



cDNA, genomic sequence cloning and overexpression of ribosomal protein gene L9 (*rpL9*) of the giant panda (*Ailuropoda melanoleuca*)

W.R. Hou, Y.L. Hou, G.F. Wu, Y. Song, X.L. Su, B. Sun and J. Li

Key Laboratory of Southwest China Wildlife Resources Conservation,
Ministry of Education, College of Life Science,
China West Normal University, Nanchong, China

Corresponding author: W.R. Hou
E-mail: hwr168@yahoo.com.cn

Genet. Mol. Res. 10 (3): 1576-1588 (2011)

Received December 18, 2010

Accepted January 15, 2011

Published August 5, 2011

DOI <http://dx.doi.org/10.4238/vol10-3gmr1159>

ABSTRACT. The ribosomal protein L9 (RPL9), a component of the large subunit of the ribosome, has an unusual structure, comprising two compact globular domains connected by an α -helix; it interacts with 23 S rRNA. To obtain information about *rpL9* of *Ailuropoda melanoleuca* (the giant panda), we designed primers based on the known mammalian nucleotide sequence. RT-PCR and PCR strategies were employed to isolate cDNA and the *rpL9* gene from *A. melanoleuca*; these were sequenced and analyzed. We overexpressed cDNA of the *rpL9* gene in *Escherichia coli* BL21. The cloned cDNA fragment was 627 bp in length, containing an open reading frame of 579 bp. The deduced protein is composed of 192 amino acids, with an estimated molecular mass of 21.86 kDa and an isoelectric point of 10.36. The length of the genomic sequence is 3807 bp, including six exons and five introns. Based on alignment analysis, *rpL9* has high similarity among species; we found 85% agreement of DNA and amino acid sequences with the other species that have been analyzed. Based on topology predictions, there are two N-glycosylation sites, five protein kinase C phosphorylation sites, one

casein kinase II phosphorylation site, two tyrosine kinase phosphorylation sites, three N-myristoylation sites, one amidation site, and one ribosomal protein L6 signature 2 in the L9 protein of *A. melanoleuca*. The *rpL9* gene can be readily expressed in *E. coli*; it fuses with the N-terminal GST-tagged protein, giving rise to the accumulation of an expected 26.51-kDa polypeptide, which is in good agreement with the predicted molecular weight. This expression product could be used for purification and further study of its function.

Key words: Giant panda (*Ailuropoda melanoleuca*); Genomic sequence; Ribosomal protein gene L9 (rpL9); cDNA; Cloning; Overexpression

INTRODUCTION

The ribosomal protein L9 (RPL9), a component of the 60S subunit, encoded by the *rpL9* gene, belongs to the L6P family of ribosomal proteins. Within the context of the ribosome, the L9 protein is located in the large subunit, near proteins L1, L2, L16, L19, and L28 (Walleczek et al., 1989; Brimacombe et al., 1990; Nag et al., 1991), and is believed to be relatively well-exposed on the ribosome surface (Agafonov et al., 1997). Structurally, the L9 protein is reminiscent of calmodulin (Babu et al., 1988) and translational initiation factor IF3 (Biou et al., 1995). Previous study has shown that L9 is an elongated protein with a 34-residue α -helix that connects and orients the N- (NTL9) and C-terminal (CTL9) domains, which both contain a predicted RNA-binding site (Adamski et al., 1996; Lieberman et al., 2000). NTL9 is a 56-residue α - β globular protein with a β 1-loop- β 2- α 1- β 3- α 2 topology (Hornig et al., 2002), and CTL9 is a 92-residue α - β globular protein containing two α -helices and an unusual three-stranded mixed parallel, anti-parallel β -sheet (Hoffman et al., 1996).

Since it exhibits specific binding to naked ribosomal RNA, L9 is one of the “primary” RNA binding proteins of the ribosome (Roth and Nierhaus, 1980). The binding site of the L9 protein has been localized to the domain V region of the 23 S rRNA (Adamski et al., 1996). In addition, L9 has been implicated as a participant in a recoding event known as “ribosomal hopping”, where L9 mutations have been associated with the ability of ribosomes to bypass a 50-nucleotide region within the coding region of the bacteriophage T4 gene 60 mRNA (Herbst et al., 1994; Adamski et al., 1996). In view of its remarkably elongated structure, it has been proposed that ribosomal protein L9 may play an architectural role within the ribosome, perhaps as a “molecular strut” serving to stabilize a particular conformation of 23 S RNA (Hoffman et al., 1994, 1996; Lillemoen et al., 1997).

Ailuropoda melanoleuca, a rare species currently found only in China, has a very high ecological, scientific, economic, cultural, and aesthetic value. RPL9 is therefore an RNA binding protein of the ribosome, and the study of RPL9 is already quite rich (Mazuruk et al., 1996; Schmidt et al., 1996; Sato et al., 2000, 2001; Sato and Raleigh, 2002; Cho et al., 2004; Shan et al., 2008; Voelz et al., 2010). The sequence information about RPL9 of *A. melanoleuca* has been also reported (Li et al., 2010), but the mechanism of its action is still controversial, which has drawn our interest in the study, especially on the unreported RPL9 of giant panda. We addressed this question by describing the structure of the protein and providing new data for genetic diversity, parentage, and phylogenesis of the species.

