African Journal of Biotechnology Vol. 11(21), pp. 4752-4758, 13 March, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.3538
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

cDNA, genomic cloning and sequence analysis of ribosomal protein *S4X* gene (*RPS4X*) from *Ailuropoda melanoleuca*

Fu-yan Luo[1, 2], Wan-ru Hou[2]* and Wei-ying Chen[2]

[1]Department of Environmental Science, Sichuan University, No. 24 South Section 1, the First Ring Road, Chengdu, Sichuan 610065, China.
[2]College of Life Science, China West Normal University, No. 1 Shida Road, Nanchong, Sichuan 637009, China.

Accepted 27 February, 2012

Ribosomal protein S4X (RPS4X) is one of the 40S ribosomal proteins encoded by the *RPS4X* gene. The cDNA and the genomic sequence of *RPS4X* were cloned successfully from giant panda (*Ailuropoda melanoleuca*) using reverse transcriptase-polymerase chain reaction (RT-PCR) and touchdown-PCR technology respectively, followed by sequencing and analyzing. The cDNA of the *RPS4X* cloned is 812 bp in size, containing an open reading frame (ORF) of 792 bp encoding 263 amino acids with deduced molecular weight of 29.6 kD and isoelectric point (PI) of 10.16. The length of the genomic sequence is 5074 bp, which was found to possess 7 exons and 6 introns. There are 1 N-glycosylation site, 3 protein kinase C phosphorylation sites, 4 casein kinase II phosphorylation sites, 2 tyrosine kinase phosphorylation sites and 4 N-myristoylation sites in the RPS4X of giant panda by topology prediction. Alignment analysis indicated that the coding sequence and the deduced amino acid sequence of *RPS4X* gene shows very high homology with other reported mammalian species. The information on cDNA and the genomic sequence from giant panda *RPS4X* will provide basic information for further research of mammals *RPS4X* gene, which would be beneficial in the study of ribosomal biogenesis and would allow the elucidation of structure, organization and regulation of genes encoding ribosomal proteins in eukaryote.

Key words: *RPS4X* gene, giant panda, genomic cloning, sequence analysis.

INTRODUCTION

The ribosome, an organelle that catalyzes protein synthesis, consists of a small 40S subunit and a large 60S subunit. Together, these subunits are composed of 4 RNA species and approximately 80 structurally distinct proteins. RPS4X is one member of 40S ribosomal proteins belonging to the S4E family of ribosomal proteins that is encoded by the *RPS4X* gene which is not subject to X-inactivation (Wiles et al., 1988; Fisher et al., 1990; Watanabe et al., 1991; Zinn et al., 1994; Wool et al., 1995; Alberts et al., 2002). And ribosomal protein S4 is the only ribosomal protein known to be encoded by more than one gene, and ribosomal protein S4, Y-linked (RPS4Y). The 2 isoforms encoded by these genes are not identical, but are functionally equivalent. It has been suggested that haploinsufficiency of the ribosomal protein S4 genes plays a role in Turner syndrome; however, this hypothesis is controversial. As is typical for genes encoding ribosomal proteins, there are multiple processed pseudogenes of this gene dispersed through the genome (Wiles et al., 1988). With the continuous advancement of science and technology, the ribosomal proteins were found to play an important role in human disease (Yang and Liu, 2005). The giant panda (*Ailuropoda melanoleuca*) is known as “National Treasure of China”, belonging to national level of endangered animal. Moreover, giant panda is renowned as living fossil, which has important scientific value in exploring environment change, and research and protection of

*Corresponding author. E-mail: hwr168@yahoo.com.cn.

biodiversity (Su et al., 2010). Because of its important academic value, its survival and protection draws attention of the whole world. For years, the studies about giant panda mainly focused on macrograph such as propagation, ecology, genetic diversity, parentage, phylogenesis, etc. Even so, to study giant panda in molecular level gradually become the focus of research recently (Montali, 1990; Mather et al., 1997; Liao et al., 2003; Hou et al., 2007, 2008, 2009; Li et al., 2010; Song et al., 2010). In recent years, we have found many important functions of *RPS4X g*ene in human disease study. But there are still few reports on *RPS4X g*ene of giant panda in the literature.

