Turkish Journal of Biology

Volume 39 | Number 4 Article 1

1-1-2015

cDNA cloning, molecular characterization, and expression analyses of two novel porcine ARRDC genes-ARRDC1 and ARRDC

HAILONG HUO

PEI WANG

SHUYAN WANG

YONGWANG MIAO

HONGJIANG WEI

See next page for additional authors

cDNA cloning, molecular characterization, and expression analyses of two novel porcine ARRDC genes-ARRDC1 and ARRDC

Authors

HAILONG HUO, PEI WANG, SHUYAN WANG, YONGWANG MIAO, HONGJIANG WEI, YONGYUN ZHANG, QIAOLING ZHANG, FUQUAN LI, LIXIAN LIU, RUI WANG, WEIZHEN LI, YUE ZHAO, JINLONG HUO, and HENG XIAO

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2015) 39: 523-531

doi:10.3906/biy-1404-101

cDNA cloning, molecular characterization, and expression analyses of two novel porcine *ARRDC* genes—*ARRDC1* and *ARRDC5*

Hailong HUO[1,2,*], Pei WANG[1,3,4,*], Shuyan WANG[3,4], Yongwang MIAO[4], Hongjiang WEI[4], Yongyun ZHANG[5], Qiaoling ZHANG[6], Fuquan LI[7], Lixian LIU[2,4], Rui WANG[2], Weizhen LI[4], Yue ZHAO[2], Jinlong HUO[3,4], Heng XIAO[1,*,]**

[1]Faculty of Life Science, Yunnan University, Kunming, Yunnan, P.R. China
[2]Department of Husbandry and Veterinary, Yunnan Vocational and Technical College of Agriculture, Kunming, Yunnan, P.R. China
[3]Key Laboratory of Banna Mini-Pig Inbred Line of Yunnan Province, Kunming, Yunnan, P.R. China
[4]Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming, Yunnan, P.R. China
[5]Teaching Demonstration Center of the Basic Experiments of Agricultural Majors, Yunnan Agricultural University, Kunming, Yunnan, P.R. China
[6]College of Animal Science and Veterinary Medicine, Jilin University, Changchun, Jilin, P.R. China
[7]Department of Biotechnology, Neijiang Vocational and Technical College, Neijiang, Sichuan, P.R. China

Received: 26.04.2014 • **Accepted/Published Online:** 29.10.2014 • **Printed:** 30.07.2015

Abstract: Arrestin-domain–containing proteins belong to the arrestin family, members of which are important for regulating signal transduction within cells. In this study, two *ARRDC* genes (*ARRDC1* and *ARRDC5*) of Diannan small-ear pigs were amplified by reverse transcription polymerase chain reaction and cloned for the first time (accession nos.: KF589202 and KF589203). The full-length coding sequence of *ARRDC1* was 1284 nucleotides in length (molecular weight, 45.6 kDa; pI, 7.61) and encoded 427 amino acids. The *ARRDC5* CDS was 1014 nucleotides in length (molecular weight, 37.5 kDa; pI, 7.08) and encoded 337 amino acids. Sequence homology analysis revealed that the amino acid sequences of Diannan small-ear pig ARRDC1 had 94.1%, 93.9%, 93.4%, 92.3%, 89.4%, 88.3%, and 84.7% identity with those of *Bos taurus*, *Tursiops truncatus*, *Felis catus*, *Canis lupus*, *Capra hircus*, *Homo sapiens*, and *Mus musculus*, respectively; Diannan small-ear pig ARRDC5 had 84.9%, 86.9%, 81.2%, 86.9%, 84.3%, 81.2%, and 71.2% identity with the same species, respectively. The phylogenetic analyses based on the amino acid sequences of ARRDC1 and ARRDC5 from 7 species showed that swine have close genetic relationships with *Tursiops truncatus*. Tissue expression analyses by quantitative PCR revealed that the mRNA expression levels of *ARRDC1* and *ARRDC5* were different in various tissues of the Diannan small-ear pig. *ARRDC1* was highly expressed in the testis, thyroid, submandibular gland, and kidney, while *ARRDC5* was only expressed in the testis, thyroid, and thymus. This study will aid investigations regarding the role of porcine *ARRDC1* and *ARRDC5* in regulating signal transduction within cells.

Key words: Diannan small-ear pig, arrestin domain containing 1, arrestin domain containing 5, arrestin family, tissue expression profile

1. Introduction

The arrestins are multifunctional adaptor proteins best known for their role in regulating G-protein–coupled receptor signaling. G-protein–coupled receptors are a large family of signaling molecules that respond to a wide variety of extracellular stimuli. The arrestins were first discovered as a part of a conserved two-step mechanism for regulating the activity of G-protein–coupled receptors (GPCRs) (Wilden et al., 1986; Lohse et al., 1990; Gurevich and Gurevich, 2006). To prevent the activation of heterotrimeric G proteins by GPCRs, G-protein–coupled receptor kinases (GRKs) are phosphorylated by a class of serine/threonine kinases. Then the phosphorylated GRK-activated receptor binds to arrestin, blocks G-protein–mediated signaling, and targets receptors for internalization. In addition to GPCRs, arrestins bind to other classes of cell surface receptors and a variety of other signaling proteins (Gurevich and Gurevich, 2004).

