MOLECULAR AND CELLULAR BIOLOGY, May 1990, p. 1940–1949
0270-7306/90/051940-10$02.00/0

Vol. 10, No. 5

Selenocysteine tRNA[Ser]Sec Gene Is Ubiquitous within the Animal Kingdom

BYEONG J. LEE,[1] MALINI RAJAGOPALAN,[1] YEONG S. KIM,[1] KWANG H. YOU,[1] K. BRUCE JACOBSON,[2] AND DOLPH HATFIELD[1]*

Laboratory of Experimental Carcinogenesis, National Cancer Institute, Bethesda, Maryland 20892,[1] and Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831[2]

Received 6 October 1989/Accepted 15 January 1990

Recently, a mammalian tRNA which was previously designated as an opal suppressor seryl-tRNA and phosphoseryl-tRNA was shown to be a selenocysteyl-tRNA (B. J. Lee, P. J. Worland, J. N. Davis, T. C. Stadtman, and D. Hatfield, J. Biol. Chem. 264:9724–9727, 1989). Hence, this tRNA is now designated as selenocysteyl-tRNA[Ser]Sec, and its function is twofold, to serve as (i) a carrier molecule upon which selenocysteine is biosynthesized and (ii) as a donor of selenocysteine, which is the 21st naturally occurring amino acid of protein, to the nascent polypeptide chain in response to specific UGA codons. In the present study, the selenocysteine tRNA gene was sequenced from *Xenopus laevis*, *Drosophila melanogaster*, and *Caenorhabditis elegans*. The tRNA product of this gene was also identified within the seryl-tRNA population of a number of higher and lower animals, and the human tRNA[Ser]Sec gene was used as a probe to identify homologous sequences within genomic DNAs of organisms throughout the animal kingdom. The studies showed that the tRNA[Ser]Sec gene has undergone evolutionary change and that it is ubiquitous in the animal kingdom. Further, we conclude that selenocysteine-containing proteins, as well as the use of UGA as a codon for selenocysteine, are far more widespread in nature than previously thought.

The gene for mouse glutathione peroxidase has been sequenced and found to contain a TGA codon in an open reading frame (3). This observation has been confirmed in several laboratories by sequencing the glutathione peroxidase gene from the same (24) and other mammalian sources including humans (24, 35), bovines (24), and rats (30). The TGA codon within this gene corresponds to a selenocysteine moiety in the gene product. We have recently shown that a minor tRNA in mammalian tissues which recognized UGA in a ribosome binding assay and was identified as an opal suppressor seryl-tRNA (6) and phosphoseryl-tRNA (10, 23) is actually a selenocysteyl-tRNA (18), which is now designated as selenocysteine tRNA[Ser]Sec. These studies provide evidence that tRNA[Ser]Sec has a twofold function of serving as a carrier molecule upon which selenocysteine is synthesized and as a direct donor of selenocysteine to protein (e.g., glutathione peroxidase) in response to specific UGA codons (18). The gene encoding tRNA[Ser]Sec occurs in single copy in human (22, 27), rabbit (28), and chicken (11) genomes, and its expression is controlled by multiple 5′ extragenic regulatory elements (17). Transcription of this unique tRNA begins at the first nucleotide within the gene (16).

A similar tRNA gene whose expression is correlated with a selenocysteine moiety in formate dehydrogenase (20) that is coded by UGA (38) also occurs in *Escherichia coli*. Evidence was recently presented that this gene occurs in all enterobacteria (14). The product of this gene, which has a UCA anticodon, forms selenocysteyl-tRNA (19). This tRNA gene, like the corresponding one described above in mammalian and avian systems, codes for an isoacceptor which is responsible for inserting the 21st naturally occurring amino acid, selenocysteine, into protein in response to specific UGA codons.

Proteins containing selenocysteine have been characterized in higher vertebrates and bacteria (see reference 34 for a review). One possible means of determining whether selenocysteine may occur in proteins of other organisms is to examine genomes and/or tRNA populations for the presence of the selenocysteine tRNA gene or gene product. In the present study, we focused our attention on the animal kingdom to determine how widespread this tRNA gene may be in animals. We isolated and sequenced a tRNA[Ser]Sec gene from *Xenopus laevis*, *Drosophila melanogaster*, and *Caenorhabditis elegans*, which are three of the most thoroughly utilized organisms for scientific study below mammals and birds within the animal kingdom. In addition, we demonstrated that the tRNA[Ser]Sec gene product occurs within the seryl-tRNA population of a number of higher and lower animals and determined that the human gene can be used as a probe to identify homologous sequences in the genomic DNAs from a wide variety of higher and lower animals. We conclude from these studies that the tRNA[Ser]Sec gene, and thus the occurrence of selenocysteine in protein and the use of UGA as a codon for selenocysteine, is ubiquitous in the animal kingdom.

MATERIALS AND METHODS

Genomic libraries, isolation of tRNA[Ser]Sec genes, and DNA sequencing. A *Xenopus* genomic DNA library prepared by cloning a partial *Mbo*I digest of *Xenopus* blood genomic DNA into the *Bam*HI site of λJ1 (4) was obtained from I. Dawid, National Institute of Child Health and Human Development, Bethesda, Md.; a *Drosophila* genomic DNA library prepared as a Not-Bam-Not-CoSpeR library by J. W. Tamkun and J. Kennison was obtained from J. W. Tamkun, University of Colorado, Boulder; and a *C. elegans* genomic DNA library prepared by cloning a partial *Sau*3AI digest of *C. elegans* genomic DNA into the *Bam*HI site of λEMBL4 was obtained from S. Ward, University of Arizona, Tucson. Screening of the *Xenopus*, *Drosophila*, and *C. elegans* libraries, bacteriophage and plasmid preparation and isolation, gel electrophoresis, restriction analysis of DNA, sub-

