Available online at www.sciencedirect.com
SCIENCE DIRECT®
Gene 305 (2003) 151–160
GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES
www.elsevier.com/locate/gene

Translational machinery of channel catfish: II. Complementary DNA and expression of the complete set of 47 60S ribosomal proteins

Andrea Patterson, Attila Karsi, 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, 203, Swingle Hall, Auburn, AL 36849, USA

Received 11 September 2002; received in revised form 7 November 2002; accepted 29 November 2002
Received by W. Makalowski

Abstract

Ribosomal protein genes have become widely used as markers for phylogenetic studies and comparative genomics, but they have not been available in fish. We have cloned and sequenced a complete set of all 47 60S ribosomal protein cDNAs from channel catfish (*Ictalurus punctatus*), of which 43 included the complete protein encoding regions. Most ribosomal protein mRNAs in channel catfish are highly similar to their mammalian counterparts. However, L4, L14, and L29 are significantly shorter in channel catfish than in mammals due to deletions in the 3′ end of the gene. Two distantly related L5 cDNAs, L5a and L5b, were found in channel catfish. L5a is more similar to L5 in other vertebrates, while L5b showed significant levels of divergence, suggesting independent evolution of the two L5-encoding genes. The 47 ribosomal protein genes are generally highly expressed and together account for 11–14% of overall gene expression, depending on the tissues. Expression levels were highly variable both within a single tissue among different ribosomal protein genes, and among tissues with regard to a single ribosomal protein gene. Strong tissue preference expression was also observed for some ribosomal proteins. This set of ribosomal protein gene sequences represents one of the most complete sets from any single organism and will aid in fish phylogenetic and comparative genomic studies.

Keywords: Phylogeny; Expression; Ribosomal protein; Fish; Genomics; Catfish; Alternative polyadenylation

1. Introduction

The study of molecular evolution requires a battery of genes that are optimally informative at overlapping taxonomic levels (Simon et al., 1994; Regier et al., 1998). Ribosomes are subcellular particles that play a structural and functional role in the template-directed synthesis of proteins. As the common primordial ancestor, their basic structural and functional features have been preserved in all diverse descendants. As a result, the macromolecular components of the ribosome 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 ribosomal protein genes make them the sensible alternatives in overcoming these problems and, as a result, they have 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 to overcome the problem of GC content differences; second, there are large numbers of ribosomal proteins to provide a large set of homologous sequences for analysis; and third, with the rapid progress in genomics research, particularly in the analysis of expressed sequence tags (ESTs) (Adams et al., 1991), the availability of large numbers of ribosomal protein gene sequences from a wide spectrum of organisms is becoming a reality.

Genomic mapping of the evolutionarily conserved ribosomal proteins is also of great interest because of its value in understanding genome evolution. The sequence

Abbreviations: EST, expressed sequence tag; cDNA, complementary DNA; PCR, polymerase chain reaction; NCBI, National Center for Biotechnology Information; TIGR, The Institute of Genome Research.

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

0378-1119/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0378-1119(02)01183-6

information of these genes is the first required element for mapping them onto chromosomes. As the first complete map of the human ribosomal proteins was finally produced (Uechi et al., 2001), mapping of these genes in other vertebrates is gaining greater enthusiasm. Although fish represent the largest group of vertebrates, large numbers of ribosomal protein genes from this group have not been available. The objective of this study was to produce ribosomal protein coding sequences from channel catfish (*Ictalurus punctatus*), making them available for phylogenetic analysis in fishes. The sequence information will also allow development of mapping strategies for the ribosomal genes in catfish for comparative genomics. Following our previous report of all 32 40S ribosomal proteins from channel catfish and as part of our channel catfish genome research concerning genome structure, organization, and evolution, here we report the complete set of 47 60S ribosomal protein cDNAs, their expression, and evolutionary analysis.

2. Materials and methods

2.1. Construction of cDNA libraries

Directional cDNA libraries were made from mRNAs isolated from head kidney (Cao et al., 2001), brain (Ju et al., 2000), and skin (Karsi et al., 2002a) using the pSPORT-1 SuperScript Plasmid Cloning System (GIBCO/BRL, Bethesda, MD). Construction of the cDNA library followed GIBCO/BRL's instructions, except that the library was electroporated into ElectroMax DH12S cells. These cells are highly adapted to efficient electroporation and production of single-stranded phagemids (GIBCO/BRL), features advantageous to the development of normalized cDNA libraries.

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 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 (LI-COR, Inc., Lincoln, NE).

2.3. Sequence analysis

BLAST searches (NCBI, Bethesda, MD) were conducted to determine gene identity using BLASTN and BLASTX with all parameters at the defaults. Matches were considered to be significant only when the probability (P) was less than 1×10^{-4}. ESTs with significant similarities to ribosomal protein genes in BLASTN searches were considered orthologues of ribosomal protein genes after manual examination of the sequence alignments for extended length of identity. For most ribosomal protein genes, multiple clones were encountered in the EST analysis. Sequences were first edited using the EDITSEQ program of the DNASTAR software package (DNAStar, Inc., WI). Vector sequences were manually removed. Overlapping clones of each ribosomal protein cDNA were aligned using the MEGALIGN program of the DNASTAR package. Gap penalty and gap length penalty was set at 10. Regions with discrepancies were double checked by two individuals to verify the correct sequences. Complete sequences of the cDNAs for each ribosomal protein were then obtained by sequencing one of the clones containing the complete cDNA insert. One reference ribosomal protein cDNA from a mammal (human, rat, or mouse) was used to guide the proper open reading frames. In the absence of a complete open reading frame, the partial cDNA sequences were translated in all three frames and compared to human or rat amino acid sequences in order to determine the proper open reading frame.

