Total DNA transcription *in vitro*: A procedure to detect highly repetitive and transcribable sequences with tRNA-like structures

(newt/tortoise/tRNA$_1^{Lys}$ gene/tRNA^{Glu} gene/evolution)

HIROSHI ENDOH AND NORIHIRO OKADA*

Institute of Biological Sciences, University of Tsukuba, Sakura-mura, Ibaragi-ken 305, Japan

Communicated by Motoo Kimura, September 3, 1985

ABSTRACT   Total DNAs from various animals were transcribed *in vitro* in a HeLa cell extract, and it was found that one to several discrete RNAs were transcribed by RNA polymerase III. With tortoise (*Geoclemys reevessi*) and newt (*Cynops pyrrhogaster*), distinct 6.5S and 8S RNAs were transcribed from these respective DNAs. Representative phage clones carrying the 6.5S and 8S RNA genes were isolated from genomic libraries of these animals, and the sequences of these genes were determined. The 5' parts of highly repetitive and transcribable sequences of tortoise and newt were found to have close resemblance to tRNA$_1^{Lys}$ (rabbit) gene (78% homology) and a tRNA^{Glu} (*Drosophila*) gene (74% homology, not counting the aminoacyl stem region), respectively. The homologies extended to secondary structures, homologous nucleotides being located on similar secondary structures. It is proposed that many, if not all, highly repetitive and transcribable sequences detected by total DNA transcription have specific tRNA genes as their progenitors.

Interspersed throughout virtually all eukaryotic genomes are families of repeated sequences. In higher eukaryotes, these sequences have been classified into two categories based on size: long interspersed repetitive families (LINEs), which include proviral type sequences and *L1* sequences, and short interspersed repetitive families (SINEs), such as the primate *Alu* and rodent type 1 or 2 *Alu* families (for review, see ref. 1).

Usually, one animal species contains several kinds of SINEs, however in one animal, there seems to be only one kind of highly repetitive and interspersed sequences that are transcribed by RNA polymerase III. Sakamoto *et al.* (2) demonstrated that the rodent type 2 *Alu* family can be detected as a distinct 6S RNA by total DNA transcription *in vitro*. Although *in vitro* transcription of total DNA isolated from a variety of animal species does not always result in a discrete sized RNA, this procedure has been shown to be very useful for detection and analysis of highly repetitive and transcribable sequences in some animal species. We propose to call the highly repetitive and transcribable sequences detected by total DNA transcription Hirt sequences. Recently, Sakamoto and Okada (3) found that the rodent type 2 *Alu* family, the rat *ID* sequences, the rabbit *C* family, and the bovine or goat 73-base-pair repeat, have remarkable sequence homology with a few specific tRNAs, and we suggested that these repetitive sequences may have evolved from specific tRNA genes. These results raise the question of whether the Hirt sequences that will be characterized have sequence homology with specific tRNA genes. In this communication, we present evidence that the newly sequenced repetitive families in tortoise and newt have close structural resemblance with a tRNA$_1^{Lys}$ (rabbit) gene and a tRNA^{Glu}

(*Drosophila*) gene, respectively. The results presented in this paper were briefly described (4).

MATERIALS AND METHODS

Materials. High molecular weight DNAs were extracted from several animals principally as described by Blin and Stafford (5). DNAs were extracted from human placenta, male BALB/c mouse liver, tortoise (*Geoclemys reevessi*) liver, frog (*Xenopus laevis*) liver, newt (*Cynops pyrrhogaster*) whole body, lung fish (*Lepidosiren paradoxa*) liver, coelacanth (*Latimeria chalumnae*) liver, salmon (*Oncorhynchus keta*) sperm, and lamprey (*Lampetra japonica*) liver.

Preparation of HeLa Cell Extract. HeLa cell extract was prepared as described by Manley *et al.* (6) with a slight modification according to Talkington *et al.* (7) to achieve a more efficient extraction of RNA polymerases.