* Corresponding author.

```
                                                                       -540        -520
Xenopus                                           GC CTTGGCAGAA CATATCCATC GGTCCGCCAT CTCCAGCAGC ACGCGGCGCA

              -500                  -480                  -460                  -440                  -420
Xenopus       TCTCGGACTG CTCCCCCTCC CATTCTCCCG CAGCATCTTT CCCGCGCTCT ATCCTTTGCT CCCTGTGACT CCTTCCCTCA TTACACTGAT TTCTCACCCC

              -400                  -380                  -360                  -340                  -320
Xenopus       CATTCTCCGT CACCTCCTCT GACGCGGAAA CTCTGTGTGA CTCAGGCTGA CCCCCATTGT CACTGTAAAG ACCAACAGCA GACCAGCGCA GTGATGTATC

C. elegans                                                                                                GCTTATC ATCATTGTCG

              -300                  -280                  -260                  -240                  -220
Xenopus       CCACATCCAC TAACAAACAG CAACCACACG CCCCTCCTCC CTCTTCCCGT TTTTCATTGA AAATAAACCG AAGCGTATTG TTATGGAAGT ACCAGCATGC

C. elegans    TCTACGGTGG TCTCATCCTC TTCTCGGCCT TCCTTCTCTA TGATACGCAG AGACTCGTCA AAAAGGCTGA AAATCACCCA CATTCATCTC AATTGTACGG

              -200                  -180                  -160                  -140                  -120
Xenopus       CTCGCGCGCG TGTATGCTAT GCGATCTTGT TTGATTTCCC TCGATTTCAC GTTAGACTAG TCGGGGTATG TAAGCGGCGA TACGTTTAAC TAGAAAAGGA
Drosophila                                                                                       TG CATGCCTGCA GGTCGACTCT
C. elegans    TTCCGACATG CAAATCCCGC CATTTGACCC AATCAACGCG TGAGTTTTGG GTTTCAACAC ACACACACTC TGATGTCTCC TCGATTCGTG TACTCTCTCG

              -100                  -80                   -60                   -40                   -20
Xenopus       ATAGTCAGAT TTTTGATACC AGAAACAATT TTTTGTCACC CCAATATATA ATATAATGGT GGGAGGGGGT ATAAAAGGAA ATGGGAGTAC TGAGGTATTT
Drosophila    AGAGGATCTA CTGGAATTTG CTCTACACTT ACTCTTAAAA TTCCCAAGTG CTTATTCAGC ACAACGTGTA TATAATTAGG GAAACTCTAG CACACTAAAT
C. elegans    TTGCCGTGCC ACAGGCCTCC ATGGTGCAAT GCGAAACCGT GAGCGTCGGT TGCTTGAATG CCAGCGTGTT TATAAGCGAC ACCTCGATGG AGTTTTGATC

                               +20                   +40                   +60                   +80                  +100
Xenopus       GCCCGGATGA CCCTCAGTGG TCTGGGGTGC AGGCTTCAAA CCTGTAGCTG TCTAGCGACA GAGTGGTTCA ATTCCACCTT TCGGGCGAGG GCTTCTCCTT
Drosophila    GCCCCACTGA ACTTCGGTGG TCCGGGGTGC GGACTTCAAA TCCGTAGTCG ATTTGCGTCG AAGTGGTTCG ATTCCACCTG GGGGGCGAAT ATTAATGGAG
C. elegans    GCCCGGATGA ACCATGGCGG TCTGTGGTGC AGACTTCAAA TCTGTAGGCG GTTAGCGCCG CAGTGGTTCG ACTCCACCTT TCGGGTGTTT TAGCATATTT

                               +120                  +140                  +160                  +180                  +200
Xenopus       TTTCATCTCT TCTATCCTCT GATAAATAGC TCATCTCAGT GGCTATATAA TGCGCTTAGC AATATTCTTT ACCGAAACGG CTTTCTATTT TACATTCAGT
Drosophila    TTGCATTCTT AAACTTTTTT AAGCATGGTC TTTATTTTCC AAAATCTGGT TCATATGGCA TTGGCGCGCC TTGACTATCC AAGGTCGCAG AATCGCCACA
C. elegans    TTTGTTAAAT GTTTAATTTA AATATAAAAA TCTTATTAAT TATTAAATTT TTTCAGACAA ATGTCAATCT ACATGGATGT CCTGAATGTC TTTATGCGAC

                               +220                  +240                  +260                  +280                  +300
Xenopus       CACAGTATAC AGTATCCATA ACAACCAGGT TCCTTTTATT ATGTTTGTCA TTTCTGATTA CTTTAGTACA TGGGTAACTC CAGGTGGCCA GTCTATCAGT
Drosophila    TCGGCATAGA CACCGGGGTA GTTGTCGTCC GCACACCGGT AACCCCAGAT ACAATGCCCA CCAGCTGGCT ACTGGCCACC AAAGGACCTC CAGAGTCTCC
C. elegans    TTGTAATGAT TATGGGAGGA ATGAACGGAA ACCGAAGAAG TAGATCATCA ATCGATAAGC TTTAAT

                               +320                  +340                  +360                  +380                  +400
Xenopus       CGTAAGCTGC TCCCTTATAT CTTGCTCAAT GTGAACTTAA TATATCAGTA ACAGTTAATA CAGTTAAAAG GGAATAAAAA AAACATATGT GCAGTACATG
Drosophila    CGATCTAGGC ATCTGCATCA GTGCTGGCAG CGCAAATTGT TTCCTCGCCC ACAAAATCCG CACCGCACCG TAGCCAAACT TTTGCGAGGC ACACTCTCCG

                               +420                  +440                  +460                  +480                  +500
Xenopus       ATCACAGTTA TCTTATTGAG TGTAAGATTC TTTTCTGGAA GTTGCATTTT CCTTTAGTTT AATTCTATTA TGGACTAATT GTGCTTGTAT TTACAGGTAT
Drosophila    CGATCCCCGG GTACCGAGCT CGAA

Xenopus       GGGATC
```

FIG. 1. Anticoding sequences of the *Xenopus*, *Drosophila*, and *C. elegans* tRNA[Ser]Sec genes along with their 5'- and 3'-flanking regions. Bold letters (bases +1 to +87) show the gene, boxed bases (+36 to +38) show the region corresponding to the anticodon of the mature gene product, italicized letters in the 3' flank show the termination signal for RNA polymerase III transcription, and the underlined bases in the 5' flank near −30 show the TATA boxes. The sequences of the primary transcript (16) and of the first 100 nucleotides upstream of the *Xenopus* gene (17) have been presented elsewhere.

cloning of fragments, Southern blotting, and labeling of probes with ^{32}P were as previously described (11, 17, 27). The *Xenopus* library was screened with an RNA probe (11), and a 1,250-base-pair (bp) *Ava*I-*Pvu*II fragment was isolated from digests of recombinant DNA from a positive phage and subcloned between the *Ava*I-*Pvu*II site of pBR322 for further restriction analysis and sequencing. The *Drosophila* library was screened with a synthetic DNA probe (CTTCAAACCT GTAGCTGTCTAGCGACAGAGTGGTTCAATTCCAC CTTTCGGGCG; synthesized by Genetic Designs, Inc., Houston, Tex.), and a 546-bp *Sau*3AI fragment was isolated from digests of recombinant DNA from a positive phage and subcloned into the *Bam*HI site of the multiple cloning region of pUC18 for sequencing. The *C. elegans* library was screened with a 193-bp fragment encoding the human tRNA[Ser]Sec gene (22, 27), and a 620-bp *Cla*I fragment was isolated from digests of recombinant DNA from a positive phage and subcloned into the *Cla*I site of pBR322 for sequencing. The *Xenopus* gene (87 bp) and 209 bp of its 5'-flanking region and 171 bp of its 3'-flanking region were sequenced by the Maxam and Gilbert procedure (21) as described previously (11, 27), while 343 additional upstream base pairs of 5'-flanking region and 248 additional downstream base pairs of 3'-flanking region were sequenced by the dideoxy procedure of Sanger et al. (31). The *Drosophila* and *C. elegans* genes and their flanking regions were sequenced by the dideoxy technique (31).

Isolation of tRNA, its aminoacylation and chromatography, and ribosome binding studies. tRNA was isolated and aminoacylated with [^{3}H]serine (specific activity, 30 Ci/mmol; Amersham Corp., Arlington Heights, Ill.), the resulting [^{3}H]seryl-tRNA was fractionated over a reverse-phase chromatographic (RPC-5) column, and column fractions were pooled, isolated, and prepared for coding studies (12, 13).