In this paper, the primers were designed according to the related information to clone the cDNA and genomic sequence of *RPS4X* gene, and the sequence characteristics of these sequences, along with the deduced protein were analyzed, which is of significance to provide correct and scientific data for studying giant panda at molecular level.

MATERIALS AND METHODS

Materials used for DNA and RNA extraction were skeletal muscle tissue collected from a dead giant panda at the Wolong Conservation Center of the Giant Panda, Sichuan, China. And it was frozen in liquid nitrogen.

DNA and RNA isolation

The genomic DNA of giant panda was isolated from skeletal muscle tissue using the CTAB DNA isolation method and was dissolved in TE buffer at -20°C. Total RNAs were isolated from about 500 mg of muscle tissue using the Total Tissue/Cell RNA Extraction Kits (Waton Inc., Shanghai, China) according to the manufacturer's instructions, then dissolved in DEPC-treated water and stored at -70°C.

Primers design, RT-PCR and cloning of cDNA

The PCR primers were designed by Primer Premier 5.0, according to the mRNA sequences of *RPS4X* genes from *Homo sapiens* (NC_000023), *Bos taurus* (NC_007331), *Mus musculus* (NC_000086) and *Monodelphis domestica* (NC_008809). The specific primers of cDNA sequence are as follows:
Pd- *RPS4X* -F: 5′--ACGAGAAAGGCACGGATCGCGTCGG --3′; Pd-*RPS4X* -R: 5′--TCCAATCATGTCTCCTAGAGACCCATT --3′.

Total RNAs were synthesized into the first-stranded cDNAs using a reverse transcription kit with Oligo dT as the primers according to the manufacturer's instructions (Promega). The total reaction volume was 20 µl for first-strand cDNA synthesis reaction system. Then, the first-strand cDNA was used as a template to synthesis the second strand. The total reaction volume for the double-stranded cDNA synthesis was 25 µl. The PCR products were analyzed by electrophoresis in 1.5% agarose gel and purified from the gel using a DNA harvesting kit (Omega, China), and the cDNA were ligated into plasmid pUC18 at 22°C for 12 h. Then, the recombinant molecules were transformed into *E. coli* competence cells (JM109) overnight, and spread on the LB-late containing 50 µg/ml ampicillin, 200 mg/ml IPTG and 20 mg/ml X-gal. Plasmid DNA was isolated and digested by *Pst*I and *Sac*II to verify the insert size. Plasmid DNA was sequenced by Huada Zhongsheng Scientific Corporation (Beijing, China).

Cloning of the genomic sequence of *RPS4X*

The PCR primers were the same as the *RPS4X* -F and *RPS4X* -R presented above. The genomic sequence of the *RPS4X* gene was amplified using Touchdown-PCR with the following conditions: 94°C for 30 s, 62°C for 45 s, 72°C for 4 min in the first cycle and the anneal temperature decreased to 1°C per cycle; after 15 cycles, conditions changed to 94°C for 30 s, 52°C for 45 s, 72°C for 1 min for another 20 cycles. The fragment amplified was also purified, ligated into the clone vector and transformed into the *Escherichia coli* competence cells. Finally, the recombinant fragments were sequenced by Sangon (Shanghai, China).

Data analysis

The sequence data were analyzed by GenScan software. Homology research of giant panda *RPS4X* gene when compared with the gene sequences of other species were performed using Blast 2.1 (Li et al., 2010; Su et al., 2010). ORF of the DNA sequence was searched using ORF Finder software (Hou et al., 2007). Multiple sequence alignment was performed by DNAMAN 6.0. Protein structure and feature of the sequence cloned were deduced using Predict Protein software. The prediction of RPS4X functional sites and biochemical characteristics was depended on the software ExPASy Proteomics Server. Secondary structure prediction of the RPS4X sequence was based on the software (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html).

