Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 10432–10435, October 1995
Genetics

Transfer RNA editing in land snail mitochondria

(genome organization/acceptor stems/discriminator base/polyadenylylation/RNA circularization)

SHIN-ICHI YOKOBORI AND SVANTE PÄÄBO*

Institute of Zoology, University of Munich, Luisenstrasse 14, P.O. Box 202126, D-80021 Munich, Germany

Communicated by Walter Gilbert, Harvard University, Cambridge, MA, July 21, 1995 (received for review June 20, 1995)

ABSTRACT **Some mitochondrial tRNA genes of land snails show mismatches in the acceptor stems predicted from their gene sequences. The majority of these mismatches fall in regions where the tRNA genes overlap with adjacent downstream genes. We have synthesized cDNA from four circularized tRNAs and determined the sequences of the 5′ and 3′ parts of their acceptor stems. Three of the four tRNAs differ from their corresponding genes at a total of 13 positions, which all fall in the 3′ part of the acceptor stems as well as the discriminator bases. The editing events detected involve changes from cytidine, thymidine, and guanosine to adenosine residues, which generally restore base-pairing in the stems. However, in one case an A-A mismatch is created from an A-C mismatch. It is suggested that this form of RNA editing may involve polyadenylylation of the maturing tRNAs as an intermediate.**

RNA editing is the post- or cotranscriptional modification of genetic information such that a mature transcript will come to differ from what would conventionally be inferred from its genomic sequence (1–3). Whereas most cases of RNA editing have been reported for organelles of plants and unicellular organisms (1–3), only two cases of RNA editing have been reported for metazoan mitochondria. (*i*) The polyadenylylation of mRNAs in several cases creates the termination codon UAA (1, 4–6). (*ii*) The anticodon of the $tRNA^{Asp}$ in the mitochondria of marsupials is altered posttranscriptionally (7).

Recently one complete and two partial sequences of land snail mitochondrial (mt-) genomes have become available (8–11). They are small (14.1–14.5 kb) when compared to other metazoans, yet the complete sequence of the mtDNA of the land snail, *Cepaea nemoralis*, shows that it encodes the same 13 proteins and 2 rRNAs typical of other metazoan mitochondria (8, 9). Furthermore, 22 tRNA genes can be predicted from the genomic sequence of the land snails *Euhadra herklotsi* from Japan (N. Yamazaki, R. Ueshima, M. Kaifu, S.Y., T. Ueda, K. Nishikawa, and K. Watanabe, unpublished work) and *Cepaea nemoralis* from Britain (9). However, these tRNA genes as well as those of a Greek land snail, *Albinaria turrita*, show some unusual features. The most prominent of these is that several tRNA genes exhibit one to four mismatches in their acceptor stems. We have analyzed cDNA sequences of four tRNAs of the Japanese land snail *E. herklotsi*. The results show that a hitherto unknown RNA editing mechanism exists in the mitochondria of land snails.

MATERIALS AND METHODS

tRNA and DNA Preparation. Total RNA was prepared by the guanidine thiocyanate method (12) from the fresh hepatopancreas of one individual of *E. herklotsi* collected from Tsunoshima, Yamaguchi, Japan. The total RNA was applied to a DEAE-cellulose column (DE 52, Whatman), and tRNAs were eluted with a buffer containing 20 mM Tris·HCl (pH 7.5), 10 mM $MgCl_2$, and 800 mM NaCl. Total DNA was prepared (13) from *E. herklotsi* hepatopancreas from the same individual used to prepare RNA.

tRNA Circularization. The ligation condition used was modified from Nishikawa (14). Forty micrograms of total tRNA was ligated in 100 μl of a solution containing 50 mM Hepes (pH 8.3), 10 mM $MgCl_2$, 3.5 mM dithiothreitol (DTT), 10 mg of bovine serum albumin (BSA) per ml, 3.3 mM ATP, and 2000 units of T4 RNA ligase per ml (New England Biolabs), at 37°C for 2 hr. Subsequently, one phenol extraction and one chloroform extraction were followed by an ethanol precipitation. As controls, total DNA treated with RNase A and total DNA not treated with RNase were processed as above.

