The Journal of Agricultural Science

cambridge.org/ags

Nucleotide variation in the ovine *KRT31* promoter region and its association with variation in wool traits in Merino-cross lambs

W. Chai[1,2], H. Zhou[1,2], H. Gong[1,2], J. Wang[2], Y. Luo[2] and J. G. H. Hickford[1,2]

[1]Gene-Marker Laboratory, Faculty of Agricultural and Life Science, Lincoln University, Lincoln 7647, New Zealand and [2]International Wool Research Institute/Faculty of Animal Science and Technology, Gansu Agricultural University, Lanzhou 730070, China

Animal Research Paper

Cite this article: Chai W, Zhou H, Gong H, Wang J, Luo Y, Hickford J G H (2019). Nucleotide variation in the ovine *KRT31* promoter region and its association with variation in wool traits in Merino-cross lambs. *The Journal of Agricultural Science* **157**, 182–188. https://doi.org/10.1017/S0021859619000406

Received: 13 January 2019
Revised: 29 April 2019
Accepted: 13 May 2019
First published online: 10 June 2019

Key words:
KRT31; sheep; variation; wool traits

Author for correspondence:
J. G. H. Hickford,
E-mail: jon.hickford@lincoln.ac.nz

Abstract

Keratins are the main structural proteins of wool fibres, and it is thought that variation in the keratins may affect wool fibre characteristics. Polymerase chain reaction-single stranded conformational polymorphism (PCR-SSCP) analyses were used to investigate four regions of the ovine keratin gene *KRT31* including a portion of the promoter, the exon 1, exon 3 and exon 7 regions. Initially, in a screening panel of 300 New Zealand Romney, Merino and White Dorper sheep obtained from 26 farms, three, two, two and two PCR-SSCP banding patterns were observed for these four regions, respectively. The promoter region, the exon 1 and exon 3 regions contained two single nucleotide polymorphisms (SNPs) and the exon 7 region contained one SNP. The effect of the variation found in the promoter region on wool traits was subsequently investigated in 485 Southdown × Merino-cross lambs from seven sire-lines. The three variants identified in the original 300 sheep (named *A*, *B* and *C*) were observed with frequencies of 56, 29 and 15%, respectively. The presence of *A* and *B* had no significant effect on wool traits, but the presence of *C* was found to be associated with an increase in greasy fleece weight (GFW), clean fleece weight (CFW) and mean staple length (MSL). There was an effect of genotype on CFW and MSL, with *BC* sheep producing wool of higher CFW and MSL than *AA*, *AB*, *AC* and *BB* sheep. These results suggest that ovine *KRT31* might be a useful candidate gene for improving wool traits.

Introduction

In wool fibres, the main structural proteins are hard α-keratins. They are assembled in a highly organized fashion into keratin intermediate filaments (KIFs) and enveloped by an inter-filamentous matrix consisting of keratin-associated proteins (KAPs) (Popescu and Höcker, 2007). The KIF proteins are typically low-sulphur proteins when compared to the other wool proteins, and they are grouped into type-I and type-II families, according to their site of expression, function and pKa (whether they are acidic or neutral/basic). They can be further subdivided into epithelial cytokeratins and hair keratins (Heid *et al.*, 1986, 1988; Lynch *et al.*, 1986).

The α-keratin K31 belongs to the type I family and is found in the cortex of wool fibres, along with K85 and K38 (Yu *et al.*, 2009). The K31 gene (*KRT31*, formerly known as *KRT1.1*) has been mapped to ovine chromosome 11q25-q29 (Hediger *et al.*, 1991) and is described at ENSOARG00000016473, Oar_v3.1:CM001592.1; Chromosome 11:41 105 175–41 108 801. Eighteen *KRTs* and *KRTAPs* have been mapped on the same chromosome near *KRT31*, and in a cluster (Sumner *et al.*, 2013; Gong *et al.*, 2016).

A number of quantitative trait loci (QTLs) and markers for wool traits, including greasy fleece weight (GFW), clean fleece weight (CFW), mean staple strength (MSS) and coefficient of variation of fibre diameter (CVFD) have been reported on ovine chromosome 11 (Rogers *et al.*, 1994; Roldan *et al.*, 2010), but there are only a small number of studies describing genetic variation in the chromosome 11 *KRT* genes.

