

Full Length Research Paper

# Isolation, sequence identification and tissue expression profile of a novel ribokinase gene (*RBKS*) from Chinese Banna mini-pig inbred line (BMI)

Jinlong Huo<sup>1,2,3</sup>, Pei Wang<sup>2,3</sup>, Yongwang Miao<sup>3</sup>, Hailong Huo<sup>1,4</sup>, Lixian Liu<sup>4</sup>, Yangzhi Zeng<sup>2,3</sup> and Heng Xiao<sup>1\*</sup>

<sup>1</sup>Faculty of Life Science, Yunnan University, Kunming 650091, Yunnan, China. <sup>2</sup>Key Laboratory of Banna Mini-pig Inbred Line of Yunnan Province, Kunming 650201, Yunnan, China. <sup>3</sup>Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, Yunnan, China. <sup>4</sup>Department of Husbandry and Veterinary, Yunnan Vocational and Technical College of Agriculture, Kunming 650031, China.

Accepted 23 November, 2018

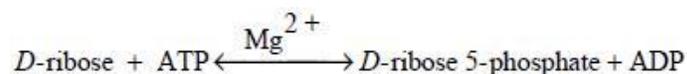
The complete expressed sequence tag (CDS) sequence of Banna mini-pig inbred line (BMI) ribokinase gene (*RBKS*) was amplified using the reverse transcription-polymerase chain reaction (RT-PCR) based on the conserved sequence information of the cattle or other mammals and known highly homologous swine ESTs. This novel gene was then deposited into NCBI database and assigned to accession number JF944892. Sequence analysis revealed that the BMI *RBKS* encodes a protein of 323 amino acids that has high homology with the ribokinase proteins of seven species: cattle (99%), horse (99%), orangutan (99%), human (89%), monkey (89%), rat (88%) and mouse (80%). The phylogenetic tree analysis revealed that the BMI *RBKS* gene has a closer genetic relationship with the *RBKS* genes of bovine and horse than with those of orangutan, human, monkey, rat and mouse. Analysis by RT-PCR showed that BMI *RBKS* gene was over-expressed in ovary and lung, moderately expressed in spleen, nerve fiber, large intestine and diencephalon, weakly expressed in heart, skin, muscle, small intestine, midbrain, kidney and fat, while almost silent in other five tissues. Four microRNA target sites were predicted in the CDS of BMI *RBKS* mRNA for further study of this gene in the future. The 3D structure of the *RBKS* by homology modeling was similar to that of human ribokinase (2fv7). Our experiment will establish a foundation for further insight into this swine gene.

**Key words:** Banna mini-pig inbred line (BMI), pig, ribokinase gene (*RBKS*), tissue expression analysis, homology modeling.

## INTRODUCTION

Ribokinase (*RBKS*) is a key enzyme which catalyzes the phosphorylation of ribose to ribose 5-phosphate (R-P-5) using ATP as a phosphate donor according to the

reaction:



\*Corresponding author. E-mail: xiaoheng@ynu.edu.cn. Fax: +86-871-5033732.

**Abbreviations:** *RBKS*, Ribokinase gene; **EST**, expressed sequence tag; **CDS**, coding sequence; **RT-PCR**, reverse transcription-polymerase chain reaction; **RMSD**, root mean square deviation.

Exogenous ribose plays a number of roles in the biological metabolism. For example, ribose addition can help maintain higher levels of ATP in rat heart and dog kidney during transplantation experiments (Müller et al., 1998) and it leads to improvement of neurological symptoms in patients with adenylosuccinase deficiency (Salerno et al., 1999). In order to trap them inside the cell

and prepare them for further chemical reactions, ribose must first be converted into ribose-5-phosphate (Bork et al., 1993). The phosphorylated ribose can then enter the pentose phosphate pathway (PPP) for energy production or be used as a carbon source for the synthesis of nucleotides, histidine and tryptophan (Anderson and Cooper 1969; Lopilato et al., 1984). Sequence analyses showed that the enzyme *RBKS* belongs to the PfkB family of carbohydrate kinases (Bork et al., 1993; Wu et al., 1991). As another member, *RBKS* contains two highly conserved sequence motifs, a glycine-rich area served as a substrate binding site near the N-terminus and a motif involved in ATP binding and the catalytic center near the C-terminus (Sigrell, et al., 1998, 1999). Particularly, *RBKS* is implicated in tissue-protective mechanisms against various ischemic insults.

Based on the above description of *RBKS* gene, it is necessary to isolate this gene from pig for it is associated with energy metabolism, health and other important biological functions of animals. But until today, the porcine *RBKS* has not been reported. Surprisingly, the Banna mini-pig Inbred line (BMI) was exploited by Yunnan Agricultural University from 1980s based on the small-ear pigs at Xishuangbanna, Yunnan Province, China. A pair of progenitors was a sow and her son. Then, the propagation was conducted by means of highly full sibling or parent-offspring inbreeding and each generation underwent the strict selection. As heterozygotic genes were separated and recombined in the process of inbreeding, BMI has already owned six families and 18 substrains with different phenotypes and genotypes. Due to their consistent genetic background and minor interindividual differences, BMI is considered as an ideal model organism for biological studies (Crabbe et al., 2005; Yu et al., 2004; Zeng and Zeng, 2005).

The objective of this study was to isolate the full length coding sequence of BMI *RBKS* gene according to the conserved sequence information of cattle or other mammals and highly homologous swine ESTs sequence information, conduct sequence analysis and some necessary function analysis of established nucleotide sequence, and finally examine the expression in a range of BMI tissues. These will provide a primary foundation for further research on this porcine gene.

