African Journal of Biotechnology Vol. 10(36), pp. 6852-6856, 18 July, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.073
ISSN 1684–5315 © 2011 Academic Journals

Full Length Research Paper

Six SNPs and a TTG indel in sheep desmoglein 4 gene are in complete linkage disequilibrium

Ge Zhang[1], Yue-hui Ma[1], Yuan-yue Zhang[2], Wei-jun Guan[1], Qian-jun Zhao[1], Hui-jing Zhao[1] and Jian-lin Han[1,3,4]*

[1]Key Laboratory of Farm Animal Genetic Resources and Utilization of Ministry of Agriculture, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China.
[2]Department of Animal Science and Technology, Hunan Agricultural University, Changsha 410128, China.
[3]International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi 00100, Kenya.
[4]CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China.

Accepted 13 June, 2011

Desmoglein 4 (DSG4) plays an important role in the regulation of growth and differentiation of hair follicles in mammals. In this study, a 755 bp long segment of *DSG4* was screened in 544 sheep sampled from nine Chinese indigenous breeds and two Western breeds using PCR-SSCP assay with three different pairs of primers. Two of the three fragments showed polymorphisms with genotypes defined as AA, AB, BB and BC, and DD, DE, and EE, respectively. Interestingly, polymorphisms in these two fragments were in strong linkage disequilibrium. Only three haplotypes were found, of which haplotype AD determined by alleles A and D was the major one in all breeds, while haplotype BE was only found in Chinese breeds that possess divergent frequencies ranging from 0.02 to 0.43; haplotype CD was very rare and present in only one Chinese sheep. Sequences of the three haplotypes showed seven single nucleotide polymorphisms (SNPs) and a TTG insertion/deletion (indel), leading to five amino acid substitutions and a glycine indel. Our study provides valuable genetic markers in evaluating the impact of the *DSG4* gene on wool traits in sheep.

Key words: Sheep, *DSG4* gene, single-strand conformational polymorphism (SSCP), variation, linkage disequilibrium.

INTRODUCTION

Wool traits are extremely important to the sheep industry. The study of candidate genes is one of the primary methods to identify specific genes responsible for wool traits in sheep (Purvis and Franklin, 2005). In hair follicles, the cell-cell adhesion is mediated in part by desmosomes which are composed of multi-proteins including desmogleins and desmocollins (Green and Jones, 1996). The human *desmoglein 4* (*DSG4*) is 37 kb long, and consists of 16 exons and 15 introns (GenBank accession no. AY177663). *DSG4* presents a very specific tissue expression in salivary gland, testis, prostate, and skin (Whittock and Bower, 2003) and is regulated by transcription factors such as HOXC13, LEF1 and FOXN1 (Bazzi et al., 2009). In the skin, *DSG4* is expressed in suprabasal epidermis and hair follicles (Kljuic et al., 2003; Bazzi et al., 2006). Kljuic et al. (2003) identified a deletion encompassing exons 5-8 of human *DSG4* in families with localized autosomal recessive hypotrichosis (LAH), and a single nucleotide insertion in exon 7 as well as a missense mutation in exon 6 in mice with lanceolate hair (lah). Subsequently, a number of mutations within the *DSG4* associated with LAH (Messenger et al., 2005; Wajid et al., 2007) and monilethrix hairs in humans (Schaffer et al., 2006; Shimomura et al., 2006; Zlotogorski et al., 2006) and lah in rats (Bazzi et al., 2004; Jahoda et al., 2004; Meyer et al., 2004) were reported, which further demonstrated the function of *DSG4* as a key mediator of keratinocyte cell adhesion in hair follicles. These studies strongly imply that *DSG4* is a

*Corresponding author. E-mail: h.jianlin@cgiar.org. or hanjl@iascaas.net.cn. Tel/Fax: +(86-10) 62829175.

candidate gene that may contain polymorphic variation affecting wool traits in sheep. Based on the comparative mapping of OAR23 linkage and RH maps against HSA18 physical and cytogenetic maps (Tetens et al., 2007), the sheep *DSG4* is probably located between the two microsatellite markers of DIK4464 and AGLA269 on chromosome 23. BLAST analysis indicated that several sheep ESTs (GenBank accession nos. EE752650 and GO691893) were homologous with the exon 16 and 3'UTR of bovine *DSG4* (GenBank accession no. XM_617938); therefore the aim of this study was to identify possible polymorphisms within this region in sheep *DSG4*.