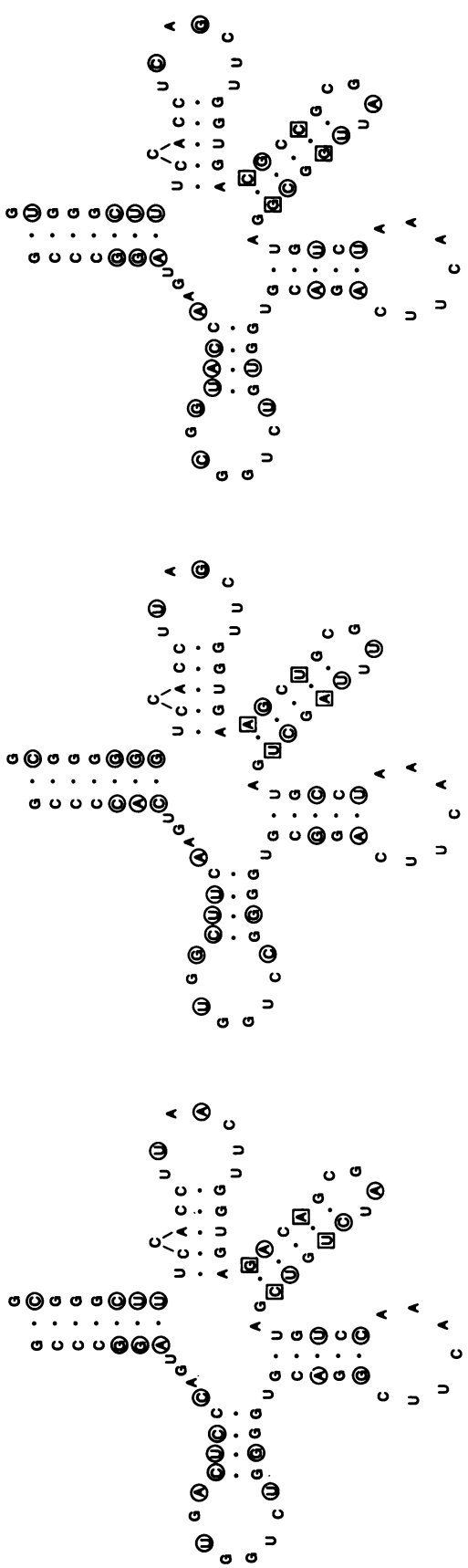

FIG. 2. Cloverleaf models of transcripts of and evolutionary changes in the *Xenopus*, *Drosophila*, and *C. elegans* $tRNA^{[Ser]Sec}$ genes. Although the $tRNA^{[Ser]Sec}$ gene products can be arranged in different cloverleaf models (6), base pairings in the dihydrouracil stem of *C. elegans* favor the model shown in the figure (see text). Circled and boxed letters are explained in the text.

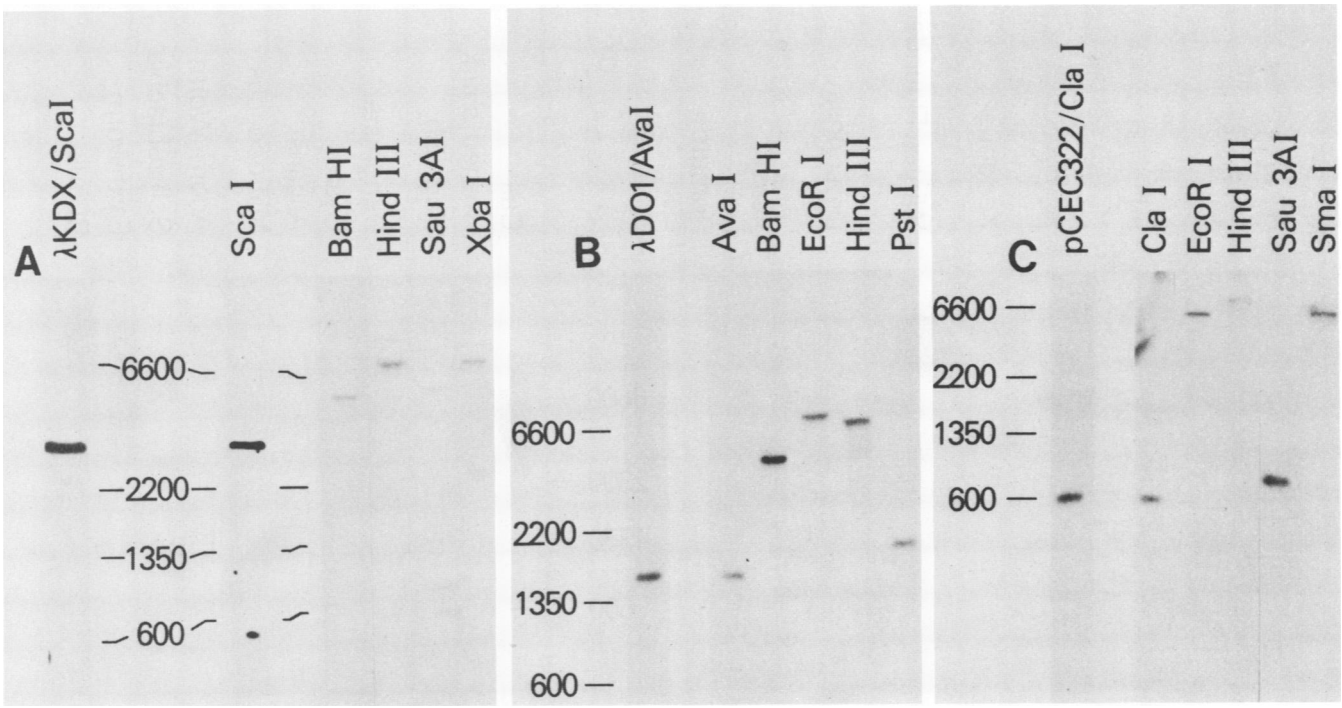

FIG. 3. Hybridization of total *Xenopus*, *Drosophila*, and *C. elegans* DNA with the corresponding tRNA[Ser]Sec genes. (A) The first lane represents the digestion of 1 μg of phage DNA encoding the *Xenopus* gene, the second lane represents 20 μg, and the third through sixth lanes represent 5 μg of *Xenopus* genomic DNA. (B) The first lane represents the digestion of 1 μg of phage DNA encoding the *Drosophila* gene and the other lanes represent 3.5 μg of *Drosophila* genomic DNA. (C) The first lane represents the digestion of 1 μg of phage DNA encoding the *C. elegans* gene and the other lanes represent 4.0 μg of *C. elegans* genomic DNA. The restriction enzymes used are shown in the figure. Approximately 10^8 cpm of ^{32}P-labeled probe of a 193-bp *Ava*I-*Hin*dIII fragment encoding the human gene, of a 210-bp *Xba*I-*Ssp*I fragment encoding the *Drosophila* gene, and of a 240-bp *Cla*I-*Sau*3AI fragment encoding the *C. elegans* gene were used in hybridization assays with the *Xenopus*, *Drosophila*, and *C. elegans* DNAs, respectively. Autoradiograms represent 24-h exposures. Numbers show number of nucleotides.

Drosophila tRNA was prepared from adult flies (1) and aminoacylated with [^{3}H]serine, the resulting aminoacyl-tRNA was fractionated on an RPC-5 column (see above), and column fractions were then denatured with glyoxal and bound to a nitrocellulose filter (0.2-μm pore size) by the procedure of E. Utz and R. Trewyn (Ohio State University, Columbus, Ohio; personal communication), using a Manifold II slot-blot apparatus from Schleicher & Schuell, Inc. (Keene, N.H.). The resulting filters were hybridized with a 210-bp *Xba*I-*Ssp*I fragment encoding the *Drosophila* gene. Organisms and tissues used for tRNA isolation and their sources were as follows: rainbow trout (*Oncorhynchus mykiss*) liver, W. B. Schill, National Fish Health Research Laboratory, Kearneysville, W.V.; marine snail (*Busycon carica*) kidney, J. Harasewych, Smithsonian Institution, Washington, D.C.; earthworm (*Lumbricus terrestris*), local bait store; *D. melanogaster*, Y. S. Kim, National Institutes of Health, Bethesda, Md.; horseshoe crab (*Limulus polyphemus*), acorn worm (*Saccoglossus kwoalevskii*), rock crab (*Cancer borealis*), surf clam (*Spisula solidissima*), ribbon worm (*Cerebratulus lacteus*), and the tips and stalks of *Bugula simplex*, J. Valois, Marine Biological Laboratory, Woods Hole, Mass. Binding of [^{3}H]seryl-tRNA to *E. coli* ribosomes in the presence and absence of UGA was done by the procedure of Nirenberg and Leder (26).