Tertiary structure prediction of the RPS4X sequence was dependent on ExPASy Proteomics Server.

RESULTS

Analysis of the cDNA of *RPS4X* from giant panda

The amplified cDNA fragment from giant panda was 812 bp in length. Blast research showed that the cDNA sequence cloned shares a high homology with the *RPS4X* gene from other mammals reported, including *H. sapiens, B. taurus, M. musculus* and *M. domestica* (Table 2). On the basis of the high identity, we concluded that the cDNA isolated is the cDNA encoding giant panda RPS4X protein. The *RPS4X* gene sequence has been submitted to Genbank (accession number: HQ318009). An ORF of 792 bp encoding 263 amino acids was found in the cDNA sequence (Figure 1). The initiation codon of *RPS4X* gene is ATG, and terminator codon is TGA. The average levels of the bases sequence are: 27.0% A, 22.7% C, 25.4% G and 24.9% T.

Analysis of the genomic sequence of *RPS4X* from giant panda

The cloned DNA fragment with primers *RPS4X* -F and *RPS4X*-R is 5074 bp in length from the giant panda. The genomic sequence of the *RPS4X* has been submitted to Genbank (accession number: HQ318008). The nucleotide

Table 1. Comparison of *RPS4X* genomic sequence among 5 mammalian species.

Item	Length (bp)	Number of exons	Join sites in the CDS	Accession Numbers
Pa	5074	7	14 to 16, 921 to 998, 1525 to 1705, 2102 to 2199, 3504 to 3675, 4251 to 4408, 4973 to 5074	HQ318008
Ho	4689	7	97 to 99, 1058 to 1135, 1568 to 1748, 2142 to 2239, 3320 to 3491, 3903 to 4060, 4520 to 4621	NC_000023
Bo	4424	7	24 to 26, 937 to 1014, 1215 to 1395, 1618 to 1715, 2847 to 3018, 3597 to 3754, 4264 to 4365	NC_007331
Mu	4364	7	7 to 9, 936 to 1013, 1305 to 1485, 1855 to 1952, 2740 to 2911, 3647 to 3804, 4176 to 4277	NC_000086
Mo	2203	6	1 to >79, 596 to 776, 897 to 994, 1146 to 1317, 1608 to 1765, 2102 to 2203	NC_008809

Ho: *H. sapiens*; Bo: *B. Taurus*; Mu: *M. musculus*; Mo: *M. domestica*.

sequences comparison of the genomic and cDNA sequences indicated that the cloned gene contains seven exons and six introns with Genscan software. The distribution sites of the seven exons are: 14 to 16, 921 to 998, 1525 to 1705, 2102 to 2199, 3504 to 3675, 4251 to 4408, 4973 to 5074 (Table 1).

Prediction of protein functional sites in RPS4X protein of giant panda

The molecular weight of the putative RPS4X protein of giant panda is about 29.6 kD with a theoretical pI of 10.16 by primary structure analysis, and it contains 49 positively charged amino acid residues (Arg and Lys), 25 negatively charged amino acid residues (Asp and Glu) and 189 uncharged residues. The Ile (I) possesses the optimal content in this protein, and Gln (Q) presents the least. Topology prediction showed that there are one N-glycosylation site, three protein kinase C phosphorylation sites, four casein kinase II phosphorylation sites, two tyrosine kinase phosphorylation sites and four N-myristoylation sites in the RPS4X protein of giant panda (Figure 2).

Secondary structure analysis of RPS4X from giant panda

The secondary structure prediction of the RPS4X protein sequence indicated that 27.00% of the protein sequence is extended strand, 23.57% is alpha helix, and 49.43% is random coil for giant panda.

Tertiary structure analysis of RPS4X from giant panda

The protein tertiary structure of RPS4X protein from giant panda demonstrated that the conserved sequence is located between 43rd and 242nd amino acid residues. Further analysis revealed that the tertiary structure of RPS4X protein from giant panda was entirely identical when compared with other four animals, which suggested that the subtle changes of the primary structure of RPS4X protein from *H. sapiens, B. taurus, M. musculus* and *M. domestica* have no effect on their tertiary structure (Figure 3).

