GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Complementary DNA cloning, sequence analysis, and tissue transcription profile of a novel *U2AF2* gene from the Chinese Banna mini-pig inbred line

S.Y. Wang[1,2]*, J.L. Huo[1,2]*, Y.W. Miao[2]*, W.M. Cheng[1,2] and Y.Z. Zeng[1,2]

[1]Key Laboratory of Banna Mini-Pig Inbred Line of Yunnan Province, Kunming, Yunnan, China
[2]Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming, Yunnan, China

*These authors contributed equally to this study.
Corresponding author: Y.Z. Zeng
E-mail: bnwxzjjx@yahoo.com.cn

Genet. Mol. Res. 12 (2): 925-934 (2013)
Received July 27, 2012
Accepted January 15, 2013
Published April 2, 2013
DOI http://dx.doi.org/10.4238/2013.April.2.9

ABSTRACT. U2 small nuclear RNA auxiliary factor 2 (*U2AF2*) is an important gene for pre-messenger RNA splicing in higher eukaryotes. In this study, the Banna mini-pig inbred line (BMI) *U2AF2* coding sequence (CDS) was cloned, sequenced, and characterized. The *U2AF2* complete CDS was amplified using the reverse transcription-polymerase chain reaction (RT-PCR) technique based on the conserved sequence information of cattle and known highly homologous swine expressed sequence tags. This novel gene was deposited into the National Center for Biotechnology Information database (Accession No. JQ839267). Sequence analysis revealed that the BMI *U2AF2* coding sequence consisted of 1416 bp and encoded 471 amino acids with a molecular weight of 53.12 kDa. The protein sequence has high sequence homology with *U2AF65* of 6 species - *Homo sapiens* (100%), *Equus caballus* (100%), *Canis lupus* (100%), *Macaca mulatta* (99.8%), *Bos taurus*

(74.4%), and *Mus musculus* (74.4%). The phylogenetic tree analysis revealed that BMI *U2AF65* has a closer genetic relationship with *B. taurus U2AF65* than with *U2AF65* of *E. caballus*, *C. lupus*, *M. mulatta*, *H. sapiens*, and *M. musculus*. RT-PCR analysis showed that BMI *U2AF2* was most highly expressed in the brain; moderately expressed in the spleen, lung, muscle, and skin; and weakly expressed in the liver, kidney, and ovary. Its expression was nearly silent in the spinal cord, nerve fiber, heart, stomach, pancreas, and intestine. Three microRNA target sites were predicted in the CDS of BMI *U2AF2* messenger RNA. Our results establish a foundation for further insight into this swine gene.

Key words: Banna mini-pig inbred line pig; China; U2AF65; U2 small nuclear RNA auxiliary factor 2; Tissue expression analysis

INTRODUCTION

The U2 small nuclear RNA auxiliary factor 2 gene (*U2AF2*) encodes the U2AF65 protein, large subunits of U2 auxiliary factor (U2AF). U2AF is a heterodimer of two protein subunits, U2AF65 and U2AF35, which bind to the Py tract and the downstream AG dinucleotide, respectively (Zamore et al., 1992; Zorio and Blumenthal, 1999; Merendino et al., 1999; Wu et al., 1999). U2AF belongs to the splicing factor serine-arginine family, which is characterized by the presence of a ribonucleoprotein (RNP)-type RNA-binding motif and a carboxyl-terminal arginine-serine-rich (RS) domain that plays important roles in both the assembly of spliceosomes and the regulation of alternative splicing (Shen and Green, 2004; Das et al., 2007; Lin and Fu, 2007). U2AF recognizes consensus 3'-splice site sequences in pre-messenger RNA (mRNA) and coordinates the initial states of spliceosome assembly. The U2AF65 protein contains a sequence-specific RNA-binding region with three RNA recognition motifs (RRMs) and an RS domain necessary for splicing pre-mRNA (Zamore et al., 1992). The RRM-mediated interaction of U2B'' and U2A' is required for specific RNA binding, and the helical surface of the noncanonical RRM of U2AF35 interacts with the N-terminus of U2AF65 in the U2AF heterodimer (Price et al., 1998; Kielkopf et al., 2001). An RRM protein interface is required for the recognition of capped RNA by the CBP20 RRM in the cap-binding complex (Mazza et al., 2001). Stable association of the U2 small nuclear RNP requires an N-terminal, RS domain of U2AF65 (Valcarcel et al., 1996; Shen and Green, 2004), and RNP-unwindases such as the U2AF-associated-protein-56 kDa (UAP56) (Fleckner et al., 1997).

