ELSEVIER

Gene 291 (2002) 177–186

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression

Attila Karsi, Andrea Patterson, Jinian Feng, Zhanjiang Liu*

The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences, Auburn University, 201 Swingle Hall, Auburn, AL 36849, USA

Received 14 November 2001; received in revised form 3 March 2002; accepted 29 March 2002
Received by S. Salzberg

Abstract

Ribosomal protein (RP) genes have become widely used as markers for phylogenetic studies and comparative genomics. However, they have not been available for evolutionary studies in fish although teleosts are the largest group of vertebrates with more than 23,000 species. Using a transcriptomic approach, we have cloned and sequenced 32 40S RP complementary DNAs (cDNAs) from channel catfish (*Ictalurus punctatus*), making them one of the most complete sets of 40S RP gene sequences from a single organism. Most 40S RPs in channel catfish are highly similar to their orthologues in mammalian species, but S19, S21, and S25 are highly divergent. Only one type of cDNA was found for all RP genes except S26 and S27, for which two cDNAs were found in channel catfish. Alternatively spliced transcripts for the S3 and alternatively polyadenylated transcripts for S19 and S21 were found. The 32 40S RP genes are generally highly expressed and together they account for 5.33–11.42% of expression depending on the tissues. Expression levels of the RP genes were highly variable both within a single tissue among different RP genes and among tissues with regard to a single RP gene. Taken together, these data strongly suggest post-transcriptional regulation of RP gene expression, particularly in consideration of the stoichiometry of their representation in ribosomes.

Keywords: Transcriptome; cDNA; EST; Ribosomal protein; Gene expression; Teleost

1. Introduction

Ribosomes are responsible for protein synthesis in all cells and thereby link transcriptomes with proteomes. Because of their fundamental importance, numerous studies have been conducted concerning the structures of ribosomal proteins (RPs) and their interactions with RNAs (Wool, 1979, 1986; Wool et al., 1995, 1996; Draper and Reynaldo, 1999). Such understanding of structures is critical for a coherent molecular account of the functions. RPs have also been a model for the study of gene regulation, particularly in understanding the post-transcriptional regulation of gene expression (Perry and Meyuhas, 1990; Aloni et al., 1992; Meyuhas et al., 1996; Meyuhas, 2000).

Ribosomes have been an important resource for the study of molecular evolution because they are present in a wide spectrum of organisms ranging from bacteria to higher animals. As a result, the macromolecular components of ribosomes have been useful for evolutionary studies. Initial work was carried out using the small subunit ribosomal RNA (rRNA, Sogin, 1991), but it was soon revealed that rRNA-based phylogenies were sensitive to problems derived from drastic differences in GC content among taxa (Loomis and Smith, 1990; Galtier and Gouy, 1995). Several features of RP genes have made them potential alternatives in overcoming these problems and they have thus been used in recent phylogenetic studies (Liao and Dennis, 1994; Veuthey and Bittar, 1998; Yang et al., 1999). First, the DNA sequences can be translated into amino acid sequences overcoming the problem of GC content differences; second, there are large numbers of RPs that provide a large set of homologous sequences for analysis.

Mammalian ribosomes are composed 79 proteins and four RNAs (Wool et al., 1995 and personal communication; Warner and Nierras, 1998). The 60S ribosome is composed of three rRNAs and 47 RPs whereas the 40S ribosome is composed of the 18S rRNA and 32 RPs (Wool, 1979; Wool et al., 1995). The large number of RPs has complicated the quest for a complete understanding of their gene structures and genomic organizations. This was further complicated by

Abbreviations: RP, ribosomal protein; cDNA, complementary DNA; EST, expressed sequence tag; LB, Luria–Bertani; PCR, polymerase chain reaction; NCBI, National Center for Biotechnology Information; ORF, open reading frame; UTR, untranslated region

* Corresponding author. Tel: +1-334-844-4054; fax: +1-334-844-9208.
E-mail address: zliu@acesag.auburn.edu (Z. Liu).

0378-1119/02/$ - see front matter
PII: S0378-1119(02)00595-4

the fact that all RPs studied to date have large pseudogene families (Wool et al., 1996). Because of these difficulties, it was not until recently that all 79 human RP genes were mapped (Uechi et al., 2001). However, recent progress in genomics has provided alternative means for the systematic analysis of RP genes. RP genes are highly expressed in cells and, therefore, are highly represented in cDNA libraries constructed for expressed sequence tag (EST) analysis (Adams et al., 1991; Ju et al., 2000; Cao et al., 2001; Karsi et al., 2002). Due to this high representation, it is reasonable to efficiently compile a complete set of RP cDNA sequences by EST analysis.

Fish species make up the largest group of vertebrates. However, the primary structures of RPs and their encoding genes have not been completely available from any teleost species, including the model organism zebrafish, although teleosts are the largest group of vertebrates with more than 23,000 species (Helfman et al., 1997). To fill this gap, we have taken a transcriptomic approach to clone and analyze all the RP cDNAs from channel catfish (*Ictalurus punctatus*). Here we report 32 40S RP cDNAs, their structural features, and expression. In spite of high levels of evolutionary conservation of the 40S RP genes, the channel catfish RP genes exhibit several unique features including alternative polyadenylation, differential splicing, and multifunctional genes encoding for two RP mRNAs.

2. Materials and methods

2.1. cDNA library construction

Directional cDNA libraries were made from mRNAs isolated from the brain, head kidney, and skin tissues using the pSPORT-1 SuperScript Plasmid Cloning System from (Gibco BRL). Construction of the cDNA libraries followed the manufacturer's instructions. The brain, head kidney, and skin libraries have already been described (Ju et al., 2000; Cao et al., 2001; Karsi et al., 2002).

2.2. Plasmid preparation and sequencing

Each plasmid cDNA library was plated to a density appropriate for picking individual colonies. Random clones were grown in 1.5 ml LB medium overnight in 12×75-mm culture tubes. Plasmid DNA was prepared by the alkaline lysis method (Sambrook et al., 1989) using the Qiagen Spin Column Mini-plasmid kits. Three microliters of plasmid DNA (about 0.5–1.0 μg) were used in sequencing reactions. Chain termination sequencing (Sanger et al., 1977) was performed using cycleSeq-farOUT polymerase (Display Systems Biotech, Vista, CA). The PCR profiles were: 95 °C for 30 s, 55 °C for 40 s, 72 °C for 45 s for 30 cycles. An initial 2 min denaturation at 96 °C and a 5-min extension at 72 °C were always used. Sequences were analyzed on an automatic LI-COR DNA Sequencer Long ReadIR 4200 or LI-COR DNA Analyzer Gene ReadIR 4200.