2.4. Phylogenetic analysis

Phylogenetic dendrograms were constructed using the MEGALIGN program of the DNASTAR software package. In this program, sequences were first edited using the EDITSEQ program and subsequently entered into the MEGALIGN program. Sequences were then aligned using the clustal method with gap penalty and gap length penalty at 10. Phylogenetic relations of the sequences were displayed as dendrograms. The MEGALIGN generated dendrograms were confirmed using PAUP using either neighborhood joining or UPGMA methods. Protein similarities were made using the MEGALIGN program, distance tables were generated, and similarity percentages were taken from the distance tables.

2.5. Gene expression studies

Gene expression studies were conducted by EST profiling. In the process of development for ESTs (Ju et al., 2000; Cao et al., 2001; Karsi et al., 2002a), a total of 5338 ESTs were sequenced including 1201 clones from the brain, 2228 clones from the head kidney, and 1909 clones from the skin. Frequency of each ribosomal protein was recorded. Expression data were calculated as encountered clones per one thousand sequenced clones. These data were

transferred into Microsoft Excel program for the production of graphic depiction.

3. Results and discussion

3.1. Sequences of 47 60S ribosomal protein cDNAs

The 60S ribosomes in eukaryotic organisms contain 47 proteins (Wool et al., 1996). Here we report the complete set of 47 60S ribosomal protein (RP) cDNAs from channel catfish, 43 of which included complete coding sequences. Based on identity comparisons with their orthologues, four RP cDNAs were partial. L7, L13a, and L23a were all missing just one amino acid after the initiation codon AUG, while L6 was missing a large portion of the 5′ sequences, as inferred from sequence comparisons with those of other organisms. The cDNA sequences have been deposited in the

Table 1
Structural characteristics of the 47 large ribosomal proteins from channel catfish (*Ictalurus punctatus*)

Genes	Accession numbers	cDNA (bp)	ORF (bp)	Observed 5′-UTR (bp)	3′-UTR (bp)	Termination	Poly(A)$^+$ signals	bp from Poly(A)$^+$ signal to tail
P0	AF401551	1081	954	73	54	UAA	AATAAA	13
P1	AF401552	493	342	122	29	UAA	AATAAA	12
P2	AF401553	442	348	47	47	UAA	AATAAA	17
L3	AF401554	1299	1209	5	85	UGA	AATAAA	12
L4	AF401555	1245	1158	57	30	UAG	AATAAA	17
L5a	AF401556	1015	891	69	55	UAG	AATAAA	15
L5b	AF401557	1038	900	61	77	UAA	AATAAA	19
L6	AF401558	>813	>784	N/D	29	UAA	AATAAA	20
L7	AF401559	>840	>790	N/D	50	UGA	AATAAA	12
L7a	AF401560	859	801	5	53	UAA	ATTAAA	15
L8	AF401561	821	774	16	31	UAA	AATAAA	12
L9	AF401562	738	579	6	153	UAG	CATAAA	13
L10	AF401563	738	648	9	81	UAA	AATAAA	19
L10a	AF401564	719	651	24	44	UAG	AATAAA	13
L11	AF401565	585	537	3	45	UAA	AATAAA	15
L12	AF401566	628	504	96	28	UAA	AATAAA	9
L13	AF401567	703	636	30	37	UAG	AATAAA	10
L13a	AF401568	>652	>607	N/D	40	UGA	AATAAA	14
L14	AF401569	>456	426	N/D	30	UAA	AATAAA	15
L15	AF401570	691	615	22	54	UAA	AATAAA	17
L17	AF401571	600	555	16	29	UAA	AATAAA	18
L18	AF401572	635	567	4	63	UAA	AATAAA	21
L18a	AF401573	612	531	21	60	UAA	AATAAA	16
L19	AF401574	640	591	16	39	UAA	AATAAA	12
L21	AF401575	549	483	35	31	UAA	AATAAA	13
L22	AF401576	440	387	11	42	UAA	AATAAA	16
L23	AF401577	>483	423	N/D	45	UGA	AATAAA	13
L23a	AF401578	>507	>463	N/D	44	UAA	AATAAA	12
L24	AF401579	526	474	9	43	UAA	AATAAA	14
L26	AF401580	488	438	17	33	UGA	AATAAA	8
L27	AF401581	468	441	19	38	UAG	AATAAA	14
L27a	AF401582	492	447	7	38	UAA	AATAAA	12
L28	AF401583	485	414	13	58	UAA	AATAAA	21
L29	AF401584	295	195	34	66	UAA	AATAAA	13
L30	AF401585	414	351	15	48	UAG	AATAAA	16
L31	AF401586	417	378	N/D	39	UAA	AATAAA	11
L32	AF401587	477	408	24	45	UAA	AATAAA	19
L34	AF401588	16	354	38	39	UAA	AATAAA	13
L35	AF401589	430	372	13	45	UAA	AATAAA	21
L35a	AF401590	460	333	60	67	UAA	AATAAA	16
L36	AF401591	379	318	22	39	UAA	AATAAA	17
L36a	AF401592	411	321	39	51	UAA	AATAAA	6
L37	AF401593	330	291	13	29	UAA	AATAAA	12
L37a	AF401594	400	279	16	105	UAA	AATAAA	17
L38	AF401595	268	213	9	46	UAG	AATAAA	14
L39	AF401596	334	156	25	153	UAA	AATAAA	13
L40	AF401597	466	387	19	60	UAA	AATAAA	12
L41	AF401598	349	78	23	248	UAA	AATAAA	17

