Full Length cDNA, Genomic Organizations and Expression Profiles of the Porcine Proteasomal ATPases *PSMC5* Gene

Y. F. Wang, M. Yu, B. Liu, B. Fan, H. Wang, M. J. Zhu and K. Li*
Lab of Molecular Biology and Animal Breeding, School of Animal Science and Technology
Huazhong Agricultural University, Wuhan 430070, P. R. China

ABSTRACT : *PSMC5* subunit, which belongs to the 26S proteasomal subunit family, plays an important role in the antigen presentation mediated by MHC class I molecular. Full-length cDNA of porcine *PSMC5* was isolated using the *in silico* cloning and rapid amplification of cDNA ends (RACE). Amino acid was deduced and the primary structure was analyzed. Results revealed that the porcine *PSMC5* gene shares the high degree of sequence similarity with its mammalian counterparts at both the nucleotide level and the amino acid level. The RT-PCR was performed to detect the porcine *PSMC5* expression pattern in seven tissues and the result showed that high express level was observed in spleen, lung, marrow and liver while the low express level was in muscle. The full-length genomic DNA sequence of porcine *PSMC5* gene was amplified by PCR and the genomic structure revealed that this gene was comprised by 12 exons and 11 introns. Best alignment of the cDNA and genomic exon DNA sequence presents 4 mismatches and this information potentially bears further study in gene polymorphisms. ***(Asian-Aust. J. Anim. Sci. 2004. Vol 17, No. 7 : 897-902)***

Key Words : Sequences Analysis, Expression Profiles, Porcine, *PSMC5*

INTRODUCTION

It is firmly established that the 26S proteasome plays key role in degrading proteins, which marked by the ubiquitin chains (Hershko and Ciechanover, 1998). 26S composed by a core proteasome complex named as 20S and two associated regulators termed as PA700 and PA28 (Ma et al., 1992,1994). The PA700, or 19S regulator, can be subdivided into two groups, one is AAA (ATPases Associated with diverse cellular Activities) family, containing 6 ATPases, which encoded by the homologous genes and shared the high degree of conservation evolutionarily (DeMartino et al., 1994). Six putative ATPases are *PSMC1, PSMC2, PSMC3, PSMC4, PSMC5* and *PSMC6* and they all bear a highly conserved ATPase domain (AAA domain) and the leucine zipper-like domain. But, to a surprising extent, these ATPases are functionally non-redundant and have diverse cellular functions respectively (Rubin et al., 1998). For example, yeast Rpt1 (porcine *PSMC2* gene homologue) mutant displayed a G_1 cell-cycle defect and was strongly growth defective (Rubin et al., 1998) while *PSMC3* was responsible for mediating inhibition of the cellular proliferation and transformation of erbB-inhibited cells (Park et al., 1999). Mutation of the members of AAA family may be is associated with the disease (Tsukamoto et al., 1995). Another group is non-ATPase family, which consisted of at least 15 subunits. Most of them differ in structure and their function is still elusive.

Human *PSMC5* gene, also known as the thyroid hormone receptor-interacting protein (*TRIP1*), was identified in yeast two-hybrid to isolate the proteins that mediated the transcriptional response of the thyroid hormone receptor (Lee et al., 1995). Although the sequence data of this gene has been published, its potential biological function has so far not been studied except that it involved in the ubiqutin-proteasome pathway and the antigen presentation mediated by MHC class I molecular. Human *PSMC5* gene was mapped to HAS17q24-25 by fluorescence in situ hybridization (FISH) (Hoyle et al., 1997), while Tanahashi et al demonstrated that the location of *PSMC5* was HSA17q23.1-q23.3 (Tanahashi et al., 1998). By PCR amplification of a partial sequence of *PSMC5* in a panel of pig and Chinese hamster cell hybrid (IMpRH), the porcine *PSMC5* gene is located on chromosome 12p14 (Wang et al., 2003), the same region as the glial fibrillary acidic protein (*GFAP*) and *SW60* (Yu et al., 2001; Liu et al., 2002), which is in agreement with the high level of evolutionary conservation between SSC12 and HAS17.

Association studies between gene polymorphism and some performance traits have been widely used for the study of gene potential function. However, doing this work must be on the basis of the gene sequences. In present study, we report the cDNA cloning, genomic organization and expression profile in seven tissues of porcine *PSMC5* gene. This study found the basis for further investigating some new genetic variants, even the biological and physiologic function of *PSMC5* gene.

삭제됨: .