2.3. Bioinformatic analysis

BLAST searches (http://www.ncbi.nlm.nih.gov) were conducted to determine gene identity. Matches were considered to be significant only when the probability (P) was less than 1×10^{-4} using BLASTN and BLASTX with all parameters set at the defaults. ESTs with significant similarities to RP genes in BLASTN searches were considered orthologues of RP genes after visual examination of the sequence alignment for extended length of identity. For most RP genes, multiple clones were encountered in the EST analysis (Fig. 2). Sequences were first edited using the Editseq program of the DNAStar software package. Sequences of each RP cDNA were aligned using the Megalign program of the DNAStar package. Sequencing errors were obvious when multiple sequences were aligned together. Regions with discrepancies were double-checked by converting sample files to chromatogram by SCF file creation program and analyzed by SCS file ViewIR program (Li-Cor, Lincoln, NE). Complete sequences of the cDNAs for each RP were then obtained by sequence analysis of multiple RP clones or by re-sequencing RP clones when necessary. One or more reference RP cDNAs from the human, rat, or mouse were used to guide the analysis of the proper open reading frames (ORFs).

3. Results and discussion

3.1. Systematic analysis of ribosomal protein genes using a transcriptomic approach

We used a transcriptomic approach for the systematic analysis of RP genes and their expression (Fig. 1). In the process of EST analysis from brain, head kidney, and skin tissues of channel catfish, 5338 cDNA clones were sequenced (Ju et al., 2000; Cao et al., 2001; Karsi et al., 2002). Bioinformatic analysis of these sequences revealed the candidate clones for all 40S RP genes, with 1–91 clones for each of the RP genes (Fig. 2). We strictly followed the annotation of the rat (Wool et al., 1996) because of the numerous synonyms existing in the GenBank databases as well as in the literature. For instance, the Sa RP protein was also known as the 40 kDa RP protein; RP S3a was also termed v-fos transformation effector; and S27a and S30 were derived from ubiquitin fusion proteins. Through the years, via a tour de force of technology and persistence, Wool and Chan (e.g. Chan et al., 1993) along with their colleagues have completed the sequences of the entire set of mammalian RPs. Following their nomenclature should avoid much of the current nomenclature confusion.

The transcriptomic approach is an efficient way for the systematic analysis of abundantly expressed genes. In this case, the RP genes were generally highly expressed and therefore, highly represented in the cDNA libraries. Using the routine molecular biological approaches, one has to obtain the cDNA clones for each RP gene by screening

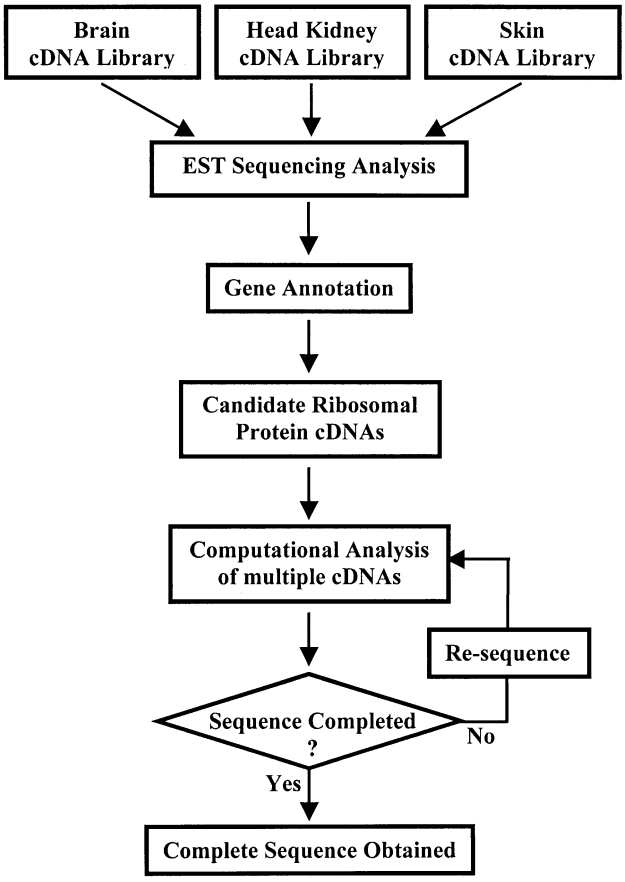

Fig. 1. Schematic representation of the transcriptomic approach to the systematic analysis of ribosomal protein cDNAs.

cDNA libraries, which would take a relatively long time and great efforts to obtain all the clones of RP cDNAs. In contrast, using the transcriptomic approach during the development of EST resources, we have obtained all 79 RP cDNAs from channel catfish including the 40S RP cDNAs reported here and the 47 60S RP cDNAs (Patterson, 2001).

3.2. Characteristics of the 40s ribosomal proteins

cDNA sequences of the 32 40S RPs were obtained by the alignments of multiple ESTs or by re-sequencing the entire insert of the clone containing the RPs. Of the 32 RP cDNAs, 30 had complete ORFs, while S3a and S15a had partial sequences missing the upstream sequences. All cDNA sequences have been deposited in the GenBank with consecutive accession numbers from AF402808 to AF402841 (Table 1).

The characteristics of the cDNAs including their coding regions, 5′- and 3′-untranslated regions (UTRs), and poly(A)$^+$ signals are shown in Table 1. All 32 40S RPs initiate at the first AUG. This is similar to the situation in the rat where all but L5 initiate at the first AUG. UAA is the most frequently used termination codon in the catfish RP genes (73.53%), which is different from the most frequently used termination codons UGA in vertebrates (Cavener and Ray, 1991). The 3′-UTRs are highly AT-rich. Of the 34 mRNAs (including two different mRNAs for S26 and S27), 33 have the typical AAUAAA polyadenylation signals (Proudfoot, 1991). Only S11 has AAUUAA poly(A)$^+$ signal. The poly(A)$^+$ tract starts 8–21 nucleotides from the poly(A)$^+$ signal (Table 1). The exception was S8 which had the typical AAUAAA poly(A)$^+$ signal 49 bases upstream of the poly(A) sites. However, a non-typical AUUAAA was found 16 bases from the poly(A)$^+$ sites. It is possible that this second poly(A)$^+$ signal was used.