The arrestins can be subdivided into alpha-arrestins (aArrs), beta-arrestins (bArrs), and visual arrestins (vArrs). Mammals have six aArrs: arrestin domain-containing 1 through 5 (*ARRDC1–5*) and thioredoxin-interacting protein or vitamin D-upregulated protein 1 (*TXNIP* or *VDUP1*). Different arrestins, including visual arrestin

* These authors contributed equally to this study.
** Correspondence: xiaoheng@ynu.edu.cn

(or Arrestin 1), beta-arrestin 1 (or Arrestin 2), and beta-arrestin 2 (or Arrestin 3), can reduce the activity of their target GPCRs in several ways. The protein architecture of the ancestral Vps26/aArr has been preserved in the bArrs, which have two homologous arrestin domains (Arr-N and C) separated by a small hinge and followed by a C-terminal tail domain, showing minimal secondary structure (Alvarez, 2008). One mammalian aArrs mediates ubiquitylation and lysosomal trafficking of an activated plasma membrane receptor (Léon and Haguenauer-Tsapis, 2009). This alpha-arrestin is also involved in endocytic trafficking and can combine with beta-arrestin to maintain the optimal complement and function of the cell surface proteins according to the cellular physiological context and external signals (Shea et al., 2012).

Arrestin-domain–containing protein 1 (ARRDC1) exists as multiple alternatively spliced isoforms with ARRDC2. ARRDC1 is able to form the ARR-1-MPZ-1-DAF-18 complex with the phosphatase and tensin homolog (PTEN) ortholog DAF-18, which can regulate IGF-1R signaling and longevity (Palmitessa and Benovic, 2010). It is packaged into secreted microvesicles, and two highly conserved PPxY motifs at the C-terminal domain of ARRDC1 have been implicated in ubiquitination and ESCRT-mediated, PPxY-dependent retroviral budding (Hurley and Emr, 2006; Rauch and Martin-Serrano, 2011). Ubiquitin modification of ARRDC1 may be required for ARMM release (Nabhan et al., 2012). ARRDC1 can also act as a negative regulator of Notch signaling by heterodimerization with β-arrestin to promote the degradation of nonactivated Notch receptor (Puca et al., 2013).

The function of ARRDC5 is uncertain. *ARRDC2* and *ARRDC5* genes map to human chromosome 19, which consists of over 63 million bases, contains approximately 1400 genes, and is recognized for having the greatest gene density of all the human chromosomes. The *ARRDC1* gene is located on chromosome 9. Hereditary hemorrhagic telangiectasia, which is characterized by harmful vascular defects, and familial dysautonomia are both associated with chromosome 9.

The unprecedented degree of endocytic membrane trafficking regulation via mammalian GPCRs is due to their specificity and plasticity. These features have fundamental implications for GPCR pharmacology and suggest new mechanisms that could be exploited in GPCR-directed pharmacotherapy (Hanyaloglu and von Zastrow, 2008). Studies of the functional mechanisms of arrestin proteins, which are regulators of GPCRs, are of considerable theoretical significance.

The Diannan small-ear pig is a typical local breed in the southern areas of Yunnan Province, China. Pigs are of great practical value in the field of life science research because they are very similar to humans in anatomy, physiology, and the mechanism of disease. Miniature pig characteristics include inherent small size, early sexual maturity, rapid breeding, and ease of management (Yu et al., 2003). It is an ideal model organism for biological studies. Many different types of miniature pigs have been bred in China, and the Diann small-ear pig is one type of miniature pig used for science research.

Although arrestin-domain–containing proteins are essential for the regulation of transmembrane signal transduction, the sequences of porcine *ARRDC1* and *ARRDC5* genes have not yet been reported. Thus, we isolated the coding sequences of the porcine *ARRDC1* and *ARRDC5* genes, conducted bioinformatics analysis, and examined the gene expression patterns in 30 tissues by qPCR for the first time. This study will aid investigations regarding the role of porcine *ARRDC1* and *ARRDC5* genes in regulating signal transduction proteins and pathways.

2. Materials and methods

2.1. Sample collection, RNA extraction, and cDNA synthesis

Fresh tissue samples were obtained from ten male pigs. Thirty types of tissues, including brain, cerebellum, hypothalamus, brainstem, spinal cord, pituitary, heart, liver, spleen, lung, kidney, skin, muscle, duodenum, jejunum, ileum, colon, cecum, rectum, stomach, pancreas, esophagus, lymph nodes, testis, epididymis, submandibular gland, thyroid, adrenal gland, sublingual gland, and thymus, were immediately dissected and frozen in liquid nitrogen until RNA extraction could be performed. All experimental procedures were approved by the Yunnan Agricultural University Committee of Laboratory Animal Care. Total RNA was extracted using the RNAiso Plus kit (TaKaRa, Dalian, China) according to the manufacturer's instructions. To limit genomic DNA contamination, total RNA was digested with RNase-free DNase I (TaKaRa, Dalian, China). Four micrograms of RNA were reverse-transcribed with oligo $(dT)_{18}$ primers and M-MLV reverse transcriptase (Invitrogen, USA). Electrophoresis on 2% agarose gel was used to analyze the efficiency of reverse transcription. The concentrations of cDNA samples obtained from different types of tissues were also measured using a UV-Vis spectrophotometer (NanoDrop 2000).

2.2. Isolation of the porcine *ARRDC1* and *ARRDC5* genes

The cattle *ARRDC1* sequence (accession no.: NM_001080365) and the highly homologous pig-expressed sequence tag sequences (accession nos.: EW103216 and BQ604520) were used to design a primer pair using Primer Premier 5.0 software. The primers were 5′-ATGGGGCGGGTRCAGCTYTTCGAG-3′(forward) and 5′-CCCGTCCAGAGAAGGYGRCAT-

3′(reverse). Similarly, the primers for pig *ARRDC5* gene isolation were designed based on the *ARRDC5* expressed sequence tag sequences for pigs (accession nos.: CX061584 and BI343733). The primers were 5'-CATGTCTGTGGTGAAGTCGAT-3' (forward) and 5'-GAGTCTGATA-AAGCTTTTAATA-3' (reverse).