MATERIALS AND METHODS

Samples

Ear tissue samples of 544 sheep from nine Chinese indigenous and two Western breeds were collected and used in the study, including 48 Langkazi sheep sampled from Tibetan Autonomous Region, 96 Tan sheep from Ningxia Hui Autonomous Region, 40 Ujumqin sheep from Inner Mongolian Autonomous Region, 45 Duolang sheep, 47 Bashibai sheep and 48 Kazakh Fat-Rumped sheep from Xinjiang Uygur Autonomous Region, 48 Small-Tailed Han sheep from Shandong, 46 Zhaotong sheep from Yunnan and 48 Hu sheep from Zhejiang Provinces, 48 Polled Dorset and 30 Suffolk kept on the farm of Institute of Animal Science of CAAS in Beijing. DNA extraction was performed following the routine phenol-chloroform protocol (Sambrook et al., 1989).

Primer design

The two sheep ESTs isolated from skin tissue (EE752650 and GO691893) which share 97% homology with bovine *DSG4* mRNA sequence predicted from a Hereford cow genome (XM_617938, Liu et al., 2009) were used to design three pairs of PCR primers using Primer program version 5.0 (Premier Biosoft International, Palo Alto, CA, USA). Primer pairs P1 (5'-CCACGAAGGAGTAGGGTCT-3' and 5'-CCAAGTAAAGGGAGGTCAGT-3') and P2 (5'-GCACTGACCT CCCTTTACTT-3' and 5'-CAACCATCACTCACGCTACT-3') were designed to amplify partial exon 16. Primer P3 (5'-TAGTAGCGTG AGTGATGGTT-3' and 5'-ATGTTGGTGATTACAAGGTG-3') was designed to amplify partial exon 16 and 3'UTR. The amplicons P1, P2, and P3 were 195, 364, and 239 bp, respectively. The homologous nucleotide positions of the three amplicons at bovine *DSG4* mRNA (XM_617938) were ranged from 2529 to 2723 bp, from 2702 to 3065 bp, and from 3045 to 3286 bp, respectively.

PCR-SSCP analysis

Polymerase chain reaction (PCR) and single-strand conformational polymorphism (SSCP) were used to screen for polymorphisms within the amplified fragments. The PCR was performed in a 15 μL reaction mixture containing 100 ng of genomic DNA, 1×PCR reaction buffer, 0.2 μM of each primer, 200 μM each dNTP, and 0.75 U *Taq* polymerase (Tiangen Biotech, Beijing, China). The PCR conditions were 94°C for 5 min, followed by 35 cycles of 30 s at 94°C, 30 s at 63°C for P1 or 64°C for P2 or 62°C for P3 and 30 s at 72°C, and a final extension at 72°C for 7 min. SSCP conditions were as follows: 2 μL PCR products were mixed with 8 μL. formamide loading dye for denaturation at 98°C for 10 min and then imbedded in ice for 10 min. The samples were finally run on a 12% non-denaturing acrylamide gel (Acr:Bis = 29:1) for 12 to 14 h in 1×TBE at 4°C and 160-200 V, and the gel was silver stained.

DNA sequencing analysis

After the polymorphisms were detected, DNA samples with representative genotypes were amplified using the forward primer of P1 and the reverse primer of P3. The new PCR conditions were 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 40 s, and a final extension at 72°C for 7 min. PCR products of the 755 bp segment were purified and subsequently sequenced using both PCR primers on an ABI 3730 DNA sequencer.

Statistic analysis

Observed heterozygosity (H_O), expected heterozygosity (H_E), Hardy-Weinberg equilibrium (HWE) test and linkage disequilibrium (LD) analysis were performed using Arlequin version 3.1 (Excoffier et al., 2005). DNA sequences were aligned and analyzed using MEGA version 4.0 (Tamura et al., 2007). The predictions of functional impact for each nonsynonymous substitution were conducted using Polyphen (Ramensky et al., 2002; http://genetics.bwh.harvard.edu/pph/).