In this research, we generated sequences for all the 40S RP genes in channel catfish. Along with our recent analysis of 47 60S RP cDNAs in channel catfish (Patterson, 2001), these 32 40S RP cDNAs complete the set of all 79 RP cDNAs in channel catfish. These sequence data should be of significant value for filling the existing gap in fish molecular biology. As the largest group of vertebrates, fish are important for an understanding of evolutionary biology as well as general molecular biology. Because of the ubiquitous nature of RP genes, the RP sequences should be of significant value for phylogenetic and mapping studies in fish. With the complete set of RP sequences, it is now possible to form peptidic ‘mega-sequence’ joining of many RPs to produce powerful phylogenetic markers (Veuthey and Bittar, 1998).

3.3. Comparison of the channel catfish 40s ribosomal proteins with those of mammals

The large majority of the 40S channel catfish RP proteins are highly similar to their mammalian counterparts as

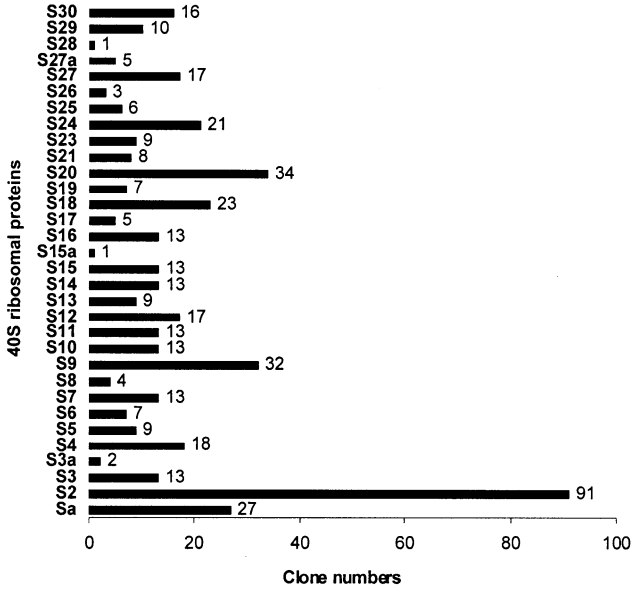

Fig. 2. Frequency of clones sequenced for each 40S ribosomal protein cDNAs. Ribosomal protein identities are shown on the left with the nomenclature of the rat. Numbers indicate numbers of clones analyzed for each ribosomal protein.

Table 1
Structural characteristics of the cDNAs encoding 32 ribosomal proteins of channel catfish (*Ictalurus punctatus*)[a]

Gene	Clone for complete sequences	Accession no.	Coding region	Available 5′-UTR	3′-UTR	Poly(A) from poly(A)$^+$ signal
Sa	IpHdk00494	AF402808	954	67	50	12
S2	IpHdk00495	AF402809	834	9	39	14
S3	IpHdk00297	AF402810	738	5	56	11
S3a	IpHdk02015	AF402811	801	ND	45	13
S4	IpHdk01279	AF402812	792	8	61	16
S5	IpSkn00600	AF402813	612	31	47	12
S6	IpHdk01450	AF402814	750	24	45	20
S7	IpSkn01931	AF402815	585	22	34	14
S8	IpSkn01233	AF402816	627	11	50	49 (AATAAA) 16 (ATTAAA)
S9	IpBrn00892	AF402817	585	14	78	15
S10	IpHdk01210	AF402818	501	23	36	11
S11	IpHdk01138	AF402819	480	7	74	11 (AATTAA)
S12	IpHdk01274	AF402820	399	42	22	9
S13	IpHdk03186	AF402821	456	17	38	13
S14	IpSkn00945	AF402822	456	16	39	12
S15	IpSkn00392	AF402823	438	41	34	13
S15a	IpBrn00086	AF402824	393	ND	41	16
S16	IpSkn01936	AF402825	441	6	68	13
S17	IpHdk02148	AF402826	405	12	42	16
S18	IpSkn02205	AF402827	459	12	42	12
S19	IpBrn01136	AF402828	444	13	153	8
S20	IpSkn01436	AF402829	360	83	73	15
S21	IpSkn00068	AF402830	252	47	54	15
S23	IpBrn00202	AF402831	432	30	42	16
S24	IpSkn00226	AF402832	396	19	66	11
S25	IpSkn01055	AF402833	375	19	60	13
S26-1	IpSkn00663	AF402834	348	10	53	13
S26-2	IpSkn01753	AF402835	348	6	56	21
S27-1	IpBrn01443	AF402836	255	10	246	12
S27-2	IpSkn01028	AF402837	255	20	202	16
S27a	IpHdk03163	AF402838	471	35	38	15
S28	IpBrn02110	AF402839	210	7	101	16
S29	IpSkn01504	AF402840	186	15	97	13
S30	IpBrn01069	AF402841	402	28	45	14

[a] Unless otherwise indicated, all poly (A)$^+$ signals were typical AAUAAA. Coding region, UTRs and poly(A) distance from poly(A)$^+$ signal are in bases. ND, not determined.

analyzed by deduced amino acid sequences (Table 2). The numbers of amino acids were highly conserved throughout evolution. Of the 32 40S RP cDNAs, 21 had ORFs with identical numbers of amino acids as in the rat; two had one extra amino acid; three had one fewer amino acid; and four had two extra amino acids (Table 2). Sa and S2 had significantly longer amino acid sequences than that of the rat, but S2 was shorter than that of the mouse.

The overall similarity of the channel catfish 40S RP proteins to those of the mammals was 94.3%. The most conserved 40S RPs were S18, S14, and S23; and the most divergent ones were S19, S21, and S25.

3.4. Channel catfish has two types of mRNAs for S26 and S27

All the channel catfish 40S RPs except S26 and S27 appeared to have only one type of mRNA. However, two types of cDNAs were found for both S26 and S27 in channel catfish. For the convenience of presentation, we named them S26-1 and S26-2 for the two types of cDNAs of S26, and S27-1 and S27-2 for the two types of cDNAs of S27, respectively.