This study was conducted using reverse transcription-polymerase chain reaction (RT-PCR) to amplify the cDNA of the *rpL9* gene from total RNA, which was extracted from the skeletal muscle of the giant panda. The genomic sequence of *rpL9* was cloned successfully from the giant panda using touchdown-PCR. We also analyzed the sequence characteristics of the protein encoded by the cDNA and compared it to those of human and other mammalian species reported. The study is of significance in providing sequence data for the giant panda.

MATERIAL AND METHODS

Materials

Skeletal muscle was collected from a dead giant panda at the Wolong Conservation Center of the Giant Panda, Sichuan, China. The collected skeletal muscle was frozen in liquid nitrogen and then used for DNA and RNA isolation.

DNA and RNA isolation

The genomic DNA of *A. melanoleuca* was isolated from muscle tissue based on literature (Sambrook et al., 1989). The DNA obtained was dissolved in sterile water and kept at -20°C. Total RNA was isolated from about 400 mg muscle tissue using the Total Tissue/Cell RNA Extraction kits (Waton Inc., Shanghai, China), according to manufacturer instructions, dissolved in DEPC (diethylpyrocarbonate) water, and kept at -70°C.

DNA and RNA sample quality was checked by electrophoresis on a 1% agarose gel and quantification was performed spectrophotometrically.

RT-PCR, cloning of cDNA sequence

The PCR primers were designed by Primer Premier 5.0, according to the mRNA sequence of RPL9 from *Homo sapiens* (NM_000661), *Bos taurus* (NM_001024469), *Pongo abelii* (NM_001132412), *Mus musculus* (NM_011292), and *Rattus norvegicus* (NM_001007598). The specific primers of cDNA sequence are as follows: RPL9-F: CTTTGC[C/T] GC [A/G] TCTACTGCGA; RPL9-R: GTTTCTGTAGCCAGGTA ACT.

Total RNAs were synthesized into the first-stranded cDNAs using a reverse transcription kit with Oligo dT as the primers, according to manufacturer instructions (Promega). A 20- μ L volume of first-strand cDNA synthesis reaction system included 1 μ g total RNA, 5 mM MgCl₂, 1 mM dNTPs, 0.5 μ g Oligo dT₁₅, 10 U/ μ L RNase inhibitor, and 15 U AMV reverse transcriptase, and was incubated at 42°C for 60 min.

The first-strand cDNA synthesized was used as a template. The total reaction volume for DNA amplification was 25 μ L. Reaction mixtures contained 1.5 mM MgCl₂, 200 μ M each of dATP, dGTP, dCTP, and dTTP (Omega), 0.3 μ M of each primer, and 5.0 U Taq Plus DNA polymerase (Sangon Co., Shanghai, China). DNA amplification was performed using an MJ Research thermocycler, Model PTC-200 (Watertown, MA, USA) with a program of 4 min at 94.0°C, followed by 30 cycles of 1 min at 94.0°C, 0.5 min at 45°C and 1.5 min at 72.0°C, and then ended with a final extension for 10 min at 72.0°C. After amplification, PCR products were separated by electrophoresis on a 1.5% agarose gel with 1X TAE (Tris-acetate-

EDTA) buffer, stained with ethidium bromide and visualized under UV light. The expected fragments of PCR products were harvested and purified from the gel using a DNA harvesting kit (Omega, China), and then ligated into a pET28a vector at 22°C for 12 h. The recombinant molecules were transformed into *Escherichia coli* complete cells (JM109), and then spread on an LB-plate containing 50 µg/mL ampicillin, 200 mg/mL IPTG (isopropyl-beta-D-thiogalactopyranoside), and 20 mg/mL X-gal. Plasmid DNA was isolated and digested by *Pst*I and *Sca*II to verify the insert size. Plasmid DNA was sequenced by Huada Zhongsheng Scientific Corporation (Beijing, China).

Cloning the genomic sequence of RPL9

The PCR primers were the same as RPL9-F and RPL9-R shown above. The genomic sequence of the RPL9 gene was amplified using touchdown-PCR with the following conditions: 94°C for 30 s, 62°C for 45 s, and 72°C for 4 min in the first cycle and the annealing temperature decreased 1°C per cycle; after 2 cycles, the conditions changed to 94°C for 30 s, 52°C for 45 s, and 72°C for 4 min for another 20 cycles. The fragment amplified was also purified, ligated into the clone vector and transformed into the *E. coli* competent cells. Finally, the recombinant fragment was sequenced by Sangon (Shanghai, China).

Overexpression of recombinant RPL9

PCR fragment corresponding to the RPL9 polypeptide was amplified from the *RPL9* cDNA clone with the forward primer, 5'-ACTGGATCCATGAAGACCATTC (*Bam*HI)-3' and reverse primer, 5'-GACAAGCTTTTATTCATCAGCC (*Hind*III)-3', respectively. PCR was performed at 94°C for 3 min; 35 cycles of 30 s at 94°C, 45 s at 53°C and 1 min at 72°C, and then finally 10 min at 72°C. The amplified PCR product was cut and ligated into the corresponding site of pET28a vector (Stratagene). The resulting construct was transformed into *E. coli* BL21 (DE3) strain (Novagen) and used for the induction by adding IPTG at an OD₆₀₀ of 0.6 and cultivating further for 4 h at 37°C, using the empty vector transformed BL21 (DE3) as a control. The recombinant protein samples were induced after 0, 0.5, 1.5, 2.5, 3.5, and 4 h and then separated by SDS-PAGE and stained with Coomassie blue R250.