cDNA Synthesis. Primers used for cDNA synthesis were Eh-G-FW (5′-AAGGTAACTGTACTTAT-3′) for $tRNA^{Gly}$, Eh-H-FW2 (5′-TAGATTATATATCTTAAT-3′) for $tRNA^{His}$, Eh-K-FW2 (5′-TGGCCATACTCGGCCATA-3′) for $tRNA^{Lys}$, and Eh-Y-FW2 (5′-CGCTTTAATTAAGCTATT-3′) for $tRNA^{Tyr}$. Annealing of the primers to the circular tRNAs was performed in 23 μl of TE (10 mM Tris·HCl/1 mM EDTA, pH 8.0) containing 2 μg of circular tRNA and 1 pmol of cDNA primer. After heating to 90°C for 2 min, the reaction was left at room temperature for 15 min and subsequently on ice for 15 min. After adding 3 μl of 10-fold concentrated reverse transcription buffer (500 mM Tris·HCl, pH 8.4/80 mM $MgCl_2$/300 mM KCl), 0.3 μl of 1 M DTT, 1.5 μl of 1.25 mM dNTPs, and 2.4 units of avian myeloblastosis virus (AMV) reverse transcriptase (Stratagene) or 5 units of Moloney murine leukemia virus (MMuLV) reverse transcriptase (New England Biolabs), the volume was adjusted to 30 μl. The reaction was incubated for 45 min at 37°C for MMuLV reverse transcriptase or at 45°C for AMV reverse transcriptase. In control experiments, total linear tRNA, RNase-treated DNA with RNA ligase treatment, and RNase-treated DNA without RNA ligase treatment were used as substrate for the cDNA synthesis.

PCR, Cloning, and Sequencing. For PCR amplification of cDNAs, the following primers were used: Eh-G-FW and Eh-G-RV (5′-CAAGTAAGTAGTCTAATG-3′) for $tRNA^{Gly}$; Eh-H-FW2 and Eh-H-RV (5′-TGGCATCCTAAAAGCTCT-3′) for $tRNA^{His}$; Eh-K-FW2 and Eh-K-RV (5′-TAATCTA-ATTACGGGTTA-3′) for $tRNA^{Lys}$; and Eh-Y-FW2 and Eh-Y-RV (5′-AATCCTATGTTGGTTTAA-3′) for $tRNA^{Tyr}$. For PCR of genomic regions primers were Eh-ND2 (5′-ACAGTATTAAATTGGCCAGT-3′) and Eh-COI (5′-CCAATATCCTTATGGTTAGT-3′) for the $tRNA^{Lys}$ gene, Eh-COII (5′-GGGCAATGTTCAGAAATTTG-3′) and Eh-Y-FW2 for the $tRNA^{Gly}$ and $tRNA^{His}$ genes, and Eh-H-RV and Eh-W-RV (5′-TTCAAAGTTGGAAATGTAC-3′) for the $tRNA^{Tyr}$ gene. One and one-half microliters of cDNA solution was used for 50-μl amplifications containing 70 mM Tris·HCl (pH 8.8 at 25°C), 10 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 1 mM DTT, 100 mg of BSA per ml, 0.1% Triton X-100, 200

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: mt-, mitochondrial.

*To whom reprint requests should be addressed.

μM dNTPs, and 20 units of *Taq* DNA polymerase per ml. The 40 cycles performed consisted of 92°C for 1 min, 40–50°C for 1–1.5 min, and 72°C for 1–1.5 min. The PCR products were checked by agarose gel (4% MetaPhor agarose, FMC) electrophoresis in 40 mM Tris-acetate buffer containing 1 μg of ethidium bromide per ml. PCR products were cloned using a TA cloning kit (Invitrogen) according to the manufacturer's protocol. Sequencing was performed using Sequenase Version 2 (United States Biochemical) and [α-^{32}P]- or [α-^{33}P]dATP (Amersham).