Although *U2AF2* is important in pre-mRNA splicing reactions, porcine *U2AF2* has not yet been reported. The Banna mini-pig inbred (BMI) line was developed by Yunnan Agricultural University in the 1980s, based on small-ear pigs in Xishuangbanna, Yunnan Province, China. After more than 30 years of highly full-sibling or parent-offspring inbreeding and strict selection for each generation, the BMI line has formed an inbred group with a clear genetic background, high homozygosity, and stable inheritance. The BMI line is considered an ideal model organism for biological studies (Yu et al., 2004; Crabbe et al., 2005; Zeng and Zeng, 2005).

The objective of this study was to clone the full-length *U2AF2* gene coding sequence from BMI according to the conserved sequence information of cattle or other mammals and highly homologous swine-expressed sequence tag information, conduct sequence function

analyses of established nucleotide sequences, and examine the expression in a range of BMI tissues. Collectively, these results lay the foundation for further studies of this porcine gene.

MATERIAL AND METHODS

Sample collection, RNA extraction, and complementary DNA (cDNA) synthesis

RNA was extracted from the brain, spinal cord, nerve fiber, heart, liver, spleen, lung, kidney, stomach, pancreas, intestine, ovary, muscle, and skin of female BMI pigs by using RNAiso Plus (TaKaRa, Dalian, China) and reverse transcribed with oligo-(dT)18 primer and a High Fidelity PrimeScript RT-PCR Kit (TaKaRa). The efficiency of reverse transcription was checked on 2% agarose gels stained with ethidium bromide.

Isolation of the BMI *U2AF2* gene

The GenBank *U2AF2* sequences for *Bos taurus* (accession No. NM_001075336) and the highly homologous pig-expressed sequence tag sequences FS666330, HX234528, and EW073881 were used to design a primer pair to amplify the complete coding sequence of *U2AF2* using the Primer Premier 5.0 software. The primers for the BMI *U2AF2* gene were 5'-TCAGCATGTCGGACTTCGACG-3' and 5'-CCGCCGCCTCTACCAGAAGT-3'. Reverse transcription-polymerase chain reaction (RT-PCR) was performed to isolate BMI *U2AF2* using the pooled cDNAs from the tissues listed above. The 25-µL reaction system contained the following: 2.0 µL 25 ng/µL cDNA, 2.0 µL 2.5 mM mixed deoxyribonucleotide triphosphates, 2.5 µL 2X GC buffer (TaKaRa), 0.5 µL 10 µM forward primer, 0.5 µL 10 µM reverse primer, 0.25 µL 5 U/L Taq DNA polymerase (TaKaRa), and 17.25 µL sterile water. The initial reaction conditions for PCR were as follows: denaturation at 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 60°C for 45 s, and 72°C for 2 min, and finally, an extension at 72°C for 10 min and then at 4°C to terminate the reaction. After PCR, the gene product was cloned into a pMD18-T vector and sequenced bidirectionally by the commercial fluorometric method. At least five independent clones were sequenced.