***In Vitro* RNA Synthesis and Fingerprinting of Oligonucleotides.** *In vitro* RNA synthesis was performed principally as described by Manley *et al.* (6). The RNA was isolated by phenol extraction and subjected to electrophoresis on an 8% or a 10% polyacrylamide denaturing gel (8). Individual bands were excised and RNA was extracted (9). Fingerprints were prepared (10) using thin-layer homochromatography on polyethylenimine (PEI) cellulose sheets for the second dimension (11).

Isolation of Phage Clones and DNA Sequencing. Phage clones were isolated from a tortoise or newt genomic library by hybridization by using as a probe ^{32}P-labeled RNA transcribed from respective total DNA. The region of the tortoise or newt Hirt sequence was located by Southern blot hybridization (12) and then restriction enzyme fragment was subcloned in pBR322. The subfragment was ligated into bacteriophage vector M13mp8, -mp9, -mp10, or -mp11 replicative-form DNA (13) and sequenced by the dideoxychain-terminator method (14).

RESULTS

***In Vitro* Transcription of Total DNAs Isolated from Various Animals.** Sakamoto *et al.* (2) demonstrated that a discrete 6S RNA transcribed *in vitro* from mouse total DNA (Fig. 1, lane 2) is identical to transcripts from the type 2 *Alu* family. Therefore, products transcribed *in vitro* from total DNAs seemed not to be artifacts but to reflect the repetitive "genes" in the genome. We isolated total DNAs from a variety of animals and these DNAs were transcribed in a HeLa cell extract (Fig. 1). Total DNAs from some animals (tortoise, newt, salmon, and lamprey) produce a very distinct RNA with smeared background that varies in extent while total DNAs from other animals (frog, lung fish, and coelacanth) produce several distinct RNAs. Human total
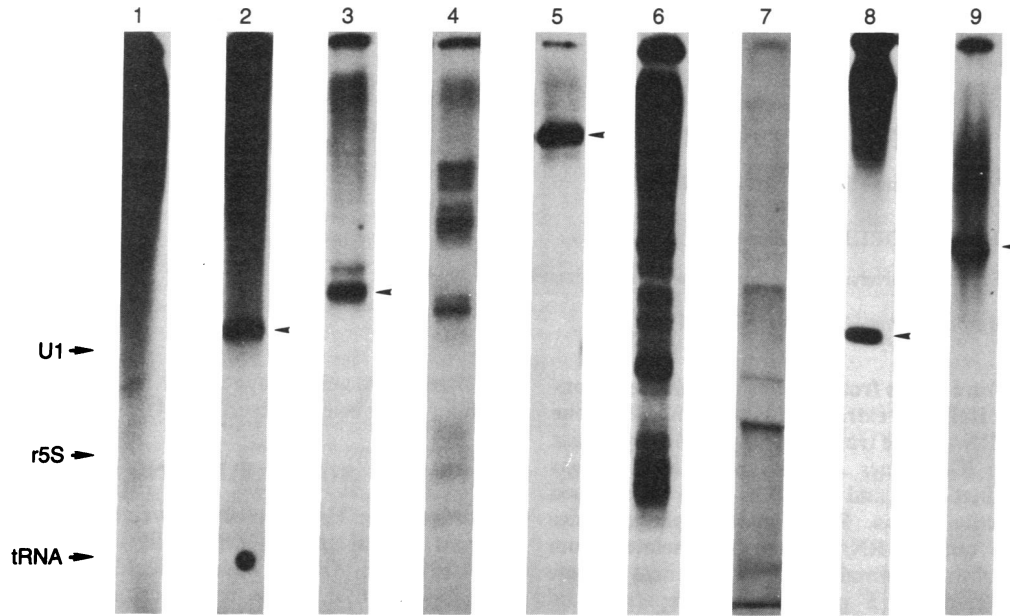
*To whom correspondence should be addressed.

FIG. 1. Analysis of *in vitro* transcripts from total DNAs isolated from various animals. Total DNAs used as templates were isolated from human (lane 1), mouse (lane 2), tortoise (lane 3), frog (lane 4), newt (lane 5), lung fish (lane 6), coelacanth (lane 7), salmon (lane 8), and lamprey (lane 9). The size markers used were ^{32}P-labeled rat snRNAs. The transcripts were analyzed by electrophoresis in 10% polyacrylamide denaturing gel at 600 V for 3 hr. An arrowhead shows a major discrete transcript.