Hybridization studies with genomic DNA. Genomic DNAs were digested with *Bam*HI, the digests were electrophoresed on 1% agarose gels, and DNA was transferred to nitrocellulose filters. The filters were hybridized with 8.1×10^7 to 2.07×10^8 cpm of ^{32}P-labeled human DNA probe encoding the tRNA[Ser]Sec gene (22), and hybridization and washing of filters were as described previously (11, 27). Genomic DNAs were prepared by the procedure of Blair et al. (2) after quick-freezing organisms and flaking them in a Waring blender in the presence of liquid nitrogen. Organisms and their sources were as follows: acorn worm, ribbon worms (*Cerebratulus lacteus* and *Lineus socialis*), *Bugula simplex* (only the tips and stalks were used), and *Tubularia crocea* (a hydroid from which only heads that were parasite-free were used) from J. Valois, and earthworms from a local bait shop. Wheat germ was obtained from a local health food store, and an extract was prepared, genomic DNA was isolated (2), and the resulting DNA was banded on CsCl before use. *Mnemiopsis leidyi* (jellyfish) was obtained from the Marine Biological Laboratory frozen in seawater. It was then lyophilized, redissolved in 0.1 M Tris chloride (pH 8.0), and dialyzed against 0.1 M NaCl–0.01 M Tris chloride (pH 8.0)–0.001 M EDTA. Genomic DNA was isolated as described above and banded on CsCl before use. Genomic DNAs from other organisms and their sources are given in the figure legends.

RESULTS

We used three strategies in examining the distribution of the tRNA[Ser]Sec gene within the animal kingdom. Initially, we obtained genomic DNAs from a variety of organisms both in and outside the animal kingdom (as summarized in Fig. 5 below) and examined them by hybridization studies for sequences homologous to the human tRNA[Ser]Sec gene. Second, we isolated and sequenced the tRNA[Ser]Sec gene