Data analysis

The sequence data were analyzed by GenScan (<http://genes.mit.edu/GENSCAN.html>). Homology research of the *A. melanoleuca rpL9* compared with the gene sequences of other species was performed using Blast 2.1 (<http://www.ncbi.nlm.nih.gov/blast/>). The open reading frame (ORF) of the DNA sequence was searched using ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Protein structure of the *rpL9* sequence cloned was deduced using the Predict Protein software (<http://cubic.Bioc.columbia.edu/predictprotein/>) and (<http://swissmodel.expasy.org/>). Multiple sequence alignment was performed by DNAMAN 6.0. The prediction of protein functional sites and biochemical characteristics depended on the ExPASy Proteomics Server software. The construction of evolutionary tree relied on MEGA 4.0.

RESULTS

Analysis of the cDNA of RPL9 of *Ailuropoda melanoleuca*

The cDNA of RPL9, about 600 bp, was amplified from *A. melanoleuca* with primers RPL9-F and RPL9-R. The sequencing results showed that the cloned fragment is 627 bp, containing an ORF of 579 bp, encoding 192 amino acids. The start codon of *rpL9* is ATG, and the stop codon is TAA (Figure 1). The average levels of the base sequence are: A, 31.8%; C, 18.7%; G, 23.1%; T, 26.4%. Blast search showed that the cloned cDNA sequence of *rpL9* is highly homologous with *H. sapiens* and some other species reported. On the basis of the high identity, we concluded that the cDNA isolated is the cDNA encoding the *A. melanoleuca* RPL9 protein. The *rpL9* sequence has been submitted to GenBank (accession No. HM153525).

```

1  ctttgctgcgtctactgcgaga  atg aag acc att ctc agc aat cag act gtc gac att
1                               M  K  T  I  L  S  N  Q  T  V  D  I
58  cca gaa aat gtc gac atc act ctg aag gga cgt aca gtt att gtg aag ggc ccc
13  P  E  N  V  D  I  T  L  K  G  R  T  V  I  V  K  G  P
112 aga gga acc ctg cga agg gac ttc aat cac atc aat gta gaa ctc agt ctt ctt
31  R  G  T  L  R  R  D  F  N  H  I  N  V  E  L  S  L  L
166 gga aag aaa aaa aag agg ctc cga gtt gac aaa tgg tgg gga aat aga aag gaa
49  G  K  K  K  K  R  L  R  V  D  K  W  W  G  N  R  K  E
220 ctg gct act gtt cgc act atc tgt agt cat gta cag aac atg atc aag ggt gtt
67  L  A  T  V  R  T  I  C  S  H  V  Q  N  M  I  K  G  V
274 aca ctg ggc ttc cgt tac aag atg agg tcc gtg tat gct cat ttc ccc atc aac
85  T  L  G  F  R  Y  K  M  R  S  V  Y  A  H  F  P  I  N
328 gtt gtt att cag gag aat ggt tct ctt gtt gaa atc cga aat ttc tta ggt gaa
103 V  V  I  Q  E  N  G  S  L  V  E  I  R  N  F  L  G  E
382 aaa tac atc cgc agg gtt cgg atg agg cca ggt gtt gct tgt tct gta tct caa
121 K  Y  I  R  R  V  R  M  R  P  G  V  A  C  S  V  S  Q
436 gcc caa aaa gat gag tta att ctt gaa gga aat gac att gaa ctt gta tca aac
139 A  Q  K  D  E  L  I  L  E  G  N  D  I  E  L  V  S  N
490 tca gct gct ttg att cag caa gcc aca act gtt aaa aac aag gat atc aga aaa
157 S  A  A  L  I  Q  Q  A  T  T  V  K  N  K  D  I  R  K
544 ttt ttg gat ggt atc tat gtt tct gaa aaa gga aca gtt cag cag gct gat gaa
175 F  L  D  G  I  Y  V  S  E  K  G  T  V  Q  Q  A  D  E
598 taa gatctgagttacctggctacagaaac
      *
```

Figure 1. Nucleotide sequence and putative amino acid sequence of *rpL9* cDNA from *Ailuropoda melanoleuca*. The asterisk represents the termination codon.

Analysis of the genomic sequence of *rpL9* of *Ailuropoda melanoleuca*

Electrophoresis showed that DNA fragment of about 4000 bp was amplified with primers RPL9-F and RPL9-R. The sequencing showed that the length of the DNA fragment cloned is 3807 bp. The genomic sequence of the *RPL9* has been submitted to GenBank (accession No. HM153524). A comparison of the nucleotides of the genomic and cDNA sequences indicated that the gene cloned contains six exons and five introns with the Genescan software. The distribution of six exons is as follows: 23-68; 158-273; 889-985; 2206-2341; 3380-3454; 3673-3781 (Table 1).

Table 1. Comparison of the *rpL9* gene and encoding sequences of *Ailuropoda melanoleuca* with human and six other species (%).