DNA produces a wide variety of transcripts. We focused our attention on two animals that produce only one distinct transcript, tortoise and newt.

Characterization of the Tortoise and Newt Hirt Sequences. A discrete 6.5S RNA was transcribed *in vitro* from tortoise total DNA and a discrete 8S RNA from newt total DNA (Fig.
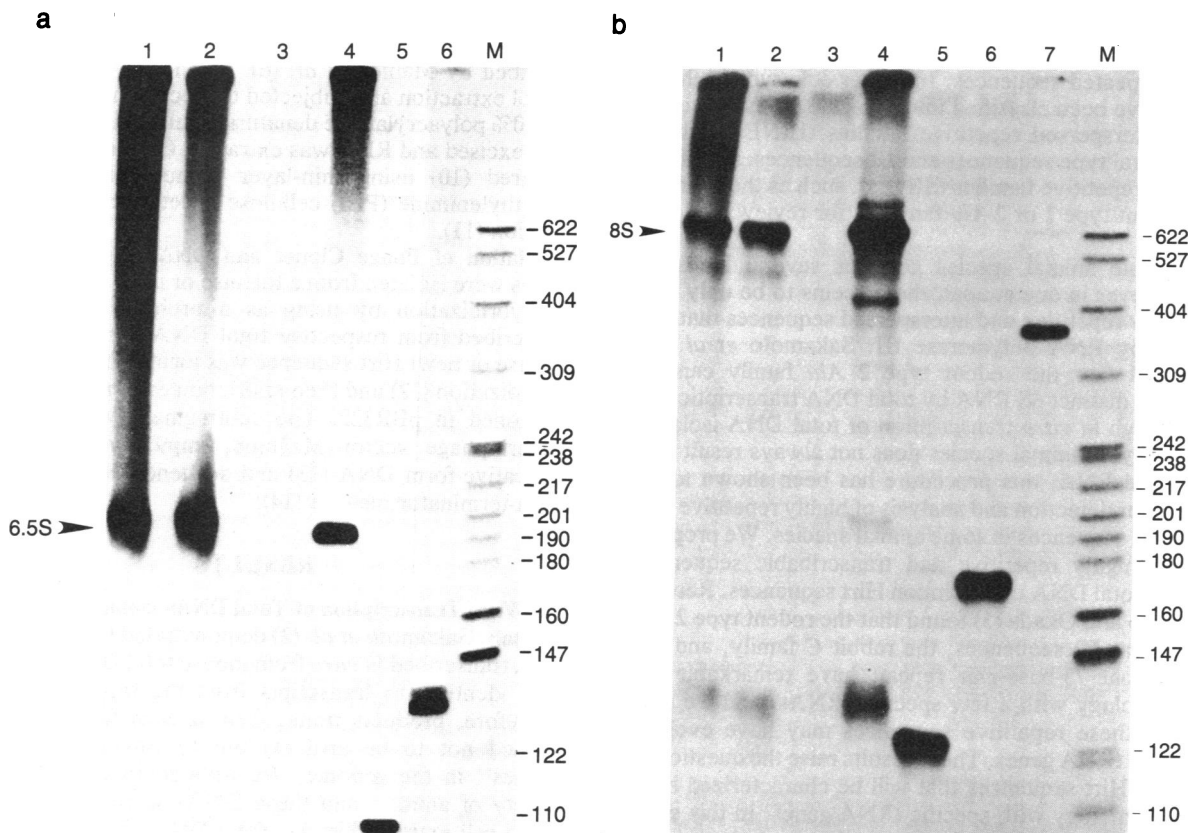