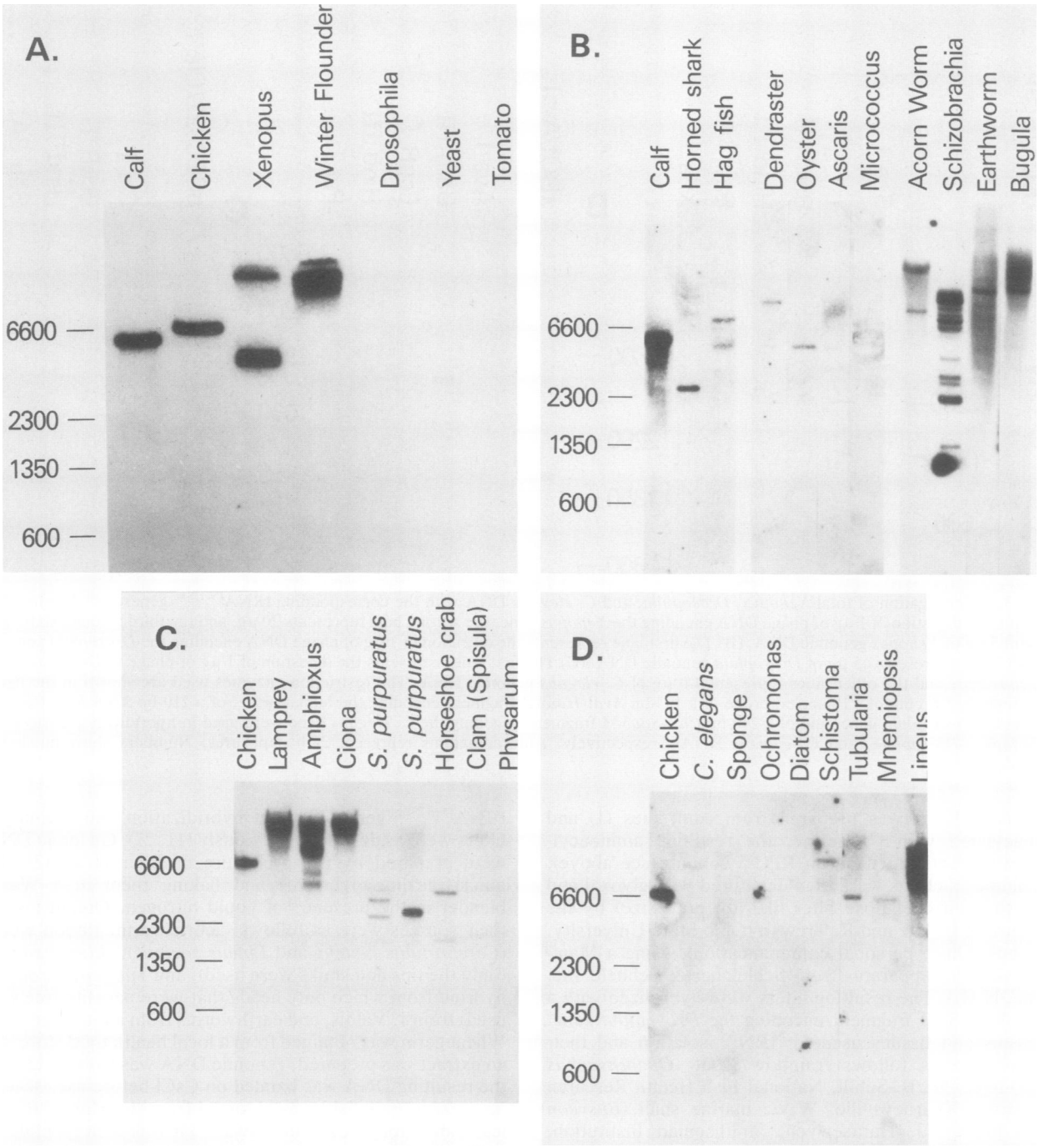

FIG. 4. Southern hybridization with the human tRNA$^{[Ser]Sec}$ gene of genomic DNAs from various organisms representing the five major kingdoms. Genomic DNAs were digested and electrophoresed, the DNA was transferred to nitrocellulose filters, and hybridization assays were done as described in the text. The autoradiograms were exposed for 24 h (A), 36 h (B), 32 h (C), and 40 h (D). Sources of genomic DNA and the amounts of digested DNA used for blotting were as follows: (A) calf thymus, 30 μg (ultrapure DNA from Sigma Chemical Co., St. Louis, Mo.); chicken, 20 μg (11); *Xenopus* erythrocyte cell, 60 μg (I. Dawid); winter flounder, 22.5 μg (R. Huang, The Johns Hopkins University, Baltimore, Md.); *D. melanogaster*, 10 μg (S. Haynes and I. Dawid); yeast, 5 μg (R. Wickner, National Institutes of Health); and tomato, 36 μg (T. Diener, Department of Agriculture, Beltsville, Md.); (B) calf, same as above; horned shark, 12 μg, hagfish, 9.8 μg (G. Litman, Showa University Research Institute, St. Petersburg, Fla.); *Dendraster excentricus* (sanddollar), 11.5 μg, oyster (*Crassostrea gigas*), 9.7 μg (A. H. Whiteley, University of Washington, Seattle); *Ascaris lumbricoides*, 8.0 μg (S. Ward); *Micrococcus lysodeikticus*, 7.4 μg (Sigma); acorn worm (*Saccoglossus kwoalevskii*), 11.0 μg (see text); *Schizobranchia insignis* (tubiculous polychaet), 11.2 μg (A. H. Whiteley); earthworm, 10.3 μg, and *Bugula simplex*, 12.3 μg (see text); (C) chicken, same as above; lamprey, 12.7 μg, amphioxus, 11 μg (R. Doolittle, University of California at San Diego); *Ciona intestilis* (tunicate), 10 μg (T. Meedle, Marine Biological Laboratory, Woods Hole, Mass.); *S. purpuratus*, DNA from different organisms, 10.2 μg, and from a single organism, 11.3 μg (R. Britten, California Institute of Technology, Corona Del Mar); horseshoe crab, 8.5 μg (S.-M. Cheng, National Institutes of Health); clam (*Spisula solidissima*), 9 μg (J. Ruderman, Duke University Medical Center, Durham, N.C.); and *Physarum polycephalum*, 11.1 μg (E. Johnson, Mt. Sinai Hospital, New York, N.Y.); and (D) chicken, same as above; *C. elegans*, 7.3 μg (T. Blumenthal, Indiana University, Bloomington); sponge (*Microciona*

from a higher vertebrate (*X. laevis*) and from organisms whose DNA responded weakly (*C. elegans*) or not at all (*D. melanogaster*) to the human gene. Finally, we examined the tRNA populations of a variety of higher and lower animals for the presence of the tRNA[Ser]Sec isoacceptor. In presenting the results of these studies, we begin with sequences of the tRNA[Ser]Sec genes, then examine hybridization of genomic DNAs with the human gene, and finally discuss the occurrence of the tRNA[Ser]Sec gene product within the tRNA population of a variety of animals.

Sequence of tRNA[Ser]Sec genes. The sequences of a 1,058-bp fragment of *Xenopus* DNA, a 546-bp fragment of *Drosophila* DNA, and a 583-bp fragment of *C. elegans* DNA encoding the tRNA[Ser]Sec gene from each organism were determined (Fig. 1). The regions encoding the tRNA genes which are 87 nucleotides long are shown in bold letters. Each gene contains a TCA sequence (designated by the boxed area) that corresponds to the anticodon in the tRNA product, which demonstrates that the isoacceptor decodes UGA codons. The termination sequence for transcription, which consists of a T cluster in the 3′-flanking region of the anticoding strand of tRNA genes, is shown in italics. Each gene contains a TATA box in its 5′-flanking region near −30 which is underlined. The TATA box is a regulatory site involved in transcription of the tRNA[Ser]Sec genes (17) as has also been shown in several other RNA polymerase III-transcribed genes (see references 5, 7, 25, 37, and references therein).

Evolutionary changes in tRNA[Ser]Sec gene. Comparison of the tRNA[Ser]Sec gene sequences showed that they have undergone evolutionary change. The differences can be more easily visualized in the transcript of each tRNA gene when they are presented in a cloverleaf model (Fig. 2). The bases within the transcripts which manifest a single difference between *X. laevis*, *D. melanogaster*, and *C. elegans* are circled in the figure, while those which are different in all three organisms are boxed. The most highly conserved regions of the tRNAs are the TψC stem and loop and the anticodon loop. Substantial evolutionary changes have occurred elsewhere in each gene. There are 22 base changes between *D. melanogaster* and vertebrates, 18 between *C. elegans* and vertebrates, and 20 between *D. melanogaster* and *C. elegans*. Interestingly, the sequence of the *Xenopus* gene product is identical to that of chickens (11) and differs from those of humans (27) and rabbits (28) by a single pyrimidine transition at position 11. Thus, the gene is highly conserved in higher vertebrates.

Alternative cloverleaf models of the selenocysteine tRNAs have been presented previously (6). However, the *C. elegans* tRNA does not fully base pair in the dihydrouracil stem in models other than that shown in the figure. Thus, this model may represent the more stable secondary structure of these tRNAs. Note that the two bases within the dihydrouracil loop which are next to the stem in each tRNA can also base pair, leaving a six-membered loop instead of the eight-membered loop shown.

tRNA[Ser]Sec genes occur in single copy. We have previously shown that the human (22, 27), rabbit (28), and chicken (11) tRNA[Ser]Sec genes occur in single copy within their respective genomes. The tRNA[Ser]Sec gene copy number was examined within the genomic DNA of *X. laevis*, *D. melanogaster*, and *C. elegans* (Fig. 3A, B, and C, respectively). The first lane of each graph is a digest of recombinant λ or plasmid DNA containing an insert from each organism encoding the tRNA[Ser]Sec gene. The second lane is genomic DNA from each organism digested with the same endonuclease as that used in first lane. In each case, the positive restriction fragment in phage DNA is identical to that found in genomic DNA, providing evidence that the homologous region encoding the tRNA[Ser]Sec gene from each organism was cloned and sequenced. Digests of genomic DNA with other restriction enzymes showed only a single positive band in *D. melanogaster* and *C. elegans*, suggesting the occurrence of a single-copy gene in these organisms. In *Xenopus* DNA, however, two positive bands were obtained in *Sau*3AI digests and apparently incomplete cleavage occurred in *Hin*dIII and *Xba*I digests. The *Sau*3AI digest suggests that the tRNA[Ser]Sec gene is polymorphic and/or that a pseudogene also occurs in *X. laevis*. A pseudogene has been observed in the human (27) and rabbit (28) genomes. We propose that the *Xenopus* gene is not polygenic since a single-copy gene has been observed in the genomes of humans, rabbits, and chickens and most certainly in *D. melanogaster* and *C. elegans*.

Southern hybridization of genomic DNA from organisms representing the five living kingdoms within the human tRNA[Ser]Sec gene. Genomic DNAs from a variety of organisms within the animal, plant, fungus, protist, and monera kingdoms were tested for their ability to hybridize with a DNA probe encoding the human tRNA[Ser]Sec gene. DNAs within the animal kingdom gave a positive signal or signals with the probe (Fig. 4A to D, and see the legend to Fig. 4). Some yielded a single positive signal (e.g., calf, chicken, horned shark, dendraster, oyster), while others gave multiple positive signals (e.g., hagfish, schizobrachia, horseshoe crab). It is not known whether the multiple positive signals may represent DNA that is polymorphic, polygenic, and/or may contain a high number of pseudogenes. Pseudogenes of tRNA[Ser]Sec have been observed in mammalian systems (27, 28). In *Strongylocentrotus purpuratus*, which contains 4 to 5% polymorphism (8), DNA from a single organism yielded one positive band when digested with *Bam*HI, while DNA