The cDNA obtained from the tissues was used to isolate the swine genes *ARRDC1* and *ARRDC5*, which were detected using RT-PCR. The 25-μL reaction mixture comprised 12.5 μL of 2X GC buffer I (TaKaRa), 2.0 μL of 2.5 mM mixed deoxyribonucleotide triphosphates (dNTP), 1.0 μL of 50 ng/μL cDNA, 0.5 μL of 10 μM forward primer, 0.5 μL of 10 μM reverse primer, 0.25 μL of 5 U/μL Ex Taq DNA polymerase (TaKaRa), and 8.25 μL of sterile water. The PCR program was initiated with denaturation at 95 °C for 2 min followed by 34 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 2 min. Next, a 72 °C extension step was performed for 10 min, and finally the temperature was set to 4 °C to terminate the reaction. PCR amplification was detected by 2% agarose gel electrophoresis, and PCR products were sequenced using a commercial fluorometric method. Then the products were cloned into the pMD18-T vector to obtain independent clones. Five independent clones were selected for sequencing.

2.3. Bioinformatics analysis

The swine *ARRDC1* and *ARRDC5* cDNA sequences were compared with the *Sus scrofa* genomic database (http://www.ncbi.nlm.nih.gov/blast) using BLAST analysis. The nucleotide and protein sequences of *ARRDC1* and *ARRDC5* genes were examined, edited, and predicted using DNASTAR software. Sequence alignments were performed using online software from the NCBI. The complete coding sequences of the porcine *ARRDC1* and *ARRDC5* genes have been deposited in the NCBI database. The genome BLAST database was used to map genes to chromosomes. The molecular weight (Mw) and theoretical isoelectric point (pI) were calculated using Compute Mw/pI (http://web.expasy.org/compute_pi/). The signal peptide was predicted using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/). Protein transmembrane topology analysis was conducted using TMHMM-2.0 software (http://www.cbs.dtu.dk/services/TMHMM-2.0/; Moller et al., 2001). The subcellular localization of the proteins was determined with PSORT (http://psort.hgc.jp/).

The predicted phosphorylation and glycosylation sites were identified using the NetPhos 2.0 server (http://cbs.dtu.dk/services). The software miRBase (http://mirbase.org/) was used to identify and analyze potential microRNA (miRNA) target sequences involved in the regulation of *ARRDC1* and *ARRDC5* gene expression.

The secondary structures of the deduced amino acid sequences were predicted using SOPMA (http://npsa-pbil.ibcp.fr/). Prediction and analysis of the conserved domains of ARRDC1 and ARRDC5 were conducted using the Conserved Domain Architecture Retrieval Tool of BLAST, accessed via the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST).

Homology analysis based on the amino acid sequences of some species was conducted using MegAlign software. ARRDC1 and ARRDC5 amino acid sequence alignments were constructed using MegAlign, CLUSTAL X 2.0, BioEdit, and MEGA 5 software. The neighbor-joining phylogenetic tree was constructed based on the ARRDC1 and ARRDC5 amino acid sequences using the MEGA 5 program (Kumar et al., 2008). The statistical significance of groups within phylogenetic trees was evaluated using the bootstrap method with 10,000 replications (Reza, 2013).

2.4. Expression profile analysis by RT-PCR

We conducted qPCR to determine the *ARRDC1* and *ARRDC5* mRNA expression in 30 porcine tissues. Considering that *18S* is stably expressed in most body tissues, we selected the housekeeping gene *18S* (AY265350) as the endogenous control. The primers used for testing the control gene were 5′-CTAAGGGCATCACAGACC-3′ (forward) and 5′-TTCCGATAACGAACGAGA-3′ (reverse). The porcine *ARRDC1* primer sets were 5′-CACCATTGACACGCCACGTTT-3′ (forward and 5′-TGGAGGCCACATTGGGTTGCT-3′ (reverse), and the primer sets of *ARRDC5* were 5′-TGGAGGCTGAGAAGAAGAT-3′ (forward and 5′-CGTAGAGGGCAAAGATGA-3′(reverse).

Relative transcript quantification was performed using standard curves generated for the *ARRDC1*, *ARRDC5*, and *18S* genes from five-fold serial dilutions of cDNA. In this experiment, the efficiencies of the *ARRDC1*, *ARRDC5*, and *18S* primers were all in the ideal range (from 90% to 105%) in each sample. The 20 μL reaction included 10 μL of Power SYBR Green PCR Master Mix, 8 μL of sterile water, 1 μL of cDNA, 0.5 μL of 10 μM forward primer, and 0.5 μL of 10 μM reverse primer. Amplification was performed with the default settings, which were as follows: denaturation at 95 °C for 2 min followed by 40 cycles of 94 °C for 10 s and 60 °C for 45 s. The amplification was followed by a melting curve stage. Optical data were collected at the end of each extension step, and the relative expression of PCR products was determined using the $2^{-\Delta\Delta Ct}$ method (Wu et al., 2014).

3. Results

3.1. Sequence analysis of *ARRDC1* and *ARRDC5*

The *ARRDC1* and *ARRDC5* PCR products amplified from different swine tissue cDNAs were 1309 and 1031 bp, respectively (Figure 1). Sequence alignment revealed that the *ARRDC1* and *ARRDC5* gene sequences obtained in this study were not homologous to any of the known swine

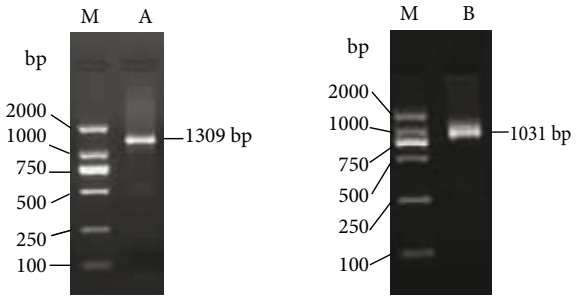

Figure 1. PCR results for swine *ARRDC* genes. M, DL2000 DNA marker; A, PCR product for the swine *ARRDC1* gene; B, PCR product for the swine *ARRDC5* gene.

genes. The CDSs for the *ARRDC1* and *ARRDC5* genes were 1284 bp and 1014 bp, which encoded 427 and 337 amino acids, respectively. The sequences of the porcine *ARRDC1* and *ARRDC5* genes have been deposited in the NCBI database and were assigned nucleotide accession numbers KF589202 and KF589203. The amino acid accession numbers for the two proteins are AHB08948 and AHB08949, respectively. The CDSs and their deduced amino acid sequences are presented in Figures 2A and 2B.