Bioinformatics analysis

Sequence analysis of BMI *U2AF2* was performed using the software available from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) and ExPaSy (http://www.expasy.org). The cDNA sequence was predicted using the GenScan software (http://genes.mit.edu/GENSCAN.html). We predicted protein theoretical molecular weight (Mw) and isoelectric point (pI), signal peptide, subcellular localization, and transmembrane topology using the Compute pI/Mw Tool (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. A web-based microRNA (miRNA)-predicting program was used to locate conserved potential miRNA targets (http://www.mirbase.org/). The protein Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST) and Conserved Domain Architecture Retrieval Tool were used to search for similar proteins and conserved

domains, respectively. The alignments of the nucleotide sequences and deduced amino acid sequences were computed using ClusterX, and the phylogenetic tree was computed using the ClustalX and Molecular Evolutionary Genetics Analysis 4.0 software with standard parameters. Secondary structures of deduced amino acid sequences were predicted with the self-optimized prediction method with alignment (http://npsa-pbil.ibcp.fr/).

Semi-quantitative RT-PCR

To characterize the *U2AF2* gene further, we conducted RT-PCR to determine its expression in 14 BMI pig tissues. We selected the housekeeping gene 18S ribosomal RNA (NR_002170) as a positive control to eliminate the effect of cDNA concentration. The control primers used were 5'-TCAGCATGTCGGACTTCGACG-3' and 5'-CCGCCGCCTCTACCAGAAGT-3'. The BMI *U2AF2* primers used to perform semi-quantitative RT-PCR for tissue expression profile analysis were the same as those used for the isolation RT-PCR described above. PCRs were optimized for a number of cycles to ensure product intensity within the linear phase of amplification.

RESULTS

Cloning and identification of BMI *U2AF2* cDNA

RT-PCR with pooled tissue cDNAs for BMI *U2AF2* yielded a PCR product of 1430 bp (Figure 1) that included a complete 1416-bp coding sequence and partial 5'-untranslated region (UTR) and 3'-UTRs.

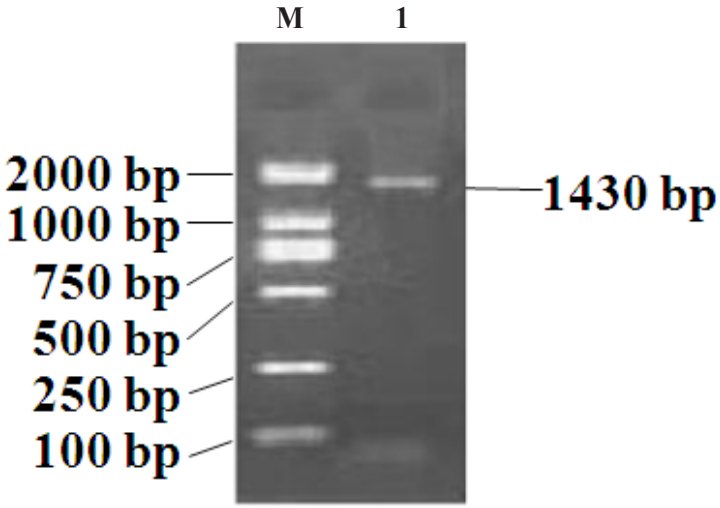

Figure 1. Reverse transcription-polymerase chain reaction results for the Banna mini-pig inbred U2 small nuclear RNA auxiliary factor 2 gene (*U2AF2*). *Lane M* = DL2000 DNA marker and *lane 1* = polymerase chain reaction product.

The BLAST software revealed that BMI *U2AF2* was not homologous to any known porcine genes and was therefore deposited into GenBank database under accession No.

JQ839267. The sequence prediction was carried out using the GenScan software, and the results showed that the 1416-bp coding sequence represented a single gene encoding 471 amino acids (Figure 2).