	<i>H. sapiens</i>	<i>B. taurus</i>	<i>P. abelii</i>	<i>M. musculus</i>	<i>R. norvegicus</i>	<i>D. rerio</i>	<i>D. melanogaster</i>
CDS similarity (%)	94.65	93.78	94.30	88.60	90.85	74.74	60.10
Amino acid similarity (%)	100	99.48	98.96	98.96	99.44	88.60	63.54
Molecular mass (kDa)	21.86	21.88	21.79	21.88	21.92	21.91	21.39
Isoelectric point	10.39	10.39	10.34	10.39	10.43	10.39	10.10

Prediction and analysis of protein functional sites

Primary structure analysis revealed that the molecular mass of the putative RPL9 protein is 21.86 kDa with a theoretical isoelectric point (pI) of 10.39. Furthermore, topology prediction showed that there are seven different sites: two N-glycosylation sites, five protein kinase C phosphorylation sites, one casein kinase II phosphorylation site, two tyrosine kinase phosphorylation sites, three N-myristoylation sites, one amidation site, and one ribosomal protein L6 signature 2 (Figure 2).

Overexpression of the *rpL9* gene in *E. coli*

The *rpL9* gene was overexpressed in *E. coli*, using pET28a plasmids carrying strong promoter and terminator sequences derived from phage T7. Therefore, the *rpL9* gene was amplified individually by PCR and cloned in a pET28a plasmid, resulting in a gene fusion coding for a protein bearing a His-tag extension at the N terminus. Expression was tested by SDS-PAGE analysis of protein extracts from recombinant in *E. coli* BL21 strains (Figure 3). The results indicated that the protein *RPL9* fusion with the N-terminally His-tagged form gave rise to the accumulation of an expected 26.51 kDa polypeptide that formed inclusion bodies. Apparently, the recombinant protein was expressed after 30 min of induction, and the highest level was reached after 2.5 h. These results suggested that the protein is active and that it is just the protein encoded by the *RPL9* of *A. melanoleuca*. The expression product obtained could be used to purify the protein for further study of its function.

Comparison of sequence similarity of the *rpL9* gene sequence and encoding sequence

The DNA sequence, CDS and deduced amino acid sequence were compared with the corresponding sequences of *rpL9*. As a result, CDS and coding sequence of *rpL9* of *A. melano-*

leuca share high similarity with the eight other eukaryotes (Table 1). The molecular weight and pI of RPL9 also showed a close comparison. A few changes were found through the comparison of functional sites. *B. taurus* has one less protein kinase C phosphorylation site than the other mammals due to the change of residue 19 (T→N). At the same time, in the other mammals compared, amino acid changes were also found in *P. abelii*, *M. musculus* and *R. norvegicus* (Figure 2). *Danio rerio* and *Drosophila melanogaster* have one more casein kinase II phosphorylation site, *Xenopus laevis* does not have N-glycosylation site, and *D. melanogaster* does not have amidation site but has two additional protein kinase C phosphorylation sites (Figure 2).