Table 2
Evolutionary comparison of ribosomal proteins as determined from the cDNA sequences[a]

Ribosomal protein genes	Number of amino acids in catfish	Number of amino acids in rat	Similarity to that ofhuman (%)	Similarity to that of rat (%)	Similarity to that of chicken (%)	Similarity to that of *Xenopus laevis* or as noted (%)	Similarity to that of zebrafish or as noted (%)
P0	317	316	81.1	82.0	80.7	82.9 *R. sylvatica*	82.7
P1	113	114	77.2	76.3	79.8	71.9 *R. esculenta*	75.0 *G. mirabilis*
P2	115	115	76.5	76.5	NA	NA	83.6 *P. olivaceus*
L3	402	403	90.3	90.6	NA	NA	NA
L4	385	421	77.5	77.5	NA	77.7	90.0 *O. mykiss*
L5a	296	296	86.5	86.2	86.9	86.8	NA
L5b	299	296	84.5	83.5	84.2	85.1	NA
L6	>260	297	75.1	75.9	NA	NA	69.0 *G. mirabilis*
L7	>261	260	79.4	75.4	NA	77.7 *R. sylvatica*	NA
L7a	266	265	89.1	89.1	89.5	NA	93.2 *T. rubripes*
L8	257	257	94.6	94.6	NA	93.8	NA
L9	192	192	89.1	89.1	NA	NA	NA
L10	215	214	91.6	91.1	90.5	NA	NA
L10a	216	217	91.7	91.7	NA	NA	NA
L11	178	178	93.8	92.7	NA	NA	91.5 *O. latipes*
L12	167	165	87.9	87.3	NA	NA	NA
L13	211	211	84.8	85.3	86.3	NA	88.6
L13a	>201	203	85.6	84.6	NA	NA	89.9 *S. trutta*
L14	141	214	74.6	74.6	80.9	NA	84.7 *T. rubripes*
L15	204	203	95.6	95.6	95.2	NA	NA
L17	184	184	95.1	95.7	NA	NA	95.7 *P. olivaceus*
L18	188	187	84.6	85.6	NA	84.6	90.4 *O. niloticus*
L18a	176	176	87.5	87.5	NA	NA	92.6 *S. salar*
L19	196	196	93.9	93.9	NA	NA	NA
L21	160	159	83.1	81.9	NA	NA	NA
L22	128	128	94.5	93.8	94.5	89.8	84.8 *G. morhua*
L23	140	140	100	100	NA	NA	98.6 *G. mirabilis*
L23a	>153	156	90.3	90.3	NA	NA	99.0 *O. mykiss*
L24	157	157	88.5	88.5	NA	NA	82.7 *G. mirabilis*
L26	145	145	95.9	95.2	98.4	NA	NA
L27	136	135	93.4	93.4	93.4	NA	79.4 *G. mirabilis*
L27a	148	147	83.8	85.1	NA	85.1	88.9 *O. mykiss*
L28	137	136	76.6	75.9	NA	75.8	NA
L29	64	155	84.6	84.6	NA	NA	NA
L30	116	114	93.0	93.0	93.0	NA	NA
L31	125	124	92.8	92.8	NA	NA	91.9 *P. olivaceus*
L32	135	135	93.3	93.3	NA	NA	NA
L34	117	117	94.0	93.2	NA	NA	NA
L35	123	122	87.8	88.6	87.0	NA	NA
L35a	110	109	89.1	90	NA	87.3	NA
L36	105	104	91.4	90.5	90.5	NA	NA
L36a	106	105	98.1	97.2	NA	NA	NA
L37	96	96	88.7	88.7	NA	NA	NA
L37a	92	91	97.8	97.8	96.7	NA	NA
L38	70	69	100	100	NA	NA	NA
L39	51	50	98.0	98.0	96.0	NA	NA
L40	128	128	97.7	97.7	96.1	NA	98.4 *O. mykiss*
L41	25	25	92.3	92.3	NA	NA	NA

[a] Noted species are as follows: fish-*Gadus morhua* (Atlantic cod), *Gillichthys mirabilis* (long-jawed mudsucker), *Oncorhyncus mykiss* (rainbow trout), *Oreochromis niloticus* (Nile tilapia), *Oryzias latipes* (Japanese medaka), *Paralichthys olivaceus* (bastard halibut), *Salmo salar* (Atlantic salmon), *Salmo trutta* (brown trout), and *Takifugu rubripes*; Amphibians-*Rana sylvatica* (wood frog), *Rana esculenta* (edible frog).

GenBank with consecutive accession numbers from AF401551 to AF401598 (Table 1).