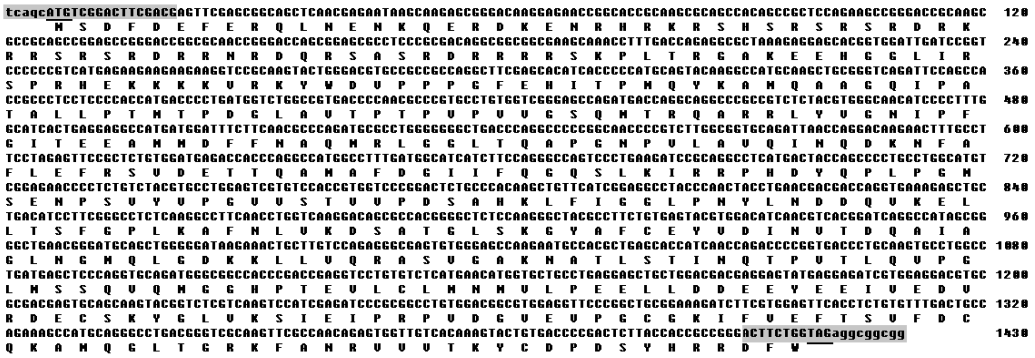

Figure 2. Complete complementary DNA and amino acid sequences of the protein encoded by *U2AF2* (GenBank accession No. JQ839267). ATG = start codon; TAG = stop codon; capital letters = complete coding DNA and amino acid sequences; gray highlighted nucleotide sequence = primers.

Physical and chemical characteristics of BMI *U2AF2*

The theoretical pI and Mw of BMI *U2AF2* were 9.103 and 53120.67 Da, respectively. SignalP revealed that U2AF65 did not contain an N-terminal signal peptide and that it is a non-secretory protein. Using a hidden Markov model algorithm, we predicted transmembrane topology with the TMHMM program (Moller et al., 2001), which showed that BMI *U2AF2* was a potential extrinsic protein. For subcellular localization analysis, the amino acid sequence was submitted to the PSORT II program, and the Reinhardt method showed that BMI *U2AF2* was localized in the nuclei with up to 87.0% probability.

Prediction and analysis of structures and conserved domains of BMI *U2AF2*

Proteins often contain several domains, each of which has their own evolutionary origins and functions. The Conserved Domain Architecture Retrieval Tool of BLAST indicated that BMI U2AF65 contained three RRM-specific hits that had RNA/DNA-binding sites and RRM dimerization sites (Figure 3). Then, putative protein was analyzed using the PROSITE software (http://expasy.org/prosite/). Seven types of sites were found: casein kinase II phosphorylation (2-SdfD-5, 204-SvdE-207, and 434-SvfD-437), adenosine-3',5'-cyclic monophosphate- and cyclic guanosine monophosphate-dependent protein kinase phosphorylation (25-RKrS-28, 38-KRrS-41, and 59-RRrS-62), protein kinase C phosphorylation (79-SpR-81, 223-SlK-225, and 446-TgR-448), N-myristoylation (114-GQipAT-119, 141-GSqmTR-146, 248-GVvsTV-253, 264-GGlpNY-269, 297-GLskGY-302, 322-GMqlGD-327, 338-GAknAT-343, 358-GLmsSQ-363, 367-GGhpTE-372, and 406-GLvkSI-411), N-glycosylation (311-NVTD-314 and 341-NATL344), amidation (446-tGRK-449), and tyrosine kinase phos-

phorylation (458-Kyc.DpdsY-465). The prediction of secondary structure with the self-optimized prediction method with alignment indicated that the deduced BMI U2AF65 contained 135 alpha helices, 61 extended strands, 23 beta turns, and 252 random coils (Figure 4).

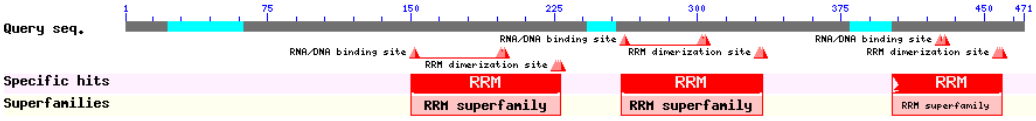

Figure 3. Putative domains of the protein encoded by BMI U2AF65. RRM = recognition motif.

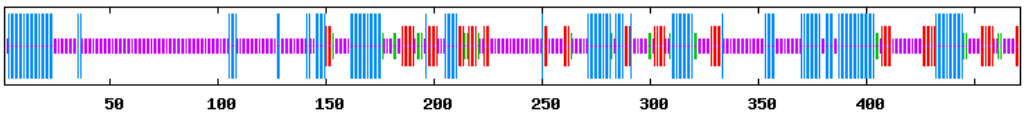

Figure 4. Secondary structure of the BMI U2AF65 protein predicted with the self-optimized prediction method with alignment. Helices, extended strands, beta turns, and random coils are indicated with the longest, second longest, second shortest, and shortest vertical lines, respectively.