RESULTS

Identification of polymorphisms

The partial exon 16 and 3'UTR of sheep *DSG4* were amplified by the P1, P2, and P3 primer pairs separately. After the SSCP analyses, the products of P1 showed no polymorphism. For P2, four different patterns were detected and named genotypes AA, BB, AB, and BC (Figure 1). For P3, three genotypes were found and named DD, EE and DE (Figure 2). The comparison of new sequences of genotypes AADD and BBEE (GenBank accession nos. HM030496 and HM030497) showed six single nucleotide polymorphisms (SNPs) and a TTG insertion/deletion polymorphism (indel), with four SNPs 248T>A, 260G>A, 287A>G, and 508C>T in the P2 fragment, and 583C>T, 709T>C, and 629_630insTTG in the P3 fragment. The sequence of genotype AADD showed 100% sequence identity with the above two ESTs. The new sequence of genotype BCDE possessed another SNP 297T>C (GenBank accession no. HM030498 for CD haplotype) located in the P2 fragment.

Prediction of the impact of polymorphisms

Because the complete coding sequence (CDS) of sheep *DSG4* is still unclear so far, we determined whether these nucleotide variations led to amino acid changes according to the *DSG4* CDS of human (GenBank accession no. AY227350), mouse (AY227349), rat (AY314982), and cattle (XM_617938). The nucleotide positions of all

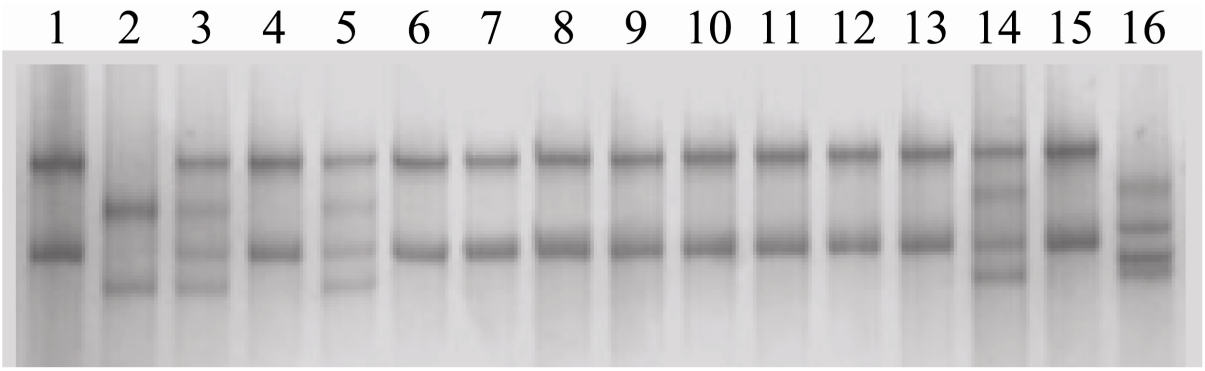

Figure 1. SSCP analysis of PCR products of P2 fragment. 1, 4, 6-13 and 15: AA; 2: BB; 3, 5 and 14: AB; 16: BC.

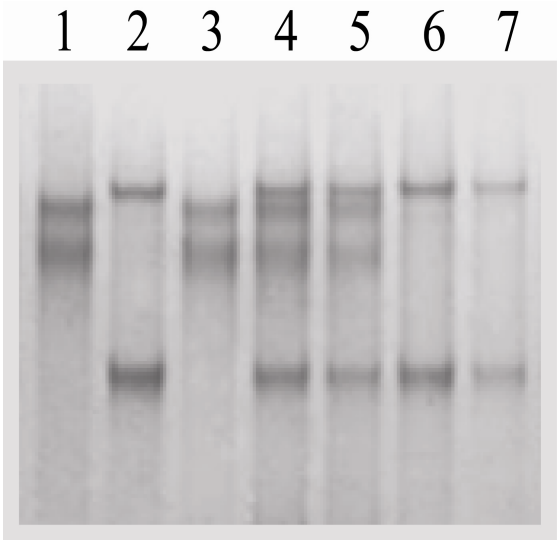

Figure 2. SSCP analysis of PCR products of P3 fragment. 1 and 3: EE; 2, 6 and 7: DD; 4 and 5: DE.

variations in the homologous CDS region differed by one or two codons among these five species (plus sheep). Of these variations, the four SNPs 248T>A, 260G>A, 287A>G, and 297T>C were nonsynonymous, the two SNPs 508C>T and 583C>T were synonymous, the 709T>C was located at the 3'UTR, and the 629_630insTTG was located at coding region leading simultaneously to an amino acid substitution and also a glycine insertion. The amino acid substitutions caused by the 248T>A, 287A>G, and the 629_630insTTG possibly had an impact on the structure and function of sheep *DSG4* protein, as predicted by the Polyphen program (Table 1).