DISCUSSION

Cisplatin is one of the most important chemotherapeutic drugs in clinical practice. At present, it is widely used in the treatment of many tumors, particularly in ovarian cancer (Jordan and Carmo-Fonseca, 2000). But chemotherapy is curative in only a fraction of these patients (Parker et al., 1996). Resistance to available chemotherapeutic drugs is the main obstacle to effective potentially curative chemotherapy in this cancer (Yang et al., 2002). But some research suggests that the *RPS4X*/YB-1 complex is a significant potential target to counteract cisplatin resistance in breast cancer (Garand et al., 2011). Thus, *RPS4X*/YB-1 complex can be used in all cancer therapy. Some researches have found that haploinsufficiency of *RPS4X* is not the cause of Ullrich-Turner syndrome (UTS) (Geerkens et al., 1996), and *RPS4X* is not the most prominent candidate gene for Turner syndrome (Just et al., 1992). And there are more researches on *RPS4X* gene in recent years (Jung et al., 2011). So, ribosomal protein S4X plays an important role in human disease research, and not only involved in several steps in protein synthesis.

In this research, cDNA and genomic sequence were cloned from giant panda. The former size is 812 bp, containing an ORF of 792 bp. Deduced protein contains 263 amino acids with molecular weight of about 29.6 kD and PI of 10.16. Alignment analysis indicates that the nucleotide sequence of the coding sequence shows a high homology to those of *H. sapiens, B. taurus, M.*