During the process of EST analysis of 5338 clones (Ju et al., 2000; Cao et al., 2001; Karsi et al., 2002a), 708 clones (13.3%) were identified as representatives of large ribosomal proteins. Sequence analysis indicated that these clones represented all 47 unique ribosomal proteins of channel catfish. In addition, two types of cDNAs were found to be related to L5 (see below). The nomenclature of ribosomal proteins in different organisms is quite confusing because of many synonyms for the same gene in different organisms (Wool et al., 1996). This was particularly true for L10 (also known as ribosomal protein QM, laminin receptor, human Wilm's tumor-related protein, chicken Jun-binding protein), L13 (also known as activin B), and L13a (also known as transplantation antigen in *Salmo trutta*, and 23 Kda highly basic protein in human). To avoid further confusion, we have followed the nomenclature of the rat (Wool et al., 1996). The genes were annotated using identity comparisons with the rat ribosomal protein gene sequences.

Channel catfish ribosomal protein mRNAs were comparatively short, as is the case with the rat (Wool et al., 1996). Individual lengths of cDNAs, their coding regions, the observed 5′-UTR, 3′-UTRs, and poly(A)$^+$ signals are shown in Table 1. All of the 43 ribosomal protein mRNAs containing complete open reading frames initiated at the first AUG. This is similar to the situation in rat where all but L5 initiate at the first AUG. UAA was the most frequently used termination codon in the catfish ribosomal proteins (Table 1), different from the most frequently used termination codon in vertebrate proteins in general, UGA (Cavener and Ray, 1991). The 3′-UTRs were highly AT-rich. Of the 47 mRNAs, 45 had the typical AAUAAA polyadenylation signals (Proudfoot, 1991), one (L7a) had AUUAAA, and one (L9) had CAUAAA. The poly(A)$^+$ tract started 6–21 nucleotides from the poly(A)$^+$ signal hexamer (Table 1). We must point out that the actual lengths of 5′-UTRs were not determined. The lengths of the 5′-UTRs listed in Table 1 were observed in the cDNA clones and may be useful for future studies, but longer 5′-UTR sequences are possible.

In this research, we generated the complete set of 47 cDNA sequences for ribosomal protein genes in channel catfish, 43 of which had complete coding regions. These sequences filled a large gap that previously prohibited phylogenetic analysis using ribosomal proteins in fishes. Along with our recent analysis of 32 40S ribosomal proteins in channel catfish (Karsi et al., 2002b), this set of ribosomal proteins represents one of the most complete lists of ribosomal proteins in animals. These sequence data should be of significant phylogenetic value because of the ubiquitous nature of ribosomes. In addition, further studies in genome mapping of these ribosomal protein genes in fish will allow comparative analysis of genome evolution. Ribosomal protein genes should serve as one of the