LD analysis and HWE test

Interestingly, polymorphisms in these two fragments were in strong linkage disequilibrium (LD) in all of the 11 sheep populations (D' = 1). The A allele of P2 fragment was always linked to the D allele of P3 fragment, and the B allele of P2 fragment was always linked to the E allele of P3 fragment, while the C allele of P2 fragment (present only in one sheep) was linked to the D allele. The genotypes between P2 and P3 loci formed into only four combinations of AADD, ABDE, BBEE, and BCDE, leading to the identification of three haplotypes of AD, BE, and CD. The possible combination genotypes of AAEE, AADE, ABDD, ABEE, BBDD, and BBDE were not present in any of the samples (n = 544).

Table 2 shows the genotype and haplotype frequencies, the observed (H_O) and expected (H_E) heterozygosities, and the P values of Hardy-Weinberg equilibrium (HWE) test for *DSG4* in the 11 sheep populations. Of the three haplotypes, CD, determined by C and D alleles, was very rare and present once in Tan sheep; AD was the major haplotype in nine Chinese breeds and fixed in Polled Dorset and Suffolk; the frequency of haplotype BE was the highest in Duolang sheep (0.43), the lowest in Hu sheep (0.02) but similar among Tan, Ujumqin, and Small-Tailed Han sheep. The genotype BBEE was not detected in Langkazi, Zhaotong, and Hu sheep, probably due to the low frequency of haplotype BE. The values of H_O and H_E for *DSG4* in nine Chinese sheep breeds were divergent ranging from 0.04 to 0.69 and from 0.04 to 0.50, respectively. The *DSG4* polymorphisms were in HWE in all breeds except in Tan (P = 0.03) and Duolang sheep (P = 0.01).

Validation of the complete LD among these variations

To further confirm the complete linkage disequilibrium among these variations (except 297T>C, of which the C allele was present in only one sheep), 16 individuals with AADD genotype including four Tan sheep, four Small-Tailed Han sheep, four Polled Dorset, and four Suffolk, as well as six individuals with BBEE genotype including four Tan sheep and two Small-Tailed Han sheep were selected for PCR using the P1 forward and P3 reverse primers and sequencing using the P1 forward primer

Table 1. Homologous positions and functional predictions of sheep *DSG4* variations.

Name[1]	Position (bp)[2]	Codon change	AA change	Polyphen predicting
248T>A	2671	TTC→ATC	Phe→Ile	Possibly damaging
260G>A	2683	GTG→ATG	Val→Met	Benign
287A>G	2710	AGG→GGG	Arg→Gly	Probably damaging
297T>C	2720	GTA→GCA	Val→Ala	Benign
508C>T	2931	GAC→GAT	synonymous	
583C>T	3006	ATC→ATT	synonymous	
629-630insTTG	3053-3055	AGT→ATTGGT	Ser→Ile Gly insertion	Possibly damaging Undetermined
709T>C	3135	3'UTR		

[1]Described based on the sequence of genotype AADD (GenBank accession no. HM030496). [2]Homologous positions based on bovine *DSG4* mRNA sequence excluding 5'UTR (GenBank accession no. XM_617938).

Table 2. Genotype and haplotype frequencies, H_O and H_E, and P values of HWE test for *DSG4* in 11 sheep breeds.

Breed	Genotype frequency				Haplotype frequency			H_O	H_E	P value of HWE test
	AADD	ADBE	BBEE	BCDE	AD	BE	CD			
Tan	0.60 (57)*	0.30 (29)	0.09 (9)	0.01 (1)	0.74	0.25	0.01	0.31	0.38	0.03
Small-Tailed Han	0.65 (31)	0.31 (15)	0.04 (2)		0.80	0.20		0.31	0.32	1.00
Ujumqin	0.65 (26)	0.27 (11)	0.08 (3)		0.79	0.21		0.28	0.34	0.33
Duolang	0.22 (10)	0.69 (31)	0.09 (4)		0.57	0.43		0.69	0.50	0.01
Bashibai	0.72 (34)	0.26 (12)	0.02 (1)		0.85	0.15		0.26	0.26	1.00
Kazakh	0.88 (42)	0.10 (5)	0.02 (1)		0.93	0.07		0.10	0.14	0.21
Langkazi	0.88 (42)	0.12 (6)			0.94	0.06		0.13	0.12	1.00
Zhaotong	0.87 (40)	0.13 (6)			0.93	0.07		0.13	0.12	1.00
Hu	0.96 (46)	0.04 (2)			0.98	0.02		0.04	0.04	1.00
Suffolk	1 (30)				1					
Polled Dorset	1 (48)				1					

*The number of individuals with each genotype is shown in appropriate bracket.

alone. The results showed no additional polymorphism within this region.