S26-1 and S26-2 are highly divergent at the nucleotide level sharing only 72.3% identity (Fig. 3A), but encode highly similar proteins with 94.8% identity. Similar to the situation of S26, S27-1 and S27-2 are even more divergent at nucleotide level with only 55.3% identity (Fig. 3B), but they encode proteins that differ by only one amino acid. These different cDNAs for both the S26 and S27 are suspected to represent the transcripts of two different genes, respectively, since single nucleotide polymorphisms (SNPs) within coding regions are generally around 1–2% in channel catfish (Liu unpublished data). However, two functional genes are rare for RP genes in mammalian species. Although there are multiple copies of most mammalian RP genes, in no instance with the exception of the separate alleles encoding human S4 (Fisher et al., 1990), has it

Table 2
Comparison of the 32 channel catfish 40S ribosomal proteins with those of the rat and human

Gene	Number of amino acids			Similarity to (%)	
	Channel catfish	Rat	Human	Rat	Human
Sa	317	295	295	88.8	87.8
S2	277	250	221	92.4	95.0
S3	245	243	243	96.3	95.9
S3a	266	264	264	92.8	92.8
S4	263	263	263	92.8	92.8
S5	203	204	204	95.1	95.6
S6	249	249	249	94.4	94.4
S7	194	194	194	96.4	96.4
S8	208	208	208	90.4	82.2
S9	194	194	194	95.9	95.4
S10	166	165	165	89.7	90.9
S11	159	158	158	90.5	90.5
S12	132	130	132	94.6	95.5
S13	151	151	151	97.4	97.4
S14	151	151	151	98.7	99.3
S15	145	145	145	97.2	97.2
S15a	130	130	130	97.7	97.7
S16	146	146	146	97.3	97.3
S17	134	135	135	97.8	99.3
S18	152	152	152	98.7	98.7
S19	147	145	145	84.8	84.8
S20	119	119	119	98.3	98.3
S21	83	83	83	86.7	91.6
S23	143	143	143	98.6	98.6
S24	131	133	130	89.3	90.0
S25	124	125	125	86.3	86.3
S26-1	115	115	115	95.7	94.8
S26-2	115	115	115	98.3	97.4
S27-1	84	84	84	92.9	92.9
S27-2	84	84	84	92.9	92.9
S27a	156	156	156	98.1	98.7
S28	69	69	69	95.7	95.7
S29	56	56	56	98.2	98.2
S30	59	59	59	94.9	94.9

been shown that more than one of the genes is functional (Wool et al., 1996). The presumption is that the others are retroposon pseudogenes. In the case of human S4, there are two transcribed genes, one on the X chromosome and the other on the Y chromosome (Fisher et al., 1990). We recently found that two different genes encode channel catfish L5 (Patterson, 2001). Together, these data indicated that at least three RPs might be encoded by two genes in channel catfish.

RP S26-1 and S26-2, and S27-1 and S27-2 appeared to be expressed at an approximately equal rate. In the EST analysis of 5338 clones in the three tissues, two clones of S26-1 and one clone of S26-2 were sequenced. Similarly, seven clones of S27-1 and ten clones of S27-2 were sequenced. These expression patterns are different from that of the human S4X and S4Y, the only RP encoded by two functional genes in human. In male human ribosomes, S4X accounts for 90% and S4Y for 10% of S4 RP (Zinn et al., 1994). No tissue specificity was observed for the two types of S26 and S27 transcripts among the three tissues from which the ESTs were analyzed.

3.5. S19 has alternative polyadenylated transcripts

In this research, we used a transcriptomic approach that allowed the collection of large sets of cDNA clones for each RP mRNA from different tissues making it possible to align large numbers of cDNAs from a given gene. In this process, we found cDNA clones with alternative poly(A)$^+$ sites for S19 (Fig. 4). S19 has two typical poly(A)$^+$ signals of AAUAAA in the 3′-UTR located 126 bases apart. Analysis of cDNA clones revealed that both poly (A)$^+$ signals were used, generating S19-1 and S19-2 mRNAs.

In eukaryotes, two main sequence components are essential for effective polyadenylation. A conserved AAUAAA and a variable GU-rich element located 10–40 nucleotides from the 3′ end (Wahle and Keller, 1996; Colgan and Manley, 1997; Beaudoing et al., 2000). A coupled reaction of endonucleolytic cleavage followed by poly(A)$^+$ synthesis results in mature mRNAs. A significant fraction of mRNAs contains multiple poly(A)$^+$ (Gautheret et al., 1998). This research indicated that alternative polyadenylation might occur more frequently than assumed. We recently also found that RP L31 also has three types of mRNAs derived from differential polyadenylation. Such alternative polyadenylation may play a role in the post-transcriptional regulation of the RP gene expression by affecting the stabilities of the RP mRNAs (for review, see Shatkin and Manley, 2000; Macdonald, 2001).

3.6. RP S3 has multiple alternative spliced transcripts

Alignments of 13 clones of the S3 cDNA clones revealed the presence of three different types of S3 cDNAs (Fig. 5). They were named as S3-1, S3-2, and S3-3. S3-1 is the longest version of the S3 cDNA, while S3-2 and S3-3 are differentially spliced forms. As compared to S3-1, S3-2 has a deletion of 48 bp, and S3-3 has a deletion of 85 bp. Typical AG and GT splicing junction sequences exist in the S3-2, but are absent from S3-3.

S3-1 and S3-2 have proper ORFs whereas S3-3 is out of frame due to a deletion of 85 bp (non-integral of triplet codon). S3-1 encodes a protein of 245 amino acids, while S3-2 encodes a protein of 229 amino acids. In channel catfish, the major spliced form was S3-1. Among 13 clones analyzed, only one clone was found for S3-2 and S3-3, both of which were from head kidney library; the other 11 clones were S3-1. It is unknown at present if the shorter version of the S3 is biologically active, but it is clear that the production of the aberrant transcripts may reduce the overall expression of the S3 protein and thereby serve as one mechanism of posttranscriptional regulation of its expression. We suspect that alternative splicing may be generally underestimated simply because the methodologies used do not provide such opportunities (Mironov et al., 1999).

