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

Polymorphisms of KAP6, KAP7, and KAP8 genes in four Chinese sheep breeds

Y.X. Liu[1], G.Q. Shi[2], H.X. Wang[1], P.C. Wan[2], H. Tang[2], H. Yang[2] and F. Guan[1,3]

[1]College of Life Sciences, China Jiliang University, Hangzhou, China
[2]The Key Sheep Breeding and Reproduction Biotechnology Laboratory of Xinjiang Production and Construction Group, Shihezi, China
[3]Zhejiang Provincial Key Laboratory of Biometrology and Inspection & Quarantine, Institute of Plant Environmental Biology and Measurement Technology, China Jiliang University, Hangzhou, China

Corresponding author: F. Guan
E-mail: guanfengzgjl@163.com

Genet. Mol. Res. 13 (2): 3438-3445 (2014)
Received February 6, 2013
Accepted July 4, 2013
Published April 30, 2014
DOI http://dx.doi.org/10.4238/2014.April.30.5

ABSTRACT. High glycine-tyrosine proteins (HGTPs), also known as keratin-associated proteins (KAPs), play a key role in the major structures and mechanical properties of wool fiber. Sheep HGTPs consist of three multigene families: KAP6, KAP7, and KAP8 genes. Polymorphisms of these three genes have been proposed to have important effects on wool fiber traits. The aim of the present study was to identify polymorphisms of the KAP6, KAP7, and KAP8 genes in four sheep breeds, including Chinese Merino superfine wool sheep, Hu sheep, a Merino x Hu crossed breed, and Romney sheep. Polymerase chain reaction (PCR) product direct sequencing, PCR-single-strand conformation polymorphism, and cloned sequencing methods were used to find genetic variation and identify polymorphisms in these genes. The Mutation Surveyor v3.97 software was used to analyze the sequences. These methods revealed six different sequences of the KAP6 gene, two different sequences of the KAP7 gene, and five different sequences of the KAP8 gene. Accordingly,

three (with frequencies >1%) single nucleotide polymorphisms (SNPs) of the KAP6 gene, one SNP of the KAP7 gene, and five SNPs of the KAP8 gene were detected. Interestingly, some of these sequences were present in only certain sheep breeds, thereby suggesting that these special allele sequences could be used as candidate genes of wool characteristics in further studies.

Key words: Sheep; Keratin-associated protein; Polymorphism; Wool fiber diameter

INTRODUCTION

Wool fiber is a highly organized structure, whose main histological components include the cuticle, the cortex (orthocortex, paracortex, and mesocortex), and the medulla (Plowman et al., 2009). Almost all of the fiber (90%) is composed of the cortex, which consists of keratin intermediate filaments embedded in a matrix of keratin-associated proteins (KAPs) (Kuczek and Rogers, 1987; Itenge-Mweza et al., 2007). KAPs cross-link the keratin intermediate filaments with inter-chain disulfide-bonding, and form the wool fiber after keratinization (Gong et al., 2011). Wool fiber characteristics, such as diameter, crimps, and length, are essential parameters of the wool trait, as well as important indications of the spinning efficiency of the wool (Plowman et al., 2009). Previous studies have shown that variations in the protein compositions of the orthocortex and paracortex, which make up nearly the whole wool fiber, were highly related with wool trait parameters (Marshall et al., 1991; Plowman et al., 2000). Particularly, the formation of KAP composites gives the wool special mechanical attributes of strength, inertness, and rigidity (Parry and Steinert, 1992; Powell et al., 1994; Plowman, 2003; Plowman et al., 2009).

Sheep KAPs are characterized based on their high proportions of cysteine residues or glycine and tyrosine residues (Powell and Rogers, 1997; Gong et al., 2011). According to their amino acid compositions, the wool matrix of KAPs are divided into three groups: the high-sulfur proteins, the ultra-high-sulfur proteins, and the high-glycine-tyrosine proteins (HGTPs) (Cockett et al., 2001). The amount of protein from different groups varies both spatially and developmentally within the wool fiber (Powell and Rogers, 1990). Specifically, the most significant variation was found with respect to HGTP contents, which varied both within and between species, ranging from less than 3% in human hair and in the wool of Lincoln sheep, to 13% in Merino sheep, and up to 30-40% in echidna quill (Gillespie and Broad, 1972; Parsons et al., 1994).

