription process. Comparisons such as those in Table I show the evolutionary separation

of tRNAs spans a range from two to six base differences to a stage where as much as 69% of sequential difference exists between two tRNAs, as in the case of yeast Asp and E. coli Tyr tRNAs.

ACKNOWLEDGMENT. - This work was supported by grants from NASA (NGR 05-003-460) and NIH (HE - 11553).

References

1. Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R. and Zamir, A., Science, 147:1462 (1965).
2. Jukes, T. H., and Gatlin, L. L., Progress in Nuc. Acid Research and Molecular Biology, 11:303 (1971).
3. Jukes, T. H., Biochem. Biophys. Res. Commun., 24:744 (1966).
4. Nishimura, S., Progress in Nuc. Acid Research and Molecular Biology, 12:49 (1972).
5. Squires, C., and Carbon, J., Nature New Biology, 233:274 (1971).
6. Smith, C. J., Ley, A. N., D'Obrenan, P. and Mitra, S. K., J. Biol. Chem., 246:7817 (1971).
7. Weissenbach, J., Martin, R. and Direheimer, A., Presented at tRNA Conference Princeton University, May 1, 1972.
8. Cory, S., Marcker, K. A., Dube, S. K. and Clark, B.F.C., Nature, 220:1039 (1968).
9. Yaniv, M. and Barrell, B. G., Nature, 222:278 (1969).
10. Yaniv, M. and Barrell, B. G., Nature New Biology, 233:113 (1971).
11. Yarus, M. and Barrell, B. G., Biochem. Biophys. Res. Commun., 43:729 (1971).
12. Singer, C. E., Smith, C., Cortese, R. and Ames, B. N., Nature New Biology, 238:72 (1972).