DISCUSSION

DSG4 is essential for the growth and differentiation of hair follicles in mammals. The human *DSG4* comprises of five homologous extracellular domains, a transmembrane domain, and a carboxy-terminal cytoplasmic tail (Whittock and Bower, 2003). All the missense mutations associated with hair disorders including Y196S in mouse (Kljuic et al., 2003), A129S, P267R, and S192P in human (Messenger et al., 2005; Shimomura et al., 2006; Zlotogorski et al., 2006), and E288V in rat (Jahoda et al., 2004), are located in the extracellular domains. In this study, we identified five amino acid substitutions and a glycine indel located in the cytoplasmic domain of sheep *DSG4*. Notably, the glycine indel and four out of the five amino acid substitutions form into only two types of *DSG4* (as nonsynonymous substitutions 248T>A, 260G>A, 287A>G, and the 629_630insTTG formed into only two haplotypes of AD and BE). Of these four amino acid substitutions, three possibly had an impact on the function of the sheep *DSG4*. It would be very interesting to investigate whether these *DSG4* polymorphisms are associated with wool traits in sheep.

The complete linkage disequilibrium determined by the six SNPs and the TTG indel was present in all sheep breeds, suggesting that they may have originated from different ancestral genetic backgrounds. The haplotype BE was only found in Chinese breeds that possess divergent frequencies. Based on the difference in the distribution of mitochondrial DNA (mtDNA) lineages among Chinese and European sheep, example the exclusive presence of lineage C in Chinese sheep (Guo et al., 2005), we postulate that the *DSG4* haplotype BE was probably derived from the ancestral stocks carrying the

mtDNA lineage C. Although, Duolang, Bashibai, and Kazakh sheep are all distributed in Xinjiang, haplotype frequency divergence between Duolang and Kazakh or Bashibai was quite obvious (Table 2). The Duolang sheep (once known as Maigaiti) was developed by crossing a local population with the fat-rumped sheep introduced from Afghanistan (Liu et al., 1991), whereas the Kazakh sheep was one of the most primitive breeds in China (Zheng, 1988), and the Bashibai sheep was developed from the Kazakh sheep (Jia et al., 2003). The frequency of haplotype BE was similar among Tan, Ujumqin, and Small-Tailed Han sheep, probably due to their common origin (Zheng, 1988). Investigation of the distribution of *DSG4* haplotype BE in additional sheep breeds (especially from Middle East) may help to understand its origin.

ACKNOWLEDGEMENTS

This work was supported by the National Key Project of Scientific and Technical Supporting Programs Funded by Ministry of Science and Technology of China during the 11th 'Five-Year' Plan (2008BADB2B01), National Special Construction Fund of Modern Agricultural Industrial Technology System (nycytx-40-5) and the National GMO Cultivation Major Project of New Varieties (2008ZX08009-003).

REFERENCES

Bazzi H, Kljuic A, Christiano AM, Panteleyev AA (2004). Intragenic deletion in the desmoglein 4 gene underlies the skin phenotype in the Iffa Credo hairless rat. Differentiation, 72: 450-464.

Bazzi H, Getz A, Mahoney MG, Ishida-Yamamoto A, Langbein L, Wahl JK, Christiano AM (2006). Desmoglein 4 is expressed in highly differentiated keratinocytes and trichocytes in human epidermis and hair follicle. Differentiation, 74: 129-140.

Bazzi H, Demehri S, Potter CS, Barber AG, Awgulewitsch A, Kopan R, Christiano AM (2009). Desmoglein 4 is regulated by transcription factors implicated in hair shaft differentiation. Differentiation, 78: 292-300.

Excoffier L, Laval G, Schneider S (2005). Arlequin version. 3.0: An integrated software package for population genetics data analysis. Evol. Bioinform. Online 1: 47-50.

Green KJ, Jones JC (1996). Desmosomes and hemidesmosomes: structure and function of molecular components. J. FASEB. 10: 871-881.

Guo J, Du LX, Ma YH, Guan WJ, Li HB, Zhao QJ, Li X, Rao SQ (2005). A novel maternal lineage revealed in sheep (Ovis aries). Anim. Genet. 36: 331-336.