Location of potential miRNA targets

miRNAs are non-coding single-stranded RNA molecules of 17-24 nucleotides that can regulate gene expression by binding to the coding region of target mRNAs (Zeng and Cullen, 2003; Bartel, 2004). We used the web-based miRNA-predicting program miRBase (http://www.mirbase.org/) to locate conserved potential miRNA targets. The results showed that 3 *Sus scrofa* mRNAs (ssc-miR-376a, ssc-miR-652, and ssc-miR-487b) had the predicted target sites (4-gauucuccuucuaugaguac-23, 5-auggcacccucuccuaggg-23, and 3-gaacaggacagggauaacc-21, respectively) in the BMI *U2AF2* sequence.

Analyses of sequence identity and evolutionary relationships of BMI *U2AF2*

The deduced protein sequence of BMI *U2AF2* was submitted to generate BLAST reciprocal best hits, and similarity comparison revealed that the BMI U2AF65 protein had high sequence homology with the U2AF65 protein of 6 other species: *Homo sapiens* (100%), *Equus caballus* (100%), *Canis lupus* (100%), *Macaca mulatta* (99.8%), *B. taurus* (74.4%), and *Mus musculus* (74.4%; Figure 5). To evaluate the evolutionary relationships of BMI U2AF65 with other species, we constructed a phylogenetic tree based on the U2AF65 amino acid sequences using DNAStar, Cluster, and Molecular Evolutionary Genetics Analysis softwares. The phylogenetic tree analysis revealed that BMI U2AF65 had a closer genetic relationship with the U2AF65 of *B. taurus* than with those of *H. sapiens*, *E. caballus*, *C. lupus*, *M. mulatta*, and *M. musculus* (Figure 6).

mRNA tissue-specific expression profile

To check the relative expression levels of *U2AF2* mRNA in various porcine tissues, we performed semi-quantitative RT-PCR analyses in 14 BMI tissues. The continuously expressed

gene, 18S, was used as an endogenous reference to determine targeted mRNA profiles. The results revealed that BMI *U2AF2* was expressed most highly in the brain; moderately expressed in the spleen, lung, muscle, and skin; weakly expressed in the liver, kidney, and ovary; and it was nearly silent in the spinal cord, nerve fiber, heart, stomach, pancreas, and intestine (Figure 7).

```
BMI pig         MSDFDEFERQ LNENKQERDK ENRHRKRSHS RSRSRDRKRR SRSRDRRNRD QRSASRDRRR RSKPLTRGAK EEHGGLIRSP RHEKKKKVRK YWDVP [ 95]
Homo sapiens    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[ 95]
Mus musculus    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[ 95]
Bos taurus      .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[ 95]
Equus caballus  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[ 95]
Canis lupus     .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[ 95]
Macaca mulatta  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[ 95]

BMI pig         PPGFEHITPM QYKAMQAAGQ IPATALLPTM TPDGLAVTPT PVPVVGSQMT RQARRLYVGN IPFGITEEAM MDFFNAQMRL GGLTQAPGNP VLAVQ[190]
Homo sapiens    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[190]
Mus musculus    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[190]
Bos taurus      .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[190]
Equus caballus  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[190]
Canis lupus     .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[190]
Macaca mulatta  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[190]

BMI pig         INQDKNFAFL EFRSVDETTQ AMAFDGIIFQ GQSLKIRRPH DYQPLPGMSE NPSVYVPGVV STVVPDSAHK LFIGGLPNYL NDDQVKELLT SFGPL[285]
Homo sapiens    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[285]
Mus musculus    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[285]
Bos taurus      .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[285]
Equus caballus  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[285]
Canis lupus     .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[285]
Macaca mulatta  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[285]

BMI pig         KAFNLVKDSA TGLSKGYAFC EYVDINVTDQ AIAGLNGMQL GDKKLLVQRA SVGAKNAT-- --LSTINQTP VTLQVPGLMS SQVQMGGHPT EVLCL [380]
Homo sapiens    .......... .......... .......... .......... .......... .........-- --......... .......... .......... ..... [380]
Mus musculus    .......... .......... .......... .......... .......... .........LV SLP....... .......... .......... ..... [380]
Bos taurus      .......... .......... .......... .......... .......... .........LV SP........ .......... .......... ..... [380]
Equus caballus  .......... .......... .......... .......... .......... .........-- --......... .......... .......... ..... [380]
Canis lupus     .......... .......... .......... .......... .......... .........-- --......... .......... .......... ..... [380]
Macaca mulatta  .......... .......... .......... .......... .......... .........-- --......... .......... .......... ..... [380]

BMI pig         MNMVLPEELL DDEEYEEIVE DVRDECSKYG LVKSIEIPRP VDGVEVPGCG KIFVEFTSVF DCQKAMQGLT GRKFANRVVV TKYCDPDSYH RRDFW [475]
Homo sapiens    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[475]
Mus musculus    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[475]
Bos taurus      .......... .......... ......G... .......... .......... .......... .......... .......... .......... .....[475]
Equus caballus  .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[475]
Canis lupus     .......... .......... .......... .......... .......... .......... .......... .......... .......... .....[475]
Macaca mulatta  .......... .......... E......... .......... .......... .......... .......... .......... .......... .....[475]
```