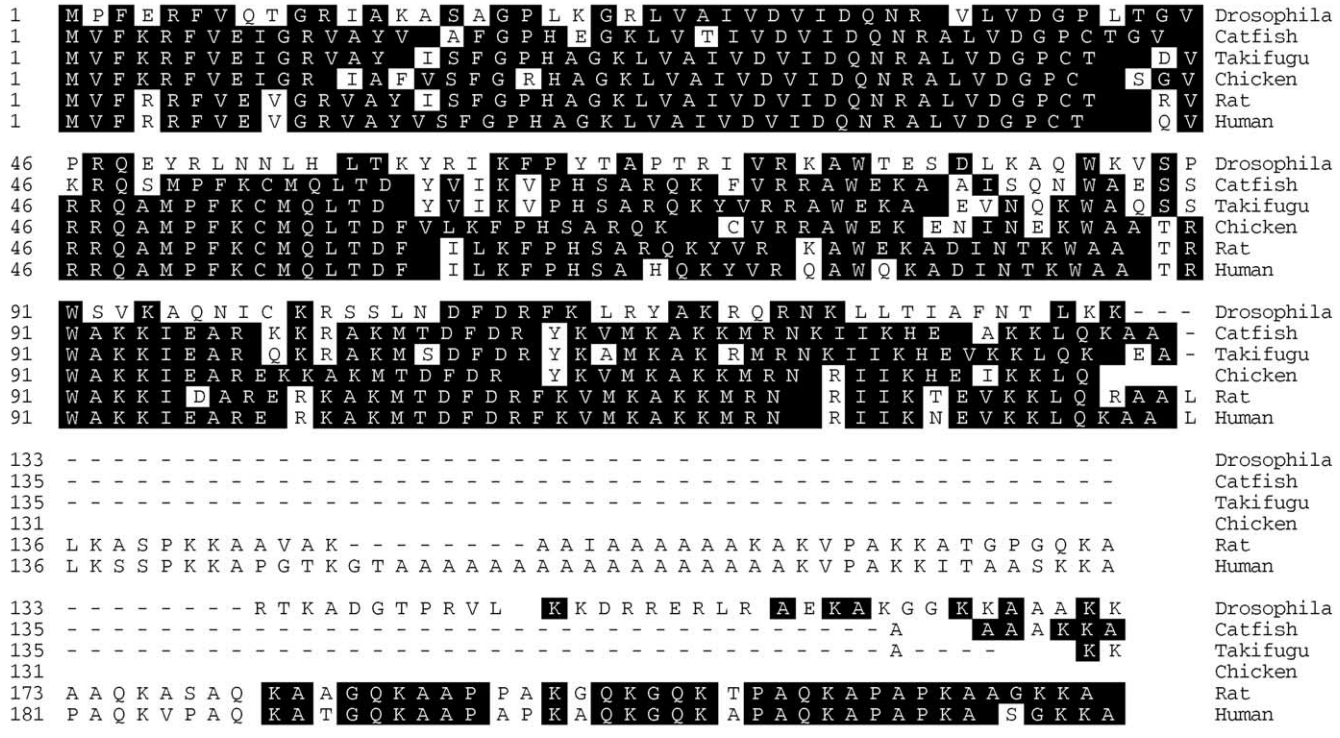

Fig. 1. Alignment of deduced amino acid sequences of L14. With exception of the channel catfish sequences, all other sequences used in the alignment were from the GenBank with accession numbers of: *Drosophila melanogaster* (CAA71121); *Takifugu rubripes* (AJ310911); chicken (AB046393); rat (NM022949); and human (XP044190). Note the 3′ deletion in all species except for rat and human.

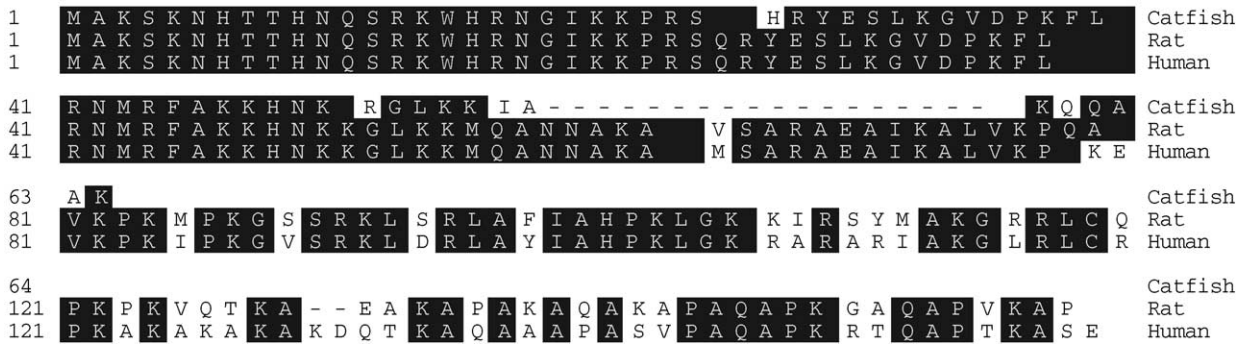

Fig. 2. Alignment of deduced amino acid sequences of the channel catfish ribosomal protein L29 with those of rat (NP058846) and human (BC008926). Note that the channel catfish L29 is only 65 amino acids in length while the rat and human L29 are 155 amino acids in length because of the deletion at the 3′ end.

strongest anchors for syntenic grouping in comparative genomics.

3.2. Evolutionary comparison of the ribosomal proteins

Most channel catfish ribosomal proteins were highly similar to their mammalian counterparts (Table 2). The numbers of amino acids and the amino acid sequences were highly conserved throughout evolution. Of the 43 ribosomal proteins with complete ORF, 18 had ORFs with identical numbers of amino acids as compared with the rat; 17 had one extra amino acid; three had one fewer amino acid; three had two to three extra amino acids (Table 2). Only three ribosomal proteins, L4, L14, and L29, had significantly different numbers of amino acids (Table 2). In all three cases, channel catfish had many fewer amino acids than their mammalian counterparts.

L14 of channel catfish had 141 amino acids, 73 amino acids shorter than the rat L14. Alignment of deduced amino acid sequences of L14 indicated that only the mammalian L14 were much longer. L14 of chicken, *Drosophila*, *Takifugu*, and channel catfish were all shorter than the rat L14 (Fig. 1). The difference was mostly caused by deletions near the 3′ end in the lower vertebrates. In addition, it appeared that expansion of the alanine tract in the rat and human, particularly in human, accounted for some of the differences.

Similarly, L29 of channel catfish was only 64 amino acids long as compared to 155 amino acids in the rat. Sequence alignment indicated that the 3′ end deletion was the cause for the difference (Fig. 2), just as in L14. In the case of L4, the channel catfish L4 mRNA encoded 388 amino acids, 33 amino acids shorter than the rat L4. Once again, deletions in the 3′ end were the major cause of the difference (Fig. 3). In all three cases of L4, L14, and L29, highly conserved amino termini coupled with deletions at the 3′ ends suggested that the functional portion of these proteins is not located in the carboxyl regions.

