Molecular cloning, tissue expression pattern, and copy number variation of porcine SCUBE3


Key Laboratory of Farm Animal Genetic Resources and Germplasm Innovation of Ministry of Agriculture of China, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China

*These authors contributed equally to this study.
Corresponding author: L.X. Wang
E-mail: iaswlx@263.net

Genet. Mol. Res. 15 (1): gmr.15017010
Received August 17, 2015
Accepted October 26, 2015
Published February 5, 2016
DOI http://dx.doi.org/10.4238/gmr.15017010

ABSTRACT. The signal peptide CUB EGF-like domain-containing protein 3 (SCUBE3) gene is a member of SCUBE gene family and plays important roles in bone cell biology and the determination of limb bone length. In this study, the full-length transcript of porcine SCUBE3 was cloned using reverse transcription-polymerase chain reaction and rapid amplification of cDNA ends. The full-length sequence of porcine SCUBE3 cDNA was 4131 base pairs and included 21 exons. The SCUBE3 gene contained a 2895-base pair open reading frame that encoded a peptide of 965 amino acids. Comparison of the deduced amino acid sequences of porcine SCUBE3 with those of human, mouse, zebrafish, and rat showed 96, 95, 73, and 95% identities, respectively. Porcine SCUBE3 mRNA expression levels were highest in the backfat, bone marrow, and cartilage tissues. Copy number variation was detected in porcine SCUBE3 and validated by real-time quantitative polymerase chain reaction. Different copy number variations were present in randomly selected individuals and may, therefore, be a
good marker for identifying phenotypic traits. Our findings provide a basis for further investigation of the functions and regulatory mechanisms of SCUBE3 in pigs.

**Key words:** Cloning; Copy number variation; mRNA expression; Porcine; Signal peptide CUB EGF-like domain-containing protein 3 gene

**INTRODUCTION**

The signal peptide CUB EGF-like domain-containing protein (SCUBE) gene family, which encodes secreted cell-surface proteins, contains 3 members (SCUBE1-3) and is evolutionarily conserved from zebrafish to humans (Wu et al., 2004; Woods and Talbot, 2005; Haworth et al., 2007; Xavier et al., 2013). All three human SCUBE proteins have been found to homodimerize, while SCUBE1 has additionally been shown to heterodimerize with other SCUBE proteins (Wu et al., 2004; Johnson et al., 2012).

Although the SCUBE genes are expressed in the vertebrate embryo in reciprocal and complementary manners, the individual genes show different expression profiles. Transcripts of the SCUBE genes have been detected in a range of tissues in mouse, rat, and zebrafish embryos (Hollway et al., 2006; Haworth et al., 2007; Xavier et al., 2010). The SCUBE1 protein is present at high levels in human endothelial cells, and in the central nervous system, somites, and gonads of mouse embryos (Grimmond et al., 2000; Yang et al., 2002; Tu et al., 2006; Xavier et al., 2009). SCUBE2 has been found in the zebrafish you-type mutant and in human endothelial smooth-muscle, fibroblast, and renal mesangial cells, as well as in the forebrain, trunk neuroepithelium, and anterior hindbrain of mice (Grimmond et al., 2001; Haworth et al., 2007). SCUBE3, which is a secreted glycoprotein, forms either homo- or hetero-oligomers that are tethered to the cell surface and can act either locally or distantly in a paracrine or endocrine manner (Wu et al., 2004). Human SCUBE3 is highly expressed in primary osteoblasts and in the humerus and femur, but shows low expression in cardiac tissue and umbilical vein endothelial cells (Wu et al., 2004; Yang et al., 2007). In mice, SCUBE3 is expressed initially in neuroectoderm and then in ectodermal, endodermal, and mesodermal tissues, including the somites, limb buds, and neural tubes (Haworth et al., 2007; Xavier et al., 2010). SCUBE3 has been mapped to human chromosome 6 and is linked with the locus for Paget’s disease, a metabolic bone disease (Wu et al., 2004). The unique expression and function of SCUBE3 in the bones and osteoblasts suggest that this gene plays an important role in bone cell biology.