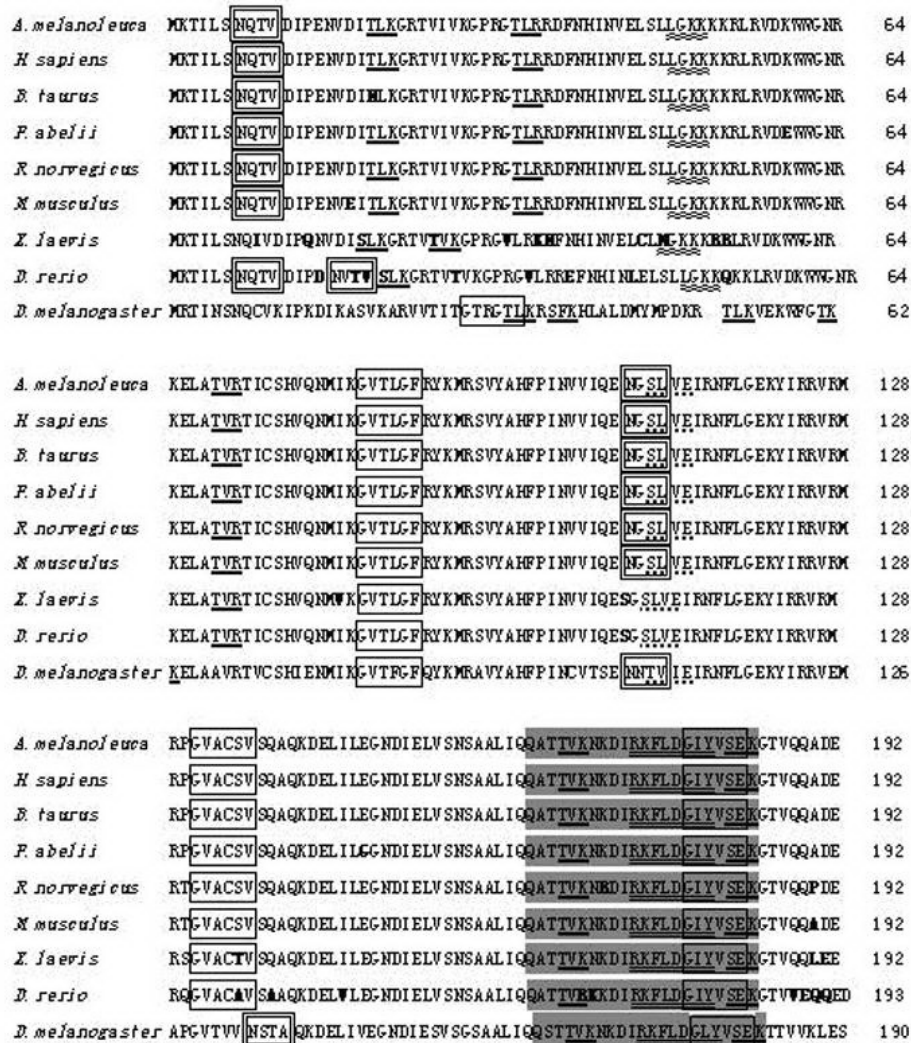


Figure 2. Comparison of functional sites based on the RPL9 amino acid sequences among the different species. Double squares = N-glycosylation site; single underlines = protein kinase C phosphorylation site; dotted lines = casein kinase II phosphorylation site; double straight underlines = tyrosine kinase phosphorylation site; rectangles = N-myristoylation site; double wave underlines = amidation site; shaded letters = ribosomal protein L6 signature 2; bold = polymorphic site.

The ribosomal protein gene *rpL9* was compared, and it was found that there are differences in the length of genes, 5'UTR, 3'UTR and introns among the eight species. The largest intron of the *rpL9* gene from *A. melanoleuca* is 1220 bp, while the length ranges from 84 bp (*H. sapiens* and *D. rerio*) to 2509 bp (*B. taurus*). The intron length was found to decide the length of the gene (Table 2).

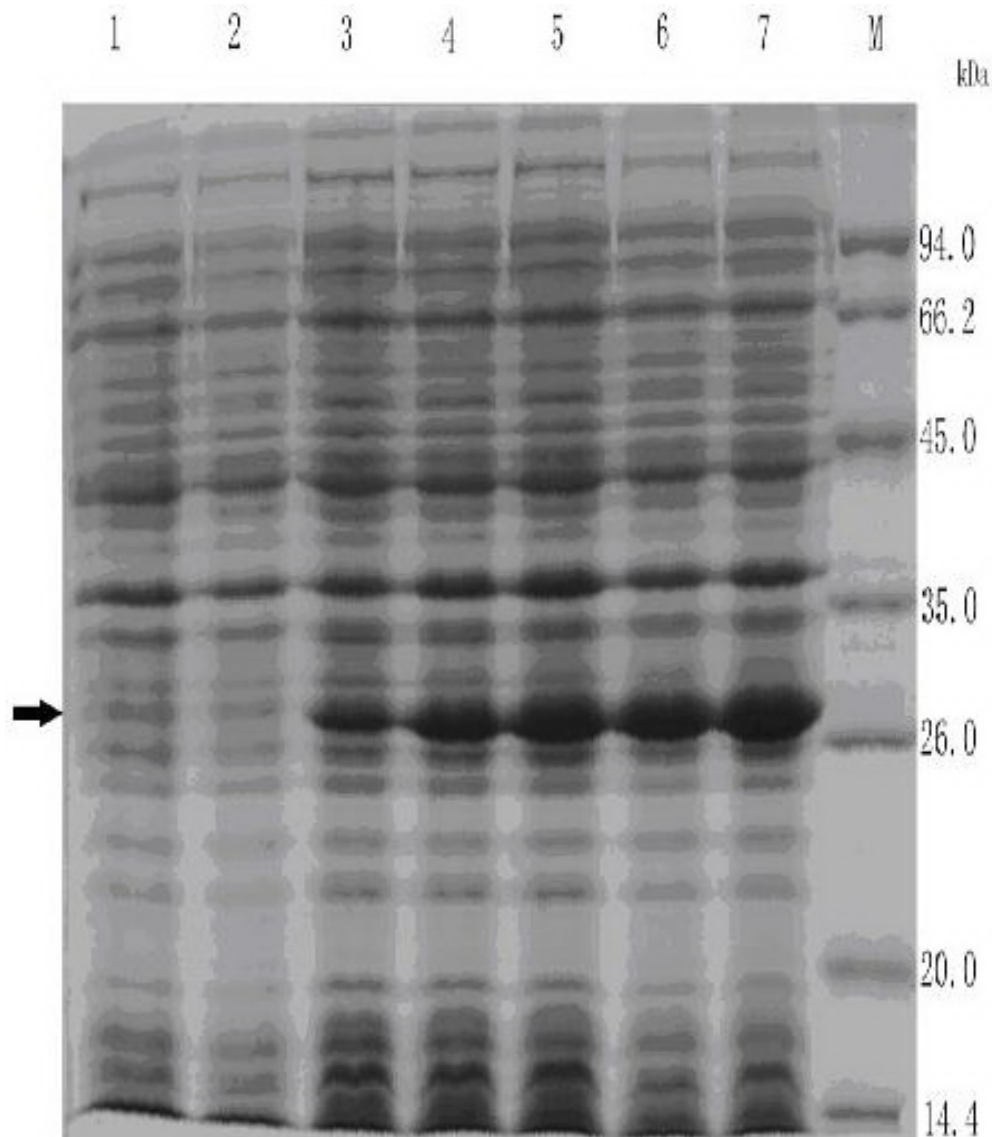


Figure 3. Proteins extracted from recombinant *Escherichia coli* BL21 strains, analyzed by SDS-PAGE gel stained with Coomassie blue R250. Numbers on the right side show the molecular masses in kDa and arrow indicates the recombinant protein bands induced by IPTG at 0, 0.5, 1.5, 2.5, 3.5, and 4 h (lanes 2-7), respectively. Lane 1 represents the products of the *E. coli* strains with the empty vectors; lane M = molecular weight marker.