HGTPs are the smallest of the sheep wool keratin proteins (Mr = 6000-9000), and consist of three gene families: KAP6, KAP7, and KAP8. KAP genes are 0.6-1.5 kb in size and do not contain introns (Cockett et al., 2001). KAP7 and KAP8 are each coded by a single gene, whereas KAP6 is encoded by a multigene family (Kuczek and Rogers, 1987). To date, abundant polymorphisms have been detected in sheep KAP6, KAP7, and KAP8 genes, which showed possible associations with variations in fiber diameter, staple strength, and brightness (Powell et al., 1983; Beh et al., 2001, 2002).

The objective of this study was to investigate the polymorphisms of the KAP6, KAP7, and KAP8 genes in four Chinese sheep breeds, which may be helpful in the process of identifying candidate genes for wool traits and in cultivating fine wool sheep breeds.

MATERIAL AND METHODS

Sample collection and DNA extraction

In this study, a total of 200 samples of Chinese Merino sheep, Romney, and a Merino x Hu crossed breed were collected from the Key Laboratory of Sheep Breeding and Reproduction Biotechnology in the Xinjiang production and construction group, and Hu sheep samples were collected from the Zhejiang Province. Notably, the Chinese Merino breed is a superfine wool sheep breed, which has a smaller wool fiber diameter than the other three breeds.

Skin samples were collected from sheep ears and transported to the laboratory at 4°C. Genomic DNA was extracted using the standard phenol-chloroform method.

Primers and polymerase chain reaction (PCR) amplifications

Primers used in this study were designed by the Primer5.0 software, according to the published nucleic acid sequences of KAP6, KAP7, and KAP8 genes (M95719, X05638, X05639). Amplified fragments covered the whole 5'-untranslated region (UTR), the whole coding sequence (CDS) region, and part of the 3'-UTR of these three genes. All primers were synthesized by Invitrogen Biotechnology Co., Ltd.

PCR amplifications of all detected loci were performed in a total 20-μL reaction system, containing 2 μL 10X buffer, 1.6 μL 2.5 mM dNTP mixture, 0.4 μL 10 μM of each primer, 0.2 μL *Taq* polymerase, 1 μL DNA template, and 14.8 μL ddH_2O. The thermal profile consisted of the following: 5 min at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at 54°-57°C, and 40 s at 72°C, with a final extension for 10 min at 72°C. Amplifications were carried out in S1000 (Bio-Rad, Hercules, CA, USA). The primer sequences and some PCR conditions are shown in Table 1.

Table 1. Sequences of primers and amplification conditions.

Gene name	Primer sequence (5'-3')	Annealing temperature (°C)	Amplicon length (bp)
KAP6	Forward: GGCTACTACGGAAACTACTATGGCG Reverse: ATAGCCACAGCCCAGCCTGC	55	498
KAP7	Forward: GCGGAAGCTACTTCCCAGGCTA Reverse: TGCTAAAGCTGCTTCCACCGAAACC	57	681
KAP8	Forward: CCAGCACCGTCTTCCCAGGTT Reverse: CATAGCCGAAGCCATAGCCCAC	54	525

PCR-single-strand conformation polymorphism (SSCP) analysis

Vertical 10% polyacrylamide (37.5:1 acrylamide:bis-acrylamide) gels (18 x16 cm, 1 mm thick spacers, 28-well comb) were prepared containing 0.5X TBE (44.5 mM Tris, 44.5 mM orthoboric acid, 1 mM Na_2EDTA, pH 8.0) and 0.2% glycerol, which were polymerized using 10% ammonium persulfate and N,N,N',N'-tetramethylethylenediamine. Gels were pre-electrophoresed for 30 min at 60 V. Three-microliter aliquots of each amplification were mixed with 15 μL loading dye (95% formamide, 10 mM Na_2EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol), denatured by heating at 95°C for 5 min, and immediately placed on ice water before loading the whole aliquots. The gels were then electrophoresed at 250 V for 18 h at 4°C with 0.5X TBE running buffer, followed by silver staining, according to the

method described in Byun et al. (2009).