## MATERIALS AND METHODS

### Samples collection, RNA extraction and first-strand cDNA synthesis

Three matured female BMI were slaughtered for sampling. Fresh tissues (lymph node, midbrain, ovary, diencephalon, cerebrum, liver, kidney, spleen, heart, lung, nerve fiber, stomach, small intestine, large intestine, pancreas, skin, muscle and fat) were snap frozen in liquid nitrogen and stored at -80°C before use. The total RNA was extracted using the RNeasy Plus (TaKaRa, Dalian) according to the manufacturer's instructions. To remove genomic DNA contamination, total RNA was digested with RNase-free DNase I (TaKaRa, Dalian). Three micrograms of RNA were reverse

transcribed with oligo (dT)<sub>18</sub> primer and M-MLV reverse transcriptase (Invitrogen, USA). The efficiency of reverse transcription was checked on 2% agarose gels stained with ethidium bromide.

### Isolation of the BMI *RBKS* gene

The GenBank *RBKS* sequences for human (accession no. NM\_022128), cattle (accession no. NM\_001191271) and their highly homologous pig ESTs sequences: CX065544, CN154557, CN165374, CK461198, BX924148 and CN165735 were used to design a primer pair to amplify the complete coding sequence of *RBKS* by using Primer Premier 5.0 software. The primers for BMI *RBKS* gene were: 5'- AAT GGC CGC GTC TGG GGA AC -3' and 5'-CAG TCA AAA CAG GTA AAG GGC C -3'. Reverse transcription-polymerase chain reaction (RT-PCR) was performed to isolate the BMI *RBKS* using the pooled cDNAs from different tissues listed above. The 25 µl reaction system was: 2.0 µl cDNA (25 ng/µl), 2.0

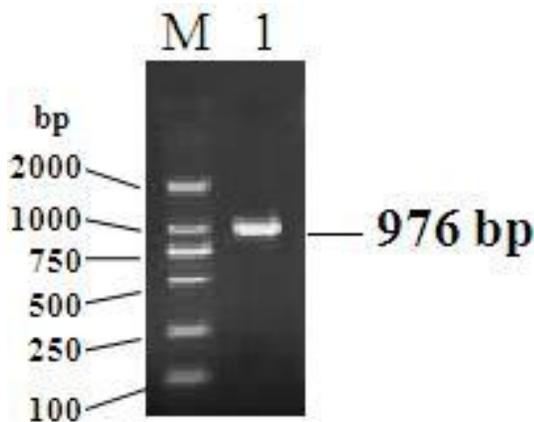
µl 2.5 mM mixed dNTPs (TaKaRa, Dalian), 2.5 µl 10-Taq DNA polymerase buffer (Mg<sup>2+</sup> Plus), 0.5 µl 10 µM forward primer, 0.5 µl, 10 µM reverse primer, 0.25 µl Taq DNA polymerase (5 U/µl, TaKaRa, Dalian), and 17.25 µl sterile water. The PCR program initially started with 94°C denaturation for 2 min, followed by 35 cycles of 94°C /30 s, 55°C /40 s, 72°C /1.5 min then 72°C extension for 10 min, and finally 4°C to terminate the reaction. After the PCR, the gene product was cloned into pMD18- T vector (TaKaRa, Dalian) and sequenced bidirectionally with the commercial fluorometric method. At least, five independent clones were sequenced.

### Bioinformatics analysis

Sequence analysis of BMI *RBKS* gene was performed using softwares in NCBI (<http://www.ncbi.nlm.nih.gov>) and ExPaSy (<http://www.expasy.org>). The cDNA sequence was predicted using the online GenScan software (<http://genes.mit.edu/GENSCAN.html>). Putative protein theoretical molecular weight (Mw) and isoelectric point (pI) prediction, signal peptide prediction, subcellular localization prediction and transmembrane topology prediction were performed using the Compute pI/Mw Tool ([http://us.expasy.org/tools/pi\\_tool.html](http://us.expasy.org/tools/pi_tool.html)), SignalP 4.0 server (<http://www.cbs.dtu.dk/services/SignalP/>), PSORT II (<http://psort.hgc.jp/>) and TMHMM-2.0 server (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>), respectively. Web-based microRNA (miRNA) predicting program was used to locate conserved potential miRNA targets (<http://www.mirbase.org/>). The Blastp program and Conserved Domain Architecture Retrieval Tool were used to search for similar proteins and conserved domain, respectively (<http://www.ncbi.nlm.nih.gov/Blast>). The alignment of the nucleotide sequences and deduced amino acid sequences were computed using ClusterX, and the phylogenetic tree was computed using the ClustalX and MEGA 4.0 softwares with standard parameters. Secondary structures of deduced amino acid sequences were predicted with SOPMA (<http://npsa-pbil.ibcp.fr/>). The 3D structures were predicted using Discovery Studio 3.1 software.

### Semi-quantitative RT-PCR

To characterize the *RBKS* gene further, RT-PCR was conducted to determine its expression in 18 BMI tissues. To eliminate the effect of cDNA concentration, we repeated the RT-PCR five times using 1, 2, 3, 4 and 5 µl cDNAs as templates. We selected the housekeeping gene 18S rRNA (NR\_002170) as a positive control. The control primers used were: 5'- GGACATCTAAGG-GCATCACAG -3' and 5'- AATTCCGATAACGAACGAGACT -3'. The



**Figure 1.** RT-PCR result for BMI *RBKS* gene. M, DL2000 DNA marker; 1, PCR product.

BMI *RBKS* primers which were used to perform the semi-quantitative RT-PCR for tissue expression profile analysis were the same as the primers for isolation RT-PCR above. The PCR reactions were optimized for a number of cycles to ensure product intensity within the linear phase of amplification.