Jahoda CA, Kljuic A, O'Shaughnessy R, Crossley N, Whitehouse CJ, Robinson M, Reynolds AJ, Demarchez M, Porter RM, Shapiro L, Christiano AM (2004). The lanceolate hair rat phenotype results from a missense mutation in a calcium coordinating site of the desmoglein 4 gene. Genomics, 83: 747-756.

Jia B, Chen J, Zhao RQ, Luo QJ, Yan GQ, Chen J (2003). Microsatellite analysis of genetic diversity and phylogenetic relationship of eight sheep breeds in Xinjiang. Acta Genetica Sinica, 30: 847-854.

Kljuic A, Bazzi H, Sundberg JP, Martinez-Mir A, O'Shaughnessy R, Mahoney MG, Levy M, Montagutelli X, Ahmad W, Aita VM, Gordon D, Uitto J, Whiting D, Ott J, Fischer S, Gilliam TC, Jahoda CA, Morris RJ, Panteleyev AA, Nguyen VT, Christiano AM (2003). Desmoglein 4 in hair follicle differentiation and epidermal adhesion: Evidence from inherited hypotrichosis and acquired pemphigus vulgaris. Cell, 113: 249-260.

Liu DT, Cui YZ, Wang LX, Yang XM, Liu X, Zhao ZK, Zhang SC (1991). Survey report on the Maigaiti fat-rumped sheep. China Sheep Goat Farming, 4: 9-10.

Liu Y, Qin X, Song XZH, Jiang H, Shen Y, Durbin KJ, Lien S, Kent MP, Sodeland M, Ren Y, Zhang L, Sodergren E, Havlak P, Worley KC, Weinstock GM, Gibbs RA (2009). *Bos taurus* genome assembly. BMC Genomics 10: 180.

Messenger AG, Bazzi H, Parslew R, Shapiro L, Christiano AM (2005). A missense mutation in the cadherin interaction site of the desmoglein 4 gene underlies localized autosomal recessive hypotrichosis. J. Invest. Dermatol. 125: 1077-1079.

Meyer B, Bazzi H, Zidek V, Musilova A, Pravenec M, Kurtz TW, Nurnberg P, Christiano AM (2004). A spontaneous mutation in the desmoglein 4 gene underlies hypotrichosis in a new lanceolate hair rat model. Differentiation, 72: 541-547.

Purvis IW, Franklin IR (2005). Major genes and QTL influencing wool production and quality: a review. Genet. Sel. Evol. 37: 97-107.

Ramensky V, Bork P, Sunyaev S (2002). Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30: 3894-3900.

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press (second edition), New York, USA.

Schaffer JV, Bazzi H, Vitebsky A, Witkiewicz A, Kovich OI, Kamino H, Shapiro LS, Amin SP, Orlow SJ, Christian AM (2006). Mutations in the desmoglein 4 gene underlie localized autosomal recessive hypotrichosis with monilethrix hairs and congenital scalp erosions. J. Invest. Dermatol. 126: 1286-1291.

Shimomura Y, Sakamoto F, Kariya N, Matsunaga K, Ito M (2006). Mutations in the desmoglein 4 gene are associated with monilethrix-like congenital hypotrichosis. J. Invest. Dermatol. 126: 1281-1285.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA 4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596-1599.

Tetens J, Goldammer T, Maddox JF, Cockett NE, Leeb T, Drögemüller C (2007). A radiation hybrid map of sheep chromosome 23 based on ovine BAC-end sequences. Anim. Genet. 38: 132-140.

Wajid M, Bazzi H, Rockey J, Lubetkin J, Zlotogorski A, Christiano AM (2007). Localized autosomal recessive hypotrichosis due to a frameshift mutation in the desmoglein 4 gene exhibits extensive phenotypic variability within a Pakistani family. J. Invest. Dermatol. 127: 1779-1782.

Whittock NV, Bower C (2003). Genetic evidence for a novel human desmosomal cadherin, desmoglein 4. J. Invest. Dermatol. 120: 523-530.

Zheng PL (1988). Sheep and goat breeds in China. Shanghai Scientific and Technical Press, Beijing, China.

Zlotogorski A, Marek D, Horev L, Abu A, Ben-Amitai D, Gerad L, Ingber A, Frydman M, Reznik-Wolf H, Vardy DA, Pras E (2006). An autosomal recessive form of monilethrix is caused by mutations in DSG4: clinical overlap with localized autosomal recessive hypotrichosis. J. Invest. Dermatol. 126: 1292-1296.