MATERIALS AND METHODS

Isolation of the full-length cDNA of porcine *PSMC5* gene

Full-length cDNA was isolated using the approach

* Corresponding Author: Kui Li. Tel: +86-27-87281306, Fax: +86-27-87280408, E-mail: lkxblghi@public.wh.hb.cn
Received November 5, 2003; Accepted March 17, 2004

Table 1. Primers for amplification of the full-length genomic DNA of porcine *PSMC5* gene

Primer name	Primer sequences (5'-3')	Binding region	Annealing temperature (°C)	Size (bp)
1PL	AAGAGGGAAGATGGCGCTTG	5'UTR and exon1	64	603
1PR	AGATAATACTGGCGGAGTCC	Exon2		
2PL	ATGGAGCTGGAAGAGGGGAA	Exon2	62	1,102
2PR	CCTGCGGAGATTCTGGCTCT	Exon3		
3PL	CTGATTGTGAATGACAAGAG	Exon3	62	989
3PR	ATCATCAGTGACACCAGTGG	Exon6		
4PL	CGTGGACAAGAACATCGACA	Exon5	61	966
4PR	CGTGCGCTGGACTTCACTGT	Exon8		
5PL	TCGACTCCATTGGCTCCTCG	Exon8	64	839
5PR	CACCTGCCTCACTTCCATAG	Exon12 and 5'UTR		

through a combination of *in silico* cloning and RACE. Blast (http://www.ncbi.nlm.nih.gov/blast/) searches were carried out with the cDNAs of human *PSMC5* gene (GenBank accession number is NM_002805) against the other-EST databases for the porcine ESTs. Those ESTs, which shared at least more than 80% homology to the corresponding human cDNA, were selected and assembled into contig for gene specific primers design. Gene specific primers designed based on the contig were 5'-GCTTGTCC AGTCCACCAATCATCTC-3' for 5' RACE and 5'-GATA TGATTGGTGGACTGGACAAGC-3' for 3' RACE. Total RNA of a mature Xiang pig (Guizhou province) was extracted from the spleen tissue using the TRIzol reagent kit (Life Technologies, Grand Island, USA). The synthesis of the first-strand cDNA and the 5', 3' RACE PCR were performed according to the standard protocols of the SMART™ RACE cDNA Amplification Kit (Clontech Inc, Palo Alto, CA, USA). RACE products were purified with Wizard PCR Preps DNA Purification System (Promega, Madison,WI, USA) and cloned into the pGEM T-easy vector (Promega, Madison, WI, USA), then several random clones were selected and sequenced commercially.

Sequences analysis

ORFs were found and the amino acid sequences were deduced with the program Seqman (DNA star, Madison, Wis). The analysis of putative sorting signal information and structural motif was obtained with the computer program PSORT II (http://www.nibb.ac.jp).

Expression pattern determination

RT-PCR was used to determine the expression pattern of *PSMC5* gene. PCR primer pairs were 5'-GCAGATGGA GCTGGAAGAGG-3' (forward) and 5'-TGCATGACCTTG GCTACGGC-3' (reverse). Total RNAs were isolated from the adult porcine skeletal muscle, heart, lung, liver, spleen, marrow and kidney and reverse transcription was performed as described earlier in detail (Pan et al., 2003). The parameter of PCR was 4 min at 94°C followed by 26 cycles of 45 s at 94°C, 45 s at 62°C, 1 min at 72°C and a final extension of 5 min at 72°C. Amplification of *GAPDH* cDNA was performed as a positive control. 10 μl PCR products were used to detect the expression profile.

Genomic DNA amplification and sequence analysis

PCR was used to amplify the genomic DNA fragments of *PSMC5* gene. Five pair primers for *PSMC5* developed from each full-length cDNA were listed in Table 1. Conditions for amplification were 4 min at 94°C followed by 35 cycles of 45 s at 94°C, 45 s at the annealing temperature (Table 1), 1 min at 72°C, and a final extension of 5 min at 72°C. PCR products were purified, cloned and sequenced as described above. DNA sequences were compiled using the DNA star software (ABI prism). Exon/intron boundaries were identified by alignment of the cDNA and DNA sequences of this gene. Repetitive sequences were determined with the Repeatmasker program (http://ftp.genome.washington.edu/RM).

RESULTS

Full-length cDNA of porcine *PSMC5* gene

Based on the bioinformatics analysis, 12 porcine ESTs (Acc. No. BI181574, BF703225, BG835422, BI399733, BQ604522, BF709102, BF189309, BG382905, BG382607, AW619546, AW618989 and BF198228, respectively) were searched and these overlapping ESTs could be assembled into one contig of 1.32 kb. The gene specific primers for RACE were designed according to the contig.