RESULTS AND DISCUSSION

Genomic Sequences of *Euhadra* tRNAs. Gene sequences for the mitochondrial tRNA^{Gly}, tRNA^{His}, tRNA^{Tyr}, and tRNA^{Lys} and their neighboring regions were determined from one individual of the Japanese land snail *E. herklotsi.* They were found to be identical to the sequence previously determined. The genes for tRNA^{Gly} and tRNA^{Lys} overlap in their inferred 3′ ends by 4 and 6 nucleotides with the downstream genes for tRNA^{His} and subunit I of cytochrome oxidase, respectively (Fig. 1). In the overlapping regions, the tRNA^{Gly} carries two mismatches, whereas the tRNA^{Lys} carries three mismatches (other than G-U base pairs) in their inferred acceptor stems. The tRNA^{His} gene overlaps in its 5′ end with the 3′ end of tRNA^{Gly} and carries two mismatches in its inferred acceptor stem. The sequence of the gene for tRNA^{Tyr} was also determined. This gene displays three mismatches in the acceptor stem but does not overlap with adjacent genes in *Euhadra* (Fig. 1).

Analysis of Circularized tRNAs. To obtain sequences of 5′ and 3′ parts of the tRNA acceptor stems, we made use of the fact that tRNAs carry monophosphate groups at their 5′ ends and hydroxyl groups at the 3′ ends. Thus, they can be circularized using RNA ligase (15). This allows the synthesis of cDNA throughout the acceptor stem as well as over its 5′ and 3′ ends. In contrast, primer extensions do not allow determination of sequences in the 3′ parts of the stems. One further advantage of this method is that the presence of 3′ CCA ends can be determined. Because mt-tRNA genes do not encode CCA sequences (4, 16), the presence of these sequences indicates that the tRNA may be functional in the translational machinery.

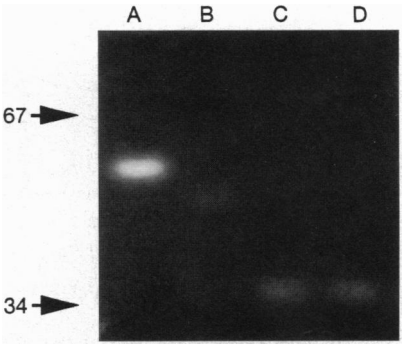

FIG. 2. Amplification of cDNA for *Euhadra* mt-tRNA^{Tyr} by PCR. cDNA of tRNA^{Tyr} from circularized tRNAs (lane A), noncircularized tRNA (lane B), circularized DNA treated by RNase A (lane C), and DNA treated by RNase A (lane D) were used for PCR. Ten microliters of 50-μl PCRs was run on a 4% agarose gel, stained with ethidium bromide (1 μg/ml).