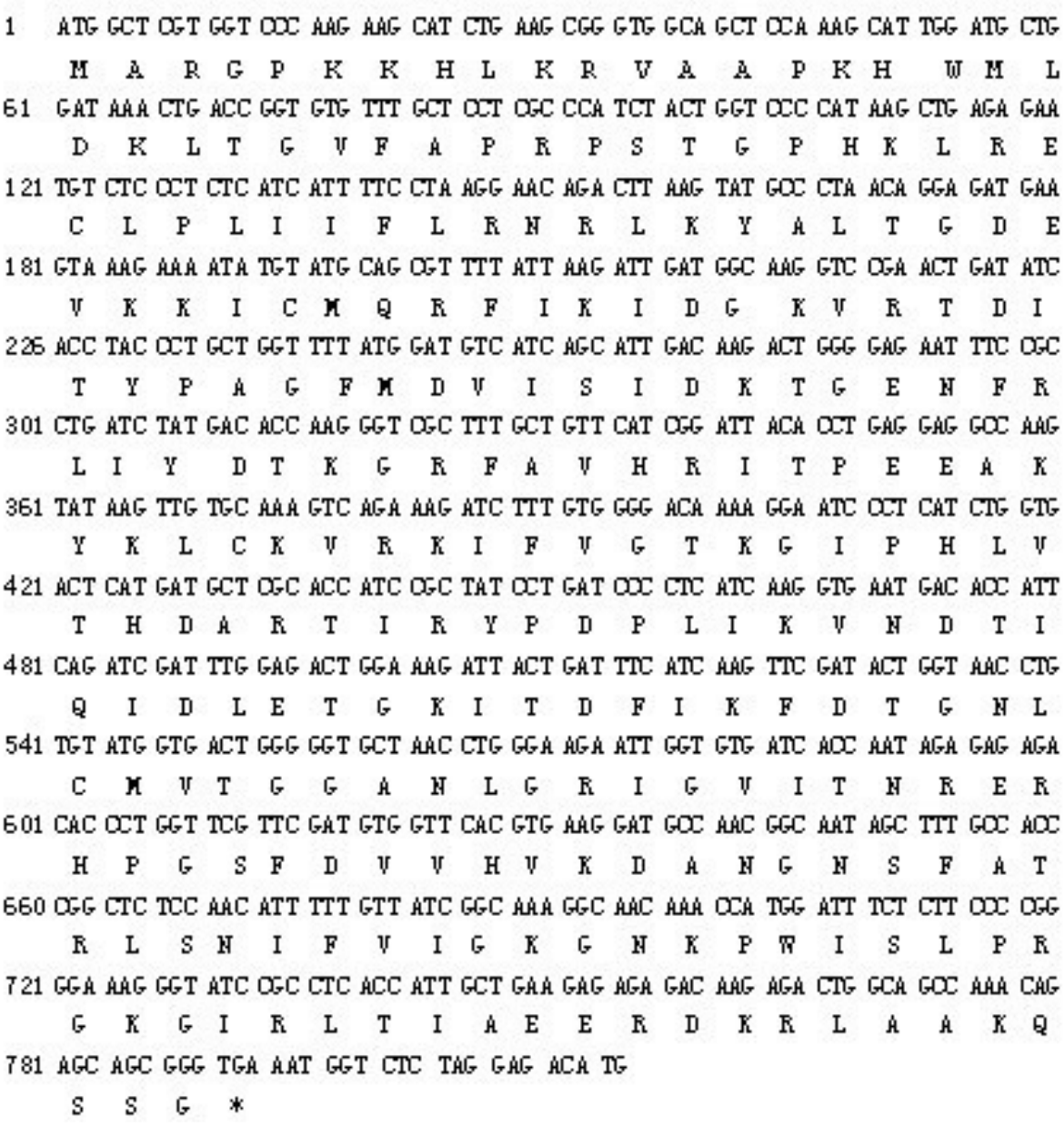

Figure 1. Nucleotide and the deduced amino acid sequence of the giant panda *RPS4X* gene (* stands for the terminator codon).

Table 2. Comparison of nucleotide and amino acid sequences of *RPS4X* between the giant panda and other 4 mammals species.

Item	Species			
	H. sapiens	***B. taurus***	***M. musculus***	***M.domestica***
cds similarity (100%)	93.06	89.52	89.2	83.84
Aa similarity (100%)	100	100	100	95.06

musculus and *M. domestica* with 93.06, 89.52, 89.2 and 83.84%, respectively. The homologies for deduced amino acid sequence are 100, 100, 100 and 95.06%, respectively. This may be caused by transformation or transition of bases, which does not result to changes in the amino acid sequences encoded. This belongs to