FIG. 2. α-Amanitin sensitivity of the transcript and mapping of the transcriptional initiation site of cloned DNAs. (*a*) Total DNA of tortoise was used as the template in the absence (lane 1) and presence of α-amanitin (2 μg/ml, lane 2 and 200 μg/ml, lane 3). For lane 4, the cloned phage DNA designated as TE6 was used as the template. For lanes 5 and 6, the TE6 DNA digested with *Ava* II or *Hin*fI was used as template. (*b*) Total DNA of newt was used as the template in the absence (lane 1) and presence of α-amanitin (2 μg/ml, lane 2 and 200 μg/ml, lane 3). For lane 4, the cloned phage DNA designated as NE1 was used as the template. For lanes 5–7, the NE1 DNA digested with *Hae* III, *Hin*fI, or *Alu* 1 was used as the template. The size markers used were the ^{32}P-labeled *Hpa* II-digests of pBR322 DNA; the sizes in bases are indicated. The transcripts were analyzed by electrophoresis in 8% polyacrylamide denaturing gel at 800 V for 7 hr.
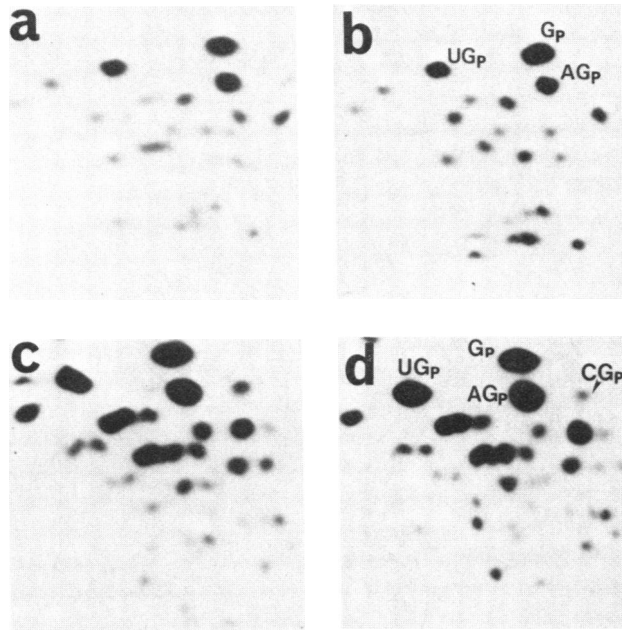
FIG. 3. Tortoise and newt Hirt sequences are presumed to be highly conserved from fingerprint analyses. Fingerprints of 6.5S RNA transcribed from tortoise total DNA (*a*) and the phage cloned DNA designated as TE6 (*b*). Fingerprints of 8S RNA transcribed from newt total DNA (*c*) and the phage cloned DNA designated as NE1 (*d*). Each gel-purified RNA species was digested with RNase T1, and the digests were separated by electrophoresis on cellulose acetate paper (horizontal dimension) followed by homochromatography on PEI thin-layer plate (vertical dimension).

1, lanes 3 and 5). In view of the sensitivity of this transcription to α-amanitin, it was concluded that RNA polymerase III transcribes these RNAs (Fig. 2*a*, lanes 1–3, and *b*, lanes 1–3). Fig. 3 *a* and *c* shows fingerprints of RNase T1-digested oligonucleotides of the 6.5S and 8S RNA transcribed from tortoise and newt total DNA, respectively. A clear pattern of oligonucleotides was obtained, suggesting that the nucleotide sequences of the genes for the tortoise 6.5S RNA or newt 8S RNA are highly conserved. Using ^{32}P-labeled transcripts as probes, we isolated and analyzed several phage clones carrying a 6.5S RNA (phage clone TE6) or an 8S RNA gene (phage clone NE1) from genomic libraries of these animals. Fig. 2 *a*, lane 4, and *b*, lane 4, shows transcripts from the cloned DNAs of tortoise and newt, respectively. Each transcript has the same electrophoretic mobility as that of the RNA transcribed from total DNA. The fingerprints of RNAs transcribed from tortoise and newt cloned DNAs are very similar to those of RNAs transcribed from total DNAs of these animals (Fig. 3), therefore, the RNA genes coded by the respective cloned DNAs are presumed to be representative.