Rogers *et al.* (1993) reported a *Msp*I polymorphism in *KRT1.2* (now called *KRT33a*) using a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) approach, and Itenge-Mweza *et al.* (2007) described five alleles of *KRT33* (now called *KRT33a*), using a PCR-single stranded conformational polymorphism (PCR-SSCP) approach. Itenge *et al.* (2010) reported that *KRTAP1.1*, *KRTAP1.3* and *KRT33a* were potential gene markers associated with a variety of wool traits, including MSS, mean staple length (MSL), fibre diameter standard deviation (FDSD), wool yellowness and lustre in Merino and Merino-cross sheep. Genes located within a region near to *KRT1.2* could also have an effect on MSS (Parsons *et al.*, 1994; Sumner *et al.*, 2013).

In the current study, PCR-SSCP was used to investigate sequence variation in *KRT31* and the associations, if any, which may occur between *KRT31* sequence variation and various wool traits.

CAMBRIDGE UNIVERSITY PRESS

Table 1. The primer sequences and PCR-SSCP conditions for screening variation in three regions of ovine *KRT31*

Gene region		Primer binding coordinates[a]	Primer sequence (5′–3′)	Predicted amplicon size (bp)	PCR annealing temperature (°C)	SSCP gel electrophoresis conditions[b]
Promoter	up	41054039-41054021	TGAGTGATAGGCAGGTGGC	457	60	200 V, 14%, 29 °C
	dn	41053583-41053602	CTCAGAAGAGTCCCTTGTCC			
Exon 1	up	41053140-41053132	GAACCTCGCTCCCTCCAG	422	55	200 V, 12%, 33.5 °C
	dn	41052719-41052736	TGAGAATCCTGGGACGCT			
Exon 3	up	41052208-41052188	ATCCAAGATCATTCCTGTGAG	271	60	320 V, 14%, 8 °C
	dn	41051938-41051956	AGGTCTGAGGCTGAGTCAG			
Exon 7	up	41049886-41049866	CTCTCTGGTTACATCACAAGC	311	60	250 V, 14%, 26 °C
	dn	41049576-41049595	CTCCAGCCATGCACATTGTG			

PCR-SSCP, polymerase chain reaction-single stranded conformational polymorphism.
[a]Refers to the Oar v4.0 sequence NC_019460.
[b]V, Voltage; %, Acrylamide gel percentage; °C, Electrophoresis temperature.

Materials and methods

Sheep and wool samples

There were two parts to the current study. The first was to ascertain the extent of genetic variation in *KRT31* in 300 New Zealand (NZ) Merino, Romney and White Dorper sheep, sourced from 26 farms. Four regions of the gene were analysed: a 457-bp fragment of the promoter, a 422-bp fragment spanning the entire exon 1, a 271-bp fragment spanning the entire exon 3 and a 311-bp fragment spanning the entire exon 7 coding sequence.

Next, the association of *KRT31* promoter variation with selected wool traits was investigated in a separate group of 485 Southdown × Merino-cross lambs derived from seven sire-lines. These lambs were all managed on the same farm and were shorn at 12 months of age.

At shearing, GFW was measured and wool samples were collected from the mid-side region for wool trait measurement at the New Zealand Wool Testing Authority Ltd (NZWTA, Napier, NZ) using the International Wool Testing Organisation (IWTO) standardized methods. This included measurement of wool yield (Yield; %), MSL (mm), MSS (N/ktex), mean fibre diameter (MFD; μm), FDSD (μm), CVFD (%), mean fibre curvature (MFC; °/mm) and prickle factor (PF; the percentage of fibres of diameter greater than 30 microns). Clean fleece weight (CFW; kg) was calculated from the GFW and Yield measurement.

Polymerase chain reaction amplification

A blood sample from each sheep was collected onto a Whatman Flinders Technology Associates (FTA) card and genomic DNA was purified using a two-step washing procedure as described in Zhou *et al.* (2006). The primers used to amplify the four regions of *KRT31* were designed based on sequences in the ovine genome assembly v4.0 (Table 1). The primers were synthesized by Integrated DNA Technologies (Coralville, IA, USA).