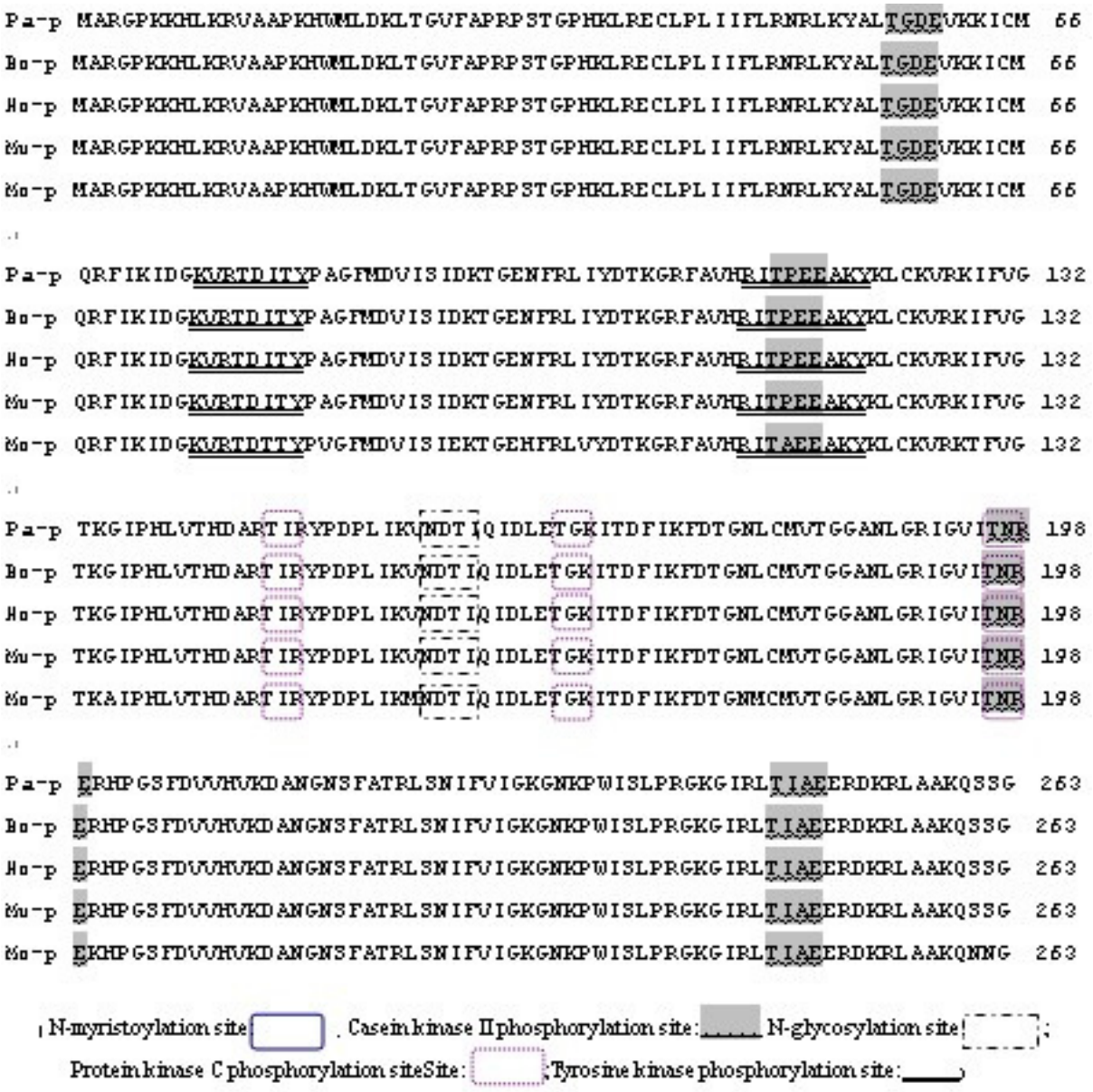

Figure 2. The predicted functional sites of the amino acid sequence encoded by *RPS4X* gene of giant panda (Ho: *H. sapiens*; Bo: *B. taurus*; Mu: *M. musculus*; Mo: *M. domestica*).

synonymous mutations. So, giant panda shares high homology in coding sequence and amino acid sequences of *RPS4X* gene with the other mammals mentioned above, and it shares the optimal homology in coding sequence with *H. sapiens* and has the least homology in amino acid with *M. domestica* (Table 2). Eventually, protein functional sites in RPS4X protein of the five mammals have the same species, positions and numbers. These results indicated that the coding sequence of *RPS4X* and the deduced amino acid sequence are highly conserved. Now, the latter (genomic sequence) length is 5074 bp, possessing 7 exons and 6 introns. When compared with *H. sapiens*, *B. taurus*, *M. musculus* and *M. domestica*, the genomic, the introns, the exons, the 5′-untranslated sequence and the 3′-untranslated sequence of *RPS4X* gene in giant panda are different in length, but the same number of the exons exists in these species, except *M. domestica* (Table 1). We also found that the molecular weight, pI, secondary structure and tertiary structure of the putative protein among the five mammalians are very approximate, so their functional sites have not been changed which suggested that the subtle changes of their primary structure have no effect on the tertiary structure. Meanwhile, the tertiary structure analysis indicated that functional zone of RPS4X protein is between the 43rd and