3.2. Bioinformatics analysis

Genome BLAST analysis indicated that the swine *ARRDC1* gene was mapped to chromosome 1, the *ARRDC5* gene was mapped to chromosome 2, and the cDNA sequence was collinear with the genomic sequence. The pIs of porcine ARRDC1 and ARRDC5 were 7.61 and 7.08, respectively. Their respective molecular weights were 45.6 kDa and 37.5 kDa. Protein signal peptide prediction and transmembrane topology analysis indicated that neither ARRDC1 nor ARRDC5 had an N-terminal signal peptide sequence or transmembrane domain. Cytoplasmic/nuclear localization analysis predicted that ARRDC1 was likely to function in the nucleus with a reliability of 52.2%, and ARRDC5 was likely to function in the cytoplasm with 52.2% reliability. The NetPhos 2.0 server was used to predict the phosphorylation and glycosylation sites in the ARRDC1 and ARRDC5 swine proteins. ARRDC1 has

```
ATGGGGCGGGTACAGCTTTTCGAGGTCTGCCTGAGCCACGGCCGCGTCGTCTACAGCCCTGGGGAGCCGCTGGCGGGGGC [80]
 M  G  R  V  Q  L  F  E  V  C  L  S  H  G  R  V  V  Y  S  P  G  E  P  L  A  G  A
CGTGCGCGTGCGCCTGGCGGCGCCGCTGCCCTTCCGAGCCATCCGGGTGACCTGTACAGGGTCCTGCAGGGTCTCCAACA [160]
  V  R  V  R  L  A  A  P  L  P  F  R  A  I  R  V  T  C  T  G  S  C  R  V  S  N
AGGCCAACGACGCGGCGTGGGTGGCGGAGGAGGGCTACTTCAACAGCGCTCTGTCCCTGGCCGACAAGGGGAGCCTGCCT [240]
K  A  N  D  A  A  W  V  A  E  E  G  Y  F  N  S  A  L  S  L  A  D  K  G  S  L  P
GCTGGAGAGCACAGCTTTCCCTTCCAGTTCCTGCTTCCTGCCACAGCGCCCACGTCCTTTGAGGGCCCTTTTGGGAAGAT [320]
 A  G  E  H  S  F  P  F  Q  F  L  L  P  A  T  A  P  T  S  F  E  G  P  F  G  K  I
CGTGCACCAGGTACGGGCCACCATTGACACGCCACGTTTTTCCAAGGATCACCAGTGCAGCCGCGTGTTCTATATCTTGA [400]
  V  H  Q  V  R  A  T  I  D  T  P  R  F  S  K  D  H  Q  C  S  R  V  F  Y  I  L
GCCCCCTGAACCTGAACAGCATCCCAGACATCGAGCAACCCAATGTGGCCTCCACCACCAGGAAGTTCTCCTACAAGCTG [480]
S  P  L  N  L  N  S  I  P  D  I  E  Q  P  N  V  A  S  T  T  R  K  F  S  Y  K  L
GTGAAGACAGGCAGCGTGGTCCTCACGGCCAGCACCGATCTCCGAGGCTACGTGGTGGGGCAAGTGCTGCGGCTGCAGGC [560]
 V  K  T  G  S  V  V  L  T  A  S  T  D  L  R  G  Y  V  V  G  Q  V  L  R  L  Q  A
TGACATCGAGAACCAGTCAGGCAAGGACACCAGCCCTGTGGTGGCCAGTCTGCTGCAGAAAGTGTCCTATAAGGCCAAGC [640]
  D  I  E  N  Q  S  G  K  D  T  S  P  V  V  A  S  L  L  Q  K  V  S  Y  K  A  K
GCTGGATTTATGACGTGCGGACCATCGCAGAGGTGGAGGGTGCCAGCGTCAAGGCCTGGAGGCGGGCTCAGTGGCAAGAG [720]
R  W  I  Y  D  V  R  T  I  A  E  V  E  G  A  S  V  K  A  W  R  R  A  Q  W  Q  E
CAGATCCTAGTGCCCGCCCTGCCCCAGTCGGCCCTGCCGGGCTGCAGCCTCATCCACGTGGACTACTACCTGCAGGTCTC [800]
 Q  I  L  V  P  A  L  P  Q  S  A  L  P  G  C  S  L  I  H  V  D  Y  Y  L  Q  V  S
CCTGAAGACTCCGGAAGCCATTGTGACCCTGCCGGTCTTCATTGGCAATATCGCTGTGAACCAAGTTCCACTGAGCCCCC [880]
  L  K  T  P  E  A  I  V  T  L  P  V  F  I  G  N  I  A  V  N  Q  V  P  L  S  P
GGCCAGGCCCAGGGCCACCTCCGGTAGTGCCCTCTGCGCCACCCCAGGAGGAGGCAGAGGCTGTGGCCAGTGGACCCCAC [960]
R  P  G  P  G  P  P  P  V  V  P  S  A  P  P  Q  E  E  A  E  A  V  A  S  G  P  H
TTCTTCTCGGACCCTGTCTCTCTCTCCACCAAGAGCCACTCGCAGCAGCAGCAGCCGCCTGCAGCCTTCGGCTCTGTGCC [1040]
 F  F  S  D  P  V  S  L  S  T  K  S  H  S  Q  Q  Q  Q  P  P  A  A  F  G  S  V  P
TGGCGCCCCTGAACCCCATCCTCCGGATGGCAGCCCCGCCCCCCACCCTCTGCCCCCTCCCTTGTGTATCTCCACAGGTG [1120]
  G  A  P  E  P  H  P  P  D  G  S  P  A  P  H  P  L  P  P  P  L  C  I  S  T  G
CCACCGTCCCCTACTTTGCAGAGGGTTCTGGAGGGCCCGTGCCCACCACCAGTACCTTGATCCTACCCCCGGAGTACAGC [1200]
A  T  V  P  Y  F  A  E  G  S  G  G  P  V  P  T  T  S  T  L  I  L  P  P  E  Y  S
TCGTGGGGCTACCCCTATGAGGCCCCACCATCCTATGAACAGAGCTGCGGCAGTGTGAACCCCTGCCTGACCCCCAGGAG [1280]
 S  W  G  Y  P  Y  E  A  P  P  S  Y  E  Q  S  C  G  S  V  N  P  C  L  T  P  R  S
CTGAccgcatgtcgccttctctggacggg                                                    [1309]
  *
```