Table 2. Length comparison of the *rpL9* gene between eight species.

Mammal	Size bp	No. of exons	No. of introns	5'UTR (bp)	3'UTR (bp)	Intron 1 (bp)	Intron 2 (bp)	Intron 3 (bp)	Intron 4 (bp)	Intron 5 (bp)	Accession No.
<i>A. melanoleuca</i>	3807	6	5	22	26	89	615	1220	1038	18	HMI53524
<i>H. sapiens</i>	4825	6	5	500	419	84	512	1047	1461	214	NC_000004
<i>B. taurus</i>	5482	6	5	449	314	89	588	2509	748	206	NC_007304
<i>C. familiaris</i>	4365	6	5	508	366	86	536	959	1122	209	NC_006585
<i>M. musculus</i>	3195	6	5	1	208	91	255	777	326	391	NC_005113
<i>R. norvegicus</i>	2628	6	5	433	208	87	425	774	314	375	NC_000071
<i>D. rerio</i>	4312	6	5	0	52	84	1994	91	110	1399	NC_007112
<i>D. melanogaster</i>	1357	2	1	438	74	273	0	0	0	0	NT_033779

Comparison of high-level structure and construction of the evolutionary tree of the RPL9 protein

We compared tertiary structure among the nine species in order to understand the functional consistency of the protein subunit RPL9 (Figure 4). The species compared were found to have identical tertiary structure.

A evolutionary tree stands for biological evolution process and genetic relationship. The MEGA 4.0 software was employed to construct the minimum evolution tree (Figure 5). The results show that the genetic distance of *A. melanoleuca* is closer to *H. sapiens* and *P. abelii*.

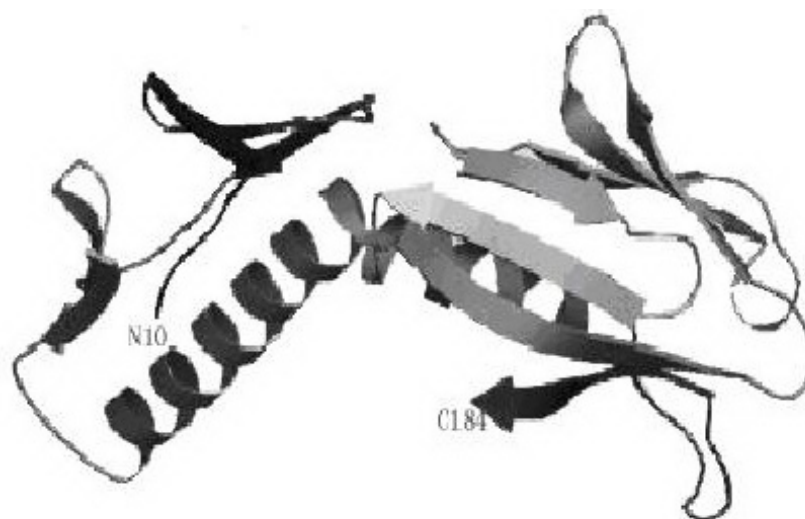


Figure 4. RPL9 protein tertiary structure.

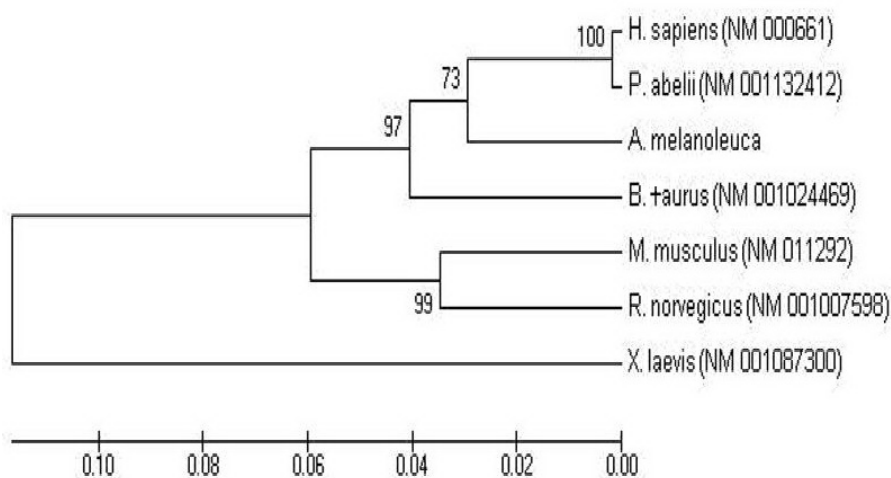


Figure 5. Minimum evolution tree constructed based on CDS of various species.

DISCUSSION

Here, we cloned the genomic sequence and cDNA sequence encoding the ribosomal protein L9 of *A. melanoleuca*. The 3807-bp sequence of the *rpL9* gene was cloned and presented six exons and five introns. Compared with some species, including *H. sapiens*, *B. taurus*, *Canis lupus familiaris*, *R. norvegicus*, *M. musculus*, *D. rerio*, and *D. melanogaster*, the number and length of introns are different. The largest intron of the *rpL9* gene of *A. melanoleuca* was found to be 1220 bp. The variations in lengths of the introns determined the lengths of the *rpL9* gene (Table 2).