The 5′ Parts of the Tortoise and Newt Hirt Sequences Have Remarkable Resemblance to a tRNA$_1^{Lys}$ Gene and a tRNA^{Glu} Gene, respectively. Fig. 4 shows the sequence of the tortoise 6.5S RNA gene. For rough mapping of initiation site of

transcription, we digested the DNA with several restriction enzymes and used the digests as templates. Fig. 2*a*, lanes 5 and 6, shows the results. From these results, we tentatively mapped the initiation site at about 10 nucleotides upstream from the 5′ terminus of the first promoter. The gene consists of about 189 nucleotides including a TTTT oligonucleotide as a transcriptional termination signal. As shown in Fig. 5, the sequence from +11 to +64 in the gene has remarkable homology with a tRNA$_1^{Lys}$ gene (ref. 15; for compilation of tRNA genes, see ref. 16), the homology between the sequences of the two being 78%. The homology starts from U8 and ends at C61 of the tRNA. The homology extends to secondary structures, and homologous nucleotides are located on nearly the same secondary structures. We also determined the sequence of newt clone NE1 (Fig. 6), and roughly mapped the transcriptional initiation site at 10 nucleotides upstream from the first promoter (Fig. 2*b*, lanes 5–7). As shown in Fig. 7, the sequence from +13 to +68 has remarkable homology with a tRNA^{Glu} gene (74% homology).

DISCUSSION

To date, total DNAs have been extracted from about 30 species of animals including invertebrates and transcribed *in vitro* in HeLa cell extracts. DNAs from some animal species produced a discrete transcript with a varying extent of smeared background (animals shown in this paper), however, there are some other animals whose DNAs do not produce any distinct RNA (for example, cat, dog, or pig; N.O., unpublished results). Since fingerprints of discrete RNAs gave very clear patterns of oligonucleotides, the transcript migrating as one band in a polyacrylamide was assumed to be homogeneous. So, we examined animals whose total DNA produce a discrete transcript. In this paper, we showed that two of the Hirt sequences from tortoise and newt may have evolved from specific tRNA genes.

Judging from the estimated frequencies of isolation of phage clones, the Hirt sequences in tortoise and newt comprise about 2–6% of the genomic DNA. Southern hybridization experiments showed that the Hirt sequences in tortoise and newt are dispersed among the genome (data not shown). These properties are similar to those of the human *Alu* family, which was recently demonstrated to have evolved from *7SL* RNA genes (17). From the close resemblance between the rodent type 2 *Alu* family and tRNA$_5^{Lys}$ (3), and the results presented in this paper, we here propose that many, if not all, highly repetitive, transcribable, and dispersed sequences in animals have evolved independently from specific tRNA genes. In this respect, we presume that the human *Alu* family is an exception.