The nucleotide sequence of the *PSMC5* cDNA and the primary structure of the *PSMC5* protein deduced from the cDNA sequence are shown in Figure 1. PCR amplification showed that the 5'RACE product was 549 bp and the 3'RACE product was 990 bp. Computer analysis of the combined nucleotide sequence revealed a 1,221 bp ORF flanked by a 21 bp 5'UTR and 118 bp 3'UTR. The putative polyadenylation signal (AATAAA) could be found in 3'-UTR. The homologous analysis revealed that the porcine *PSMC5* is 92% identical to human *PSMC5* gene (Acc. No. NM_002805), and 90% to its mouse homologue (Acc. No. NM_008950). Computer analysis showed that the *PSMC5* gene encoded 406 amino acids with a calculated molecular

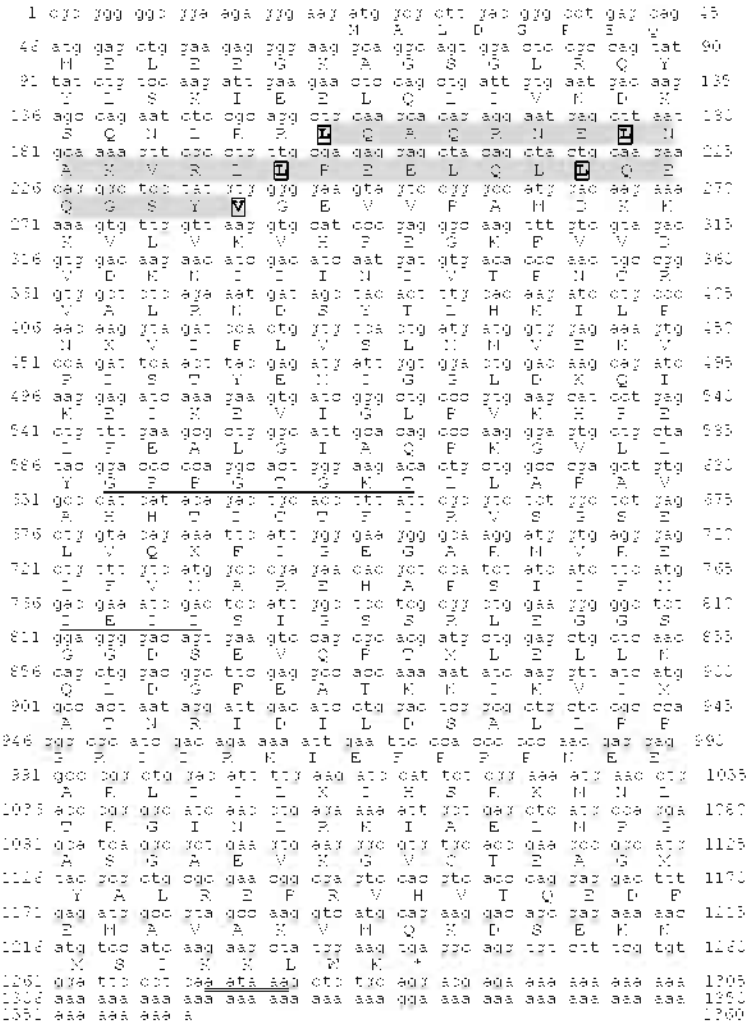


Figure 1. cDNA sequences and deduced amino acid sequences of the porcine *PSMC5* gene. The leucine zipper - like domain was shaded and the hydrophobic amino acids forming the heptad repeats was boxed. The putative consensus ATP-binding motif "GPPGTGKT" and the ATP hydrolysis motif "DEID" in the CAD domain were underlined. The double underlined sequence. aataaa, indicates the putative polyadenylation signal.

mass of 45.6 kDa and predicted isoelectric point of 7.70. Cytoplasmic/Nuclear discrimination predicted that *PSMC5* might exist predominantly in the cytoplasm with probability of 89%.

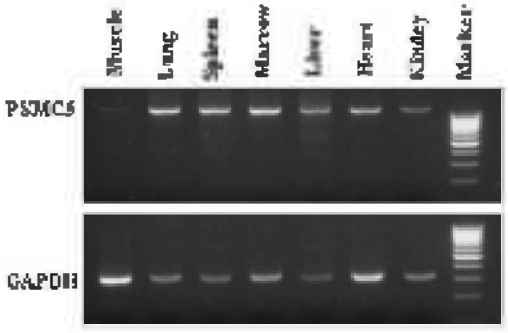
The N-terminal regions (exon3 and exon4) of porcine *PSMC5* contain the leucine zipper (LZ) domain and the exon6 to exon11 encode the conserved ATPase domain (CAD).