Sequencing and analysis

The homogenous PCR fragments of different SSCP patterns in different breeds were directly sequenced by the Shanghai Sunny Biotechnology Co., Ltd., and the heterogeneous PCR fragments were cloned into a pGEM-T vector before sequencing. Sequence alignments, translations, and comparisons were carried out using Mutation Surveyor v3.97. Single nucleotide polymorphisms (SNPs) were determined through sequence alignment. Any mutation with a frequency of less than 1% was not regarded as an SNP locus.

RESULTS AND DISCUSSION

KAP6 gene polymorphisms

The KAP6 gene is known to be a complicated and highly polymorphic family of sheep HGTPs. Multiple-protein spots were separated from total glycine/tyrosine-rich HGTPs using auto-radio graphs of two-dimensional polyacrylamide gels (Powell and Rogers, 1990). Subsequently, Northern blot analysis showed that more than one member of the KAP6 family was expressed in sheep skin (Fratini et al., 1993). Therefore, it was suggested that the KAP6 gene contained multiple-family members. Recently, five different ovine KAP6 sequences were detected, and these sequences were divided into three groups: KAP6-1, KAP6-2, and KAP6-3 (Gong et al., 2011).

The proportions and spatial arrangement of orthocortical and paracortical fibers, which are believed to be highly relevant to the crimps and curls of wool fiber, are the two major types of cortical keratinocytes (Powell and Rogers, 1990). The results of an *in situ* hybridization study suggested that the KAP6 gene family was expressed in cells of the hair shaft cortex after transcription of the hair keratin intermediate filament gene (Fratini et al., 1993; Powell and Beltrame, 1994; Powell et al., 1994). This spatial and timely arrangement in its expression pattern indicated that KAP6 family genes might play an important role in determining wool trait characteristics, such as diameter and crimps.

In the present study, six different sequences of the KAP6 gene were obtained and designated as O, A, B, C, D, and E. However, not all of the six different sequences were found in all four sheep breeds; sequences O, A, B, and C were found in Chinese Merino sheep, sequences O, A, D, and E were found in the Hu sheep breed, sequences O and A were found in Romney sheep, and sequences O, A, B, C, D, and E were found in the Chinese Merino x Hu sheep crossed breed. The results of sequence alignment revealed that sequence O was similar to the reported KAP6 gene (GenBank No. M95719). The other sequences, A-E, were different from each to varying degrees, and these five sequences were deposited into GenBank under accession Nos. JQ043155-JQ043159. Based on sequence alignment analysis, sequences O and D shared nearly the same sequence with that of the KAP6-1 allele. Sequence A was identical to that of the KAP6-2 allele within the corresponding parts, with the exception of two bases. Sequences B, C, and E showed great difference from the previously reported KAP6 gene sequences. Interestingly, sequences B and C were found in Chinese Merino sheep and in the Chinese Merino x Hu crossed breed, but not in Romney or Hu sheep. In addition, Chinese Merino x Hu crossed sheep

had smaller wool diameters than those of Hu sheep, but larger diameters than those of Chinese Merino sheep. Therefore, it was suggested that these two sequences might be unique to Chinese Merino sheep. Considering this, further study is required to determine whether sequences B and C are associated with variations in wool diameter. Comparisons with the reported sequence of M95719 revealed that whereas mutations in sequences B and C were identical in the noncoding region, they were different in the CDS region (Table 2). Statistical analysis revealed that these mutations were found in the CDS region with lower frequencies (<1%), which suggested that these mutations might be meaningless with respect to wool traits. Moreover, previous studies have shown that the 5'- and 3'-UTRs played critical roles in post-transcriptional regulation of gene expression (Köhler, 2007; Piccone et al., 2009; Zhou et al., 2011; Song et al., 2012). Thus, the effects of the mutations in the UTRs on expression of the KAP6 gene, and its association with wool need further study.