Since the ^{32}P-labeled 8S RNA probe did not hybridize to DNA of *Xenopus laevis*, sequences homologous to the 8S RNA gene family of newt seem to be restricted to the Urodela (unpublished results). Moreover, the facts that a discrete RNA is produced by total DNA transcription (Fig. 1) and that a fingerprint of the transcript gives a very clear pattern of oligonucleotides (Fig. 3) strongly suggest that amplification of the tortoise or newt repetitive family is evolutionarily a

```
              +1
              •••
CCTCTTTGTGCCAGCCATTGGGGGAGGGATAGCTCAGTGGTTTGAGCATTGGCCTGCTAAACCCAGGGTTGTGAGTTCAATCCTTAAGG  70
         ---------------------------------Ava II---------------------------Hinf I---------
GGGCCACTTAGGGATGTGGGGCAAAAAATTGGTCCTACTAGTAAAGGCAGGGGGCTGGACTCAATGACCTTTCAAGGTCCCTTCCAGTTC  160

TAGGAGCTTGGTATATCTCCAATTATTTTCTTTTTCTTTTTCATCTATCAGCTCAGACAGC
         -------------------------    ‾‾‾‾
```
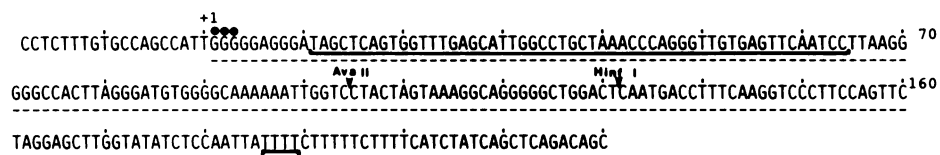
FIG. 4.   The nucleotide sequence of tortoise Hirt cloned DNA in phage TE6. The presumed initiation sites of transcription are shown by dots above the nucleotides. The region of Hirt sequences transcribed is shown by a dotted line under the nucleotides. The region homologous to tRNA is indicated by a thick line under the nucleotides. The two recognition sites of restriction enzymes used for Fig. 2 are indicated. The presumed termination site of transcription is indicated by an open box.
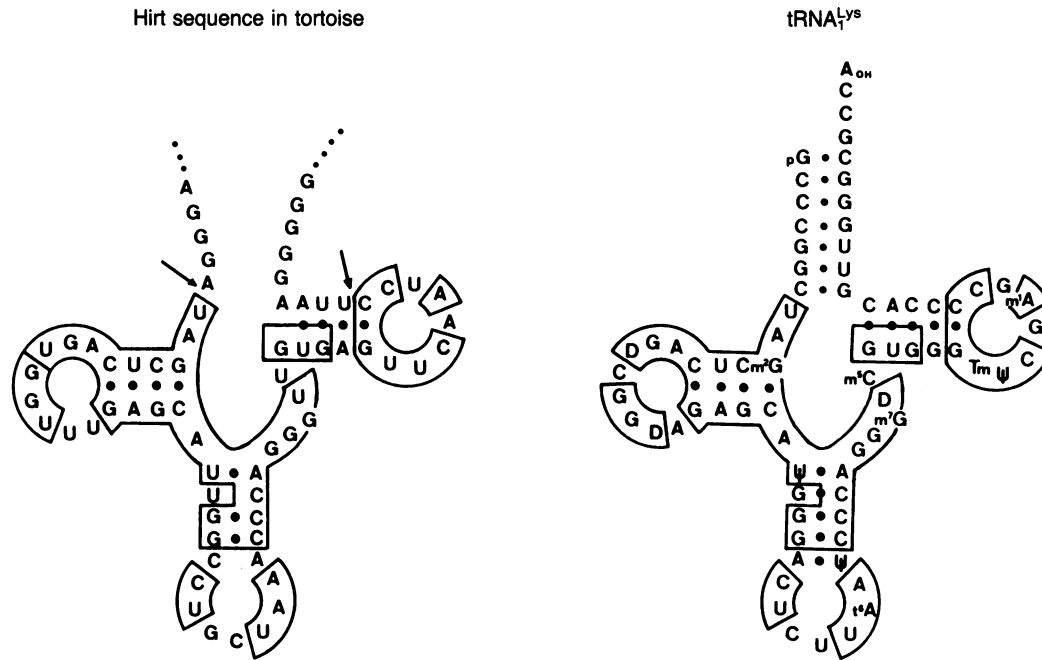
Hirt sequence in tortoise   tRNA$_1^{Lys}$

FIG. 5.   Sequence and structural homologies between the 5' part of the tortoise Hirt DNA cloned in phage TE6 and a tRNA$_1^{Lys}$ (rabbit). Identical sequences are boxed. The sequence of the tRNA is from ref. 15. Arrows indicate the sites of start and end of homology.

relatively recent event. Furthermore, we suggest that the 8S RNA family in newt, which gives a much more discrete transcript than the 6.5S RNA family, was amplified more recently than the 6.5S RNA family in tortoise. However, since the extents of homology between the two repetitive families and the two tRNAs are roughly the same, a recombinational event with a tRNA gene to yield an ancestral unit of the repetitive family may not always be immediately followed by amplification: amplification may sometimes occur coincidently with recombination and sometimes long after depending on the animal species.