In humans, limb bone length is highly correlated with body height; normal bone cell metabolism is important for human health (Marinković and Vilić, 2012). In pigs, limb length influences body height and ham yield (Andersson and Georges, 2004). Pigs are considered an important model species for research on human diseases because of the physiological similarities between the two species. Thus, investigation of SCUBE3 in pigs is not only important to the pig industry but is also potentially of value for human medical research. Previously, we carried out a genome-wide association study and found that SCUBE3 was associated with limb bone length, body length, and some carcass traits (Liu et al., 2014; Wang et al., 2014; Zhang et al., 2014). However, compared to the large number of studies in mice and humans, comparatively little information is available regarding SCUBE3 in pigs. In this study, porcine SCUBE3 cDNA was cloned and characterized, and its expression profile in different tissues was determined. We also detected and characterized copy number variation (CNV) for the gene. The results obtained in this study provide valuable insights into the function of SCUBE3 in pigs.
MATERIAL AND METHODS

Animal and tissue collection

This study was conducted in compliance with the guidelines on experimental animals established by the Council of China. Animal experiments were approved by the Science Research Department of the Institute of Animal Science, Chinese Academy of Agricultural Sciences (Beijing, China).

Three 180-day-old Large White sows were slaughtered and tissue samples were rapidly collected from the backfat, longissimus muscle, cartilage, bone marrow, heart, liver, lung, and kidney; the samples were stored at -80°C until use. The samples were used to analyze SCUBE3 mRNA expression levels. One cDNA pool was used to clone the intact, full-length cDNA. Eleven Large White x Minzhu intercross individuals were randomly selected and ear tissue samples were collected to analyze CNVs.

RNA and DNA isolation

Total RNA was extracted and purified using an RNAPrep Pure Tissue kit (TIANGEN, Beijing, China) following the manufacturer instructions, and RNA quality was evaluated by 1.5% agarose gel electrophoresis. The RNAs were stored at -80°C until use. First-strand cDNA was synthesized by reverse transcription using 1 μg total RNA and the PrimeScript™ RT reagent kit (Takara, Shiga, Japan) according to the manufacturer protocol; cDNAs were stored at -20°C until use. Genomic DNA was extracted from ear tissue samples using the phenol-chloroform method, and DNA quality was evaluated by 1% agarose gel electrophoresis; the DNA was stored at -20°C until use.

Cloning of the SCUBE3 cDNA

cDNA was produced by reverse transcription as described above. Three pairs of primers were designed using the predicted mRNA sequence of porcine SCUBE3 (GenBank accession No. XM_005665900) (Table 1) and were used in a polymerase chain reaction (PCR) to amplify the middle sequences of SCUBE3 from the cDNA template. PCR was carried out using a 50-μL reaction mixture containing 2 μL cDNA, 200 nM forward and reverse primers, 250 μM dNTPs, and 1 U Taq polymerase in 25 μL 2X GC Buffer I (Takara). The following amplification conditions were used: 94°C for 5 min; and 36 cycles of 94°C for 30 s, with an appropriate annealing temperature (Table 1) for 40 s, and 72°C for 50 s; and a final extension of 72°C for 10 min. The PCR products were ligated into the PMD-18T plasmid (TaKaRa) for DNA sequencing.

The complete 5′ and 3′ ends of the SCUBE3 cDNA were amplified using a 5′-Full RACE Kit and a 3′ RACE kit (TakaRa, Japan), respectively, according to the manufacturer instructions. The primer sequences for rapid amplification of cDNA ends (RACE) were based on the previously obtained partial sequences of the 5′ and 3′ ends (Table 1). The PCR products were ligated into the PMD-18T plasmid (Takara) for DNA sequencing.

Sequences analysis of SCUBE3

The search for the open reading frame (ORF) and the translation of nucleotide sequences into amino acids were performed using the ORF Finder at NCBI (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Exons and introns in SCUBE3 were identified using an Ensemble Blast search.
Coding sequences and deduced amino acid sequences were identified by homology searches using the multiple sequence alignment function of BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/). A phylogenetic tree of SCUBE3 genes in pig, human (NP_001290065.1), mouse (NP_001004366.1), zebrafish (NP_001247430.1), and rat (NP_001258291.1) was produced using the MEGA5.0 software. ProtParam (http://web.expasy.org/protparam/) and ProtScale (http://web.expasy.org/protscale/) in the ExPASy Proteomics Server were used to calculate physical and chemical parameters of the deduced amino acid sequences.