Figure 2A. The complete CDS of swine *ARRDC1* gene and its deduced amino acids. The nucleotide sequences of primers were underlined. An asterisk denoted the stop codon. The two conserved domains, arrestin-N domain (7-139 AA) and the arrestin-C domain (162-284 AA), were boxed.

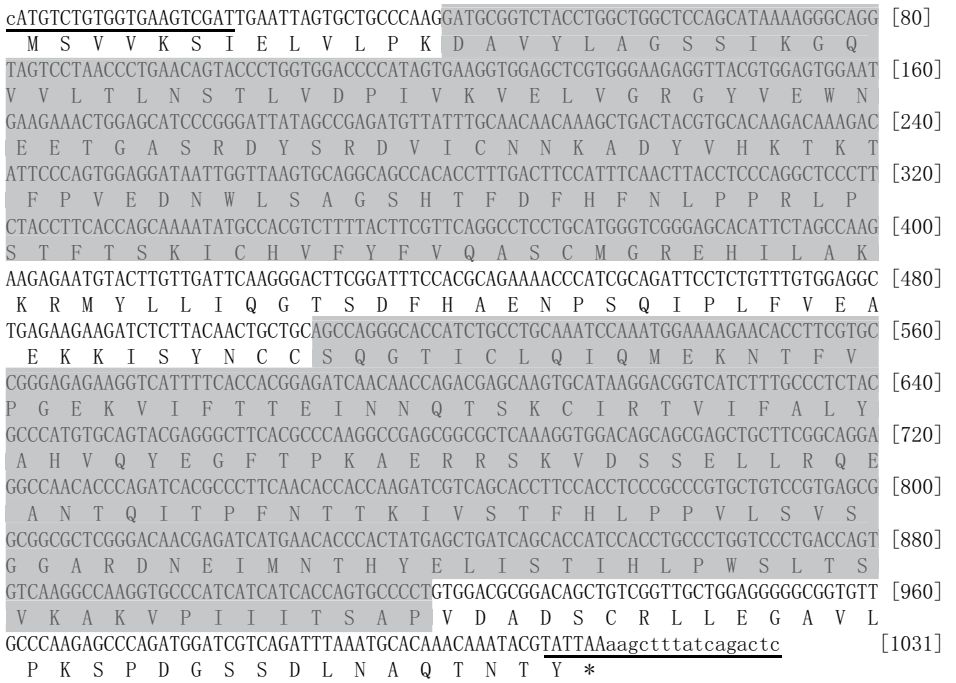

Figure 2B. The complete CDS of swine *ARRDC5* gene and its deduced amino acids. The nucleotide sequences of primers are underlined. An asterisk denoted the stop codon. Amino acids in boxes were the two conserved domains, arrestin-N domain (14-124 AA) and the arrestin-C domain (170-306 AA).

28 potential phosphorylation sites, including 17 serine, 6 threonine, and 5 tyrosine sites; ARRDC5 has 19 potential phosphorylation sites, including 11 serine, 6 threonine, and 2 tyrosine sites. ARRDC1 has 1 potential Asn glycosylation site at p.191, and ARRDC5 has 4 potential Asn glycosylation sites located at p.32, p.150, p.199, and p.249. Detailed information regarding the phosphorylation and glycosylation sites is shown in the Table.

Table. In silico predicted sites of S-, T-, and Y-phosphorylation and N-glycosylation in ARRDC1 and ARRDC5.

Gene	ARRDC1				ARRDC5			
Phosphorylated site	Serine P		Threonine P	Tyrosine P	Serine P		Threonine P	Tyrosine P
	19	209	98	66	6	311	110	74
	52	229	117	158	22	323	222	279
	72	267	152	263	63		246	
	99	292	197	399	107		250	
	121	329	221	412	111		277	
	140	332	424		229		292	
	157	334			233			
	193	411			266			
	198				293			
Glycosylated site	Asn				Asn			
	191				32 150 199 249			

We used the web-based miRNA prediction program miRBase (http://www.mirbase.org/) to locate potential conserved miRNA targets. The results showed that there were four *Sus scrofa* microRNAs in the swine *ARRDC1* sequence: ssc-miR-132 (49-gucuacagcccugg-62), ssc-miR-490-3p (786-cuaccugcaggucucccug-804), ssc-miR-1296-5p (967-ucggacccugucucucucucc987), and ssc-miR-339-5p (1264-ugccugacccccaggagcuga-1284). There were two microRNAs in the *ARRDC5* sequence: ssc-miR-708-5p (770-ccaccaagaucgucagcaccuu-749) and ssc-miR-324 (927-caccagugcccuguggacgcg-906).