Polymerase chain reaction amplification was performed in 15-μl reactions and included the genomic DNA on one 1.2 mm punch of FTA paper, 10 × reaction buffer with 0.5 U of Taq DNA polymerase (Qiagen, Hilden, Germany), a 250 nM concentration of each primer, a 150 μM concentration of each deoxyribonucleotide triphosphate (dNTP; Eppendorf, Hamburg, Germany) and a final magnesium ion (Mg^{2+}) concentration of 2.5 mM. The cycling parameters for PCR amplification consisted of denaturation at 94 °C for 2 min., followed by 35 cycles of 30 s each of denaturation at 95 °C, annealing at the temperature shown in Table 1 and extension at 72 °C; with a final extension at 72 °C for 5 min. The amplifications were undertaken in either iCyclers (Bio-Rad, Hercules, CA, USA) or S1000 thermal cyclers (Bio-Rad).

Variation screening and DNA sequencing

All SSCP analyses were carried out in 14% or 12% polyacrylamide gels (37.5:1; Bio-Rad) in 0.5 × Tris/Borate/Ethylenediaminetetraacetic acid (TBE) buffer under the electrophoresis conditions described in Table 1 for 16 h. A 0.7-μl aliquot of the product from the PCR amplification was mixed with 7 μl of loading dye (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene-cyanol), denatured at 95 °C for 5 min, then cooled rapidly on wet ice and loaded onto gels. All gels were silver-stained according to the method of Byun *et al.* (2009).

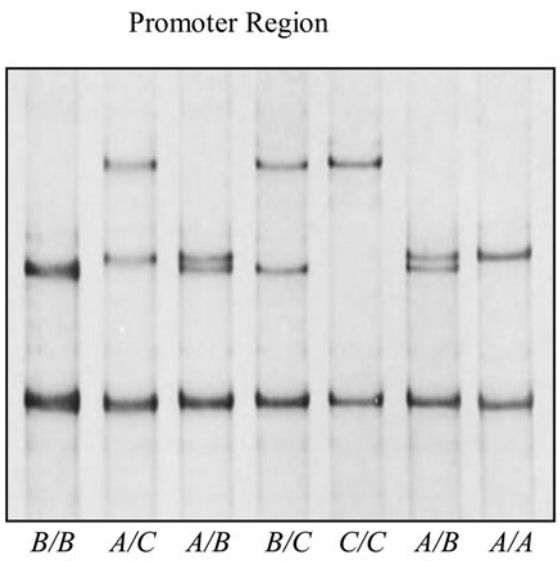
Polymerase chain reaction amplicons from sheep homozygous for each of the representative SSCP banding patterns were sequenced directly at the Lincoln University DNA Sequencing Facility, Lincoln, New Zealand. For variants that were only found in apparently heterozygous sheep, the DNA was sequenced using an approach described in Gong *et al.* (2011). Briefly, the SSCP band corresponding to the variant not found in a homozygous form was excised from the SSCP gel, macerated and used as a template for re-amplification. The product of the second amplification was then directly sequenced in triplicate.

Sequence alignments, translations (to determine presumed amino acid sequences) and comparisons were carried out using DNAMAN (version 5.2.10, LynnonBioSoft, Vaudreuil, Canada). The Basic Local Alignment Search Tool (BLAST) algorithm was used to search the National Center for Biotechnology Information (NCBI) GenBank (http://www.ncbi.nlm.nih.gov/) databases for sequence homology, in particular, homology found with the ovine genome assembly v4.0.

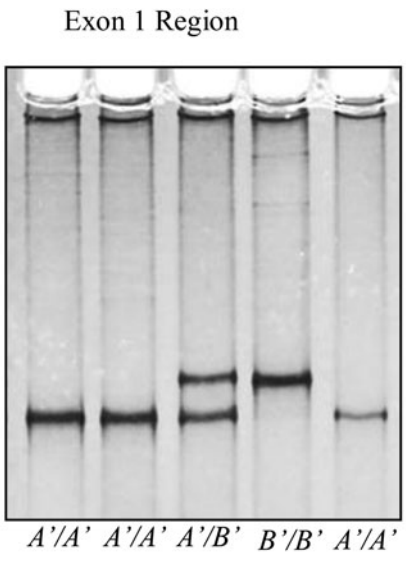
Genotyping of the promoter region

Following initial screening of the 300 NZ Romney, Merino and White Dorper sheep, an inner reverse PCR primer