### Table 1. Primers for PCR amplification of SCUBE3.

<table>
<thead>
<tr>
<th>Primer purpose</th>
<th>Primers</th>
<th>Primer sequences</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle fragment cloning</td>
<td>SCF1</td>
<td>CTGGATGTGGACGAGTGTGAG</td>
<td>903</td>
</tr>
<tr>
<td></td>
<td>SCF2</td>
<td>TCATGGCTACCTACCTATGTTGCTCA</td>
<td>1185</td>
</tr>
<tr>
<td></td>
<td>SCF3</td>
<td>GGAACCCACACACATACACACAT</td>
<td>723</td>
</tr>
<tr>
<td></td>
<td>SCF3</td>
<td>GTGCTCATGCTTCATCATAGGTAACA</td>
<td></td>
</tr>
<tr>
<td>3'-RACE</td>
<td>F1</td>
<td>GTGAGTGCTGGCTATATCGAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>GATCTTCCTGGGCTGAGG</td>
<td></td>
</tr>
<tr>
<td>5'-RACE</td>
<td>R1</td>
<td>TGACAGGGCGAGGTCTCATTAGAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R2</td>
<td>TCGGTATGGGACACAACTGTG</td>
<td></td>
</tr>
</tbody>
</table>

### Tissue expression analysis of SCUBE3

To investigate the comparative levels of expression of SCUBE3 in different tissues, we extracted total RNA from heart, liver, lung, kidney, backfat, longissimus dorsi, cartilage, and bone marrow. SCUBE3 expression level was determined using SYBR Green Real Time quantitative PCR using an ABI7500 instrument (Applied Biosystems, Inc., Foster City, CA, USA); β-Actin was used as the endogenous control gene to normalize SCUBE3 expression levels. The primers were designed using PrimerExpress3.0 (Applied Biosystems) and are listed in Table 2. The reaction mixtures contained 10 µL SYBR Select Master Mix, 0.5 µL each 10 µM forward and reverse primers, 1.5 µL cDNA, and 7.5 µL ddH₂O in a 20 µL reaction volume. After an initial denaturation step at 95°C for 2 min, the amplification cycling conditions were as follows: 40 cycles of 95°C for 10 s, 60°C for 30 s, and 72°C for 30 s. Reactions were performed in triplicate for each sample and the mean was used for further analysis. Gene expression levels were quantified using the comparative 2^{-ΔCt} method, as ΔCt = Ct_SCUBE3 - Ct_β-actin

### Table 2. Primers for qPCR amplification of SCUBE3.

<table>
<thead>
<tr>
<th>Primer purpose</th>
<th>Primers</th>
<th>Primer sequences</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA expression</td>
<td>SCU-F</td>
<td>GGCACCTACCAACCTGA</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>SCU-R</td>
<td>GGGAGCAGCTGGAGACTTT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-actinF</td>
<td>TGCGGACACATCAAGGAAAG</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>β-actinR</td>
<td>AGTTGAAGGTTGCTGCTGTG</td>
<td></td>
</tr>
<tr>
<td>CNV validation</td>
<td>SC-CNVF</td>
<td>TGAGCAAAAGGAGGAAAAAG</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>SC-CNVR</td>
<td>GGGCACGTTGTTGGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GCC-F</td>
<td>AAGCTTCAACAGGGGTACAAT</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>GCC-R</td>
<td>CAACTTTGAATGTTACCCCTAAAT</td>
<td></td>
</tr>
</tbody>
</table>
CNV detection in SCUBE3

CNV in SCUBE3 was detected and validated by SYBR Green real-time quantitative PCR using an ABI7500 instrument (Applied Biosystems); the primers for the assay are listed in Table 2. For the assay, the reaction mixtures consisted of 10 µL SYBR Select Master Mix, 0.5 µL each 10 µM forward and reverse primers, 1.5 µL cDNA, and 7.5 µL ddH₂O in a 20 µL reaction volume. The PCR cycle was as follows: 95°C for 2 min; 40 cycles of 95°C for 10 s, 60°C for 30 s, and 72°C for 30 s. The SDS2.4 software was used to analyze the results. The glucagon gene (Ballester et al., 2004) was used as the single copy control. Copy number was calculated using the 2^\(-\Delta\Delta CT\) method (Graubert et al., 2007; Wang et al., 2014), where \(\Delta CT\) is the cycle threshold (CT) of the target region minus the CT of the control region.