Figure 5. Alignment of the protein encoded by BMI U2AF65 and 6 other kinds of U2AF65 from *Homo sapiens* (NP_001012496), *Equus caballus* (XP_001496159), *Canis lupus* (XP_003638845), *Macaca mulatta* (AFI33276), *Bos taurus* (NP_001068804), and *Mus musculus* (NP_001192160).

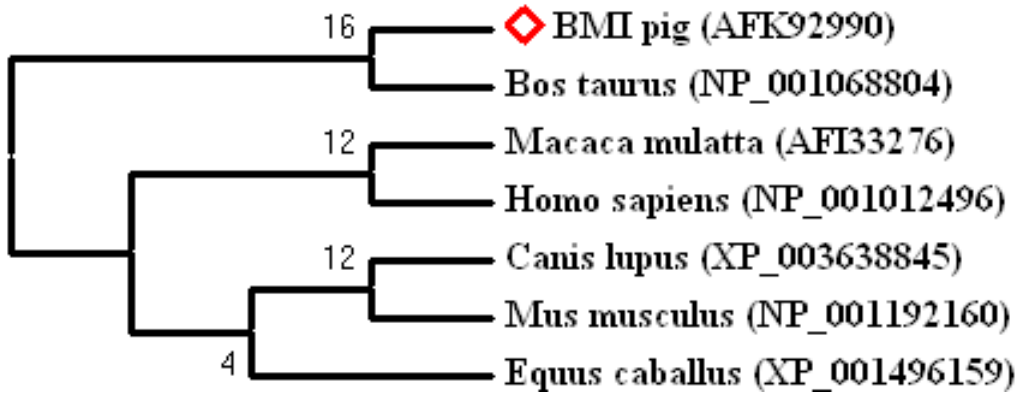

Figure 6. Phylogenetic tree for several kinds of U2AF65 from the BMI pig, *Homo sapiens*, *Equus caballus*, *Canis lupus*, *Macaca mulatta*, *Bos taurus*, and *Mus musculus*.