The length of the cloned cDNA fragment was found to be 627 bp, containing an ORF of 579 bp. The deduced protein was composed of 192 amino acids with an estimated molecular mass of 21.86 kDa and a pI of 10.86. The RPL9 gene obtained was expressed efficiently in *E. coli* with pET28a plasmids, and the fusion protein obtained was in accordance with the expected 26.51-kDa polypeptide (Figure 3). Although the ribosomal protein gene *rpL9* of *A. melanoleuca* has been reported, the sequence cloned here was different from the published sequence, which does not have termination codon and cytosine as the 50th base, as opposed to thymine in our sequence. Whether genetic diversity or other reasons account for the difference in bases needs to be further studied.

RPL9 shared high homology by determination of CDS and amino acid sequence similarity, in particular, compared with *H. sapiens* and *P. abelii* (Table 1). Alignment analysis indicated that there are many polymorphism sites in the nucleotide sequence of RPL9, most of which are due to transversion and transition in the third codon. These variations do not affect encoded amino acid sequences. RPL9 shares over 85% amino acid sequences with other eukaryotes. The pI and molecular weight were also very similar (Table 1). **Topology prediction** showed fewer changes through comparison of functional sites (Figure 2). Only *B. taurus* had one less protein kinase C phosphorylation site compared to the other mammals due to the change of residue 19 (T→N). Whether this change affects the structure and function of RPL9 needs further studies. At the same time, amino acid changes compared were also found in *P. abelii*, *M. musculus* and *R. norvegicus* (Figure 2); these are located outside the functional sites and have no effect on the structure and function of RPL9 protein. Furthermore, the tertiary structure of RPL9 protein was fully consistent (Figure 4). The RPL9 protein sequence is very conserved in eukaryotes, which is an indisputable fact, while the number of amino acids differs (Figure 2). The number of amino acids of RPL9 increases from *D. melanogaster* (190) to *D. rerio* (193), and remains stable in the other species compared (192).

L9 of the giant panda (*A. melanoleuca*) also contains two small globular 23 S RNA-binding domains separated by a 34-residue α -helix, residues 41 to 74. A ribbon diagram of L9 is shown in Figure 4. The central 17 residues of the helix are largely solvent exposed, while the first seven form part of the N-terminal domain, and the last 10 form part of the C-terminal domain. Several other proteins contain long solvent-exposed α -helices in their crystal structure. The most famous examples are probably calmodulin and troponin C. In sequences of L9 from different species, the length of the long helix is invariant and its central residues are predominantly hydrophilic. The connecting helix is remarkably rich in those amino acids (such as A, L, K, Q, E, and R), which rank high in terms of their helix-forming propensities (Chou and Fasman, 1978). Similarly, these same amino acid types are best at stabilizing helical peptides, according to host-guest studies (O'Neil and DeGrado, 1990). A large fraction of the

amino acid side-chains are charged, occurring in an order such that there are many potential salt bridges between oppositely charged residues within the sequence. In addition, the central region of the helix is densely populated with residues with long side-chains (such as K, E, Q, and R) which, although charged on the solvent-exposed end, are rich in methylene groups and are actually hydrophobic close to the helix backbone. Given the conformational stability of the connecting helix, it appears likely that the protein evolved to play an architectural role within the ribosome, with the connecting helix stabilizing the orientation and distance between the two 23 S rRNA-binding domains.

The N-terminal domain consists of a three-stranded anti-parallel β -sheet sandwiched between a small helix and the N-terminal end of the central helix. Two leucine residues from the central helix, Leu44 and Leu47, pack against the β -sheet and form part of the hydrophobic core of the N-terminal domain. Also, there is a hydrogen bond between the hydroxyl of Thr40 and the amide of Asn43. Asn43 is in the first turn of the central helix and, therefore, would normally not have a hydrogen bond partner. The C-terminal domain contains two long loops, an α -helix, and a three-stranded β -sheet, which packs against the C-terminal end of the central helix. Three residues from the central helix, Ala65, Leu68, and Leu72, form part of the hydrophobic core of the C-terminal domain, while Leu62 caps the hydrophobic core and is partially buried away from the solvent.

In summary, we cloned the cDNA and the complete coding sequence of the *RPL9* gene; the *RPL9* cDNA was efficiently expressed in a prokaryotic organism using pET28a plasmids. The fusion protein obtained is in accordance with the expected 26.51-kDa polypeptide (Figure 3). These results suggest that the protein is active and that it is the protein encoded by the *RPL9* gene of the giant panda. The complex molecular mechanisms and genetic polymorphism will be addressed in future study.

ACKNOWLEDGMENTS

Research supported by the Key Chinese National Natural Science Foundation (#30470261), Application Technology Project in Sichuan Province (#2006J13-057), Key Scientific Research Foundation of Educational Committee of Sichuan Province (#07ZA120), Key Discipline Construction Project in Sichuan Province (#SZD0420), Key Discipline of Zoology Construction Project in Sichuan Province (#404001), Application Foundation Project in Sichuan Province (#2009JY0061), Youth Fund Project of Educational Committee of Sichuan Province (#09ZB088), and Foundation Project of Educational Committee of Sichuan Province (#10ZC120).