## RESULTS

### Cloning and identification of BMI *RBKS* cDNA

Through RT-PCR with pooled tissue cDNAs for BMI *RBKS* gene, the resulting PCR product was 976 bp (Figure 1). This cDNA nucleotide sequence analysis using the BLAST software at NCBI server revealed that BMI *RBKS* gene was not homologous to any of the known porcine genes and it was then deposited into the GenBank database under accession No. JF944892. The sequence prediction was carried out using the GenScan software and results show that the 976 bp cDNA sequence represents a single gene which encoded 323 amino acids. The complete CDS and the encoded amino acids are presented in Figure 2.

### Physical and chemical characteristics of BMI *RBKS*

The theoretical isoelectric point (pI) and molecular weight (Mw) were computed using the Compute pI/Mw Tool. The pI and the molecular weight of BMI *RBKS* are 5.16 and 34532.46, respectively. Submitting the *RBKS* protein sequence to SignalP, the *RBKS* protein had no N-terminal signal peptide and is a non-secretory protein (Petersen et al., 2011). Using a hidden Markov model algorithm, transmembrane topology prediction made by TMHMM program (Moller et al., 2001), showed that BMI *RBKS* was not a potential membrane protein. For subcellular localization analysis, the amino acid sequence was submitted to the PSORT II program, and

Reinhardt's method showed that BMI *RBKS* was probably located in the cytoplasm with up to 94.1% probability (Nakai and Horton, 1999).

### Prediction and analysis of structures and conserved domains of BMI *RBKS*

Proteins often contain several domains, each of which had their own evolutionary origins and functions. Examination using the Conserved Domain Architecture Retrieval Tool of Blast at the NCBI server (<http://www.ncbi.nlm.nih.gov/BLAST>) indicated that BMI *RBKS* contains one separated conserved domain-ribokinase (from 23 to 314 amino acid residues, Figure 3). Then, putative protein was analyzed using prosite (<http://expasy.org/prosite/>) and SMART (<http://smart.embl-heidelberg.de/>) softwares. Four kinds of sites were found, which were N-myristoylation sites (23-GScmTD-28, 53-GGkgAN-58, 56-GAnqCV-61, 67-GAktSM-72, 238-GAegCV-243, 267-GAgdSF-272, 307-GTqsSY-312), Protein kinase C phosphorylation sites (33-TsR-35, 257-TeK-259), Casein kinase II phosphorylation sites (160-TsIE-163, 192-TIsD-195, 202-SeaE-205, 208-TgIE-211, 246-SrtE-249, 287-SleE-290) and N-glycosylation sites (200-NESE-203, 285-NLSL-288). The prediction of secondary structure by SOPMA indicates that the deduced BMI *RBKS* contains 105 alpha helices, 73 extended strands, 30 beta turns and 115 random coils (Figure 4).

### Homology modeling

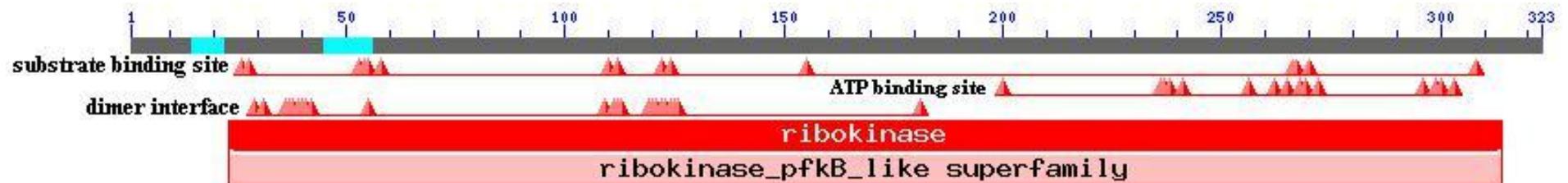
In order to better understand the detailed structures of BMI *RBKS*, the homology modeling of *RBKS* was performed to estimate its 3D structure using Discovery Studio 3.1 software. Commonly, the root mean square deviation (RMSD) between corresponding atoms of the template protein and modeled protein is a used measure of similarity between two protein structures. The smaller the RMSD is between two structures, the more similar are these two structures. In protein structure prediction, one needs the RMSD between predicted and experimental structures for which a prediction can be considered to be successful. Success is obvious only when the RMSD is as small as that for closely homologous proteins (<3 Å) (Reva et al., 1998). The RMSD value of BMI *RBKS* and 2fv7B (B chain of human *RBKS*) are 0.218 which indicates that the 3D structure of the BMI *RBKS* was similar to that of the human ribokinase (2fv7 Chain: B). Furthermore, in the ribokinase domain of *RBKS*, no differences were found among different species of animals in the shape and orientation in *RBKS* on 3D structures (Figure 5). The 3D structure analysis may provide a basis for further study of the relationship between structure and function of *RBKS*.

```

aATGGCCGCGTCTGGGGAAACCCGGGAGGCAGTGGCAGGAGCAGGTGGCGGGCGCAGTGGTGGTGGTGGGCTCCTGCATGACCGACCTGGTCAGCCTTACT 100
M A A S G E P G R Q W Q E Q V A A A V U U U V G S C M T D L V S L T 33
TCTCGTTTTGCCAAAACCTGGAGAAACCATCCATGGACAGAAGTTTTTTATCGGCTTTGGAGGGAAAGGTGCCAACCCAGTGTGTCCAGGCTGCCAGGCTTG 200
S R L P K T G E T I H G Q K F F I G F G G K G A N Q C U Q A A R L 66
GAGCCAAGACGTCATGGTGTGCAAGGTTGGCAAGATTCTTTGGCAATGATTATATAGAAAACCTTAAAGCAGAATGATATTTCTACAGAATTTACATA 300
G A K T S M U C K U G K D S F G N D Y I E N L K Q N D I S T E F T Y 100
TCAGACGGAGATGCTGGCAGGAACTGCTTCGATAATTGTCAATAATGAAGGCCAGAATATCATCGTTATAGTGGCTGGAGCAAATTTACTTTTGAAT 400
Q T E D A A T G T A S I I U N N E G Q N I I U I U A G A N L L L N 133
ACTGAAGACCTGAGGGAAGCAGCCAGTACCATCAGCAGAGCCAAAGTAATGATCTGCCAGCTAGAAGTAACTCCAGCAACTTCTTTGGAAGCCCTGACAA 500
T E D L R E A A S T I S R A K U M I C Q L E U T P A T S L E A L T 166
TAGCCACAGCAATGGAGTGAAAACACTGTTCAACCCAGCACCAGCCATTGCCGACCTGGACCCCCGGTTCTACACCCTCTCGGATGTGTTCTGCTGCAA 600
I A H S N G U K T L F N P A P A I A D L D P R F Y T L S D U F C C N 200
TGAAGCGAGGCTGAGATTCTGACCGGCCCTTGAGGTCCGCCGCCCAAGGATGCCGGGCAGGCTGGCAGGTGCTCCTGGAGAGGGGCTGCTGCGTTGTG 700
E S E A E I L T G L E U R R P K D A G Q A A Q U L L E R G C C U V 233
ATCATCACCTAGGGGCTGAAGGGTGTGTGATGCTGTACGGACAGAACCCGTCGCCGAACACATTCCCACAGAGAAAGTCAAGGCTGTGGATACTACGG 800
I I T L G A E G C U M L S R T E P U P K H I P T E K U K A U D T T 266
GTGCTGGTGACAGTTTTGTGGGAGCTTTGGCCTTCTACCTGGCTTACTACCCAAACCTGTCCTTGGGAAGAAATGATCAAGAGGTCCAATTTTCATCGCGGC 900
G A G D S F U G A L A F Y L A Y Y P N L S L E E M I K R S N F I A A 300
AGTCAGCGTCCAGGCCACAGGAACACAGTCATCTTATCCATACAAAAAAGATCTGCCCTTTACCTGTTTTGActg 976
U S U Q A T G T Q S S Y P Y K K D L P L Y L F 323