Fig. 2 shows that PCR products of the expected length were

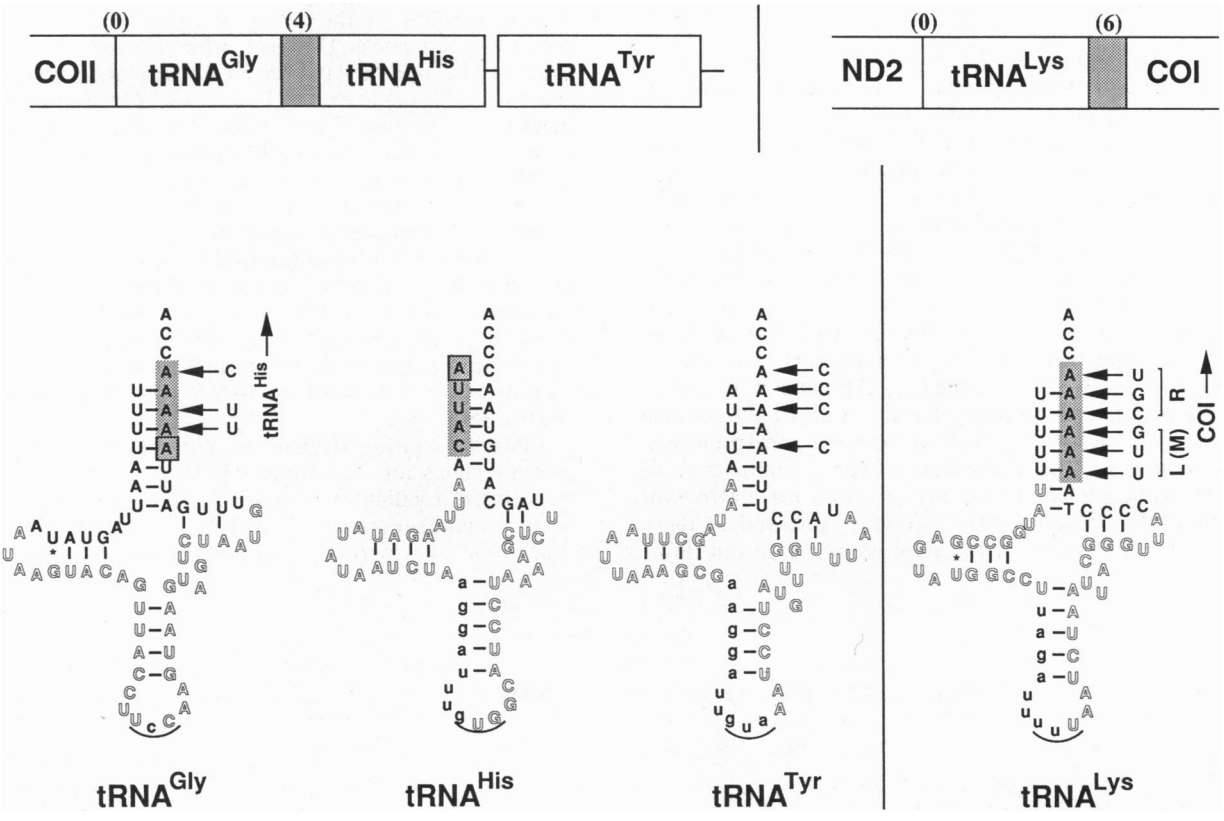

FIG. 1. (*Upper*) The gene order of tRNA^{Gly}, tRNA^{His}, tRNA^{Tyr}, tRNA^{Lys}, and their flanking regions are shown for *Euhadra*. Shadowed boxes indicate the overlaps between genes and the numbers of overlapping bases are given within parentheses. (*Lower*) The inferred secondary structures of the tRNAs are given and differences between genomic and cDNA sequences in the acceptor stems are indicated. Regions overlapping with adjacent genes are shadowed, and the −1 nucleotide of tRNA^{His} and the nucleotide in tRNA^{Gly} that might overlap with the −1 nucleotide of tRNA^{His} are shadowed and boxed. Regions used for cDNA and PCR primers are indicated by outlined letters. Other regions all given in lowercase letters.

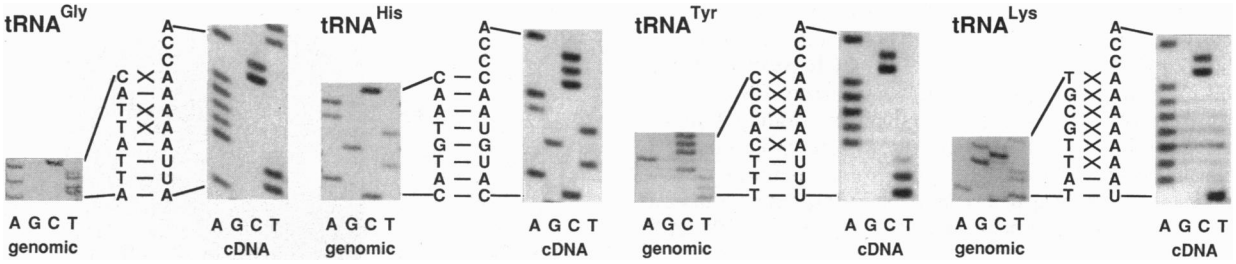

FIG. 3. Sequence comparisons between cDNA and gene sequences of each tRNA. Autoradiographs of sequence ladders are shown.

obtained only in the cases where the tRNAs were circularized. Products derived from genomic sequence are not generated since the primers used have 5′ ends facing each other and multimerization of RNA molecules is not a problem under the conditions used here (data not shown). The generation of PCR products from circularized tRNAs for Lys, His, and Tyr failed when cDNA primers with 3′ ends located upstream of the ninth nucleotide were used (not shown). This is likely due to methylations at position 9 (17) that may inhibit the cDNA synthesis.