Table 2. Alignment of the KAP6 allele sequence (A-E).

S	M95719	Alleles					
		O	A	B	C	D	E
881	G	-	A	-	-	-	-
890	T	-	A	-	-	-	-
891	C	-	T	-	-	-	-
929	G	-	A	-	-	-	-
953	***C***	**-**	***T***	**-**	**-**	***T***	***T***
991	***A***	**-**	***T***	***T***	***T***	**-**	***T***
1002	C	-	-	T	T	-	-
1006	C	-	-	-	-	-	A
1015	C	-	-	T	T	-	-
1035	G	-	-	C	C	-	-
1062	A	-	-	-	G(K→R)	-	-
1094	A	-	G(S→G)	-	-	-	-
1114	T	-	-	-	C(W→R)	-	-
1171	T	-	-	-	C(W→R)	-	-
1211	T	-	-	-	G(Y→D)	-	-
1218	A	-	G(Y→C)	-	-	-	-
1226	C	-	-	-	-	G(R→D)	-
1229	T	-	C(S→P)	-	-	-	-
1236	G	-	A(C→Y)	-	-	-	-
1240	A	-	C	-	-	-	-
1241	A	-	T(S→C)	-	-	-	-
1255	C	-	T	-	-	-	-
1298	T	-	G	-	-	-	-
1299	C	-	A	-	-	-	-
1304	G	-	A	-	-	-	-
1305	***A***	**-**	***G***	***G***	***G***	**-**	***G***
1317	C	-	T	-	-	-	-
1323	T	-	C	-	-	-	-

S = position of a substituted nucleotide in this study. The shaded area represents the coding region while the unshaded area represents the noncoding region and a dash represents the same nucleotide as the control sequence. The contents in brackets indicate that the substitution of nucleotide base results in amino acid change. The SNP sites are shown in bold and italic (the same an the following tables).

KAP7 gene polymorphisms

Sheep KAP7 is encoded by a single gene that does not have introns (McLaren et al., 1997). To date, only one KAP7 gene member has been identified in all samples investigated. In new-breeding cashmere goat, the expression of the KAP7 gene in secondary hair follicles

was found to be 1.28 times higher than that in primary follicles (Jin et al., 2011), which suggested that the KAP7 gene likely plays an important role in regulating cashmere fineness.

Only one SNP site and two different sequences were discovered in the KAP7 gene sequence in the present study (Table 3). The statistical analysis showed that the frequency of the AB genotype was 0.6607, and that of the AA genotype was 0.3303; no BB genotype was found, which was perhaps due to the relatively small sample size. The SNP site was located at base pair 489, where the substitution leads to an amino acid change. Interestingly, this SNP site was only found in Chinese Merino sheep. These results, along with the fact that the KAP7 gene has important roles in regulating goat cashmere fineness, suggests that the role of this special SNP site of the KAP7 gene should be the focus of further research attention.

Table 3. Alignment of the KAP7 allele sequence (A-B).

S	X05638	Alleles	
		A	B
278_279	/	inG	inG
290_291	/	inGC	inGC
292	G	C	C
293	G	A	A
294	A	C	C
314	T	delT	delT
323	T	G	G
324	T	G	G
489	***C***	***C(N→K)***	***A***
683	A	delA	delA
793	G	C	C

For explanations, see legend to Table 2.

KAP8 gene polymorphisms

Like KAP7, KAP8 is encoded by a single gene without introns. Only one member of the KAP8 gene family has been identified to date (McLaren et al., 1997; Gong et al., 2012). Previous studies have demonstrated abundant variations in the KAP8 gene. Wood et al. (1992) reported 11 KAP8 alleles in a group of 33 unrelated sheep from various breeds based on dinucleotide length variation in PCR amplification of the 5'-flanking region of the gene. More recently, five different nucleic acid sequences were found in the CDS region of the KAP8 gene (Gong et al., 2012).