```

**Figure 2.** The complete cDNA sequence and amino acid sequence of the protein encoded by *RBKS* (GenBank accession number: JF944892). ATG, start codon; TGA, stop codon; capital letters, complete CDS and amino acid sequence; gray highlighted nucleotide sequence, primers; pane, catalytic residues



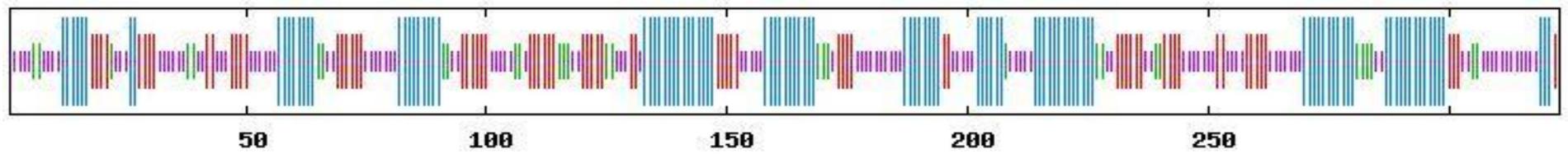
**Figure 3.** The putative domains of the protein encoded by BMI *RBKS*.

### Location of potential miRNA targets

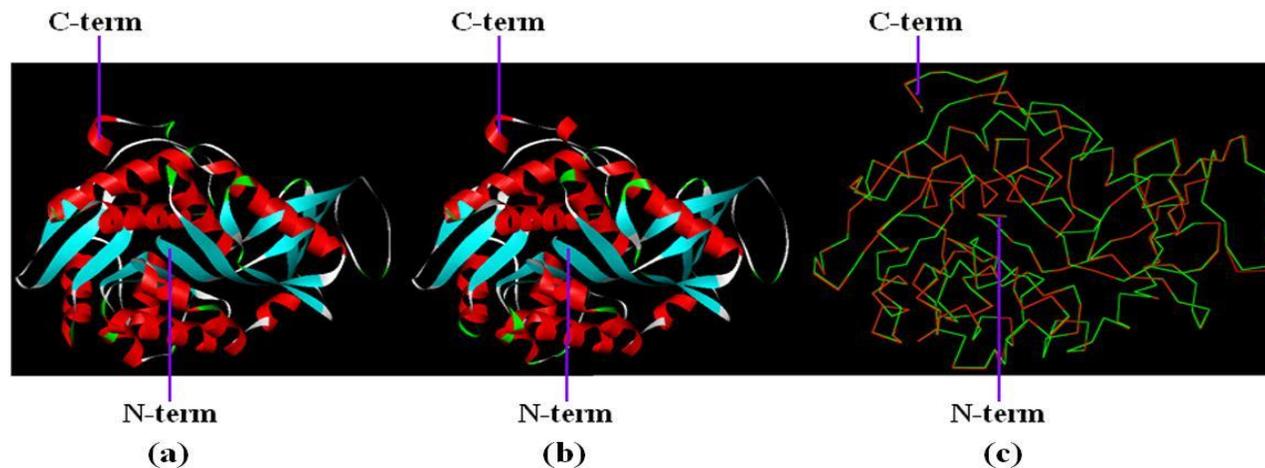
MicroRNAs are noncoding single-stranded RNA molecules of 17 to 24 nucleotides that can regulate gene expression by binding to the coding

region of target mRNAs (Bartel, 2004; Zeng et al., 2003). We used web-based microRNA (miRNA) predicting programs to locate conserved potential miRNA targets: miRBase (<http://www.mirbase.org/>). The results show that four *Sus scrofa* micro-