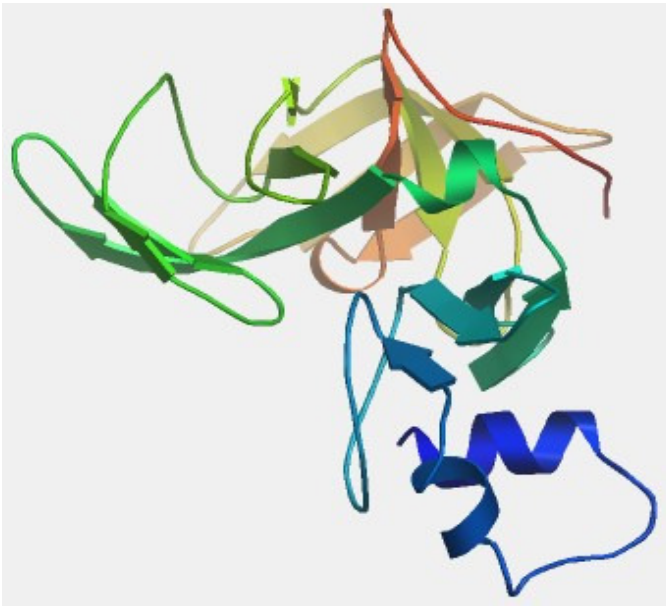

Figure 3. Tertiary structure of the five mammals RPS4X proteins.

242nd amino acid residues (Figure 3). As is typical of genes encoding ribosomal proteins, there are multiple processed pseudogenes of this gene dispersed through the genome (Wiles et al., 1988).

In summary, the complete coding and genomic sequence of *RPS4X* gene have been cloned successfully. And the characteristics of the cloned sequences, together with the deduced protein have been analyzed. The data will enrich and supplement the information on *RPS4X* gene, and it will contribute to the protection of gene resources and the study of genetic polymorphism.

ACKNOWLEDGEMENTS

I have changed the acknowledgements as following: This work was supported by the Key Chinese National Natural Science Foundation (30570275), Key Discipline Construction Project in Sichuan Province (SZD0420), Sichuan key discipline zoology construction funds subsidization project (404001), Application Foundation Project in Sichuan Province (2009JY0061), Youth Fund Project of Educational Committee of Sichuan Province (09ZB088), Foundation Project of Educational Committee of Sichuan Province (10ZC120), Application Foundation Project in Sichuan Province (2011JY0135).

REFERENCES

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002). Molecular biology of the cell. Garland Sci. p. 342.

Fisher EM, Beer-Romero P, Brown LG, Ridley A, McNeil JA, Lawrence JB, Willard HF, Bieber FR, Page DC (1990). Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. Cell, 63(6): 1205-1218.

Garand C, Guay D, Sereduk C, Chow D, Tsofack SP, Langlois M, Perreault E, Yin HH, Lebel M (2011). An integrative approach to identify YB-1-interacting proteins required for cisplatin resistance in MCF7 and MDA-MB-231 breast cancer cells. Cancer Sci. 102(7): 1410-1417.

Geerkens C, Just W, Held KR, Vogel W (1996). Ullrich-Turner syndrome is not caused by haploinsufficiency of RPS4X. Human Genet. 97(1): 39-44.

Hou WR, Chen Y, Peng ZS, Wu X, Tang ZX (2007). cDNA cloning and sequences analysis of ubiquinol_cytochrome creductase complex ubiquinone_binding protein (QP_C) from Giant Panda". Acta Theriologica Sinica, 27(2): 190-194.

Hou WR, Luo XY, Du YJ, Tian MJ (2008). cDNA Cloning and Sequences analysis of RPS15 from the Giant Panda. Recent Patent on DNA Sequence, 2(1): 16-19.

Hou YL, Hou WR, Ren ZL, Hao YZ, Zhang T (2008). cDNA, genomic sequence and overexpression of crystalline alpha_B Gene (CRYAB) of the Giant Panda. Int. J. Biol. Sci. 4(6): 415-421.

Hou YL, Du YJ, Hou WR, Zhou CQ, Hao YZ, Zhang T (2009). Cloning and sequence analysis of translocase of inner mitochondrial membrane 10 homolog (yeast) gene (*TIMM10*) from the giant panda. J. Cell Anim. Biol. 3(1): 9-14.