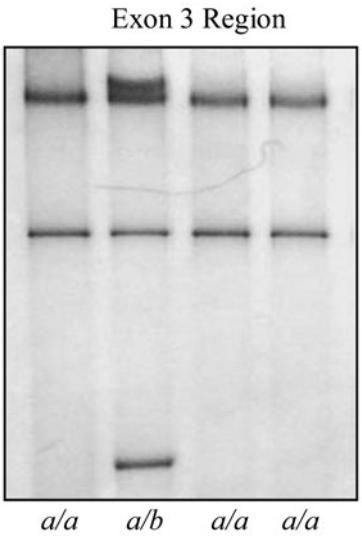
https://doi.org/10.1017/S0021859619000406 Published online by Cambridge University Press



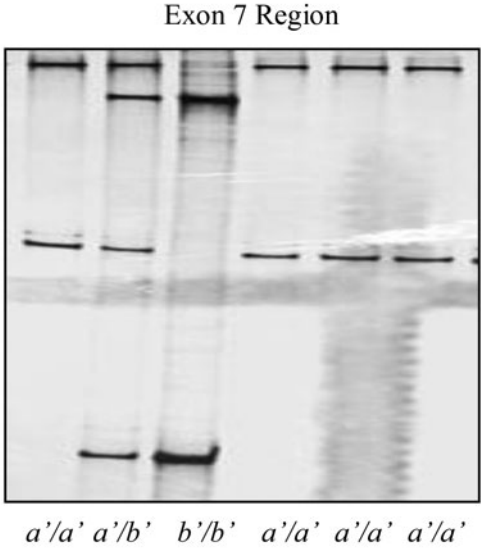
	c.-864	c.-847
Variant *A*	C	A
Variant *B*	G	A
Variant *C*	C	G



	c.55	c.348+31
Variant *A'*	A	A
Variant *B'*	C	G



	c.522	c.588+14
Variant *a*	C	C
Variant *b*	T	T



	c.*20
Variant *a'*	G
Variant *b'*	A

Fig. 1. Variation identified in ovine *KRT31*. Three PCR-SSCP patterns corresponding to three (*A, B* and *C*) variant sequences were identified in the promoter region, while two PCR-SSCP patterns responding to two (*A'* and *B'*) variant sequences were identified in the exon 1 region, two PCR-SSCP patterns responding to two (*a* and *b*) variant sequences were identified in the exon 3 region and two PCR-SSCP patterns responding to two (*a'* and *b'*) variant sequences were identified in the exon 7 region. Only the nucleotide sequences that differ between the variants are shown, and nucleotide positions refer to GenBank accession no. NC_019468 following the nomenclature described in http://www.hgvs.org/mutnomen/.

(5′-GTTAGTCCATCCAATGACATTC-3′) for the *KRT31* promoter region was designed to amplify a smaller fragment that spanned the nucleotide variations identified with the first set of primers. This primer was used along with the reverse primer at an annealing temperature of 60 °C, to type the *KRT31* promoter region in the 489 Southdown × Merino-cross lambs using PCR-SSCP in 14% polyacrylamide gels (37.5:1; Bio-Rad) under the electrophoresis conditions of 200 V, 29 °C for 16 h. The

https://doi.org/10.1017/S0021859619000406 Published online by Cambridge University Press