RNAs (ssc-miR-217, ssc-miR-196b-3p, ssc-miR-1306-3p and ssc-miR-1306-3p) were found to have the predicted target sites (313- ugcug-cgacaggaacugcuucgau-336, 718- ugaagggugugu-gaugcuguc- 738, 428- accaucagcagagccaaagu-



**Figure 4.** The secondary structure of the BMI *RBKS* protein predicted by SOPMA. Helices, extended strands, beta turns and random coils are indicated, respectively, with the longest, the second longest, the second shortest and the shortest vertical lines.



**Figure 5.** The tertiary structure of BMI *RBKS*. (a) BMI *RBKS*, (b) Human 2fv7B, (c) The superimposition figure; BMI *RBKS* (red); 2fv7B (green).

447 and 48-cggcggcagugguggug-67) in the BMI *RBKS* sequence, respectively.

#### **Analysis of sequence identity and evolutionary relationships of BMI *RBKS***

The deduced protein sequence of BMI *RBKS* was submitted to generate BLAST reciprocal best hits, and similarity comparison revealed that BMI *RBKS* protein has high homology with the *RBKS*

proteins of seven other species: cattle (91%), horse (91%), orangutan (91%), human (89%), monkey (89%), rat (88%) and mouse (86%) (Figure 6). To evaluate the evolutionary relationships of BMI *RBKS* with other species, we constructed a phylogenetic tree using DNASTar, Cluster, MEGA and DNAMAN softwares on the basis of the *RBKS* amino acid sequences. The phylogenetic tree analysis revealed that the BMI *RBKS* gene has a closer genetic relationship with the *RBKS* genes of bovine and horse than with

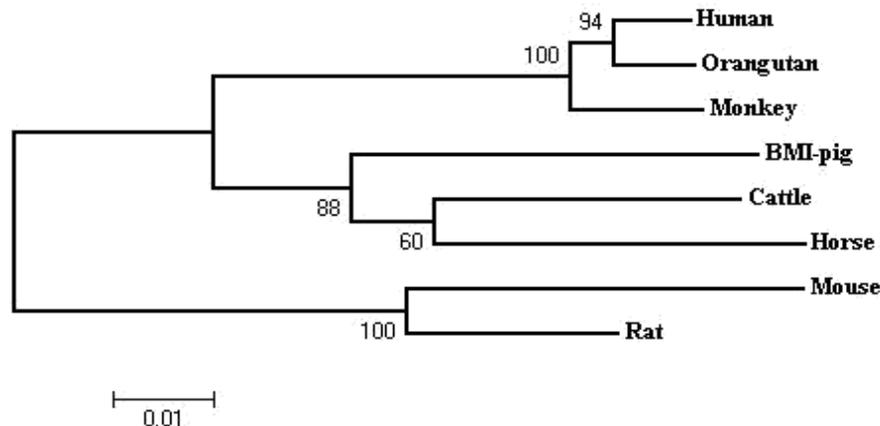
those of orangutan, human, monkey, rat and mouse (Figure 7).

#### **mRNA tissue-specific expression profile**

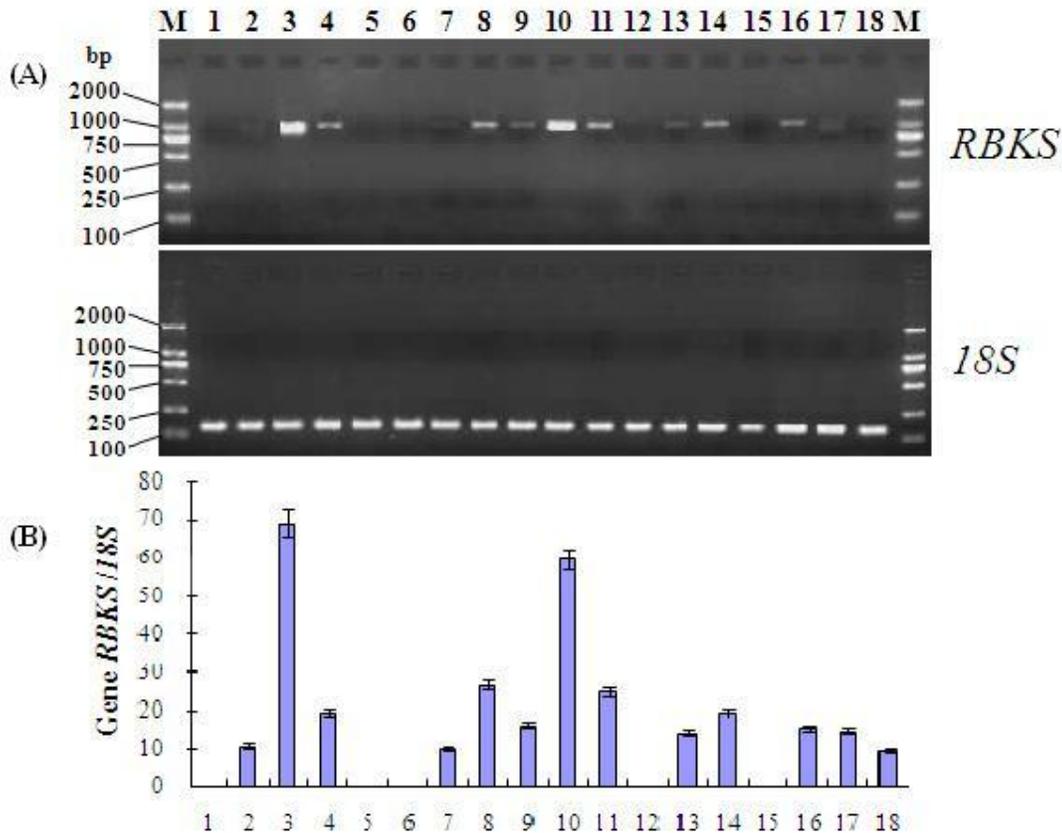
To check the relative expression levels of *RBKS* mRNA in various porcine tissues, semi-quantitative RT-PCR was performed in 18 BMI tissues mentioned above. The continuously expressed gene, 18S, was used and served as