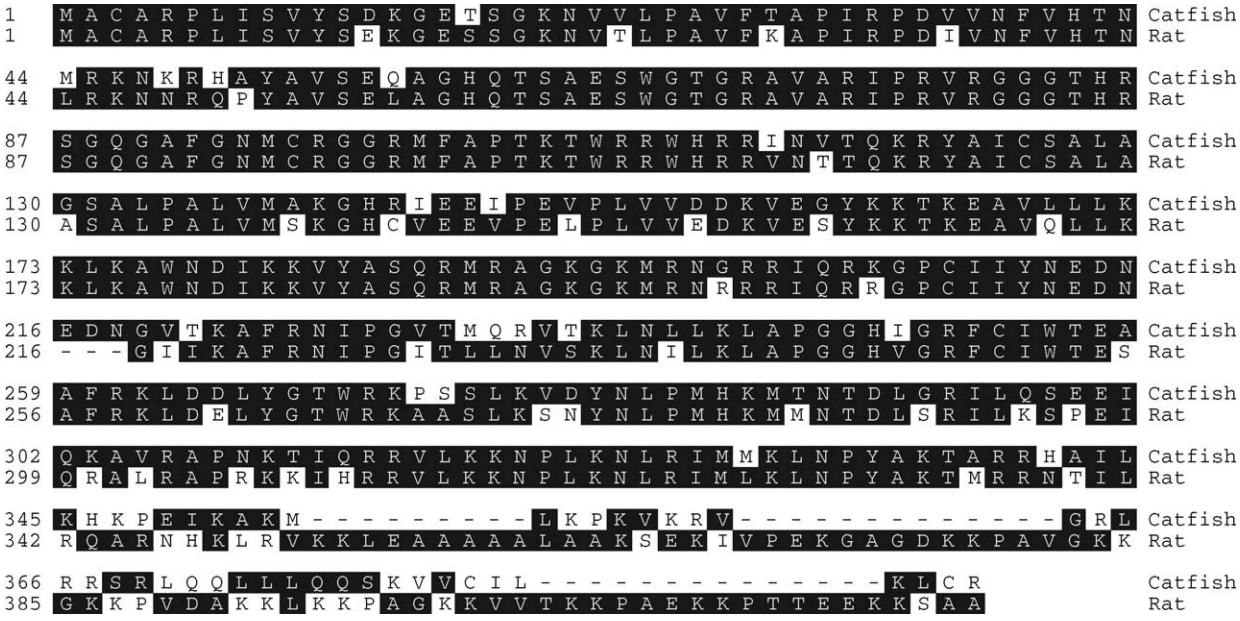

Fig. 3. Alignment of deduced amino acid sequences of the channel catfish ribosomal protein L4 with those of rat (X82180). The deleted portions of the channel catfish L4 are indicated by dashed lines.

Fig. 4. Comparison of the channel catfish ribosomal protein L5a and L5b cDNA sequences. They share only a 71.1% similarity, but encode for highly similar proteins.

The 47 channel catfish 60S ribosomal proteins exhibited high levels of similarity in amino acid sequences as compared to those of human or rat. The level of conservation varied from 74.6 to 100%. Two channel catfish ribosomal proteins, L23 and L38, were identical to the rat and human counterparts. Of the 47 RPs, 25 showed over 90% similarities to those from the rat; 16 showed 80–90% similarity; while seven were more divergent with 70–80% similarity (Table 2). Clearly, not all ribosomal proteins were similarly conserved throughout evolution, suggesting that both structural and functional constraints vary with individual RPs.

3.3. Channel catfish has two L5-related mRNAs

All but one of the channel catfish ribosomal proteins appeared to have only one type of mRNA. As shown in Table 1, large numbers of clones were analyzed for the sequences of the 47 ribosomal protein cDNAs in channel catfish. In all cases except L5, only one type of mRNA was observed. However, two L5-related mRNAs were found in channel catfish. For the convenience of presentation, we named them L5a and L5b. L5a encoded 296 amino acids, the same as L5 in the rat, whereas L5b encoded 298 amino acids. They differed significantly in the middle and down-

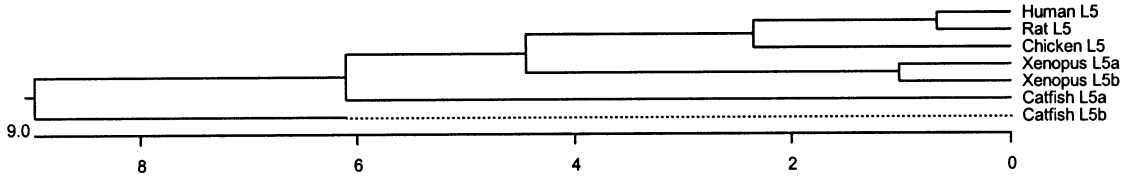

Fig. 5. Phylogenetic relationship of the selected L5 sequences. The dendrograms were generated by the MEGALIGN program of the DNASTAR package using the clustal method. With the balanced display, MEGALIGN averages the distances between ancestors in the tree. The dotted lines indicate a negative branch length introduced by averaging the tree.

stream sequences (Fig. 4). They shared only 71.1% identity at the nucleotide level. Sequence analysis indicated that L5a is more similar to other vertebrate L5 proteins than L5b. One would expect that the L5 proteins from channel catfish should be more related, but in fact L5b is the most distantly related protein in the group, suggesting that L5a and L5b could have been independently evolved (Fig. 5).

While two genes were found to encode RP L5 in *Xenopus* and catfish and to encode S26 and S27 in catfish (Karsi et al., 2002b), in no instance (with the exception of the separate alleles encoding human S4, Fisher et al., 1990) has it been shown that more than one of the genes is functional (Wool et al., 1996) in mammalian species. Ribosomal protein L5a and L5b appeared to be expressed in approximately equal ratios in catfish. In the EST analysis of 5338 clones in the three tissues, nine clones each of L5a and L5b were sequenced. This expression pattern was different from that of the human S4X and S4Y, the only ribosomal protein encoded by two functional genes in human. In male human ribosomes, S4X accounts for 90% and S4Y for 10% of S4 ribosomal protein (Zinn et al., 1994).