Compared with the reported sequence of KAP8 (X05639) (Kuczek and Rogers, 1987), five different novel DNA sequences and five SNPs were acquired from the amplified sequences of samples analyzed in the present study. These five sequences were designated as A, B, C, D, and E, and were submitted to GenBank under accession Nos. JQ043160-JQ043164. Allelic differences in the nucleotide sequences are presented in Table 4. These five sequences were highly similar to each other except for one or two base differences, verifying that they are five alleles of the KAP8 gene. Furthermore, these five SNPs were detected in all of the samples, which suggested that this gene could not reflect differences in wool traits.

In conclusion, this study revealed some new polymorphisms in the KAP6, KAP7, and KAP8 genes in four Chinese sheep breeds, which are likely to affect wool traits. Six different sequences (O, A, B, C, D, E) were identified at the KAP6 gene locus, among which sequences B and C were unique to the Chinese Merino sheep breed.

Table 4. Alignment of the KAP8 allele sequence (A-E).

S	X05639	Alleles				
		A	B	C	D	E
779_780	/	inT	inT	inT	inT	inT
840_841	/	inAT	inAT	inAT	inAT	inAT
842	***G***	-	***A***	-	-	-
849	T	A	A	A	A	A
862	**C**	-	-	-	-	***T***
883	G	delG	delG	delG	delG	delG
914	***C***	-	-	-	***T***	-
961	C	-	T	-	-	-
1017	**C**	-	-	***T(A→V)***	-	-
1026	T	C(V→A)	C(V→A)	C(V→A)	C(V→A)	C(V→A)
1095	***T***	-	-	-	-	***C(M→T)***
1130	C	T (P→S)	T (P→S)	T (P→S)	T (P→S)	T (P→S)

For explanations, see legend to Table 2.

The potential effects of the noncoding region on the expression of the KAP6 gene and its associations with wool traits need further study. Two alleles (A and B) were discovered at the KAP7 gene locus, and sequence B was only detected in Chinese merino sheep. Further study is required to determine whether sequence B affects the expression of the KAP7 gene and the conformation of wool proteins. Five alleles were found at the KAP8 gene locus, which were highly homologous to each other as well as to the reported KAP8 gene sequence. These novel SNPs detected in these genes are suggested to be potential candidate markers for wool fiber traits in sheep breeding.

ACKNOWLEDGMENTS

Research supported by the National Technological System of Wool and Cashmere Industry (#nycytx-40-06), the National High Technology Research and Development Program of China (863 Program, #2011AA100307) and the National Natural Science Foundation of China (#31360540).

REFERENCES

Beh K, Callaghan M, Leish Z and Hulme D (2001). A genome scan for QTL affecting fleece and wool traits in Merino sheep. *Wool Technol. Sheep Breed.* 49: 88-89.

Beh KJ, Hulme DJ, Callaghan MJ, Leish Z, et al. (2002). A genome scan for quantitative trait loci affecting resistance to *Trichostrongylus colubriformis* in sheep. *Anim. Genet.* 33: 97-106.

Byun SO, Fang Q, Zhou H and Hickford JG (2009). An effective method for silver-staining DNA in large numbers of polyacrylamide gels. *Anal. Biochem.* 385: 174-175.

Cockett NE, Shay TL and Smit M (2001). Analysis of the sheep genome. *Physiol. Genomics* 7: 69-78.

Fratini A, Powell BC and Rogers GE (1993). Sequence, expression, and evolutionary conservation of a gene encoding a glycine/tyrosine-rich keratin-associated protein of hair. *J. Biol. Chem.* 268: 4511-4518.

Gillespie JM and Broad A (1972). Ultra-high-sulphur proteins in the hairs of the artiodactyla. *Aust. J. Biol. Sci.* 25: 138-145.

Gong H, Zhou H and Hickford JG (2011). Diversity of the glycine/tyrosine-rich keratin-associated protein 6 gene (KAP6) family in sheep. *Mol. Biol. Rep.* 38: 31-35.

Gong H, Zhou H, Plowman JE, Dyer JM, et al. (2012). Search for variation in the ovine KAP7-1 and KAP8-1 genes using polymerase chain reaction-single-stranded conformational polymorphism screening. *DNA Cell Biol.* 31: 367-370.