Expression profiles determination

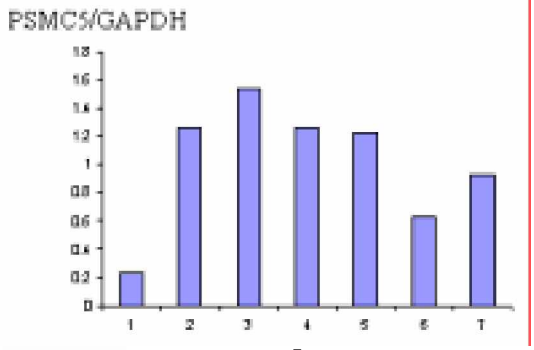
The RT-PCR was performed to detect the porcine *PSMC5* expression pattern in seven tissues and the PCR products of *PSMC5* were normalized assuming that the expression of *GAPDH* is the same level in the entire sample (according to their optical intensity value). Result showed that porcine *PSMC5* gene was ubiquitously expressed. High express level was observed in spleen, lung, marrow and liver while the low express level was in muscle (Figure 2).

Genomic structure of *PSMC5* gene

The comparison of cDNA and DNA sequences established that the *PSMC5* gene spans 4,145 bp and is made up of 12 exons, which are in the size range of 24-231 bp. The introns of *PSMC5* vary from 78 bp to 994 bp in size.



(A)



(B)

Figure 2. (A) The RT-PCR results of the porcine *PSMC5* gene. *GAPDH* was used as control for semi-quantitative analysis. (B) The bar graph of the values of *PSMC5*/*GAPDH*. 1 to7 represent the muscle, lung, spleen, marrow, liver, heart and kidney, respectively. Lane 8 is the 100 bp ladder.

Table 2 presents the conserved sequence at the boundary of the splicing sites. All of splice sites conformed to the GT/AG consensus sequence with no exception. For any genes, there are three different positions 0, 1 or 2 in relation to the triplet codon existed in the intron/exon boundaries. As for porcine *PSMC5* gene, except the intron/exon boundaries of exons 3 and 7 were class 1, the remaining

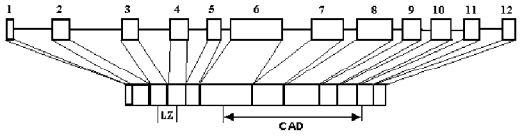


Figure 3. Schematic diagram of the genomic organization of pig *PSMC5* and the location of the LZ (leucine zipper-like) domain and CAD (conserved ATPase domain). Boxes represent the exons and lines represent introns.

boundaries are falling into class 0. Repeat sequence analysis revealed that there are two repeat sequences in this gene, one is SINE/MIR which located in intron 2 (1.167 bp-1,346 bp) and another is LINE/L2 which located in intron 6 (2.865 bp-2,968 bp). These repetitive sequences, occupied only 6.85% of the genomic sequences.

DISCUSSION

It is clearly now that the ATPase can provide the energy for unfolding the proteins in the selective proteolysis in cells and it has been suggested that the proteasomal ATPases were thought to play the role in the RNA metabolism or processing including transcriptional or post-transcriptional regulation (Makino et al., 1996). But the function of the *PSMC5* subunit is still elusive.

삭제됨: as

As we know, gene sequence is an entry point to study the gene expression and function. In the present study, we isolated the full-length cDNA and genomic DNA sequences of porcine *PSMC5* gene. Our results revealed that the porcine *PSMC5* gene shares the high sequence identity with its mammalian counterparts at both the nucleotide level and the amino acid level, which suggested the significance of their biological functions. As the members of the AAA gene family, porcine *PSMC5* gene bears the significant feature of this family that harbors the putative CAD domain and the LZ domain. There is a putative consensus ATP-binding motif "GPPGTGKT" and the ATP hydrolysis motif "DEID" in the CAD domain (Figure 3). Motif "GPPGTGKT" mutants can result in the diverse phenotypes (Rubin et al.,