3.4. Ribosomal protein L31 has three alternative polyadenylated forms of mRNA

Three types of L31 cDNAs were found with the same upstream and coding sequences (except at one single nucleotide polymorphism site) but different downstream ends. Alignments of cDNA sequences revealed that the different downstream sequences were caused by differential polyadenylation (Fig. 6). The first type of mRNA had a poly(A)$^+$ tail 11 bases downstream from the first AAUAAA signal. The second type of L31 mRNA extended its 3′-UTR for 72 bases; and the third extended its 3′-UTR for an additional 350 bases. There were no typical poly(A)$^+$ signals for the second type of mRNA, however, a long stretch of AT sequences existed 31 bases upstream of its poly(A)$^+$ tail. A typical AAUAAA poly(A)$^+$ signal existed 13 bases upstream of the poly(A)$^+$ tail of the third mRNA. The functional importance of the alternative polyadenylated mRNAs is not known at present, but they could be involved in post-transcriptional regulation of L31 expression.

3.5. Gene expression of the ribosomal proteins in the brain, head kidney, and skin

The gene expression data were obtained from expression profiles in the EST analysis. Frequencies of ribosomal protein representation in the cDNA libraries of head kidney, brain, and skin are shown in Fig. 7. Overall, ribosomal protein genes were expressed at high levels, accounting for 12.3–14.9% of all ESTs analyzed, depending on the libraries. The largest percentage of ribosomal protein representation was found in the skin library, which accounted for 14.9% of all 1909 clones sequenced, while ribosomal proteins in the brain and head kidney libraries constituted about 12.3% of the 1201 and 2228 clones sequenced, respectively. This overall greater representation of ribosomal protein cDNAs in the skin could be due to higher translational activities. One would expect that tissues with active translation would demand more ribosomal proteins for their ribosomes. A survey of the sequence context surrounding the translation start site indicated that most translational initiation codons occur in the context of (A/G)(A/C)CAUGG, a close approximation to the consensus for all vertebrates (Kozak, 1991).

Fig. 6. Alignment of the 3′ end sequences in ribosomal protein L31, beginning at nucleotide 393, showing the three alternative polyadenylation sites. The potential polyadenylation signals were boxed.