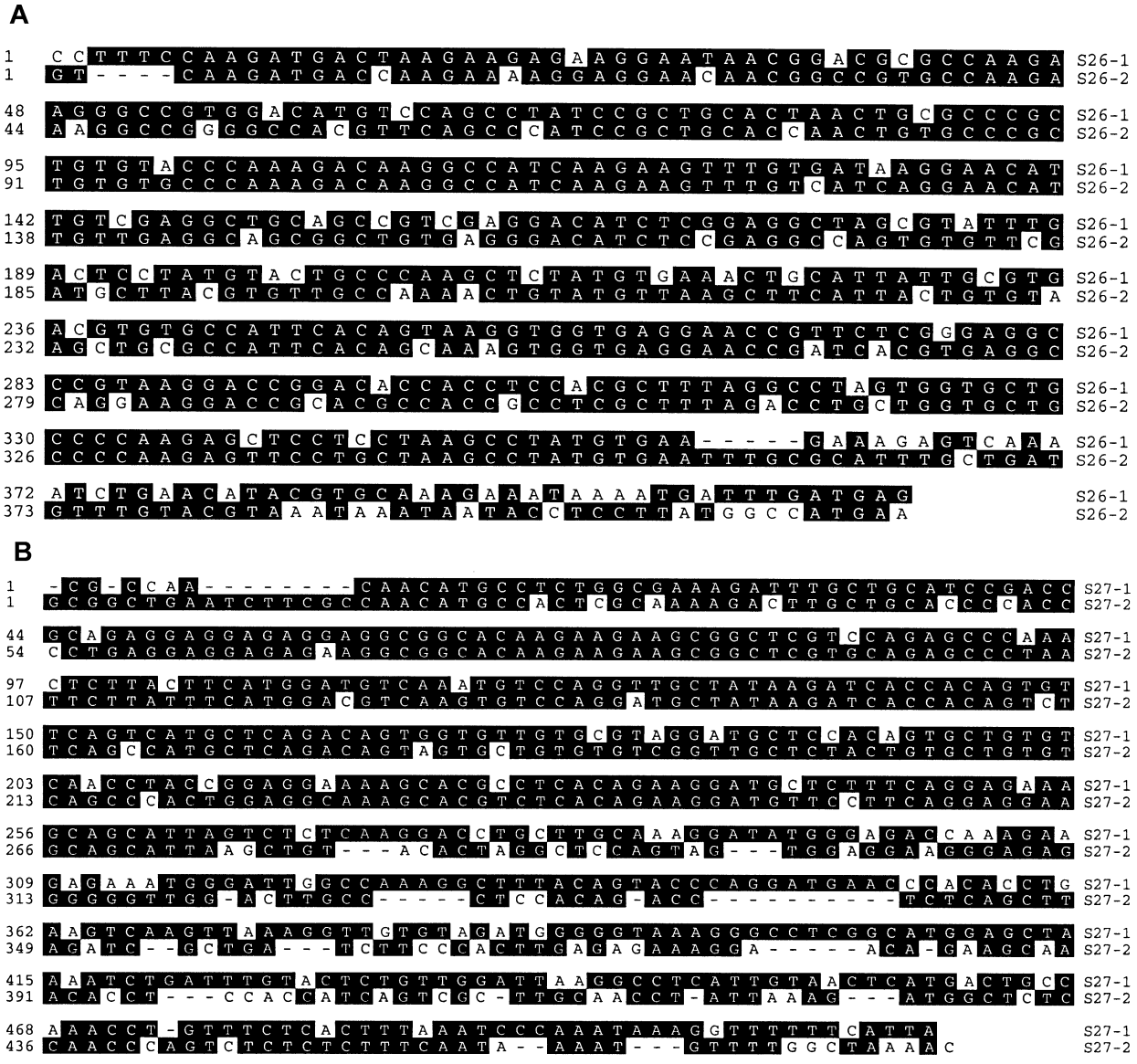

Fig. 3. (A) Alignment of the two types of S26 cDNAs showing extensive base substitutions and deletions. (B) Alignment of the two types of cDNAs of S27 showing extensive base substitutions and deletions. The sequences are only 55.3% identical.

3.7. RP S16 cDNA contains a polymorphic composite microsatellite sequences

RP S16 harbors a compound microsatellite composed of CT and CA repeats. What is interesting in this microsatellite locus is that both types of simple sequence repeats are short while both are expanding (Fig. 6) making this locus highly polymorphic. The CT repeat was detected to have alleles

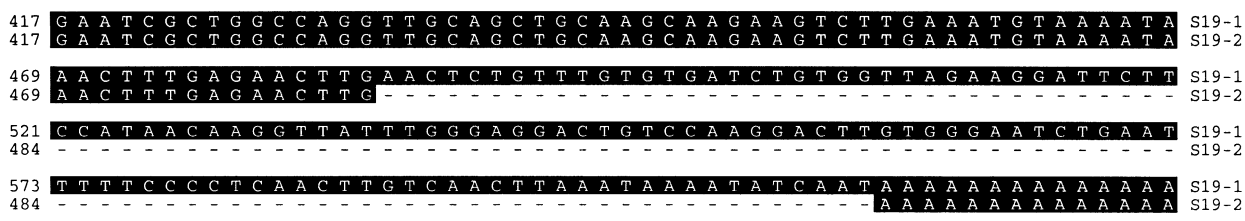

Fig. 4. Sequence alignment of the two alternative polyadenylated S19 transcripts showing only the downstream sequences close to the polyadenylation. The dashed region in S19-2 existed in the transcript of S19-1, but was absent from that of S19-2.