BMI-pig	MAASGEPGRWQEQV...AAAVVVVVGSCMTDLVSLTSRLPKTGETIHGQKFFIFGGGKGANQCVQA	63
Human	MAASGEPQRWQEEV...AAAVVVVVGSCMTDLVSLTSRLPKTGETIHGKFFIFGGGKGANQCVQA	62
Cattle	MAASGDPGRWQEEAEETAAAVVVVVGSCMTDLVSLTSRLPKTGETIHGKFFIFGGGKGANQCVQA	65
Horse	MAAAGDPGRWPEEE...EAAAVVVVVGSCMTDLVSLTSRLPKTGETIHGKFFIFGGGKGANQCVQA	64
Orangutan	MAASGEPRRWQEEV...VAVVVVVGSCMTDLVSLTSRLPKTGETIHGKFFIFGGGKGANQCVQA	62
Monkey	MAASGEPWRWQEEV...AAAVVVVVGSCMTDLVSLTSRLPKTGETIHGKFFIFGGGKGANQCVQA	62
Mouse	MAACGEPGRPWQEE...AAAVVVVVGSCMTDLVSLTSRLPKTGETIHGHEFFIFGGGKGANQCVQA	63
Rat	MAASGEPGRWLWQEE...AAAVVVVVGSCMTDLVSLTSRLPKTGETIHGKFFIFGGGKGANQCVQA	63
BMI-pig	ARLGAKTSMVCKVGDSEFGNDYIENLKQNDISTEFTYQTEDAATGTASIIVNEGQNIIVIVAGA	128
Human	ARLGAMTSMVCKVGDSEFGNDYIENLKQNDISTEFTYQTKDAATGTASIIVNEGQNIIVIVAGA	127
Cattle	ARLGAKTSMVCKVGDSEFGNDYIENLKQNDISTEFTYQTKDAATGAASIIVDNEGQNIIVIVAGA	130
Horse	ARLGAKTSMVCKVGDSEFGNDYIENLKQNDISTEFTYQTKDAATGAASIIVDNEGQNIIVIVAGA	129
Orangutan	ARLGAMTSMVCKVGDSEFGNDYIENLKQNDISTEFTYQTKDAATGTASIIVNEGQNIIVIVAGA	127
Monkey	ARLGATISMVCKVGDSEFGNDYIENLKQNGISTEFTYQTKDAATGTASIIVINGGQNIIVIVAGA	127
Mouse	ARLGAKAAIVCKVGNDSFGNDYIENLKQNHISTEFTYQTRDAATGTASIIVNEGQNIIVIVAGA	128
Rat	ARLGAKAAIVCKVGNDSFGNNDYIENLKQNHISTEFTYQTRDAATGTASIIVNEGQNIIVIVAGA	128
BMI-pig	NLLLNTEDLREAASTISRAKVMTCQLEVTIPATISLEALTIHNSGVKTLFNPAPAIADLDPRFYTL	193
Human	NLLLNTEDLRAAANVISRAKVMVCQLEITIPATISLEALTMARRSGVKTLFNPAPAIADLDPQFYTL	192
Cattle	NLLLNTEDLREAAASISRAKVMTCQLEVTIPATISLEALRIHNSGVKTLFNPAPAIADLDPQFYTL	195
Horse	NLLLNTEDLREAAANISRAKVMTCQLEVTIPATISLEALTIHNSGVKTLFNPAPAIADLDPRFYTL	194
Orangutan	NLLLNTEDLRAAANVISRAKVMVCQLEITIPATISLEALTMARRSGVKTLFNPAPAIADLDPRFYTL	192
Monkey	NLLLNTEDLRAAANVISRAKVMVCQLEITIPATISLEALTMARRSGVKTLFNPAPAIADLDPRFYTL	192
Mouse	NLLLNSEDLKKAASVISRAKVMTCQLEITSPAASLEALTMARRSGVKTLFNPAPAMADLDPQFYTL	193
Rat	NLLLNTEDLKKAHVVISRAKVMTCQLEITSPAASLEALTMARRSGVKTLFNPAPAIADLDPRFYTL	193
BMI-pig	SDVFCCNESEAEILTGLTVRRPKDAGQAAQVLLERGCQVVIITLGAEGCVMLSRTEPVPKHIPTE	258
Human	SDVFCCNESEAEILTGLTVGSAAADAGEAALVLLKRGCOVVIITLGAEGCVVLSQTEPEPKHIPTE	257
Cattle	SDVFCCNESEAEILTGLTVCSPTDAGRAALVLLERGCQVVIITLGAEGCVMLSQTEPVPKHIPTE	260
Horse	SDVFCCNETEAEILTGLTVGSPADAGKAALVLLERGCQVVIITLGAEGCVMLSQTEPVPKHIPTE	259
Orangutan	SDVFCCNESEAEILTGLTVGSAAADAGEAALVLLKRGCOVVIITLGAEGCVVLSQTEPEPKHIPTE	257
Monkey	SDVFCCNESEAEILTGLTVNSAAADAGEAALVLLKRGCOVVIITLGAEGCVVLSQTEPEPKHIPTE	257
Mouse	SSIFCCNESEAEILTGHAVSDEPTTAGKAAMILLERGCQVVIITLGAEGCVVLSQTEPVPKHIPTE	258
Rat	STVFCCNESEAEILTGHAVNDEPTTAGTAALVLLERGCQVVIITLGAEGCVVLSQTEPVPKHIPTE	258
BMI-pig	KVKAVDTTGAGDSFVGALAFYLAYYENLSLEEMIKRSNFIAAVSVQATGTQSSYPYKDLPLLYLF	323
Human	KVKAVDTTGAGDSFVGALAFYLAYYENLSLEDMLNRSNFIAAVSVQAGTQSSYPYKDLPLPLTF	322
Cattle	KVKAVDTTGAGDSFVGALAFYLAYYSHLSLEEMIKRSNFIAAVSVQVTGTQSSYPYKDLPLDLF	325
Horse	KVKAVDTTGAGDSFVGALAFYLAYYENLSLEEMIKRSNFIAAVSVQAGTQSSYPYKDLPLHLF	324
Orangutan	KVKAVDTTGAGDSFVGALAFYLAYYENLSLEDMLKRSNFIAAVSVQAGTQSSYPYKDLPLPLTF	322
Monkey	KVKAVDTTGAGDSFVGALAFYLAYYENLSLEDMLKRSNFIAAVSVQAGTQSSYPYKDLPLPLTF	322
Mouse	AVKAVDTTGAGDSFVGALAFYLAYYENLSLEEMIKRSNFIAAVSVQATGTQSSYPYKDLPLALF	323
Rat	AVKAVDTTGAGDSFVGALAFYLAYYENLSLEEMIKRSNSIAAVSVQATGTQSSYPYKDLPLALF	323