cDNA Sequences of tRNAs. For each tRNA analyzed, more than three different clones were sequenced in both directions. No discrepancies among clones were found. However, in the 3′ parts of the acceptor stems of $tRNA^{Gly}$, $tRNA^{Tyr}$, and $tRNA^{Lys}$, a total of 13 differences between the genomic sequences and their cDNA counterparts were detected (Fig. 3). The sequences and inferred secondary structures of the tRNAs are shown in Fig. 1. Together, two guanosine (G), five thymidine (T), and six cytidine (C) residues are altered to adenosine (A) residues in the cDNA sequences. In all three tRNAs, the changes involve also the discriminator bases. In the 10 cases where the changes occur in the stems, they create Watson–Crick base pairs with thymidine residues in the 5′ parts for the acceptor stems except in one case, where an A-A pair is created at the 5′ terminus of $tRNA^{Tyr}$.

The cDNA sequence of $tRNA^{His}$ is similar to its genomic sequence. It carries an A residue at the −1 position. It is a common characteristic of eukaryotic as well as prokaryotic $tRNA^{His}$ to carry such a nucleotide, although in all described cases (17) it is a G [and in one case a uridine (U)] residue rather than an A residue. In chicken mitochondria as well as in the eukaryotic cytoplasm, this −1 nucleotide is added posttranscriptionally (18). However, in the mitochondria of fungi, plants, and protozoans, as well as in prokaryotes, the −1 nucleotide is encoded by the gene (17). Although this remains to be elucidated, it seems likely that the A at the −1 position in the *Euhadra* $tRNA^{His}$ is derived from the gene sequence. Two mismatches occur at the base of the acceptor stem of $tRNA^{His}$. Although one of the primers used for cDNA synthesis and PCR included the 5′ residues involved in these mismatches, the 3′ parts of the stem sequence do not differ from the genomic sequence. Thus, $tRNA^{His}$ seems not to be affected by any editing events.

Characteristics of Land Snail tRNA Editing. Parts of the tRNAs other than acceptor stems do not show any conspicuous deviations from the expected structures. Thus, it seems likely that the editing in land snails is confined to acceptor stems. This is the case for the editing described for four mt-tRNAs of the *Acanthamoeba castellanii* (19) as well as of some mt-tRNAs of plants (20). However, whereas in these cases nucleotide substitutions occur in the 5′ parts of acceptor stems, in the three cases described here, tRNA alterations all occur in the 3′ parts of the stems.

A further difference from the other systems is that all 13 instances of editing in land snails involve changes to A residues from the other three nucleotides. In most cases, mismatches in the primary transcripts are corrected by the editing events. In one case, however, an A-A mismatch is created from an A-C mismatch.

Of the 13 editing events detected, 9 fall in regions of the tRNA genes where they overlap with downstream genes. The exception is the $tRNA^{Tyr}$ gene in *Euhadra*, which displays four edited positions in the acceptor stem yet is followed by noncoding sequences. However, when the genome organizations of the related land snails *Cepaea* and *Albinaria* are compared (Fig. 4) to that of *Euhadra* it is found that in these species the 3′ parts of the $tRNA^{Tyr}$ genes overlap with the downstream genes for $tRNA^{Trp}$ and that mismatches exist in $tRNA^{Tyr}$ gene sequences but not in the $tRNA^{Trp}$ gene sequences. Yamazaki *et al.* predicted that the gene order of *Cepaea* and *Albinaria* is ancestral to that of *Euhadra*. The mismatches and editing of the $tRNA^{Tyr}$ of *Euhadra* mitochondria may thus represent a remnant of the ancestral genome organization still present in *Cepaea* and *Albinaria*. If this is so, all cases of mismatches that are corrected by RNA editing occur at positions that in the current or previous gene organization fall in regions of the tRNA genes that overlap with downstream genes.