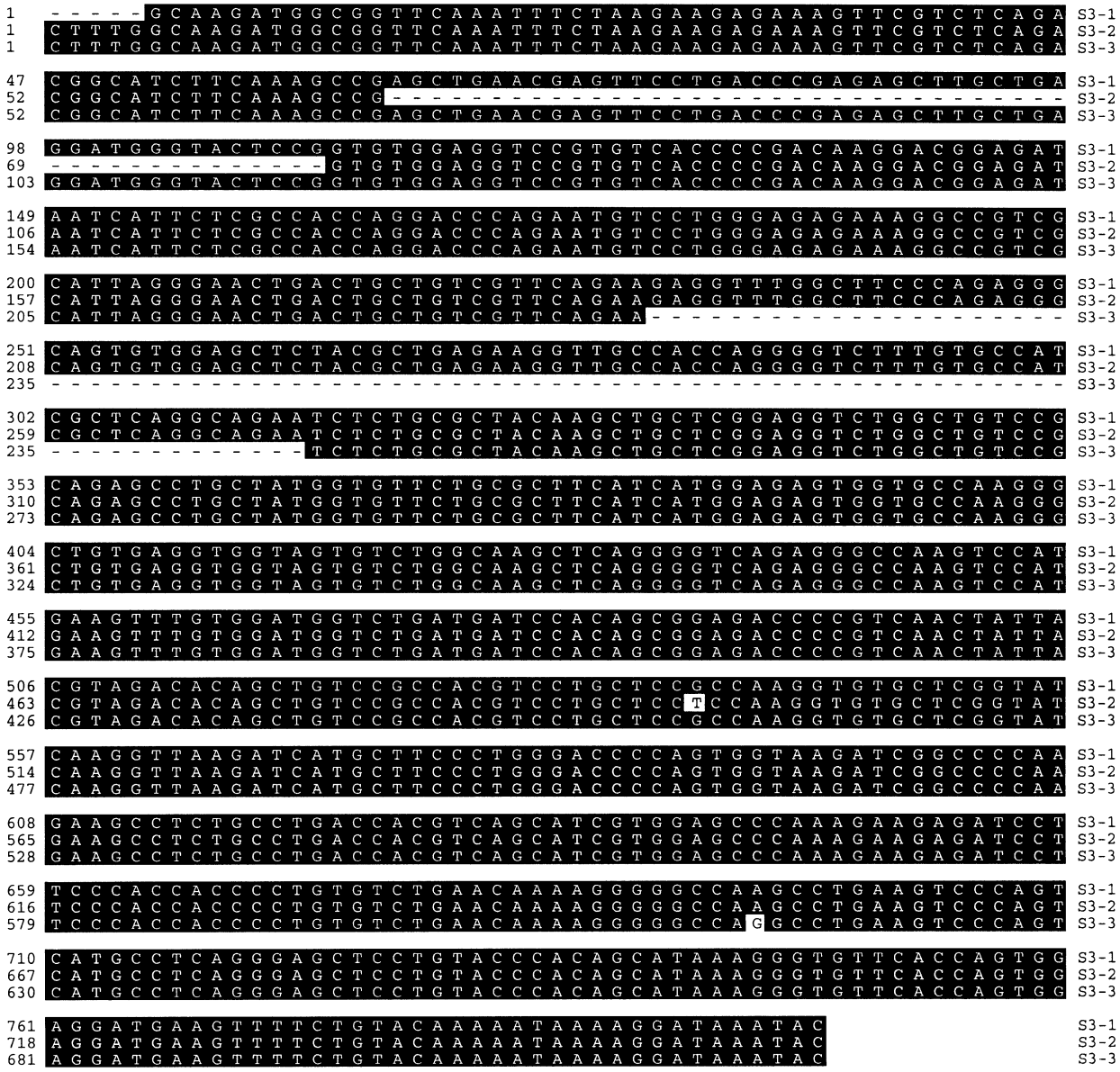

Fig. 5. Alternatively spliced transcripts of S3. The three splicing variants were named S3-1, S3-2, and S3-3. The dashed regions indicate sequences that were absent from specific spliced variants. White-boxed regions indicate single nucleotide polymorphism.

ranging from 8 to 11 repeats, while the CA repeat was found to have alleles with 4–5 repeats. This polymorphic microsatellite should be useful for the genomic mapping of the S16 gene to the genetic linkage map.

3.8. Gene expression of the 40s ribosomal proteins in brain, head kidney, and skin tissues

The gene expression data were obtained from expression profiles of the EST analysis. Expression of 40S RPs in the cDNA libraries of head kidney, brain, and skin are shown in Fig. 7. Overall, 40S RP genes were expressed at high levels accounting for 8.86% of the 5338 sequenced clones. Their expression in the three tissues ranged between 5.33 and 11.42% (Fig. 7). Skin tissue had the largest expression of the 40S RPs accounting for 11.42% of 1909 sequenced clones, while the brain tissue had the lowest overall RP expression accounting for 5.3% of 1201 sequenced clones. The head kidney exhibited an intermediate level of overall RP expression accounting for 8.6% of the 2228 sequenced clones. This overall greater representation of RP cDNAs in the skin could be caused by higher translational activities in the skin because one would expect that tissues with active translation demand more RPs for their ribosomes. The larger representation of RP cDNAs in the skin library was also found in the 60S RP cDNAs (Patterson, 2001).

Of the three analyzed libraries, on average, S2 had the highest expression representing 1.7% of overall expression.

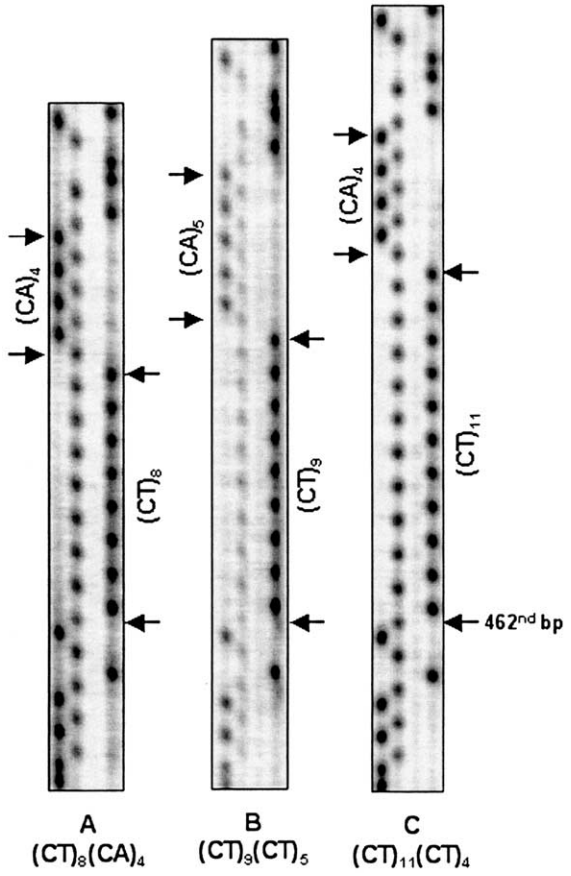

Fig. 6. Sequencing gels showing microsatellite variations within the S16. Note that both types of microsatellite repeats were variable in the number of repeat units.