RESULTS

Cloning and structural analysis of SCUBE3 cDNA and deduced amino acid sequence

A 2340-base pair (bp) fragment from the middle of the SCUBE3 cDNA sequence was amplified by PCR. The 3’ and 5’ RACE PCRs produced 1585- and 206-bp fragments, respectively. These 3 fragments comprised the full-length cDNA (4131 bp; GenBank accession No.: KP784762). SCUBE3 cDNA contained a 2895-bp ORF that coded for a 965-amino acid protein as well as a 432-bp 5’-untranslated region and an 804-bp 3’-untranslated region. The exon-intron boundaries of porcine SCUBE3 were determined using the Ensemble Blast search program: 21 exons were identified, ranging in size from 81 to 953 bp. The exon-intron boundaries followed the GT/AG consensual splicing rule. The structure of the SCUBE3 gene is depicted in Figure 1.

Figure 1. Structure of the signal peptide CUB EGF-like domain-containing protein 3 gene (SCUBE3).

Primary structure analysis revealed that the molecular weight of the putative SCUBE3 protein was 106.28 kDa, with a theoretical isoelectric point of 7.43 and an instability index of 46.95; therefore, the analysis identified SCUBE3 as an unstable protein. SCUBE3 contained most amino acids (except for Pyl and Sec), although the proportions of each amino acid differed. The amino acid composition of the protein is shown in Table 3. Gly and Cys were the most common amino acids. A comparison of the deduced amino acid sequences of pig SCUBE3 with those from humans, mice, zebrafish, and rats showed 96, 95, 73, and 95% identities, respectively. A phylogenetic tree was constructed using the MEGA5 software (Figure 2); this analysis confirmed the close genetic relationship among the proteins from different species.

Tissue expression patterns of SCUBE3

The expression of SCUBE3 was examined in 8 different tissues of 180-day-old animals.
by real-time quantitative PCR (qPCR) (Figure 3). The analysis showed that the levels of SCUBE3 mRNA differed considerably among tissues. The highest relative expression was found in the backfat and the lowest was in the heart (Figure 3). An approximately 10-fold difference in expression level was observed between backfat and the heart.

### Table 3. Amino acid composition of the protein encoded by porcine SCUBE3.

<table>
<thead>
<tr>
<th>Name</th>
<th>Count</th>
<th>Count percent</th>
<th>Name</th>
<th>Count</th>
<th>Count percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala (A)</td>
<td>51</td>
<td>5.30</td>
<td>Leu (L)</td>
<td>57</td>
<td>5.90</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>60</td>
<td>6.20</td>
<td>Lys (K)</td>
<td>55</td>
<td>5.70</td>
</tr>
<tr>
<td>Asn (N)</td>
<td>46</td>
<td>4.80</td>
<td>Met (M)</td>
<td>10</td>
<td>1.00</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>52</td>
<td>5.40</td>
<td>Phe (F)</td>
<td>32</td>
<td>3.30</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>90</td>
<td>9.30</td>
<td>Pro (P)</td>
<td>57</td>
<td>5.90</td>
</tr>
<tr>
<td>Gln (Q)</td>
<td>39</td>
<td>4.00</td>
<td>Ser (S)</td>
<td>59</td>
<td>6.10</td>
</tr>
<tr>
<td>Glu (E)</td>
<td>61</td>
<td>6.30</td>
<td>Thr (T)</td>
<td>59</td>
<td>6.10</td>
</tr>
<tr>
<td>Gly (G)</td>
<td>90</td>
<td>9.30</td>
<td>Trp (W)</td>
<td>3</td>
<td>0.30</td>
</tr>
<tr>
<td>His (H)</td>
<td>36</td>
<td>3.70</td>
<td>Tyr (Y)</td>
<td>29</td>
<td>3.00</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>35</td>
<td>3.60</td>
<td>Val (V)</td>
<td>44</td>
<td>4.60</td>
</tr>
</tbody>
</table>
Identification and validation of CNV

We determined the copy number of \textit{SCUBE3} by qPCR in 11 randomly selected individuals. One individual was selected as the standard, while the others were compared to this standard (Figure 4). The analysis identified considerable variation in relative expression levels among individuals. Compared to the standard, the relative expression levels of the others showed 0.5 to 1.6-fold increased. We concluded that the copy number of \textit{SCUBE3} varied among individuals.