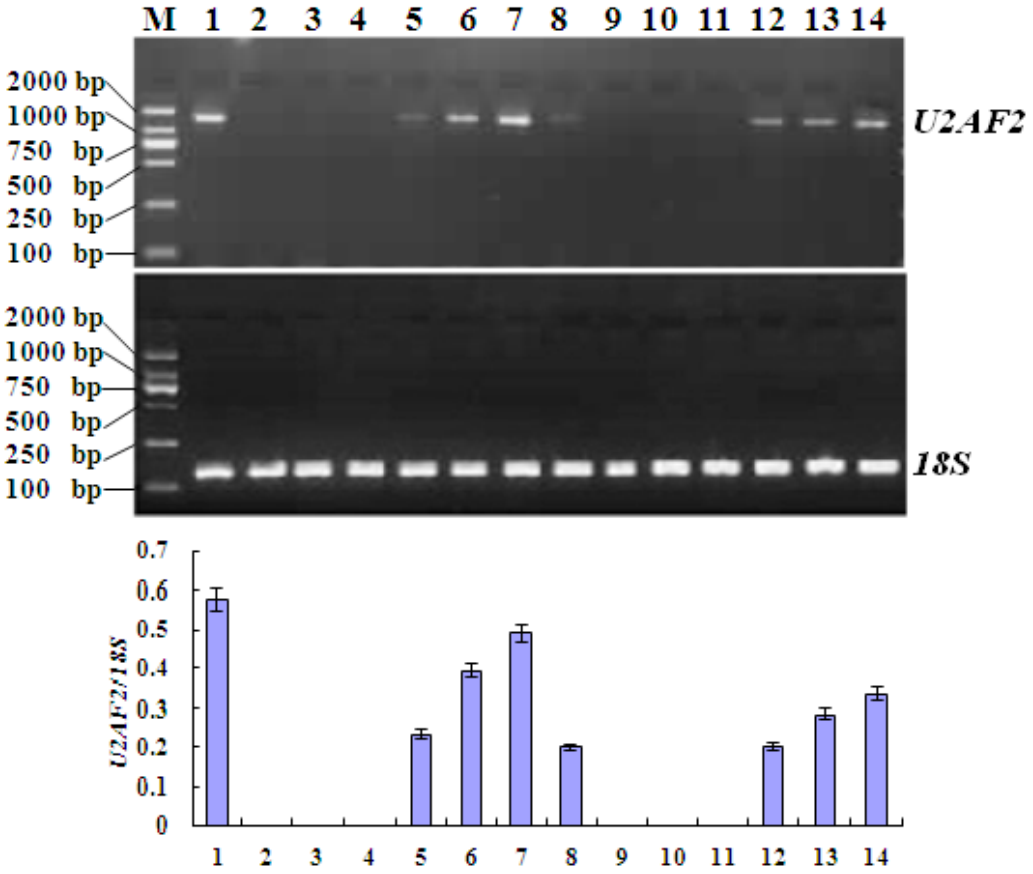

Figure 7. Tissue transcription profile of BMI *U2AF2*. *Lane M* = DL2000 DNA marker; *lane 1* = brain; *lane 2* = spinal cord; *lane 3* = nerve fiber; *lane 4* = heart; *lane 5* = liver; *lane 6* = spleen; *lane 7* = lung; *lane 8* = kidney; *lane 9* = stomach; *lane 10* = pancreas; *lane 11* = intestine; *lane 12* = ovary; *lane 13* = muscle; *lane 14* = skin. The 18S ribosomal RNA expression level was used as the internal control.

DISCUSSION

U2AF2 encoded a U2AF65 protein that contains a sequence-specific RNA-binding region with three RNA RRMs and an RS domain necessary for splicing pre-mRNA (Zamore et al., 1992). This study provides a molecular basis for an associate analysis of DNA and the amino acid sequence of pig *U2AF2*. The pig U2AF65 protein, consistent with other U2AF65 proteins, has no signal peptide or transmembrane regions and is localized in the nuclei with up to 87.0% probability.

Stable association of the U2 small nuclear RNP requires an N-terminal, RS domain of U2AF65 (Valcarcel et al., 1996; Shen and Green, 2004), and RNP-unwindases such as the UAP56 (Fleckner et al., 1997). However, the molecular mechanism by which actors bind to the U2AF65 protein remains a central unresolved problem for biochemists and pharmacologists owing to the absence of clear structural data. Our results indicate that the U2AF65 protein contains three RRM-specific hits that have RNA/DNA-binding sites and RRM dimerization sites.