That these animal species have so many tRNA-like structures in their genomes is surprising. Recently, Ames et al. (18) demonstrated that the leader mRNA of the histidine attenuator resembles tRNA^{His} and suggested that tRNA-modifying enzymes or the histidyl-tRNA synthetase might be involved in regulation of expression of the histidine operon by interacting with the tRNA-like structure. Furthermore, these authors proposed that a variety of modifying enzymes in both prokaryotes and eukaryotes might play a role in the regulation of gene expression (18). In this context, it is interesting to speculate that the expression of repetitive

families originating from specific tRNA genes may be regulated by tRNA-modifying enzymes or an aminoacyl-tRNA synthetase. Many kinds of modified nucleosides and modifying enzymes have been characterized (19, 20). Matsumoto et al. (21) recently demonstrated that a 6S RNA transcribed in vitro from salmon Hirt sequences contains pseudouridylic acid residues. The possible involvement of tRNA-related enzymes in the expression of some repetitive families is under study (23).

We proposed that six repetitive families known at present, that is, the rodent type 2 Alu family, rat ID sequence, rabbit C family, bovine or goat 73-base-pair repeat (3), tortoise Hirt sequences, and newt Hirt sequences have structural resemblances to specific tRNAs. Interestingly, in all these cases, they lack an aminoacyl stem region. Furthermore, in all cases but one, the homology starts from the site of U8. U8 is an invariant site in all tRNAs (16) and is supposed to be involved in transient covalent interaction with aminoacyl-tRNA synthetase in the aminoacylation reaction (22). Thus, it is tempting to speculate that cognate aminoacyl-tRNA synthetase played some role in a recombinational event to generate an ancestral unit gene of Hirt sequences.
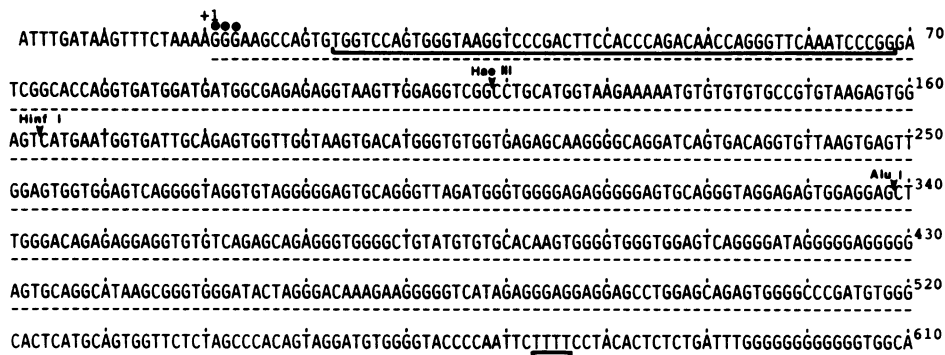

FIG. 6.   The nucleotide sequence of newt cloned DNA in phage NE1. The presumed initiation sites of transcription are shown by dots above the nucleotides. The region of Hirt sequences transcribed is shown by a dotted line under the nucleotides. The region homologous to tRNA is indicated by a thick line under the nucleotides. The three recognition sites of restriction enzymes used for Fig. 2 are indicated. The presumed termination site of transcription is indicated by an open box.
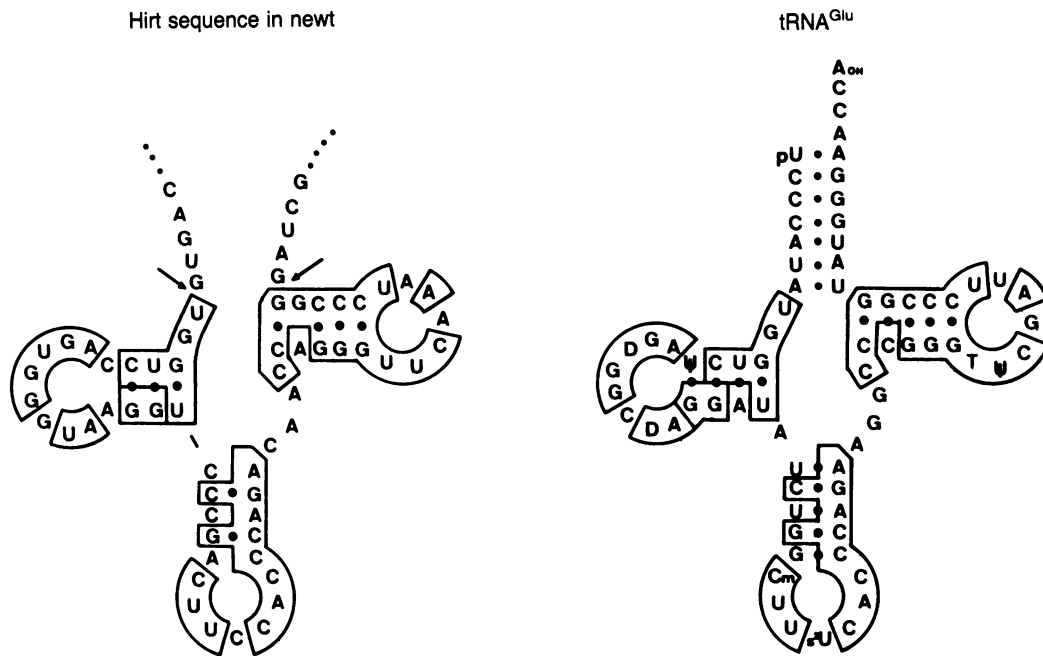
Hirt sequence in newt   tRNA^Glu