polifera), 8.7 μg (H. Boedtker, Harvard University, Cambridge, Mass.); *Ochromonas danica*, 12.6 μg (A. Coleman, Brown University, Providence, R.I.); diatom (*Cyclotella meneghiniana*), 9.9 μg (C. M. Bourne, University of Michigan, Ann Arbor); *Schistosoma mansoni* (flatworm), 9.8 μg (A. Sher, National Institutes of Health); *Tubularia crocea*, 13.1 μg, *Mnemiopsis leidyi*, 15.8 μg, *Lineus socialis*, 9.9 μg (see text). Genomic DNAs from other organisms were also examined (data not shown). Those DNAs which yielded positive signals with the human gene probe, the amounts of DNA used in hybridization assays, and their sources were *Lytechinus pictus*, 13.2 μg (P. Cserjesi, McGill University, Montreal, Quebec, Canada); *Strongylocentrotus droebachiensis*, 10.5 μg, rock crab (*Cancer productus*), 9.8 μg, freshwater snail (*Limnaea stagnalis*), 10.0 μg, *Sabellaria cementarium* (tube-dwelling polychaet), 9.6 μg (A. H. Whiteley); marine snail (*B. carica*) kidney, 10.9 μg (J. Harasewych); and *L. baileyana*, 2.1 μg, and *P. elongata*, 2.8 μg (red algae from which nuclear DNA was prepared; L. J. Goff, University of California, Santa Cruz). Those DNAs which did not hybridize with the human gene probe, the amounts of DNA used in hybridization assays, and their sources were wheat germ, 10.0 μg (see text); *Trichoderma harzianum*, 10.0 μg (R. Vilgalys, Department Agriculture, Beltsville, Md.); *L. major*, 12.1 μg (S. Beverley, Harvard Medical School, Boston, Mass.); *Paramecium tetraurelia*, 10.3 μg (J. Preer, Indiana University, Bloomington); *Tetrahymena thermophila*, 11.0 μg (M. Gorovsky, University of Rochester, Rochester, N.Y.); *E. coli*, 6.8 μg (M. Bustin, National Institutes of Health); *Clostridium sticklandii*, 8.0 μg (G. Garcia and T. Stadtman, National Institutes of Health); and *P. perforata*, 3.2 μg, and *G. lemaneiformis*, 3.5 μg (L. Goff).

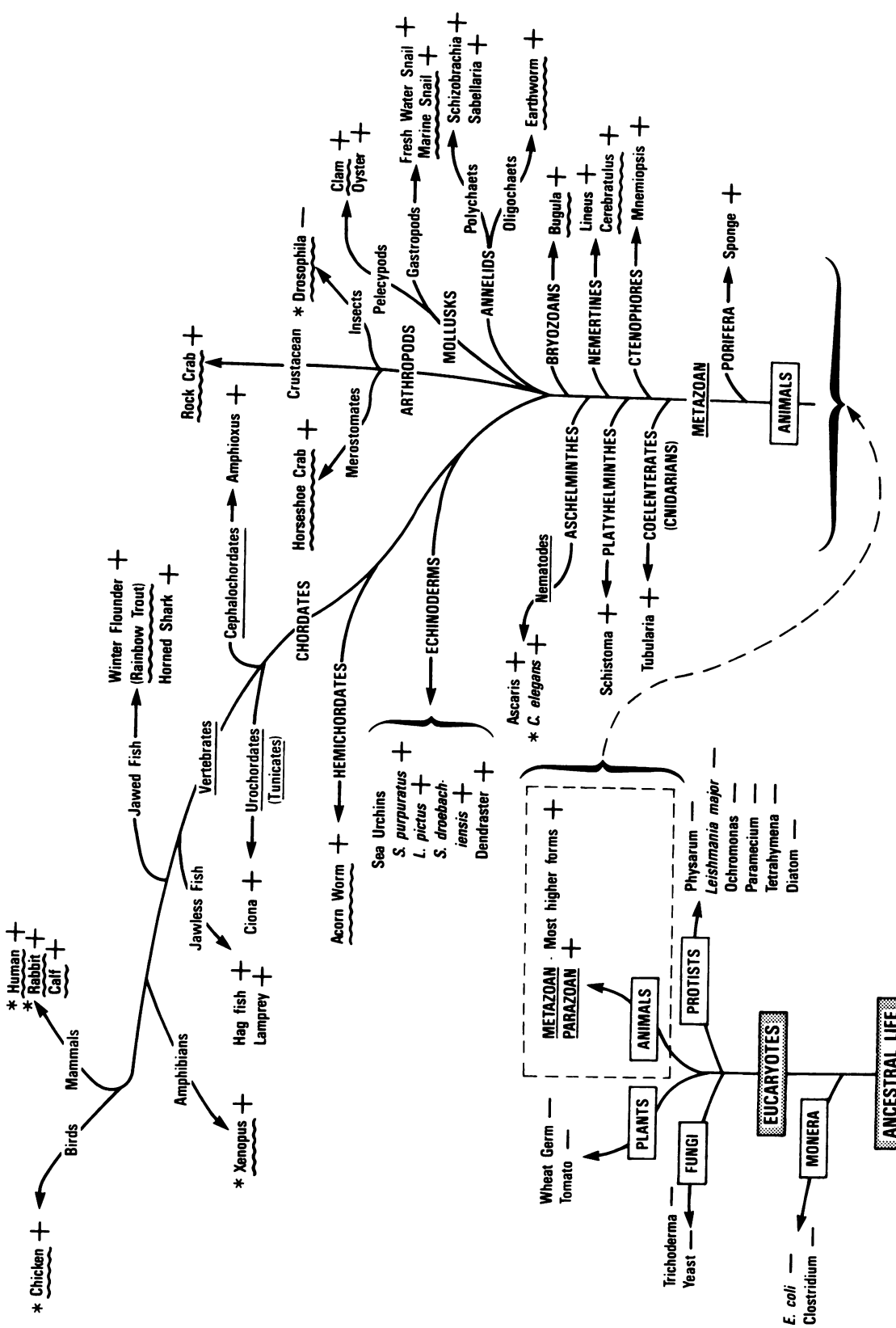

FIG. 5. Evolutionary tree showing the distribution of the selenocysteine tRNA$^{[Ser]Sec}$ gene. Kingdoms of organisms are shown in capital letters and are boxed; subkingdoms are in capital letters and are underlined; phyla are in capital letters; subphyla are in lowercase letters except for the initial capital letter and are underlined; classes are in lowercase letters except for the initial capital letter; and the organisms examined are shown at the ends of the branches of the tree. Symbols designate the following: +, positive response of restricted genomic DNA to the human tRNA$^{[Ser]Sec}$ gene probe; −, no response to the human tRNA$^{[Ser]Sec}$ gene probe; wavy underline, identification of the tRNA$^{[Ser]Sec}$ gene product within the seryl-tRNA population; and *, determination of the sequence of the tRNA$^{[Ser]Sec}$ gene. Sequences of the human (27), rabbit (28), and chicken (11) genes and isolation of the tRNA$^{[Ser]Sec}$ gene products within the human and rabbit reticulocyte (12), rabbit liver (13), calf liver (6, 10), and chicken liver (11) seryl-tRNA populations have been reported previously. All other observations in the figure are reported in this study.

pooled from different organisms yielded multiple positive bands (Fig. 4C, compare lanes 5 and 6). The lowest metazoan life forms tested, *Tubularia* and *Mnemiopsis* (Fig. 4D), hybridized with the probe, demonstrating the presence of homologous sequences to the human gene in the DNA of these organisms. Several DNAs responded weakly to the probe. For example, ascaris (Fig. 4B) and *C. elegans* (Fig. 4D), which are both nematodes, yielded weak responses. However, this positive signal represents homologous sequences between the tRNA[Ser]Sec genes of these organisms and humans, since the human gene was used as a probe to isolate the tRNA[Ser]Sec gene from *C. elegans*. The human and *C. elegans* genes have 79.3% homology to each other. Sponge DNA also hybridized weakly to the human probe, suggesting the presence of homologous sequences in this organism. The only animal DNA tested which did not hybridize with the probe was that from *D. melanogaster*. However, we were able to isolate the tRNA[Ser]Sec gene from *D. melanogaster* using a synthetic probe to screen a *Drosophila* genomic library. Thus, we conclude from these observations, and from those in which a tRNA[Ser]Sec isoacceptor was isolated from within the tRNA population from a number of higher and lower animals (see below), that the tRNA[Ser]Sec gene is ubiquitous within the animal kingdom.