Several other 40S RPs were also highly expressed including Sa, S9, and S20, all expressed at a rate of 0.5% or more; while several other RPs including S3a, S15a, and S28 were

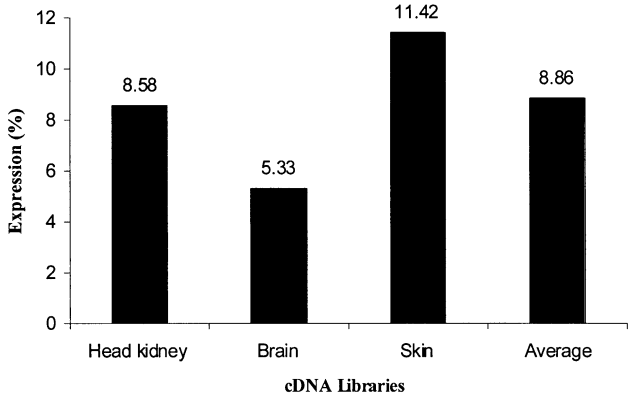

Fig. 7. Overall expression of 40S ribosomal proteins in three different tissues. The percentage of overall expression was obtained by dividing the total number of the 40S ribosomal protein clones sequenced from a specific library with the total number of clones sequenced from that library. Average overall expression was obtained by dividing the total number of 40S ribosomal protein clones sequenced from all three libraries with the total of 5338 clones sequenced.

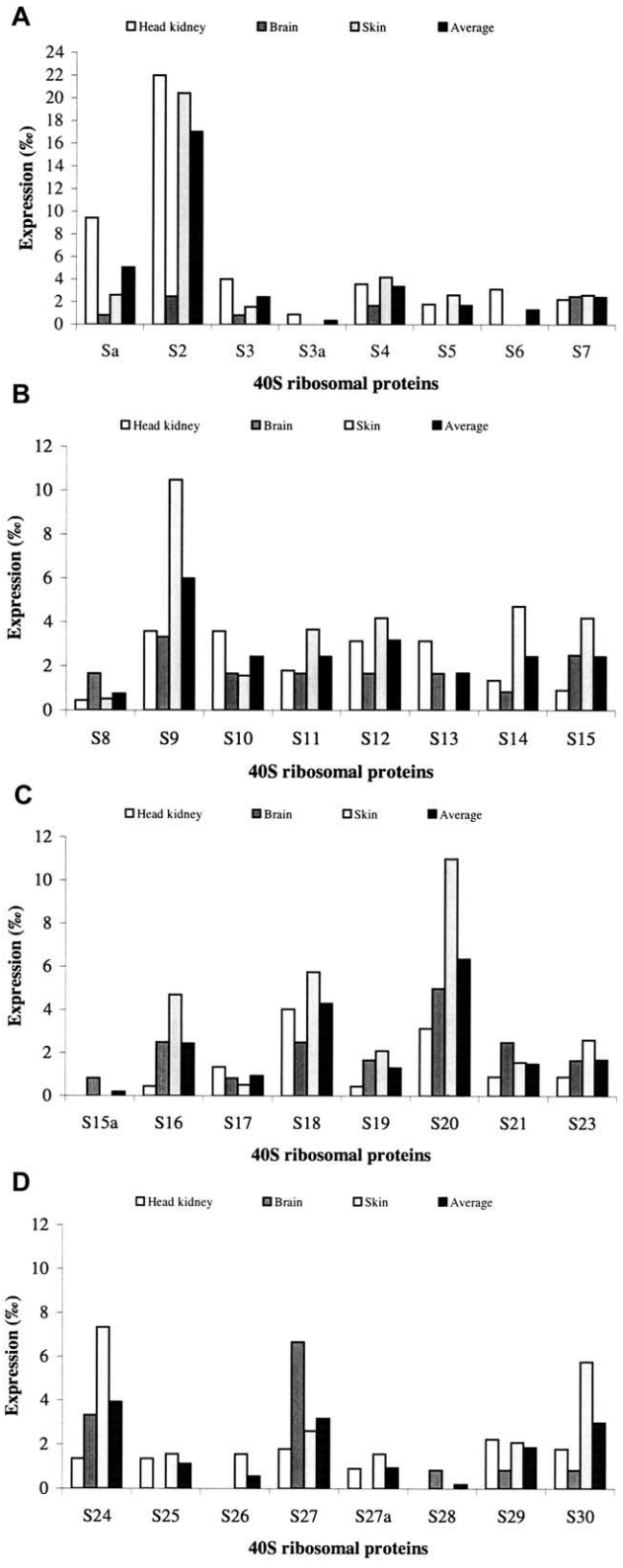

Fig. 8. Expression levels of the 40S ribosomal protein genes in three different tissues: head kidney, brain, and skin, as determined by expressed sequence tag profiling. The average was obtained from all three tissues. Expression levels are expressed as the number of 40S ribosomal protein clones per 1000 clones.

expressed at relatively low levels with 0.05% or less representation from the cDNA libraries (Fig. 8).

In spite of the specific stoichiometric ratio of RPs as required by the ribosome structure (Wool, 1979), drastically different levels of gene expression were observed within a single tissue/organ or among the tissues (Figs. 7 and 8). Strong tissue preference was observed for some of the highly expressed RPs. For instance, S2 was expressed highly in the head kidney and skin, but relatively low in the brain. Similarly, Sa was highly expressed in the head kidney; S9, S20, S24, and S30 were highly expressed in the skin. S27 was more highly expressed in the brain than in the head kidney or in the skin (Fig. 8). The highly variable mRNA levels, as represented by their cDNAs, of various RP genes within a tissue as well as the wide variations among tissues strongly suggest significant regulation at the post-transcriptional levels to meet the equimolar requirement of the RPs in the ribosomes.

Acknowledgements

This project was supported by a grant from USDA NRICGP (#98-35205-6738) and by the Auburn University Competitive BioGrant (Biogrant J. Liu 99). We appreciate the support of Auburn University Department of Fisheries and Allied Aquacultures, College of Agriculture, and the Vice President for Research for their matching funds to USDA NRI Equipment Grants (98-35208-6540, 99-35208-8512). The authors wish to thank Dr. Ira Wool at the University of Chicago for his kind advice and review of the manuscript. We would also like to thank Dr. Bernhart Kaltenback for allowing us to use his DNAStar software and Dongfeng Cao, Arif Kocabas, Zhenlin Ju, and Ping Li for their contribution in the EST analysis.

References

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., Venter, J.C., 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252, 1651–1656.