The porcine U2AF65 protein may function through these sites and domains. This should prove useful for designing studies aimed at understanding how U2AF65 interacts with U2AF35 and possibly with unknown protein partners.

Phylogenetic tree analysis revealed that the pig U2AF65 protein has the closest genetic relationship with the *B. taurus* U2AF65 protein. Similarity comparison revealed that BMI U2AF65 has high homology with the U2AF65 of 6 other species: *H. sapiens* (100%), *E. caballus* (100%), *C. lupus* (100%), *M. mulatta* (99.8%), *B. taurus* (74.4%), and *M. musculus* (74.4%), implying that pig may be a preferable animal model for the study of *U2AF2* function.

miRNAs are small non-coding RNAs. They play a role in gene expression regulation by inhibiting the translation of their target mRNAs (Zeng and Cullen, 2003; Bartel, 2004). Target predictions showed that 3 *S. scrofa* miRNAs (ssc-miR-376a, ssc-miR-652, and ssc-miR-487b) have corresponding target sites in the coding sequence of the porcine *U2AF2*. Further investigation is needed to confirm whether corresponding miRNA molecules can regulate *U2AF2* expression in pigs.

The tissue transcription profile analysis in our experiment showed that *U2AF2* was not obviously differentially expressed in some tissues. The *U2AF2* gene expressed most highly in the brain, whereas it was nearly silent in the spinal cord, nerve fiber, heart, stomach, pancreas, and intestine. Although we did not study the functions or protein levels, the results of our study indicate several possible explanations for the differential expression of porcine *U2AF2*. The most suitable explanation is that the biological activities associated with the functions of *U2AF2* were presented diversely in various tissues.

In conclusion, we first isolated porcine *U2AF2* and performed sequence and tissue transcription profile analyses. In addition, theoretical prediction revealed that several miRNAs had corresponding target sites in the coding sequence of porcine *U2AF2*. These findings established the primary foundation for further insight into the structure and function of *U2AF2*.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (#31160439).

REFERENCES

Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.

Crabbe JC, Metten P, Cameron AJ and Wahlsten D (2005). An analysis of the genetics of alcohol intoxication in inbred mice. *Neurosci. Biobehav. Rev.* 28: 785-802.

Das R, Yu J, Zhang Z, Gygi MP, et al. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* 26: 867-881.

Fleckner J, Zhang M, Valcarcel J and Green MR (1997). U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev.* 11: 1864-1872.

Kielkopf CL, Rodionova NA, Green MR and Burley SK (2001). A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell* 106: 595-605.

Lin S and Fu XD (2007). SR proteins and related factors in alternative splicing. *Adv. Exp. Med. Biol.* 623: 107-122.

Mazza C, Ohno M, Segref A, Mattaj IW, et al. (2001). Crystal structure of the human nuclear cap binding complex. *Mol. Cell* 8: 383-396.

Merendino L, Guth S, Bilbao D, Martinez C, et al. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* 402: 838-841.

Moller S, Croning MD and Apweiler R (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17: 646-653.

Price SR, Evans PR and Nagai K (1998). Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a

fragment of U2 small nuclear RNA. *Nature* 394: 645-650.
Shen H and Green MR (2004). A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Mol. Cell* 16: 363-373.
Valcarcel J, Gaur RK, Singh R and Green MR (1996). Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. *Science* 273: 1706-1709.
Wu S, Romfo CM, Nilsen TW and Green MR (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* 402: 832-835.
Yu P, Zhang L, Li S, Li Y, et al. (2004). Screening and analysis of porcine endogenous retrovirus in Chinese Banna minipig inbred line. *Transplant Proc.* 36: 2485-2487.
Zamore PD, Patton JG and Green MR (1992). Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* 355: 609-614.
Zeng R and 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 and Cullen BR (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* 100: 9779-9784.
Zorio DA and Blumenthal T (1999). Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* 402: 835-838.