**Figure 6.** The alignment of the protein encoded by the BMI *RBKS* and other seven kinds of *RBKS* from cattle (NP\_001178200), horse (XP\_001502121), orangutan (XP\_002812214), human (NP\_071411), monkey (XP\_001100564), rat (NP\_001102173) and mouse (EDL37382).



**Figure 7.** The phylogenetic tree for several kinds of *RBKS* protein from human, orangutan, monkey, BMI-pig, cattle, horse and mouse.



**Figure 8.** Tissue expression profile of BMI *RBKS* gene. The *18S* expression level is used for the internal control. M, DL2000 DNA marker; 1, lymph node; 2, midbrain; 3, ovary; 4, diencephalon; 5, cerebrum; 6, liver; 7, kidney; 8, spleen; 9, heart; 10, lung; 11, nerve fiber; 12, stomach; 13, small intestine; 14, large intestine; 15, pancreas; 16, skin; 17, muscle; 18, fat.

an endogenous reference for determination of targeted mRNA profiles. Result reveals that BMI *RBKS* gene was over-expressed in ovary and lung, moderately expressed in spleen, nerve fiber, large intestine and diencephalon, weakly expressed in heart, skin, muscle, small intestine, midbrain, kidney and fat and almost silent in lymph node, cerebrum, liver, stomach and pancreas (Figure 8).

## DISCUSSION

Comparative genomics determines the relationship of genome structure and function of different species. Researchers have learnt a great deal about the function of human genes by examining their counterparts in simpler model organisms such as the mouse and some results have revealed that virtually all (99%) of the protein-coding genes in humans align with homologues in mice, and over 80% are clear 1:1 orthologs (Hardison, 2003; Liu et al., 2008; Yu et al., 2010; Xi et al., 2011). This extensive conservation in protein-coding regions implied that the same protein-coding sequences may be expected in different mammals including pig. From the

isolation of swine *RBKS* gene, we can find that swine *RBKS* is highly homologous with *RBKS* of human, bovine and other mammals. This further validated that comparative genomics method is one useful tool to isolate the unknown genes especially the conserved coding region of genes for pig. From the alignment analyses for swine *RBKS* protein, we also found that BMI *RBKS* protein was not completely identical with human or other mammals. This implied that BMI *RBKS* will have some differences in functions with those of human, bovine and other mammals.

The phylogenetic tree analysis revealed that the BMI *RBKS* gene has a closer genetic relationship with the bovine and horse *RBKS*. Therefore, we can use bovine and horse as model organisms to study the pig *RBKS* gene or use pig as a model organism to study the bovine and horse *RBKS* gene. Most protein functions are regulated by phosphorylation/dephosphorylation, glycosylation/deglycosylation and BMI *RBKS* protein has several kinds of functional sites (such as phosphorylation sites, glycosylation sites, myristoylation sites and ribokinase protein domain, which suggest that *RBKS* protein plays important functional roles through these

sites and domain. MicroRNAs are small noncoding RNA. They play a role in gene expression regulation by inhibiting translation of their target mRNAs (Bartel et al., 2004; Zeng et al., 2003). Their target predictions showed that four *Sus scrofa* microRNAs (ssc-miR-217, ssc-miR-196b-3p, ssc-miR-1306-3p and ssc-miR-1306-3p) were found to have the corresponding target sites (313-ugcugcgacaggaacugcuucgau-336, 718-ugaagggugugugaugcuguc-738, 428-accaucagcagagccaaagu-447 and 48-cggcggcagugguggug-67) in the BMI *RBKS* coding sequence. Further investigation is needed to confirm whether corresponding miRNA molecules can regulate the *RBKS* gene expression in swine.