Table 2. Intron/exon boundaries of porcine *PSMC5* gene

#exon	intron	Exon		Intron	#intron	Size of intron (bp)
1		ATG/GCA...	24bp	gtatggcggcggctgcggag...	1	519
2	...aataattaaaactgtcttcag	ATG/GAG...	70bp	gtaagaggtcctgaggagac...	2	994
3	...ttgcctcttgtcctccacag	CTG/ATT...	70bp	gtgagggtgaagcatgggaa...	3	429
4	...tytgtatgtttcctatctag	TT/CGC...	97bp	gtaaaggcatcatgacccac...	4	160
5	...atcaagcctgtgtttctgtag	GTG/CAG...	57bp	gtgagtgcatgaggtgaggt...	5	80
6	...aactctctctctgcccag	ATG/ACA...	231bp	gtgagccggcgctctctggg...	6	269
7	...gtcctgcactgctctctag	GGA/GTG...	127bp	gtaagtggccagaagcgaaa...	7	94
8	...ccagcttggcctctgcacag	GG/GCA...	131bp	gtgtggtgggtcaggcgag...	8	89
9	...cacctctgtgtccactcag	ATT/ATC...	99bp	gtttgtgatgggcagtactg...	9	78
10	...tttttatctccatctctag	GCC/CGA...	111bp	gtaactgyagtgcccaagga...	10	90
11	...gtctctctgcctttgcccag	GGC/GTG...	87bp	gtacaggcctccctctctg...	11	104
12	...aatgctctctctctctccag	ATC/ATG...	54bp	...AAG/TGA		

1998). The lysine residue "K" of the motif "GPPGTGKT" has been characterized to interact with the phosphate groups of ATP (Walker et al., 1982). Rubin and his co-workers introduced a site-directed mutation of substituting the Lys residue with the Ser or Arg residues in yeast RPT 6 (porcine *PSMC5* homologues) genes to study its phenotypic changes in ATP binding and hydrolysis. Results showed that the conservative mutants (Lys to Arg) of the yeast RPT6 gene were viable while the non-conservative mutants (Lys to Ser) were non-viable (Rubin et al., 1998). The rat *PSMC5* gene was supposed to participate in the homo-dimerization or hetero-dimerization through the LZ domains of the N-terminal (Makino et al., 1996). As we know, seeking the single nucleotide polymorphism (SNP) of the important functional region of the candidate gene and taking the association analysis with the economic traits is the very useful tool to study the gene function. Due to these important domains of the *PSMC5* gene, we are going to scan the SNP of these motifs and expecting to develop the useful genetic markers for marker assistant selection (MAS).

RT-PCR analysis demonstrated that porcine *PSMC5* gene expressed in all seven tissues studied. But the very low expression level was obtained in muscle while the higher expression level was in spleen, lung, marrow and liver, especially in the spleen. It is now clearly that genes involved in the antigen presentation mediated by MHC class I have the higher expression in the immunological tissues, such as spleen and peripheral blood leukocytes (Li et al., 1999). Our results of that porcine *PSMC5* gene expressed strongly in spleen and marrow are consistent with the previous reports.

On the basis of the alignment of the cDNA and genomic DNA sequence, the porcine *PSMC5* gene harbors 12 exons and 11 introns. All exon-intron boundaries are in agree with the GT/AG rule (Table 2). The CLASTAL W program (Thompson et al., 1994) was used to align the cDNA sequence and the corresponding genomic DNA exon sequence of porcine *PSMC5* gene. The best alignment showed that there are 4 mismatches between them and all of mismatches located in the exon6, interestingly. These changes were conformed by sequencing the products three times and all of them occurred at the third base pairs of the triplet codon and did not lead to the amino acids exchange. This contains the great potential polymorphism information. The polymorphism detection and association studies are currently in progress in our lab.

ACKNOWLEDGEMENTS

This research was financed by National High Science and Technology Foundation of China (2001AA222241), Key Project of National Basic Research and Developmental Plan (G2000016103) of China, National Natural Science Foundation of China (30270952), the Key Project of National Natural Science Foundation of China (30330440), National Outstanding Youth Foundation of China (39925027), Backbone teacher foundation, Doctorate foundation and Key Project of Ministry of Education of China.

REFERENCES

DeMartino, G. N., C. R. Moomaw, O. P. Zagnitko, R. J. Proske, M. Chu-Ping, S. J. Afendis, J. C. Swaffield and C. A. Slaughter. 1994. PA700, an ATP-dependent activator of the 20S proteasome, is an ATPase containing multiple members of a nucleotide-binding protein family. J. Biol. Chem. 269:20878-20884.