Table 2. Association of *KRT31* promoter variants with various wool traits

				Single-variant model [a]			Multi-variant model [a]			
Trait	Variant	Absent *n*	Present *n*	Absent (mean ± S.E.)	Present (mean ± S.E.)	*P*	Other variant fitted	Absent (mean ± S.E.)	Present (mean ± S.E.)	*P*
GFW	*A*	90	395	2.3 ± 0.09	2.3 ± 0.10	0.882	*B,C*	2.5 ± 0.09	2.5 ± 0.09	0.738
(kg)	*B*	218	267	2.2 ± 0.10	2.3 ± 0.10	0.170	*C*	2.5 ± 0.09	2.5 ± 0.09	0.392
	C	400	85	**2.2 ± 0.10**	**2.3 ± 0.10**	**0.008**	*B*	**2.4 ± 0.09**	**2.6 ± 0.09**	**0.002**
CFW	*A*	90	395	1.7 ± 0.08	1.7 ± 0.08	0.342	*B,C*	1.8 ± 0.07	1.8 ± 0.06	0.763
(kg)	*B*	218	267	1.7 ± 0.08	1.8 ± 0.08	0.119	*C*	1.8 ± 0.06	1.9 ± 0.06	0.120
	C	400	85	**1.7 ± 0.08**	**1.8 ± 0.08**	**0.002**	*B*	**1.7 ± 0.06**	**1.9 ± 0.06**	**<0.001**
Yield	*A*	90	395	76 ± 1.4	75 ± 1.4	0.640	*C*	75.0 ± 0.91	74.2 ± 0.81	0.202
(%)	*B*	218	267	76 ± 1.4	76 ± 1.4	0.638	*C*	74.4 ± 0.09	74.7 ± 0.08	0.586
	C	400	85	75 ± 1.4	76 ± 1.4	0.142				
MFD	*A*	90	395	18.8 ± 0.42	19.1 ± 0.41	0.202	*B*	19.1 ± 0.28	19.2 ± 0.25	0.449
(μm)	*B*	218	267	19.1 ± 0.42	18.8 ± 0.42	0.180				
	C	400	85	18.9 ± 0.42	19.1 ± 0.42	0.425	*B*	19.2 ± 0.25	19.2 ± 0.26	0.997
FDSD	*A*	90	395	4.1 ± 0.16	4.1 ± 0.15	0.639	*C*	4.1 ± 0.11	4.2 ± 0.10	0.455
(μm)	*B*	218	267	4.1 ± 0.15	4.0 ± 0.15	0.270	*C*	*4.2 ± 0.10*	*4.1 ± 0.10*	*0.096*
	C	400	85	4.0 ± 0.15	4.1 ± 0.15	0.194				
CVFD	*A*	90	395	21.5 ± 0.61	21.4 ± 0.59	0.721				
(%)	*B*	218	267	21.5 ± 0.60	21.3 ± 0.60	0.578				
	C	400	85	21.3 ± 0.60	21.6 ± 0.60	0.329				
MSL	*A*	90	395	84 ± 2.8	82 ± 2.8	0.199	*B,C*	83 ± 1.9	85 ± 1.7	0.342
(mm)	*B*	218	267	82 ± 2.8	84 ± 2.8	0.155	*A,C*	83 ± 1.9	85 ± 1.7	0.223
	C	400	85	**81 ± 2.8**	**85 ± 2.8**	**0.006**	*A,B*	**82 ± 1.8**	**86 ± 1.8**	**0.001**
MSS	*A*	90	395	24 ± 2.0	25 ± 2.0	0.724	*B*	24 ± 1.5	23 ± 1.3	0.240
(N/ktex)	*B*	218	267	*26 ± 2.0*	*24 ± 2.0*	*0.077*				
	C	400	85	25 ± 2.0	25 ± 2.0	0.721	*B*	23 ± 1.3	24 ± 1.4	0.149
PF (%)	*A*	90	395	1.9 ± 0.75	2.2 ± 0.73	0.437				
	B	218	267	2.0 ± 0.74	2.1 ± 0.74	0.766				
	C	400	85	1.9 ± 0.74	2.2 ± 0.74	0.432				
MFC	*A*	90	395	89 ± 3.5	91 ± 3.4	0.361	*B*	95 ± 2.3	96 ± 2.0	0.264
(°/mm)	*B*	218	267	92 ± 3.4	89 ± 3.4	0.185				
	C	400	85	90 ± 3.4	91 ± 3.4	0.913	*B*	96 ± 2.1	66 ± 2.2	0.818

GFW, greasy fleece weight; CFW, clean fleece weight; Yield, wool yield; MFD, mean fibre diameter; FDSD, fibre diameter standard deviation; CVFD, coefficient of variation of fibre diameter; MSL, mean staple length; MSS, mean staple strength; PF, prickle factor (the percentage of fibres of diameter greater than 30 microns); MFC, mean fibre curvature.
[a]Estimated marginal means and standard errors (S.E.) derived from the GLMMs. $P < 0.05$ are in bold, whereas $0.05 \leqslant P < 0.10$ are italicised.

reference samples for each variant were included in each gel to facilitate precise genotyping.

Statistical analysis

For each wool trait, statistical analyses were performed using Minitab version 16. Unless otherwise indicated, all *P*-values were considered statistically significant when $P < 0.05$ and trends were noted when $0.05 \leqslant P < 0.10$.

General Linear Mixed-effects Models (GLMMs) were used to evaluate the effect, if any, of the presence/absence of the *KRT31* promoter variants, or *KRT31* genotypes, on various wool traits. In all the models, gender and birth rank were fitted as fixed factors, and sire was fitted as a random factor.