The secondary structure prediction indicated that the swine ARRDC1 protein contained 41 AA alpha helices, 101 AA extended strands, 17 AA beta turns, and 268 AA random coils, while ARRDC5 contained 51 AA alpha helices, 81 AA extended strands, 15 AA beta turns, and 190 AA random coils. Both ARRDC1 and ARRDC5 had two functional domains: arrestin N and arrestin C. The arrestin C conserved domain was found in both ARRDC1 (162-284 AA) and ARRDC5 (170-306 AA). The arrestin N domain was conserved in ARRDC1 (7-139 AA) but was not specific in ARRDC5 (14-124). The conserved domains of swine ARRDC1 and ARRDC5 proteins are displayed in Figures 2A and 2B.

The amino acid sequences of swine ARRDC1 exhibited 94.1%, 93.9%, 93.4%, 92.3%, 89.4%, 88.3%, and 84.7% identity with those of *Bos taurus* (NP_001073834), *Tursiops truncatus* (XP_004312074), *Felis catus* (XP_003996160), *Canis lupus* (XP_003435381), *Capra hircus* (XP_005687444), *Homo sapiens* (NP_689498), and *Mus musculus* (NP_848495), respectively. The swine ARRDC5 (AHB08949) amino acid sequences exhibited 84.9%, 86.9%, 81.2%, 86.9%, 84.3%, 81.2%, and 71.2% identity with those of *Bos taurus* (NP_001073834), *Tursiops truncatus* (XP_004312074), *Felis catus* (XP_003996160), *Canis lupus* (XP_003435381), *Capra hircus* (XP_005687444), *Homo sapiens* (NP_689498), and *Mus musculus* (NP_848495), respectively.

3.3. Evolutionary relationship analyses

The evolutionary relationships of swine ARRDC1 and ARDDC5 with other species are presented in Figure 3. Cluster analysis based on the amino acid sequence of ARRDC1 grouped *Sus scrofa* and *Tursiops truncatus* together and *Bos taurus* and *Capra hircus* together. The two clades were then classified into the same group. The evolutionary relationship of ARRDC5 is the same as that of ARRDC1 in the eight species.

3.4. Tissue expression profile analysis of *ARRDC1* and *ARRDC5*

To examine the differential distribution of *ARRDC1* and *ARRDC5* in swine tissues, the relative mRNA expression levels of the two genes were evaluated using qPCR. The *ARRDC1* gene was widely expressed in the tissues

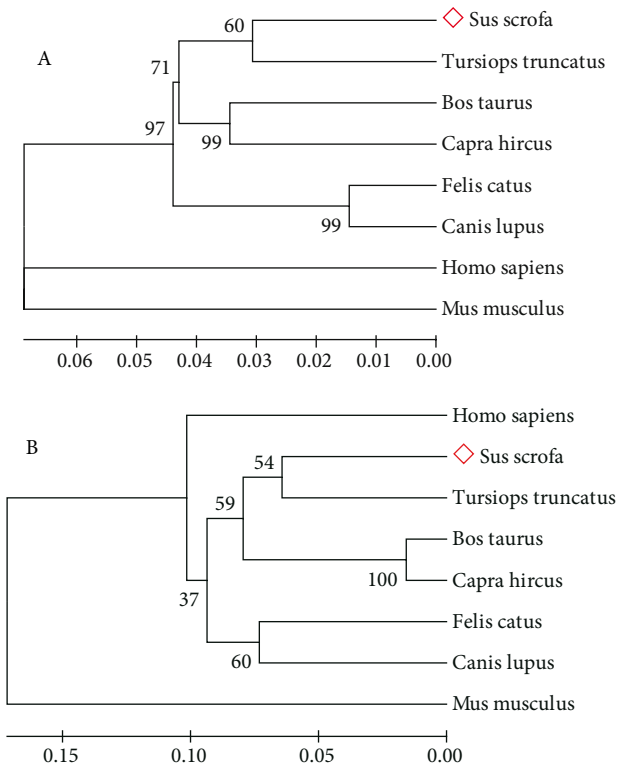

Figure 3. Phylogenetic trees based on the sequences of ARRDC1 (A) and ARRDC5 (B) between swine and other species. The trees were obtained by bootstrap analysis with the neighbor-joining method; numbers on the branches represent bootstrap values for 10,000 replications.

examined, showing different expression patterns. However, *ARRDC5* was expressed in only a few tissues (Figure 4).

ARRDC1 was highly expressed in the testis, thyroid, submandibular gland, and kidney; moderately expressed in the cerebellum, pituitary, liver, spleen, lung, adrenal gland, and sublingual gland; weakly expressed in the brain, hypothalamus, spinal cord, muscle, duodenum, jejunum, ileum, rectum, lymph nodes, epididymis, and thymus; and minimally expressed in the brain, brainstem, heart, skin, colon, cecum, stomach, pancreas, and esophagus.

ARRDC5 was highly expressed in the testis; weakly expressed in the thyroid and thymus; and minimally expressed in the brain, cerebellum, hypothalamus, brainstem, spinal cord, pituitary, heart, liver, spleen, lung, kidney, skin, muscle, duodenum, jejunum, ileum, colon, cecum, rectum, stomach, pancreas, esophagus, lymph nodes, epididymis, submandibular gland, adrenal gland, and sublingual gland.