Aloni, R., Peleg, D., Meyuhas, O., 1992. Selective translational control and nonspecific posttranscriptional regulation of ribosomal protein gene expression during development and regeneration of rat liver. Mol. Cell. Biol. 12, 2203–2212.

Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.M., Gautheret, D., 2000. Patterns of variant polyadenylation signal usage in human genes. Genome Res. 10, 1001–1010.

Cao, D., Kocabas, A., Ju, Z., Karsi, A., Li, P., Patterson, A., Liu, Z.J., 2001. Transcriptome of channel catfish (*Ictalurus punctatus*): initial analysis of genes and expression profiles from the head kidney. Anim. Genet. 32, 169–188.

Cavener, D.R., Ray, S.C., 1991. Eukaryotic start and stop translation sites. Nucleic Acids Res. 19, 3185–3192.

Chan, Y.L., Suzuki, K., Olvera, J., Wool, I.G., 1993. Zinc finger-like motifs in rat ribosomal proteins S27 and S29. Nucleic Acids Res. 21, 649–655.

Colgan, D.F., Manley, J.L., 1997. Mechanisms and regulation of mRNA polyadenylation. Genes Dev. 11, 2755–2766.

Draper, D.E., Reynaldo, L.P., 1999. RNA binding strategies of ribosomal proteins. Nucl. Acids Res. 27, 381–388.

Fisher, E.M., Beer-Romero, P., Brown, L.G., Ridley, A., McNeil, J.A., Lawrence, J.B., Willard, H.F., Bieber, F.R., Page, D.C., 1990. Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. Cell 63, 1205–1218.

Galtier, N., Gouy, M., 1995. Inferring phylogenies from DNA sequences of unequal base compositions. Proc. Natl. Acad. Sci. USA 92, 11317–11321.

Gautheret, D., Poirot, O., Lopez, F., Audic, S., Claveire, J.M., 1998. Alternate polyadenylation in human mRNAs: a large scale analysis by EST clustering. Genome Res. 8, 524–530.

Helfman, G.S., Collette, B.B., Facey, D.E., 1997. The Diversity of Fishes, Blackwell Science, Malden, MA.

Ju, Z., Karsi, A., Kocabas, A., Patterson, A., Li, P., Cao, D., Dunham, R., Liu, Z.J., 2000. Transcriptome analysis of channel catfish (*Ictalurus punctatus*): genes and expression profile from the brain. Gene 261, 373–382.

Karsi, A., Cao, D., Li, P., Ju, Z., Kocabas, A., Feng, J., Patterson, A., Mickett, K.D., Liu, Z., 2002. Transcriptome analysis of channel catfish (*Ictalurus punctatus*): Initial analysis of gene expression and microsatellite-containing cDNAs in the skin. Gene. in press.

Liao, D., Dennis, P.P., 1994. Molecular phylogenies based on ribosomal protein L11, L1, L10, and L12 sequences. J. Mol. Evol. 38, 405–419.

Loomis, W.F., Smith, D.W., 1990. Molecular phylogeny of *Dictyostelium discoideum* by protein sequence comparison. Proc. Natl. Acad. Sci. USA 87, 9093–9097.

Macdonald, P., 2001. Diversity in translational regulation. Curr. Opin. Cell Biol. 13, 326–331.

Meyuhas, O., 2000. Synthesis of the translational apparatus is regulated at the translational level. Eur. J. Biochem. 267, 6321–6330.

Meyuhas, O., Avni, D., Shama, S., 1996. Translational control of ribosomal protein mRNAs in eukaryotes. In: Hershey, J.W.B., Mathews, M.B., Sonenberg, N. (Eds.). Translational Control, Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 363–388.

Mironov, A.A., Fickett, J.W., Gelfand, M.S., 1999. Frequent alternative splicing of human genes. Genome Res. 9, 1288–1293.

Patterson, A., 2001r. Translational machinery of channel catfish: Complemetary DNA and expression of the complete set of 47 60S ribosomal proteins. M.Sc. Thesis, Auburn University, p. 57.

Perry, R.P., Meyuhas, O., 1990. Translational control of ribosomal protein production in mammalian cells. Enzyme 44, 83–92.

Proudfoot, N., 1991. Poly(A) signals. Cell 64, 671–674.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Shatkin, A.J., Manley, J.L., 2000. The ends of the affair: capping and polyadenylation. Nat. Struct. Biol. 7, 838–842.

Sogin, M.L., 1991. Early evolution and the origin of eukaryotes. Curr. Opin. Genet. Dev. 1, 457–463.

Uechi, T., Tanaka, T., Kenmochi, N., 2001. A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders. Genomics 72, 223–230.

Veuthey, A.L., Bittar, G., 1998. Phylogenetic relationships of fungi, plantae, and animalia inferred from homologous comparison of ribosomal proteins. J. Mol. Evol. 47, 81–92.

Wahle, E., Keller, W., 1996. The biochemistry of polyadenylation. Trends Biochem. Sci. 21, 247–250.

Warner, J.R., Nierras, C.R., 1998. Trapping human ribosomal protein genes. Genome Res. 8, 419–421.

Wool, I.G., 1979. The structure and function of eukaryotic ribosomes. Annu. Rev. Biochem. 48, 719–754.

Wool, I.G., 1986. Studies of the structure of eukaryotic (mammalian) ribosomes. In: Hardesty, B., Kramer, G. (Eds.). Structure, Function, and Genetics of Ribosomes, Springer, New York, pp. 391–411.

Wool, I.G., Chan, Y.L., Gluck, A., 1995. Structure and evolution of mammalian ribosomal proteins. Biochem. Cell Biol. 73, 933–947.

Wool, I.G., Chan, Y.L., Gluck, A., 1996. Mammalian ribosomes: the structure and the evolution of the proteins. In: Hershey, J.W.B., Mathews, M.B., Sonenberg, N. (Eds.). Translational Control, Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 685–732.

Yang, D., Kusser, I., Kopke, A.K., Koop, B.F., Matheson, A.T., 1999. The structure and evolution of the ribosomal proteins encoded in the *spc* operon of the archaeon (Crenarchaeota) *Sulfolobus acidocaldarius*. Mol. Phylogenet. Evol. 12, 177–185.

Zinn, A.R., Alagappan, R.K., Brown, L.G., Wool, I., Page, D.C., 1994. Structure and function of ribosomal protein S4 genes on the human and mouse sex chromosomes. Mol. Cell. Biol. 14, 2485–2492.