DNAs outside of the animal kingdom did not hybridize with the human gene probe (see yeasts and tomato in Fig. 4A; *Micrococcus* in Fig. 4B; *Physarum* in Fig. 4C; *Ochromonos* and diatom in Fig. 4D; and wheat germ, *Trichoderma*, *Leishmania major*, *Paramecium*, *Tetrahymena*, *E. coli*, and *Clostridium* in the legend to Fig. 4). These observations suggest that even though a tRNA[Ser]Sec gene may occur in organisms outside the animal kingdom, the gene does not have sufficient homology to that of humans to hybridize with it under our conditions of assay. Also note that DNA from two species of red algae, *Lomentaria baileyana* and *Polysiphonia elongata*, hybridized with the human gene probe, while DNA from two other species of red algae, *Porphyra perforata* and *Gracilaria lameneiformis*, did not (data not shown). Since these are the only DNAs tested outside the animal kingdom that gave a positive signal with the probe, we are presently examining the significance of this response.

Occurrence of tRNA[Ser]Sec gene product within seryl-tRNA population. tRNA was isolated from a variety of organisms and aminoacylated with [^{3}H]serine, and the labeled seryl-tRNA population was fractionated over a reverse-phase chromatographic column. The late-eluting serine isoacceptors were isolated, and their response to UGA in a ribosome binding assay was determined. Fish (rainbow trout), acorn worm, clam, horseshoe crab, crab, marine snail, earthworm, *Bugula*, and *Cerebratulus* were found to contain an isoacceptor which recognizes the nonsense codon, UGA (data not shown). This isoacceptor corresponds to those identified in mammalian cells (for reviews, see reference 9 and D. Hatfield, D. W. E. Smith, B. J. Lee, P. J. Worland, and S. Oroszlan, Crit. Rev. Biochem., in press), which have been shown recently to form selenocysteyl-tRNA[Ser]Sec (18).

We were unable to demonstrate serine tRNA[Ser]Sec within the seryl-tRNA population of *D. melanogaster* by the ribosome binding assay. Therefore, fractionated *Drosophila* tRNA was assayed for the tRNA[Ser]Sec gene product by hybridization with a *Drosophila* tRNA[Ser]Sec gene probe (see Materials and Methods). tRNA fractions which hybridized with the probe corresponded to a minor, late-eluting peak of seryl-tRNA (data not shown), indicating that the gene is weakly expressed.

DISCUSSION

In the present study, a selenocysteine tRNA[Ser]Sec gene was isolated from a genomic DNA library of *X. laevis*, *D. melanogaster*, and *C. elegans* and sequenced. This gene is highly conserved in higher vertebrates but has undergone substantial evolutionary change between higher and lower animals (Fig. 2). The *Drosophila* and nematode genes have 77% homology to each other and 74.7 and 79.3% homology, respectively, to the genes of higher vertebrates. An examination of the number of evolutionary changes in other tRNA genes sequenced from these organisms shows that some have evolved as rapidly as the selenocysteine tRNA[Ser]Sec gene but that most have evolved more slowly (33).

We summarized our findings on the occurrence and distribution of the tRNA[Ser]Sec gene in Fig. 5. The + and − symbols following the names of the organisms indicate the response of their DNAs to the human gene probe. A positive response indicates homology to the human gene, even though in some cases the response is relatively weak (see *C. elegans* and ascaris in Fig. 4). The fact that we used the human tRNA[Ser]Sec gene as a probe to screen and isolate the corresponding gene from *C. elegans* provides evidence that the weak responses observed in this study are due to homologous sequences between different tRNA[Ser]Sec genes. With the exception of two species of red algae (see Results above), the only positive responses of genomic DNA to the human gene probe were within the animal kingdom. And within the animal kingdom, the only DNA tested which failed to hybridize with the human gene was that from *D. melanogaster*. However, we used a synthetic probe to isolate the tRNA[Ser]Sec gene from *D. melanogaster*, which was sequenced and found to have 74.6% homology to the human gene. Thus, it appears that the stringency of our hybridization conditions is such that we can detect homology of 79%, as observed between the human and *C. elegans* genes, but not 75%, as observed between the human and *Drosophila* genes.

An asterisk before the name of an organism in Fig. 5 shows that the tRNA[Ser]Sec gene has been sequenced from that source, which now totals six genes within the animal kingdom. A wavy underline shows that the tRNA[Ser]Sec gene product has been identified within the seryl-tRNA population of that organism; and thus far the tRNA gene product has been shown to occur in six vertebrates and in nine other animals including two lower forms, bugula and cerebratulus. Isolation of the tRNA[Ser]Sec gene and of the tRNA gene product from several higher and lower animals and hybridization of the human gene with genomic DNAs from a wide variety of higher and lower animals demonstrate that the gene is ubiquitous within the animal kingdom. Böck and collaborators (14) have shown that a corresponding selenocysteine tRNA gene which has been isolated and identified in *E. coli* (20) is distributed extensively in bacteria. Furthermore, selenium has been reported to induce glutathione peroxidase in algae (29, 36) and in yeasts (15) as it is known to do in mammalian cells (see reference 32 and references therein). These observations lead us to conclude that a selenocysteine tRNA gene, and thus the use of UGA as a codon for selenocysteine, is widespread, if not ubiquitous, in nature.

ACKNOWLEDGMENTS

We express our sincere appreciation to E. Utz and R. Trewyn for the use of their technique which identified tRNA[Ser]Sec within the

seryl-tRNA population of *D. melanogaster* and to the many individuals who so generously gave us genomic DNAs, genomic DNA libraries, and/or advice during these studies, especially A. Coleman, P. Cserjesi, I. Dawid, J. Harasewych, G. Litman, J. Valois, S. Ward, and A. Whiteley.

A portion of the work was sponsored by the office of Health and Environmental Research, U.S. Department of Energy, under contract DE-AC05-840R21400 with Martin Marietta Energy Systems, Inc.

ADDENDUM

The data involving the identification of a seryl-tRNA$^{[Ser]Sec}$ in the tRNA population of a variety of organisms by ribosome binding studies and in the tRNA population of *D. melanogaster* by hybridization studies which are not shown in the present study can be obtained from D.H. The tentative location of the tRNA$^{[Ser]Sec}$ gene has been mapped within the *C. elegans* genome by A. Coulson, J. Sulston, and R. Waterson, and the location can be obtained from A. Coulson and J. Sulston.