REFERENCES

- Adamski FM, Atkins JF and Gesteland RF (1996). Ribosomal protein L9 interactions with 23 S rRNA: the use of a translational bypass assay to study the effect of amino acid substitutions. *J. Mol. Biol.* 261: 357-371.
- Agafonov DE, Kolb VA and Spirin AS (1997). Proteins on ribosome surface: measurements of protein exposure by hot tritium bombardment technique. *Proc. Natl. Acad. Sci. U. S. A.* 94: 12892-12897.
- Babu YS, Bugg CE and Cook WJ (1988). Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* 204: 191-204.
- Biou V, Shu F and Ramakrishnan V (1995). X-ray crystallography shows that translational initiation factor IF3 consists of two compact alpha/beta domains linked by an alpha-helix. *EMBO J.* 14: 4056-4064.
- Brimacombe R, Gornicki P, Greuer B, Mitchell P, et al. (1990). The three-dimensional structure and function of *Escherichia coli* ribosomal RNA, as studied by cross-linking techniques. *Biochim. Biophys. Acta* 1050: 8-13.

- Cho JH, Sato S and Raleigh DP (2004). Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J. Mol. Biol.* 338: 827-837.
- Chou PY and Fasman GD (1978). Empirical predictions of protein conformation. *Annu. Rev. Biochem.* 47: 251-276.
- Herbst KL, Nichols LM, Gesteland RF and Weiss RB (1994). A mutation in ribosomal protein L9 affects ribosomal hopping during translation of gene 60 from bacteriophage T4. *Proc. Natl. Acad. Sci. U. S. A.* 91: 12525-12529.
- Hoffman DW, Davies C, Gerchman SE, Kycia JH, et al. (1994). Crystal structure of prokaryotic ribosomal protein L9: a bi-lobed RNA-binding protein. *EMBO J.* 13: 205-212.
- Hoffman DW, Cameron CS, Davies C, White SW, et al. (1996). Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy. *J. Mol. Biol.* 264: 1058-1071.
- Hornig JC, Moroz V, Rigotti DJ, Fairman R, et al. (2002). Characterization of large peptide fragments derived from the N-terminal domain of the ribosomal protein L9: definition of the minimum folding motif and characterization of local electrostatic interactions. *Biochemistry* 41: 13360-13369.
- Li R, Fan W, Tian G, Zhu H, et al. (2010). The sequence and *de novo* assembly of the giant panda genome. *Nature* 463: 311-317.
- Lieberman KR, Firpo MA, Herr AJ, Nguyenle T, et al. (2000). The 23 S rRNA environment of ribosomal protein L9 in the 50 S ribosomal subunit. *J. Mol. Biol.* 297: 1129-1143.
- Lillemoen J, Cameron CS and Hoffman DW (1997). The stability and dynamics of ribosomal protein L9: investigations of a molecular strut by amide proton exchange and circular dichroism. *J. Mol. Biol.* 268: 482-493.
- Mazuruk K, Schoen TJ, Chader GJ, Iwata T, et al. (1996). Structural organization and chromosomal localization of the human ribosomal protein L9 gene. *Biochim. Biophys. Acta* 1305: 151-162.
- Nag B, Akella SS, Cann PA, Tewari DS, et al. (1991). Monoclonal antibodies to *Escherichia coli* ribosomal proteins L9 and L10. Effects on ribosome function and localization of L9 on the surface of the 50 S ribosomal subunit. *J. Biol. Chem.* 266: 22129-22135.
- O'Neil KT and DeGrado WF (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* 250: 646-651.
- Roth HE and Nierhaus KH (1980). Assembly map of the 50-S subunit from *Escherichia coli* ribosomes, covering the proteins present in the first reconstitution intermediate particle. *Eur. J. Biochem.* 103: 95-98.
- Sambrook J, Fritsch EF and Maniatis T (1989). *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Sato S and Raleigh DP (2002). pH-dependent stability and folding kinetics of a protein with an unusual alpha-beta topology: the C-terminal domain of the ribosomal protein L9. *J. Mol. Biol.* 318: 571-582.
- Sato S, Luisi DL and Raleigh DP (2000). pH jump studies of the folding of the multidomain ribosomal protein L9: the structural organization of the N-terminal domain does not affect the anomalously slow folding of the C-terminal domain. *Biochemistry* 39: 4955-4962.
- Sato S, Xiang S and Raleigh DP (2001). On the relationship between protein stability and folding kinetics: a comparative study of the N-terminal domains of RNase HI, *E. coli* and *Bacillus stearothermophilus* L9. *J. Mol. Biol.* 312: 569-577.
- Schmidt A, Hollmann M and Schafer U (1996). A newly identified Minute locus, M(2)32D, encodes the ribosomal protein L9 in *Drosophila melanogaster*. *Mol. Gen. Genet.* 251: 381-387.
- Shan B, Bhattacharya S, Eliezer D and Raleigh DP (2008). The low-pH unfolded state of the C-terminal domain of the ribosomal protein L9 contains significant secondary structure in the absence of denaturant but is no more compact than the low-pH urea unfolded state. *Biochemistry* 47: 9565-9573.
- Voelz VA, Bowman GR, Beauchamp K and Pande VS (2010). Molecular simulation of ab initio protein folding for a millisecond folder NTL9(1-39). *J. Am. Chem. Soc.* 132: 1526-1528.
- Walleczek J, Redl B, Stoffler-Meilicke M and Stoffler G (1989). Protein-protein cross-linking of the 50 S ribosomal subunit of *Escherichia coli* using 2-iminothiolane. Identification of cross-links by immunoblotting techniques. *J. Biol. Chem.* 264: 4231-4237.