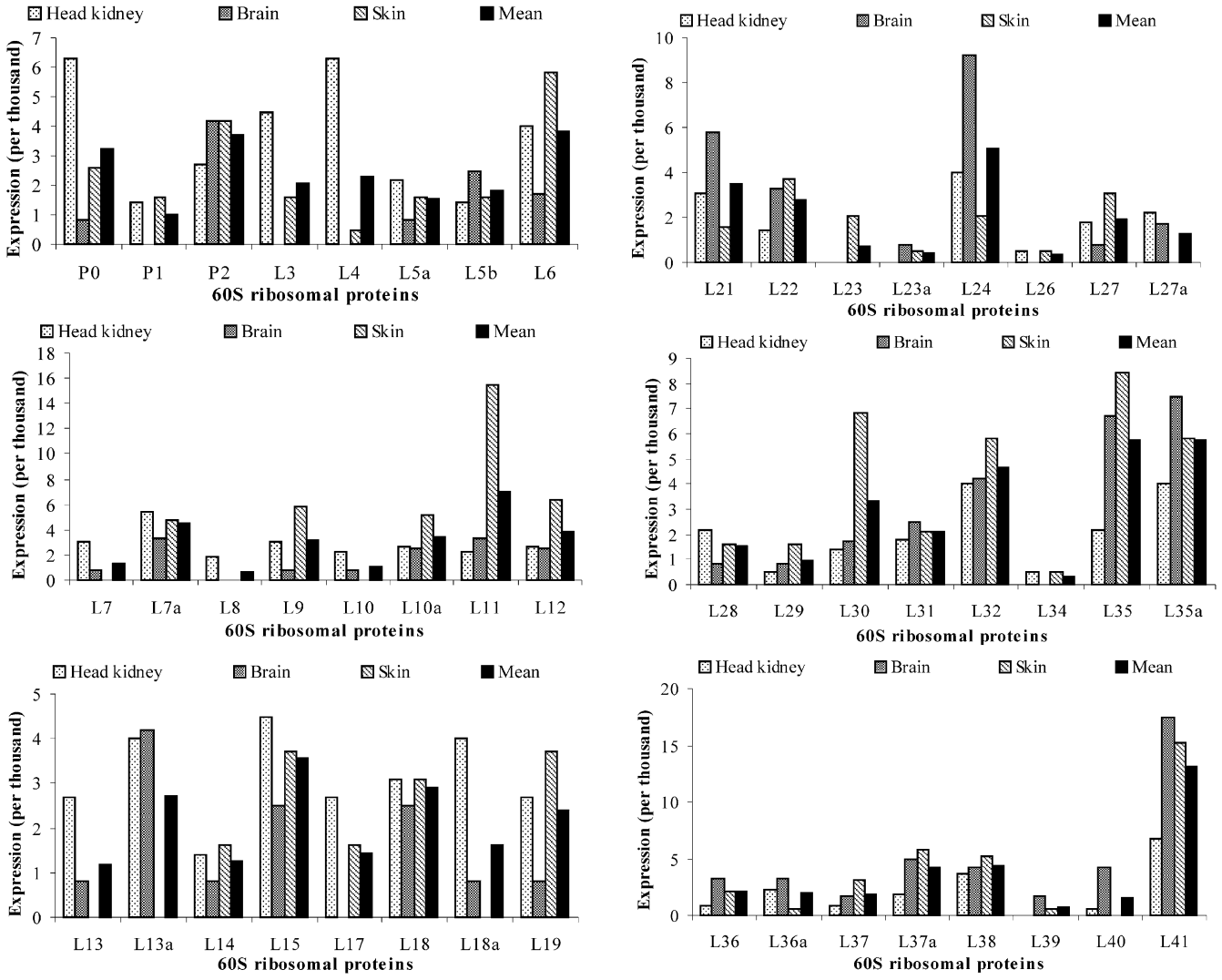

Fig. 7. Expression of the 47 large ribosomal proteins in three tissues from channel catfish (*I. punctatus*) as determined by analysis of expressed sequenced tags, expressed as clones encountered per one thousand sequenced clones. Mean expression was derived from averaging expressions in the three tissues.

In spite of the specific stoichiometric ratio of ribosomal proteins as required by the ribosome structure, drastically different levels of gene expression were observed within a single tissue/organ or among the tissues. In the head kidney, the most abundant ribosomal protein cDNAs were L41, L4, P0, and L7a, all of which were encountered over 0.5%. The most abundant cDNAs in the brain library were L41, L24, L35a, L35, L21, and L37a. In the skin, the most abundant cDNAs were L11, L41, L35, L30, L12, L35a, L32, L6, L9, L38, and L10a. Several of these highly expressed ribosomal proteins were expressed at high levels in all three tissues such as L41, L35a, L32, L7a, L15, and L24. However, strong tissue preference was observed for some of the highly expressed ribosomal proteins. For instance, L3, L4, and L18a were highly expressed in the head kidney, but their expression in the brain and skin was much lower (Fig. 7). Similarly, L11 was the most highly expressed in the skin, but expression in the head kidney and brain was moderate.

Some ribosomal protein cDNAs were rarely represented in the analysis of 5338 clones of ESTs from the three libraries including L23, L23a, L26, L34, and L39, all of which were present at a rate of 0.05% or less. Ribosomal biogenesis in mammalian cells requires the equimolar accumulation of four RNA species and 79 different ribosomal proteins (with some minor exceptions, Wool, 1979). The highly variable expression levels of various ribosomal proteins within a tissue as well as the wide variations among tissues strongly suggest significant regulation at the post-transcriptional levels. It is well known that gene regulation of ribosomal protein genes occurs at the level of translation (Perry and Meyuhas, 1990; van der Velden and Thomas, 1999; Meyuhas, 2000). Sequence analysis of the channel catfish large ribosomal protein cDNAs revealed the presence of the typical polypyrimidine stretch motifs in the majority of the ribosomal proteins (Meyuhas, 2000). This motif was suggested to be a region that binds a trans-acting factor for coordinated translational control (Perry and Meyuhas,

1990; Wool et al., 1990; Levy et al., 1991; Meyuhas et al., 1996). Although this hypothesis was not tested in this study, highly variable levels of different ribosomal protein mRNAs in the three tissues analyzed strongly suggest that protein levels and mRNA levels were not linear.

Acknowledgements

This project was supported by a grant to Z.L. from USDA NRICGP Animal Genome Basic Reagents and Tools program (USDA/NRICGP 2003-35205-12827). 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).

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.

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 of 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.

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.

Ju, Z., Karsi, A., Kocabas, A., Patterson, A., Li, P., Cao, D., Dunham, R., Liu, Z.L., 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., Patterson, A., Kocabas, A., Feng, J., Ju, Z., Mickett, K., Liu, Z.J., 2002a. Transcriptome analysis of channel catfish (*Ictalurus punctatus*): initial analysis of gene expression and microsatellite-containing cDNAs in the skin. Gene 285, 157–168.

Karsi, A., Patterson, A., Feng, J., Liu, Z.J., 2002b. Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression. Gene 291, 177–186.

Kozak, M., 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. 115, 887–903.

Levy, S., Avni, D., Hariharan, N., Perry, R.P., Meyuhas, O., 1991. Oligopyrimidine tract at the 5′ end of mammalian ribosomal protein mRNAs is required for their translational control. Proc. Natl. Acad. Sci. USA 88, 3319–3323.

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.

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.

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.

Regier, J.C., Fang, Q.Q., Mitter, C., Peigler, R.S., Friedlander, T.P., Solis, M.A., 1998. Evolution and phylogenetic utility of the period gene in *Lepidoptera*. Mol. Biol. Evol. 15, 1172–1182.

Sambrook, J., Frisch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd Edition, 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.

Simon, C., Grati, R., Beckenbach, A., Crespit, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

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.

van der Velden, A.W., Thomas, A.A., 1999. The role of the 5′ untranslated region of an mRNA in translation regulation during development. Int. J. Biochem. Cell. Biol. 31, 87–106.

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.

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

Wool, I.G., Endo, Y., Chan, Y.L., Gluck, A., 1990. Structure, function, and evolution of mammalian ribosomes. In: Hill, W.E., (Ed.), The Ribosome: Structure, Function, and Evolution, American Society for Microbiology, Washington, DC, pp. 203–214.

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.