4. Discussion

In this study, we successfully cloned and identified the *ARRDC1* and *ARRDC5* cDNA sequences from the Diannan small-ear pig. The data regarding genes that

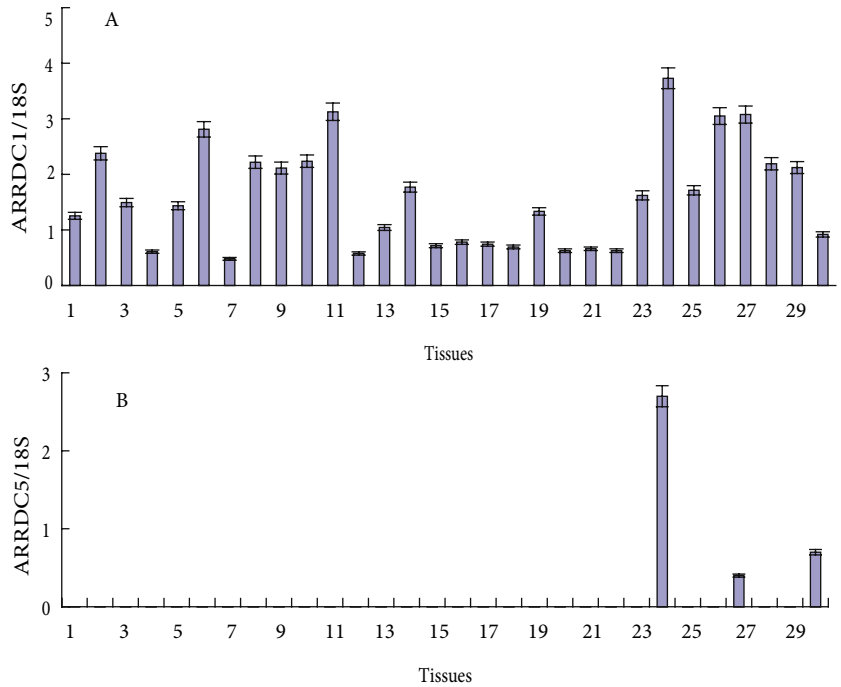

Figure 4. Tissue expression profiles of swine *ARRDC1* (A) and *ARRDC5* (B). The horizontal axis and vertical axis indicate different tissues and the $2^{-\Delta\Delta Ct}$ value (mean ± SE), respectively. The housekeeping gene *18S* was used as reference.
Lane 1 = brain; lane 2 = cerebellum; lane 3 = hypothalamus; lane 4 = brainstem; lane 5 = spinal cord; lane 6 = pituitary; lane 7 = heart; lane 8 = liver; lane 9 = spleen; lane 10 = lung; lane 11 = kidney; lane 12 = skin; lane 13 = muscle; lane 14 = duodenum; lane 15 = jejunum; lane 16 = ileum; lane 17 = colon; lane 18 = cecum; lane 19 = rectum; lane 20 = stomach; lane 21 = pancreas; lane 22 = esophagus; lane 23 = lymph nodes; lane 24 = testis; lane 25 = epididymis; lane 26 = submandibular gland; lane 27 = thyroid; lane 28 = adrenal gland; lane 29 = sublingual gland; lane 30 = thymus.

were homologous to the swine ARRDC1 and ARRDC5 sequences confirmed that the sequences we cloned encoded two new swine genes. The CDS of the two genes were 1284 and 1014 bp, which encoded 427 and 337 amino acids, respectively. The Mws of ARRDC1 and ARRDC5 were 45.6 kDa and 37.5 kDa, respectively; their respective pIs were 7.61 and 7.08.

Bioinformatics analysis suggested that ARRDC1 and ARRDC5 had no N-terminal signal peptide sequences or transmembrane regions. ARRDC1 is a hydrophilic protein located in the nucleus, while ARRDC5 is located in the cytoplasm. These results indicated that the main function of ARRDC1 is in the nucleus, the main function of ARRDC5 is in the cytoplasm, and they are both nonsecretory proteins. In addition, two conserved functional domains, arrestin N and arrestin C, were found in the two ARRDC proteins. The arrestin (or S-antigen) C-terminal domain is also an Ig-like beta-sandwich fold and is a duplication of the N-terminal domain. Arrestins comprise a family of closely related proteins that includes beta-arrestin-1 and beta-arrestin-2, which regulate the function of the beta-adrenergic receptors by binding to their phosphorylated forms, impairing their capacity to activate G proteins. The cone photoreceptors C-arrestin (arrestin-X), which binds to phosphorylated red/green opsins, and *Drosophila* phosrestins I and II, which undergo light-induced phosphorylation, most likely play a role in photoreceptor transduction (Palczewski, 1994; Granzin et al., 1998). All the ARRDC family members have arrestin-C domains, which play a critical role in activating the members.

ARRDC1 has 28 potential phosphorylation sites, and ARRDC5 has 19 potential phosphorylation sites. This may be related to their binding forms and the regulatory functions of the G(S) proteins.

miRNAs are small noncoding RNAs that regulate target-gene expression by degrading or inhibiting the translation of their corresponding target mRNAs. ARRDC1 has three potential miRNA regulatory sites: ssc-miR-490-3p (786-cuaccugcaggucucccug-804), ssc-miR-1296-5p (967-ucggacccugucucucucucc-987), and ssc-miR-339-5p (1264-ugccugacccccaggagcuga-1284). These sites are near the end of the CDS and may be important

for gene expression regulation. The ARRDC5 also has two important potential miRNA regulatory sites: ssc-miR-708-5p (770-ccaccaagaucgucagcaccuu-749) and ssc-miR-324 (927-caccagugcccuguggacgcg-906).