LITERATURE CITED

1. **Birchler, J. A., R. K. Owenby, and K. B. Jacobson.** 1982. Dosage compensation of serine-4 transfer RNA in *Drosophila melanogaster*. Genetics **102**:525–537.
2. **Blair, D., C. Cooper, M. I. Oskarson, L. Eaden, and G. Vande Woude.** 1982. New method for detecting cellular transforming genes. Science **218**:1122–1125.
3. **Chambers, I., J. Frampton, P. Goldfarb, N. Affara, W. McBain, and P. R. Harrison.** 1986. The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. EMBO J. **5**:1221–1227.
4. **Chein, Y.-H., and I. B. Dawid.** 1984. Isolation and characterization of calmodulin genes from *Xenopus laevis*. Mol. Cell. Biol. **4**:507–513.
5. **Das, G., D. Henning, D. Wright, and R. Reddy.** 1988. Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase. EMBO J. **7**:503–512.
6. **Diamond, A., B. Dudock, and D. Hatfield.** 1981. Structure and properties of a bovine liver UGA suppressor serine tRNA with a tryptophan anticodon. Cell **25**:497–506.
7. **Garcia, A. D., A. M. O'Connell, and S. J. Sharp.** 1987. Formation of an active transcription complex in the *Drosophila melanogaster* 5S RNA gene is dependent on an upstream region. Mol. Cell. Biol. **7**:2046–2051.
8. **Grula, J., T. Hall, J. Hunt, T. Giugni, G. Graham, E. Davidson, and R. J. Britten.** 1982. Sea urchin DNA sequence variation and reduced interspecies differences of the less variable DNA sequences. Evolution **36**:665–676.
9. **Hatfield, D.** 1985. Suppression of termination codons in higher eucaryotes. Trends Biochem. Sci. **10**:201–204.
10. **Hatfield, D., A. Diamond, and B. Dudock.** 1982. Opal suppressor serine tRNAs from bovine liver form phosphoseryl-tRNA. Proc. Natl. Acad. Sci. USA **79**:6215–6219.
11. **Hatfield, D., B. Dudock, and F. Eden.** 1983. Characterization and nucleotide sequence of a chicken gene encoding an opal suppressor tRNA and its flanking DNA segments. Proc. Natl. Acad. Sci. USA **80**:4940–4944.
12. **Hatfield, D., C. R. Matthews, and M. Rice.** 1979. Aminoacyl-tRNA populations in mammalian cells: chromatographic profiles and patterns of codon recognition. Biochim. Biophys. Acta **564**:414–423.
13. **Hatfield, D., F. Varricchio, M. Rice, and B. G. Forget.** 1982. The aminoacyl-tRNA population of human reticulocytes. J. Biol. Chem. **257**:3138–3188.
14. **Heider, J., W. Leinfelder, and A. Böck.** 1989. Occurrence and functional compatibility within Enterobacteriaceae of a tRNA species which inserts selenocysteine into protein. Nucleic Acids Res. **17**:2529–2540.
15. **Iizuka, M., M. Kousaku, and A. Kimura.** 1988. Induction of glutathione leakage from *Saccharomyces cerevisiae* cells by selenite. Agric. Biol. Chem. **52**:613–614.
16. **Lee, B. J., P. de la Peña, J. A. Tobian, M. Zasloff, and D. Hatfield.** 1987. Unique pathway of expression of an opal suppressor phosphoserine tRNA. Proc. Natl. Acad. Sci. USA **84**:6384–6388.
17. **Lee, B. J., S. G. Kang, and D. Hatfield.** 1989. Transcription of Xenopus selenocysteine tRNA^{Ser} (formerly designated opal suppressor phosphoserine tRNA) gene is directed by multiple 5′-extragenic regulatory elements. J. Biol. Chem. **264**:9696–9702.
18. **Lee, B. J., P. J. Worland, J. N. Davis, T. C. Stadtman, and D. Hatfield.** 1989. Identification of a selenocysteyl-tRNA^{Ser} in mammalian cells that recognizes the nonsense codon, UGA. J. Biol. Chem. **264**:9724–9727.
19. **Leinfelder, W., T. C. Stadtman, and A. Böck.** 1989. Occurrence *in vivo* of selenocysteyl-tRNA$^{Ser}_{UCA}$ in *Escherichia coli*. J. Biol. Chem. **264**:9720–9723.
20. **Leinfelder, W., E. Zehelein, M.-A. Mandrand-Berthelot, and A. Böck.** 1988. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. Nature (London) **331**:723–725.
21. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
22. **McBride, O. W., M. Rajagopalan, and D. Hatfield.** 1987. Opal suppressor phosphoserine tRNA gene and pseudogene are located on human chromosomes 19 and 21, respectively. J. Biol. Chem. **262**:11163–11166.
23. **Mizutani, T., and A. Hashimoto.** 1984. Purification and properties of suppressor seryl-tRNA:ATP phosphotransferase from bovine liver. FEBS Lett. **169**:319–322.
24. **Mullenbach, G. T., A. Tabrizi, B. D. Irvine, G. I. Bell, J. A. Trainer, and R. A. Hallewell.** 1988. Selenocysteine's mechanism of incorporation and evolution revealed in cDNAs of three glutathione peroxidases. Protein Eng. **2**:239–246.
25. **Murphy, S., C. Di Liegro, and M. Melli.** 1987. The *in vitro* transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. Cell **51**:81–87.
26. **Nirenberg, M., and P. Leder.** 1964. RNA codewords and protein synthesis. The effect of trinucleotides upon the binding of sRNA to ribosomes. Science **145**:1399–1407.
27. **O'Neill, V. A., F. C. Eden, K. Pratt, and D. Hatfield.** 1985. A human opal suppressor tRNA gene and pseudogene. J. Biol. Chem. **260**:2501–2508.
28. **Pratt, K., F. C. Eden, K. H. You, V. A. O'Neill, and D. Hatfield.** 1985. Conserved sequences in both coding and 5′ flanking regions of mammalian opal suppressor tRNA genes. Nucleic Acids Res. **13**:4765–4775.
29. **Price, N., and P. Harrison.** 1988. Specific selenium-containing macromolecules in the marine diatom *Thalassiosira pseudonana*. Plant Physiol. **86**:192–199.
30. **Reddy, A. P., B. L. Hsu, P. S. Reddy, N.-Q. Li, K. Thyagaru, C. C. Reddy, M. F. Tam, and C.-P. Tu.** 1988. Expression of glutathione peroxidase I gene in selenium-deficient rats. Nucleic Acids Res. **16**:5557–5568.
31. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
32. **Speier, C., S. S. Baker, and P. E. Newburger.** 1985. Relationships between in vitro selenium supply, glutathione peroxidase activity and phagocytic function in the HL-60 human myeloid cell line. J. Biol. Chem. **260**:8951–8957.
33. **Sprinzl, M., T. Hartmann, J. Weber, J. Blank, and R. Zeidler.** 1989. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. **17**:r1–r172.
34. **Stadtman, T. C.** 1987. Specific occurrence of selenium in enzyme and amino acid tRNAs. FASEB J. **1**:375–379.
35. **Sukenaka, Y., K. Ishida, T. Takeda, and K. Takagi.** 1987. cDNA sequence coding for human glutathione peroxidase. Nucleic

Acids Res. **15**:7178.
36. **Yokota, A., S. Shigeoka, T. Onishi, and S. Kitaoka.** 1988. Selenium as inducer of glutathione peroxidase in low-CO_2-grown *Chlamydomonas reinhardtii*. Plant Physiol. **86**:649–651.
37. **Young, L. S., N. Takahashi, and K. U. Sprague.** 1986. Upstream sequences confer distinctive transcriptional properties on gene encoding silk gland-specific $tRNA^{Ala}$. Proc. Natl. Acad. Sci. USA **83**:374–378.
38. **Zinoni, F., A. Birkmann, E. Leinfelder, and A. Böck.** 1987. Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. Proc. Natl. Acad. Sci. USA **84**:3156–3160.