*RBKS* was discovered more than 50 years ago, and there has been much progress in understanding this important enzyme within the last decade, but the molecular mechanisms of its action and target remain a central unresolved problem for biochemists and pharmacologists. For this reason, the homology modeling of the BMI *RBKS* protein was carried out using human ribokinase (2fv7 Chain: B) as the template. Our results indicate that the conserved domain (*RBKS*, 17-323AA) of pig *RBKS* protein mainly exists in the form of  $\alpha\beta$  sandwich domain and  $\beta$ -barrel. It suggests that these domains may play a pivotal role in *RBKS* activity. This first pig structural model of *RBKS* proved to be useful reference in designing studies aimed at understanding how *RBKS* interacts with unknown protein partners for explaining its various functions.

In this study, we not only cloned the CDS sequences of the BMI *RBKS* gene but also conducted the sequence analysis and tissue expression profiles analysis. From the tissue expression profile analysis, it can be seen that the gene was obviously differentially expressed in various tissues. As the researchers did not study functions at protein levels, there might be many possible reasons for differential expression of this porcine gene. The suitable explanation for this under the current conditions is that the biological activities associated with the functions of the gene were required in a different extent in different tissues at the same time.

In summary, we firstly isolated BMI *RBKS* gene and performed necessary functional analysis and tissue expression profile analysis. Furthermore, several miRNAs were found to have the corresponding target sites in the coding sequence of BMI *RBKS* by theoretical prediction.

The cDNA clone, sequence information and function analyses of BMI *RBKS* gene will be extremely important in elucidating the essential physiological function of *RBKS* protein using BMI and other pigs as experimental animal models in the future.

## ACKNOWLEDGMENT

This study was supported by the National Natural Science Foundation of China (Grant no. 31160439).

## REFERENCES

- Anderson A, Cooper RA (1969). The significance of ribokinase for ribose utilization by *Escherichia coli*. *Biochim. Biophys. Acta*, 177: 163-165.
- Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116: 281-297.
- Bork P, Sander C, Valencia A (1993). Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* 2: 31-40.
- Crabbe JC, Metten P, Cameron AJ, Wahlsten D (2005). An analysis of the genetics of alcohol intoxication in inbred mice. *Neurosci. Biobehav. Rev.* 28: 785-802.
- Hardison RC (2003). Comparative genomics. *PLoS. Biol.* 1: e58.
- Liu GY, Gao SZ, Ge CR, Zhang X (2008). Molecular characterization of the encoding regions and tissue expression analyses for three novel porcine genes: *HNRPA1*, *YIPF5*, and *UB2D2*. *Mol. Biol. Rep.* 35: 519-526.
- Lopilato JE, Garwin JL, Emr SD, Silhavy TJ, Beckwith JR (1984). D-ribose metabolism in *Escherichia coli* K-12: genetics, regulation, and transport. *J. Bacteriol.* 158: 665-673.
- Moller S, Croning MD, Apweiler R (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*, 17: 646-653.
- Müller C, Zimmer H, Gross M, Gresser U, Brotsack I, Wehling M, Pliml W (1998). Effect of ribose on cardiac adenine nucleotides in a donor model for heart transplantation. *Eur. J. Med. Res.* 3: 554-558.
- Nakai K, Horton P (1999). PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24: 34-35.
- Petersen TN, Brunak S, von Heijne G, Nielsen H (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods*, 8: 785-786.
- Reva BA, Finkelstein AV, Skolnick J (1998). What is the probability of a chance prediction of a protein structure with an rmsd of 6 Å. *Fold. Des.* 3(2): 141-147.
- Salerno C, D'Eufemia P, Finocchiaro R, Celli M, Spalice A, Lannetti P, Crifò C, Giardini O (1999). Effect of D-ribose on purine synthesis and neurological symptoms in a patient with adenylosuccinase deficiency. *Biochim. Biophys. Acta*, 1453: 135-140.
- Sigrell JA, Cameron AD, Mowbray SL (1999). Induced fit on sugar binding activates ribokinase. *J. Mol. Biol.* 290: 1009-1018.
- Sigrell JA, Cameron AD, Jones TA, Mowbray SL (1998). Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: insights into a new family of kinase structures. *Structure*, 6: 183-193.
- Wu LF, Reizer A, Reizer J, Cai B, Tomich JM, Saier MH (1991). Nucleotide sequence of the *Rhodobacter capsulatus fruK* gene, which encodes fructose-1-phosphate kinase: Evidence for a kinase superfamily including both phosphofructokinases of *Escherichia coli*. *J. Bacteriol.* 173: 3117-3127.
- Xi DM, He YD, Sun YK, Gou X, Yang SL, Mao HM, Deng WD (2011). Molecular cloning, sequence identification and tissue expression profile of three novel genes *Sfxn1*, *Snai2* and *Cno* from Black-boned sheep (*Ovis aries*). *Mol. Biol. Rep.* 38: 1883-1887.
- Yu HM, Chen SN, Xi DM, He YD, Liu Q, Mao HM, Deng WD (2010). Molecular cloning, sequence characterization and tissue transcription profile analyses of two novel genes: *LCK* and *CDK2* from the Black-boned Sheep (*Ovis aries*). *Mol. Biol. Rep.* 37: 39-45.
- Yu P, Zhang L, Li SF, Li YP, Cheng JQ, Zeng YZ (2004). Screening and analysis of porcine endogenous retrovirus in Chinese Banna minipig inbred line. *Transplant Proc.* 36: 2485-2487.
- Zeng R, Zeng YZ (2005). Molecular cloning and characterization of SLA-DR genes in the 133-family of the Banna mini-pig inbred line. *Anim. Genet.* 36: 267-269.
- Zeng Y, Yi R, Cullen BR (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl. Acad. Sci. USA.* 100: 9779-9784.