Hershko, A. and A. Ciechanover. 1998. The ubiquitin system. Annu. Rev. Biochem. 67:425-479.

Ma, C. P., C. A. Slaughter and G. N. Demartino. 1992. Identification, purification and characterization of a protein activator (PA28) of the 20S proteasome (Macropain). J. Biol. Chem. 267:10515-10523.

Ma, C. P., J. H. Vo, R. J. Proske, C. A. Slaughter and G. N. Demartino. 1994. Identification, purification, and characterization of a high-molecular weight, ATP-dependent activator (PA700) of the 20S proteasome. J. Biol. Chem. 269:3539-3547.

Rubin, D. M., M. H. Glickman, C. N. Larsen, S. Dhruvakumar and D. Finley. 1998. Active site mutants in the six regulatory particle ATPases reveal multiple roles for ATP in the proteasome. EMBO. J. 17:4909-4919.

Park, B. W., D. M. O'Rourke, Q. Wang, J. G. Davis, A. Post, X. L. Qian and M. I. Greene. 1999. Induction of the tat-binding protein 1 gene accompanies the disabling of oncogenic erbB receptor tyrosine kinases. Proc. Natl. Acad. Sci. 96:6434-6438.

Tsukamoto, T., S. Miura, T. Nakai, S. Yokota, N. Shimozawa, Y. Suzuki, T. Orii, Y. Fujiki, F. Sakai, A. Bogaki, H. Yasumo and T. Osumi. 1995. Peroxisome assemble factor-2, a putative ATPase cloned by functional complementation on a peroxisome-deficient mammalian cell mutant. Nat. Genet. 11:395-401.

Lee, J. W., H. S. Choi, J. Gyuris, R. Brent and D. D. Moore. 1995. Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor. Molec. Endocr. 9:243-254.

Hoyle, J., K. H. Tan and E. M. C. Fisher. 1997. Localization of genes encoding two human one-domain members of the AAA family: *PSMC5* (the thyroid receptor-interacting protein, TRIP1) and *PSMC3* (the tat-binding protein, TBP1). Hum. Genet. 99:285-288.

Tanahashi, N., M. Suruki, T. Fujiwara, E. Takahashi, N. Shimbara, C. H. Chung and K. Tanaka. 1998. Chromosomal localization and immunological analysis of a family of human 26S proteasomal ATPase. Biochem. Biophys. Res. Commun. 243:229-232.

Wang, Y. F., M. Yu, M. Yerle, B. Liu, S. H. Zhao, T. A. Xiong, B. Fan and K. Li. 2003. Mapping of genes encoding four ATPase and three non-ATPase components of the pig 26S proteasome. Anim. Genet. 34:393-395.

Liu, B., G. L. Jin, S. H. Zhao, M. Yu, T. A. Xiong, Z. Z. Peng and K. Li. 2002. Preparation and analysis of spermatocyte meiotic

서식 있음: 글머리 기호 및 번호 매기기

서식 있음: 글머리 기호 및 번호 매기기

pachytene bivalents of pigs for gene mapping. Cell. Res. 12(5-6):401-405.

Yu, M., S. H. Zhao, B. Liu, T. A. Xiong, P. W. Pan and K. Li. 2001. Isolation and regional localization of the porcine glial fibrillary acidic protein (*GFAP*) gene. J. Anim. Sci. 79:2754.

Pan, P W., S. H. Zhao, M. Yu, B. Liu, T. A. Xiong and K. Li. 2003. Identification of differentially expressed genes in the longissimus dorsi muscle tissue between duroc and erhualian pigs by mRNA differential display. Asian-Aust. J. Anim. Sci. 16:1066-1070.

Makino, Y., S. Yogosawa, M. Kanemaki, T. Yoshida, K. Tamano. T. Kishimoto, V. Moncollin, J. M. Egly, M. Muramatsu and T. Tamura. 1996. Structures of the rat proteasomal ATPases: determination of highly conserved structural motifs and rules for their spacing. Biochem. Biophys. Res. Commun. 220:1049-1054.

Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay. 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO. J. 1:945-951.

Li, Y. H., J. Chambers. J. Pang. K. Ngo. P. A. Peterson. W. P. Leung and Y. Yang. 1999. Characterization of the mouse proteasome regulator PA28b gene. Immunogenetics 49:149-157.

Thompson, J. D., D. G. Higgins and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic. Acids. Res. 22:4673-4680.

서식 있음: 글머리 기호 및 번호 매기기

서식 있음: 글머리 기호 및 번호 매기기