Homology analysis based on the amino acid sequences of ARRDC1 and ARRDC5 showed that the two proteins were highly conserved. The swine ARRDC1 and ARRDC5 sequences show a high level of identity with sequences from *Bos taurus*, *Tursiops truncatus*, and *Canis lupus*.

The evolutionary relationship determined based on the ARRDC1 and ARRDC5 amino acid sequences revealed that swine have close genetic relationships with *Tursiops truncatus*, *Bos taurus*, and *Capra hircus*. This result implied that the functional divergence of swine ARRDC1 and ARRDC5 from that of *Tursiops truncatus*, *Bos taurus*, and *Capra hircus* was minor. Therefore, previous studies of ARRDC1 and ARRDC5 in these animals could be used as a reference for understanding the possible functions of ARRDC1 and ARRDC5 in swine.

Tissue expression profile analysis showed that the *ARRDC1* and *ARRDC5* genes had different expression patterns in swine. *ARRDC1* was widely expressed in the tissues examined at different levels and highly expressed in the testis. The expression of *ARRDC5* was almost completely restricted to the testis. The tissue expression pattern is always closely related to the function of the gene; thus, these results implied that the *ARRDC1* and *ARRDC5* genes may be important in regulating testis function. The mRNA expression patterns of *Sus scrofa ARRDC1* and *ARRDC5* genes in our study were similar to *Mus musculus* in the Gene Expression Omnibus (GEO) database of NCBI. We speculate that ARRDC1 and ARRDC5 have analogous roles in swine testis. In future studies, the specific function of the ARRDC families in the testis should be investigated.

In summary, we isolated Diannan small-ear pig *ARRDC1* and *ARRDC5* genes, and we performed bioinformatics and tissue expression profile analyses. This study provides the primary foundation for further studies of swine *ARRDC* genes.

Acknowledgments

This study was financially supported by the National Natural Science Foundation of China (#31160439 and #31460580).

References

Alvarez CE (2008). On the origins of arrestin and rhodopsin. BMC Evol Biol 8: 222.

Granzin J, Wilden U, Choe HW, Labahn J, Krafft B, Büldt G (1998). X-ray crystal structure of arrestin from bovine rod outer segments. Nature 391: 918–921.

Gurevich VV, Gurevich EV (2004). The molecular acrobatics of arrestin activation. Trends Pharmacol Sci 25: 105–111.

Gurevich VV, Gurevich EV (2006). The structural basis of arrestin-mediated regulation of G-protein-coupled receptors. Pharmacol Ther 110: 465–502.

Hanyaloglu AC, von Zastrow M (2008). Regulation of GPCRs by endocytic membrane trafficking and its potential implications. Annu Rev Pharmacol Toxicol 48: 537–568.

Hurley JH, Emr SD (2006). The ESCRT complexes: structure and mechanism of a membrane-trafficking network. Annu Rev Biophys Biomol Struct 35: 277–298.

Kumar S, Nei M, Dudley J, Tamura K (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9: 299–306.

Lefkowitz RJ, Shenoy SK (2005). Transduction of receptor signals by beta-arrestins. Science 308: 512–517.

Léon S, Haguenauer-Tsapis R (2009). Ubiquitin ligase adaptors: regulators of ubiquitylation and endocytosis of plasma membrane proteins. Exp Cell Res 315: 1574–1583.

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25: 402–408.

Lohse MJ, Benovic JL, Codina J, Caron MG, Lefkowitz RJ (1990). β-Arrestin: a protein that regulates β-adrenergic receptor function. Science 248: 1547–1550.

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

Nabhan JF, Hu R, Oh RS, Cohen SN, Lu Q (2012). Formation and release of arrestin domain-containing protein 1-mediated microvesicles (ARMMs) at plasma membrane by recruitment of TSG101 protein. Proc Natl Acad Sci USA 109: 4146–4151.

Palczewski K (1994). Structure and functions of arrestins. Structure and functions of arrestins. Protein Sci 3: 1355–1361.

Palmitessa A, Benovic JI (2010). Arrestin and the multi-PDZ domain-containing protein MPZ-1 interact with phosphatase and tensin homolog (PTEN) and regulate *Caenorhabditis elegans* longevity. J Biol Chem 285: 15187–15200.

Puca L, Chastagner P, Meas-Yedid V, Israël A, Brou C (2013). α-Arrestin 1 (ARRDC1) and β-arrestins cooperate to mediate Notch degradation in mammals. J Cell Sci 126: 457–468.

Rauch S, Martin-Serrano J (2011). Multiple interactions between the ESCRT machinery and arrestin-related proteins: implications for PPXY-dependent budding. J Virol 85: 3546–3556.

Reza KA, Sharmin S, Moosa MM, Mahmood N, Ghosh A (2013). Identification and molecular characterization of a receptor-like protein kinase gene from *Corchorus capsularis*. Turk J Biol 37: 11–17.

Shea FF, Rowell JL, Li Y, Chang TH, Alvarez CE (2012). Mammalian α arrestins link activated seven transmembrane receptors to Nedd4 family e3 ubiquitin ligases and interact with β arrestins. PLoS One 7: e50557.

Wilden U, Hall SW, Kühn H (1986). Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-kDa protein of rod outer segments. Proc Natl Acad Sci USA 83: 1174–1178.

Wilson CJ, Applebury ML (1993). Arresting G-protein coupled receptor activity. Curr Biol 3: 683–686.

Wu CF, Liu LX, Huo JL, Li DL, Yuan YY (2014). Molecular characterization of the encoding regions and tissue expression analyses for 3 novel buffalo AKT genes, *AKT1*, *AKT2*, and *AKT3*. Turk J Biol 38: 63–75.

Yu SM, Wang CW, Zhao DM, Zhang QC, Pei DZ (2003). Raising and pathogen purification of Chinese experimental mini-pig. Lab Anim Sci Admin 20: 44–46.