Itenge-Mweza TO, Forrest RH, McKenzie GW, Hogan A, et al. (2007). Polymorphism of the KAP1.1, KAP1.3 and K33 genes in Merino sheep. *Mol. Cell. Probes* 21: 338-342.

Jin M, Wang L, Li S, Xing MX, et al. (2011). Characterization and expression analysis of KAP7.1, KAP8.2 gene in Liaoning new-breeding cashmere goat hair follicle. *Mol. Biol. Rep.* 38: 3023-3028.

Köhler A and Hurt E (2007). Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* 8: 761-773.

Kuczek ES and Rogers GE (1987). Sheep wool (glycine + tyrosine)-rich keratin genes. A family of low sequence homology. *Eur. J. Biochem.* 166: 79-85.

Marshall RC, Orwin DF and Gillespie JM (1991). Structure and biochemistry of mammalian hard keratin. *Electron. Microsc. Rev.* 4: 47-83.

McLaren RJ, Rogers GR, Davies KP, Maddox JF, et al. (1997). Linkage mapping of wool keratin and keratin-associated protein genes in sheep. *Mamm. Genome* 8: 938-940.

Parry DA and Steinert PM (1992). Intermediate filament structure. *Curr. Opin. Cell Biol.* 4: 94-98.

Parsons YM, Cooper DW and Piper LR (1994). Evidence of linkage between high-glycine-tyrosine keratin gene loci and wool fibre diameter in a Merino half-sib family. *Anim. Genet.* 25: 105-108.

Piccone ME, Pauszek S, Pacheco J, Rieder E, et al. (2009). Molecular characterization of a foot-and-mouth disease virus containing a 57-nucleotide insertion in the 3'untranslated region. *Arch. Virol.* 154: 671-676.

Plowman JE (2003). Proteomic database of wool components. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 787: 63-76.

Plowman JE, Bryson WG and Jordan TW (2000). Application of proteomics for determining protein markers for wool quality traits. *Electrophoresis* 21: 1899-1906.

Plowman JE, Deb-Choudhury S, Bryson WG, Clerens S, et al. (2009). Protein expression in orthocortical and paracortical cells of merino wool fibers. *J. Agric. Food Chem.* 57: 2174-2180.

Powell BC and Rogers GE (1990). Cyclic hair-loss and regrowth in transgenic mice overexpressing an intermediate filament gene. *EMBO J.* 9: 1485-1493.

Powell BC and Beltrame JS (1994). Characterization of a hair (wool) keratin intermediate filament gene domain. *J. Invest. Dermatol.* 102: 171-177.

Powell BC and Rogers GE (1997). The Role of Keratin Proteins and Their Genes in the Growth, Structure and Properties of Hair. In: Formation and Structure of Human Hair (Jolles P, Zahn H and Hocker E, eds.). Birkhauser Verlag, Basel, 59-148.

Powell BC, Sleigh MJ, Ward KA and Rogers GE (1983). Mammalian keratin gene families: organisation of genes coding for the B2 high-sulphur proteins of sheep wool. *Nucleic Acids Res.* 11: 5327-5346.

Powell BC, Rogers GE, Leigh I and Lane B (1994). Differentiation in Hard Keratin Tissues: Hair and Related Structures. In: The Keratinocyte Handbook (Leigh F, Watt RF and Lane B, eds.). Cambridge University Press, Cambridge, 401-436.

Song XM, Jiang JF, Zhang GZ, Shi FX, et al. (2012). DNA polymorphisms of the Hu sheep melanocortin-4 receptor gene associated with birth weight and 45-day weaning weight. *Genet. Mol. Res.* 11: 4432-4441.

Wood NJ, Phua SH and Crawford AM (1992). A dinucleotide repeat polymorphism at the glycine- and tyrosine-rich keratin locus in sheep. *Anim. Genet.* 23: 391.

Zhou JP, Zhu XP, Zhang W, Qin F, et al. (2011). A novel single-nucleotide polymorphism in the 5' upstream region of the prolactin receptor gene is associated with fiber traits in Liaoning cashmere goats. *Genet. Mol. Res.* 10: 2511-2516.