FIG. 7. Sequence and structural homologies between the 5' part of the newt Hirt cloned DNA in phage NE1 and a tRNA^Glu (*Drosophila*). Identical sequences are boxed. The sequence of the tRNA is from ref. 16. Arrows indicate the sites of start and end of homology.

Note Added in Proof. Daniels and Deininger (24) have recently reported that the highly repetitive sequences in the *Galago* genome are derived from a tRNA$_i^{Met}$ gene.

1. Singer, M. F. (1982) *Cell* **28**, 433–434.
2. Sakamoto, K., Kominami, R., Mishima, Y. & Okada, N. (1984) *Mol. Gen. Genet.* **194**, 1–6.
3. Sakamoto, K. & Okada, N. (1985) *J. Mol. Evol.* **22**, 134–140.
4. Okada, N., Endoh, H., Sakamoto, K. & Matsumoto, K. (1985) *Proc. Jpn. Acad.* **61**, Ser. B, 363–367.
5. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3**, 2303–2308.
6. Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
7. Talkington, C. A., Nishioka, Y. & Leder, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7132–7136.
8. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
9. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
10. Barrell, B. G. (1971) in *Procedures in Nucleic Acids Research*, eds. Cantoni, G. L. & Davies, D. R. (Harper, New York), Vol. 2, pp. 751–779.
11. Lerner, M. R., Andrew, N. C., Miller, G. & Steitz, J. A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 805–809.
12. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
13. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
14. Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276.
15. Raba, M., Limburg, K., Burghagen, M., Katze, J. R., Simsek, M., Heckman, J. E., RajBhandary, U. L. & Gross, H. J. (1979) *Eur. J. Biochem.* **97**, 305–318.
16. Sprinzl, M. & Gauss, D. H. (1984) *Nucleic Acids Res.* **12**, Suppl., r1–r131.
17. Ullu, E. & Tschudi, C. (1984) *Nature (London)* **312**, 171–172.
18. Ames, B. N., Tsang, T. H., Buck, M. & Christman, M. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5240–5242.
19. Nishimura, S. (1979) in *Transfer RNA: Structure, Properties and Recognition*, eds. Schimmel, P. R., Söll, D. & Abelson, J. N. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 59–79.
20. Okada, N. & Nishimura, S. (1979) *J. Biol. Chem.* **254**, 3061–3066.
21. Matsumoto, K., Murakami, K. & Okada, N. (1984) *Biochem. Biophys. Res. Commun.* **124**, 514–522.
22. Starzyk, R., Schoemaker, H. & Schimmel, P. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 339–342.
23. Sakamoto, K. & Okada, N. (1985) *Nucleic Acids Res.* **13**, in press.
24. Daniels, G. R. & Deininger, P. L. (1985) *Nature (London)* **317**, 819–822.