A Possible Editing Mechanism. Whereas the changes in the acceptor stems could be directed by the bases in the 5′ part of the stems, discriminator bases are also changed even though no obvious templates exist. For $tRNA^{Gly}$ and $tRNA^{Lys}$ it is conceivable that partially processed transcripts, extended in

	acceptor stem	D arm	anticodon stem and loop	T arm	acceptor stem
	+++++++	!!!!!	!!!!! ######	###### %%%%%	%%%%% +++++++
Euhadra	ATTTAAA TA	GCTT- AATTA	-AAGC G -AAGGA	TTGTAAA TCCTA- TGTT	GGT-- TTAAT-	--ACC TTTCACCC
Cepaea	****T*T **	****- *T*A-	-**** A -TC*AG	C****** *T*G*- A**A	*A--- CATTGT	---T* A*GGG***
Albinaria	****TGG C-	C**AG *GAG-	TT**G A TG*T*C	******* CTA*TA GT*-	*AAT- AC*T--	-GTT* CCAA*GTG
Mytilus	CA*AGG* *G	**CGA GGAA-	T*G*T * -GT*AG	C****** CT*AT- AAAC	AAGGC G***--	GCCTT *C*T*TGA
Katharina	*A*AG*G *G	**CG- *TAAA	-*G*T * -GTAA*	C****** *TTAT- *TA*	*AGG- A**T--	-CCTT C*CT*TTA
Drosophila	GA****G *G	***GA *G**-	T*G** * -*TA**	******* **TAT- *TA*	AAGA- A*T**-	-TCTT C**A*T*A

FIG. 4. Comparisons of the $tRNA^{Tyr}$ genes among three land snails (9–11), the other mollusks [*Mytilus edulis* (21) and *Katharina tunicata* (22)], and *Drosophila yakuba* (23). Nucleotides forming Watson–Crick and G-U base pairing in stem regions are indicated by single underline and double underline, respectively. Regions overlapping with adjacent genes are indicated by italics, anticodons are indicated by boldface letters, and nucleotides identical to the upper sequence are indicated by asterisks.

their 5′ ends, could serve as templates since the −1 nucleotides in these cases are T residues. However, for $tRNA^{Tyr}$ the −1 position sequence is A. Thus, even if such a precursor molecule would exist, it could not serve as template for the discriminator alteration in this case. Thus, it would seem that at least for the discriminator bases, the editing process is not directed by a template present in the transcript.

There are at least three possible mechanisms by which the editing of land snail tRNAs could occur. (*i*) Since the 3′ parts of acceptor stems that display mismatches generally overlap with adjacent genes, processing of the primary transcript is expected to remove the 3′ nucleotides of the mismatched regions. Polyadenylylation could then add A residues in a non-template-directed way. Subsequently, further processing might limit the numbers of A residues at the 3′ terminus and the CCA sequence could be added. (*ii*) A template-dependent polymerization could be envisioned, which would fill in the 3′ parts of the tRNAs after initial processing and add an A residue as a discriminator base in a non-template-directed fashion prior to CCA addition. (*iii*) By alternative processing, tRNA precursors containing mismatches might be created. An editing mechanism might then correct the mismatches by replacing the relevant bases or nucleotides, including the discriminator bases.

In view of the current data, the first mechanism seems most likely. First, in animal mitochondria, not only mRNAs but also rRNAs are known to be polyadenylylated (4). Although no report on polyadenylylation of tRNAs exists, it is conceivable that all mt-RNA species can be polyadenylylated. Polyadenylylation would explain why the nucleotide opposite to the A residue in the 5′ terminus of the $tRNA^{Tyr}$ acceptor stem is edited from a C to an A, rather than a U, residue and why discriminator bases seem to be added without templates. Furthermore, it is compatible with the fact that also bases located between the positions that differ in the tRNAs from their gene sequences are A residues. These bases would thus be added along with the other bases by polyadenylylation.

Genome Size and tRNA Editing. In general, animal mt-genomes are small in size. The number of genes encoding proteins, rRNAs, and tRNAs are limited, and introns and large noncoding sequences between genes do not exist. Furthermore, in some cases where the 3′ ends of protein-coding genes overlap with adjacent genes, termination codons are created by polyadenylylation (1, 4–6, 16). The smallest metazoan mt-genomes known are those of nematodes (24), which are 13.9–14.2 kb in size. In those cases, the gene encoding subunit 8 of the mt-ATPase has been lost from their mt-genomes and all tRNA genes lack either TΨC or DHU arms. The mt-genomes of land snails are almost as small (14.1–14.5 kb) as those of nematodes and clearly smaller than those of arthropods, echinoderms, and vertebrates (15.1–18.0 kb). It is tempting to speculate that the evolution of RNA editing in land snails is driven by the pressure to minimize the size of the mt-genome by allowing tRNA genes to overlap.