Jordan P, Carmo-Fonseca M (2000). Carmo-Fonseca M: Molecular mechanisms involved in cisplatin cytotoxicity. Cell. Mol. Life Sci. 57(8-9): 1229-1235.

Jung Y, Lee S, Choi HS, Kim SN, Lee E, Shin Y, Seo J, Kim B (2011). Clinical validation of colorectal cancer biomarkers identified from bioinformatics analysis of public expression data. Clin. Cancer Res. 17(4): 700-709.

Just W, Geerkens C, Held KR, Vogel W (1992). Expression of RPS4X in fibroblasts from patients with structural aberrations of the X chromosome. Hum. Genet. 89(2): 240-242.

Li J, Hou YL, Tian MJ, Yuan SB, Sun B, Su XL, Wu GF, Song Y, Hou WR (2010). Cloning and Sequence Analysis of Ribosomal Protein L11 Gene (RpL11) from the Ailuropoda melanoleuca. Int. Conf. Biomed. Eng. Informatics, pp. 2178-2182.

Liao MJ, Zhu MY, Zhang ZH, Zhang AJ, Li GH, Sheng FJ (2003). Cloning and sequence analysis of FSH and LH in the giant panda (*Ailuropoda melanoleuca*). Anim. Reprod. Sci. 77(1): 107-116.

Mather JP, Moore A, Li RH (1997). Activins, inhibins, and follistatins: ther thoughts on a growing family of regulators. Proc. Soc. Exp.

furBiol. Med. 215(3): 209-222.
Montali RJ (1990). Causes of neonatal mortality in giant panda. Tokyo Zoological Park Soc. pp. 83-94.
Parker SL, Tong T, Bolden S, Wingo PA (1996). Cancer statistics, CA: A Cancer J. Clin. 46(1): 5-27.
Song Y, Hou YL, Hou WR, Wu GF, Zhang T (2010). cDNA, genomic sequence cloning and sequence analysis of heat shock protein beta-1 gene (HSPB1) from the Giant Panda (*Ailuropoda melanoleuca*). Int. Conf. Bioinformatics Biomed. Eng. pp. 1-5.
Su XL, Hou YL, Yuan SB, Tian MJ, Sun Bing, Li J, Wu GF, Song Y, Hou WR (2010). cDNA, genomic sequence cloning and sequence analysis of ribosomal protein L18Agene (RPL18A) from the Giant Panda (Ailuropoda melanoleuca). Int. Conf. Biomed. Eng. Informatics, pp. 2165-2169.
Watanabe M, Furuno N, Goebl M, Go M, Miyauchi K, Sekiguchi T, Basilico C, Nishimito T (1991). Molecular cloning of the human gene, CCG2, that complements the BHK-derived temperature-sensitive cell cycle mutant tsBN63: identity of CCG2 with the human X chromosomal SCAR/RPS4X gene. Cell Sci. 100(1): 35-43.
Wiles MV, Alexander CM, Goodfellow PN (1988). Isolation of an abundantly expressed sequence from the human X chromosome by differential screening. Somatic Cell. Mol. Genet. 14(1): 31-39.
Wool IG, Chan YL, Glück A (1995). Structure and evolution of mammalian ribosomal proteins. Biochem. Cell Biol. 73(11-12): 933-947.
Yang F, Liu WP (2005). Study on the Relation of Ribosomal Protein Gene and Human Diseases. Chin. J. Clin Exp. Pathol. 21(3): 354-356.
Yang XS, Cui BX, Jiang S (2002). Molecular mechanism of cisplatin resistance. Progress in Obstetrics. Gynecol. 11(6): 475-477.
Zinn AR, Alagappan RK, Brown LG, Wool I, Page DC (1994). Structure and function of ribosomal protein S4 genes on the human and mouse sex chromosomes. Mol. Cell. Biol. 14(4): 2485-2492.