![Figure 4. Relative expression of \textit{SCUBE3} in different individuals for validating CNV.](image)

**DISCUSSION**

In this study, the entire \textit{SCUBE3} cDNA was cloned and sequenced. Although this gene has been described previously in humans, mice, zebrafish, and rats, this is the first description of the gene in pig (Wu et al., 2004; Yang et al., 2007; Johnson et al., 2012). The gene is located in the 35.85-35.87-Mb region of \textit{Sus scrofa} chromosome (SSC) 7; it was found to contain 21 exons and encode 965 amino acids. The amino acid sequence in pigs showed 96, 95, 73, and 95\% identities with the sequences in humans, mice, zebrafish, and rats, respectively. This relationship was confirmed by a neighbor-joining phylogenetic tree. These results indicate that SCUBE3 is evolutionarily conserved in mammals. This conclusion is consistent with the results of previous studies (Xavier et al., 2013).

Our qPCR analysis showed that \textit{SCUBE3} expression varied among tissues. Although \textit{SCUBE3} mRNA was present in all tissues, the levels varied considerably, with higher expression observed in backfat bone marrow and cartilage tissues and the lowest expression observed in the heart (Figure 3). The level of \textit{SCUBE3} expression is reportedly very high in human primary osteoblasts and long bones; thus, \textit{SCUBE3} has been suggested to play an important role in bone cell biology (Wu et al., 2004). According to the mouse gene expression data in MGI (http://www.informatics.jax.org/), \textit{SCUBE3} is expressed in mouse limbs. In pigs, quantitative trait loci mapping showed that the region of 12.19-121.08 Mb on SSC 7 has pleiotropic and significant effects on limb bone length (Mao et al., 2008). Furthermore, a genome-wide association study of limb bone length reported that 39 significant single nucleotide polymorphisms are located in an 11.9-Mb (31.24-43.17) region of SSC 7 and that \textit{SCUBE3} is located in this region (Zhang et al., 2014). The region from 17.05-45.42 Mb on SSC 7 was previously reported to have a significant and pleiotropic
influence on body length (Soma et al., 2011). A genome-wide association study of growth traits in pigs reported a significant region at 31.24-36.00 Mb on SSC 7; \textit{SCUBE3} was considered to be a candidate gene influencing the size of bones, body length, and body height of pigs (Wang et al., 2014). It is likely that \textit{SCUBE3} is highly expressed in cartilage, which may play an import role in limb bone length. CNVs are a significant factor in pig genomic variation and they can directly affect gene expression levels by changing the gene dosage or have an indirect effect by disrupting gene expression regulation (Lupski and Stankiewicz, 2005; Freeman et al., 2006; Stranger et al., 2007). Our real-time quantitative PCR identified CNV variation in 11 individuals. Since CNVs are considered to play important roles in normal phenotypic variability and disease susceptibility, they are good markers for identifying economic- and disease-related traits in domestic animals (Wang et al., 2012). \textit{SCUBE3} may exert its effects via CNV.

In summary, we obtained and sequenced the full-length porcine \textit{SCUBE3} cDNA, determined its tissue expression patterns, and identified CNVs. These results indicate that \textit{SCUBE3} plays important roles in bone cell biology and limb bone length and that \textit{SCUBE3} may exert these effects through CNV. \textit{SCUBE3} CNVs will provide robust markers for analyzing phenotypic traits. Our findings provide a basis for further investigation of the physiological functions and regulatory mechanisms of \textit{SCUBE3}.

ACKNOWLEDGMENTS

Research supported by the Agricultural Science and Technology Innovation Program (#ASTIP-IAS02), the Chinese Academy of Agricultural Sciences Foundation (#2014ZL006 and #2015ywf-zd-9), and the National Key Technology R&D Program of China (#2011BAD28B01), earmarked fund for the Modern Agro-industry Technology Research System.

REFERENCES


Study of porcine SCUBE3