Note Added in Proof. Since the submission of this manuscript, the complete mtDNA sequence of the land snail *Albinaria coerulea* has become available (10). In that species, RNA editing very similar to the one described here seems to be operating.

We are especially grateful to T. Yokogawa for technical advice. We are furthermore indebted to R. Ueshima for *E. herklotsi* samples, to K. Watanabe, K. Nishikawa, T. Ueda, R. Ueshima, N. Yamazaki, and M. Kaifu for permission to use unpublished data, to R. H. Thomas and J. A. Terrett for making their unpublished manuscript available to us, to M. Mörl, V. Börner, M. Dörner, and E. Prager for valuable discussions, to the Deutsche Forschungsgemeinschaft for financial support, and to the Alexander-von-Humboldt Foundation for a fellowship to S.Y.

1. Cattaneo, R. (1991) *Annu. Rev. Genet.* **25,** 71–88.
2. Gray, M. W. (1993) *Curr. Opin. Genet. Dev.* **3,** 884–890.
3. Bass, B. L. (1993) in *RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 383–418.
4. Attardi, G. (1985) *Int. Rev. Cytol.* **93,** 93–145.
5. Ojala, D., Montoya, J. & Attardi, G. (1981) *Nature (London)* **290,** 470–474.
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. & Young, I. G. (1981) *Nature (London)* **290,** 457–465.
7. Janke, A. & Pääbo, S. (1993) *Nucleic Acids Res.* **21,** 1523–1525.
8. Terrett, J. A., Miles, S. & Thomas, R. H. (1994) *Nautilus* **2,** Suppl., 79–84.
9. Terrett, J. A., Miles, S. & Thomas, R. H. (1995) *J. Mol. Evol.*, in press.
10. Hatzoglou, E., Rodakis, G. C. & Lecanidou, R. (1995) *Genetics* **140,** 1353–1366.
11. Lecanidou, R., Douris, V. & Rodakis, G. C. (1994) *J. Mol. Evol.* **38,** 369–382.
12. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
13. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
14. Nishikawa, K. (1987) in *Methods of the Study for Genes: Chemistry and Analytical Technique for Nucleic Acid*, Sequels to the Series for Experimental Technique of Biochemistry, ed. Japanese Society for Biochemistry (Tokyo Kagaku Dojin, Tokyo), Vol. 1, pp. 63–78 (in Japanese).
15. Bruce, A. G. & Uhlenbeck, O. C. (1978) *Nucleic Acids Res.* **5,** 3665–3675.
16. Wolstenholme, D. R. (1992) *Int. Rev. Cytol.* **141,** 173–216.
17. Steinberg, S., Misch, A. & Sprinzl, M. (1993) *Nucleic Acids Res.* **21,** 3011–3015.
18. L'Abbe, D., Lang, B. F., Desjardins, P. & Morais, R. (1990) *J. Biol. Chem.* **265,** 2988–2992.
19. Lonergan, K. M. & Gray, M. W. (1993) *Science* **259,** 812–816.
20. Maréchal-Drouard, L., Ramamonjisoa, D., Cosset, A., Weil, J. H. & Dietrich, A. (1993) *Nucleic Acids Res.* **21,** 4909–4914.
21. Hoffmann, R. J., Boore, J. L. & Brown, W. M. (1992) *Genetics* **131,** 397–412.
22. Boore, J. L. & Brown, W. M. (1994) *Genetics* **138,** 423–443.
23. Clary, D. O. & Wolstenholme, D. R. (1985) *J. Mol. Evol.* **22,** 252–271.
24. Okimoto, R., MacFarlane, J. L., Clary, D. O. & Wolstenholme, D. R. (1992) *Genetics* **130,** 471–498.