The *KRT31* promoter variants were coded as either present (1) or absent (0) for each animal's genotype. For each wool trait, single-variant presence/absence models were then run. Any gene variant that had an association in the single-variant models with a *P*-value <0.2, and which could thus potentially impact on the trait, was then factored into multi-variant models, such that a correction was made for this variant in the genotype.

In the third set of models, any *KRT31* promoter genotypes present at a frequency of 5% or more (thereby ensuring adequate sample size) were tested to ascertain associations with the wool

https://doi.org/10.1017/S0021859619000406 Published online by Cambridge University Press

Table 3. The effect of *KRT31* promoter genotype on various wool traits

	Mean ± S.E.[a]					
Trait	*AA* (*n* = 164)	*AB* (*n* = 183)	*AC* (*n* = 48)	*BB* (*n* = 53)	*BC* (*n* = 31)	*P*
GFW (kg)	2.2 ± 0.10	2.3 ± 0.10	2.3 ± 0.11	2.2 ± 0.11	2.4 ± 0.11	**0.015**
CFW (kg)	1.6 ± 0.08	1.7 ± 0.08	1.7 ± 0.09	1.7 ± 0.09	1.8 ± 0.09	**0.009**
Yield (%)	75 ± 1.5	75 ± 1.5	75 ± 1.6	76 ± 1.6	77 ± 1.6	0.745
MFD (μm)	19.1 ± 0.44	18.8 ± 0.45	19.3 ± 0.49	18.5 ± 0.49	18.9 ± 0.47	0.385
FDSD (μm)	4.1 ± 0.16	4.1 ± 0.16	4.4 ± 0.18	4.0 ± 0.18	4.0 ± 0.17	0.310
CVFD (%)	21.5 ± 0.63	21.5 ± 0.63	22.1 ± 0.69	21.8 ± 0.70	21.0 ± 0.66	0.291
MSL (mm)	79 ± 2.9	82 ± 2.9	83 ± 3.2	81 ± 3.2	86 ± 3.1	**0.030**
MSS (N/ktex)	26 ± 2.1	24 ± 2.1	26 ± 2.3	24 ± 2.4	245 ± 2.2	0.421
PF (%)	1.9 ± 0.79	2.2 ± 0.79	2.5 ± 0.86	1.8 ± 0.87	2.0 ± 0.83	0.608
MFC (°/mm)	92 ± 3.6	89 ± 3.7	91 ± 4.0	87 ± 4.0	90 ± 3.8	0.575

GFW, greasy fleece weight; CFW, clean fleece weight; MFD, mean fibre diameter; FDSD, fibre diameter standard deviation; CVFD, coefficient of variation of fibre diameter; MSL, mean staple length; MSS, mean staple strength; MFC, mean fibre curvature; PF, prickle factor (percentage of fibres greater than 30 microns).
[b]Estimated marginal means, standard errors (S.E.) and *P* values derived from GLMs. Bonferroni correction fitted for repetitive testing. $P < 0.05$ are in bold.

traits. Multiple pairwise comparisons between genotypes were performed using a Tukey test with Bonferroni corrections.

Results

Variants of the KRT31 *promoter, exon 1, exon 3 and exon 7 regions*

For the regions investigated (*KRT31* promoter, exon 1, exon 3 and exon 7), three, two, two and two PCR-SSCP banding patterns were observed, respectively. Sequencing of the amplicons that produced unique SSCP banding patterns for each of the four regions of *KRT31* revealed sequences that were unique, but at least 99% homologous to the ovine genome (ovine genome assembly v4.0, 41053583-41054040, 41052719-41053137, 41051938-41052208, 41049576-41049866; for the *KRT31* promoter, exon 1, exon 3 and exon 7 regions, respectively).

There were two single nucleotide polymorphisms (SNPs) identified in the *KRT31* promoter region from the 300 NZ Romney, Merino and White Dorper sheep typed: c.−864C/G and c.−847A/G (Fig. 1). These gave rise to three variants (designated *A*, *B* and *C*; Fig. 1). Two nucleotide sequences were confirmed for the *KRT31* exon 1 region amplified in the 300 NZ Romney, Merino and White Dorper sheep and these were designated *A'* and *B'* (Fig. 1). They contained two SNPs: a synonymous SNP c.55A/C and an intronic SNP c.348+31A/G. The exon 3 region that was amplified produced two nucleotide sequences (designated *a* and *b*), and two SNPs c.522C/T, c.588+14C/T were detected (Fig. 1). The nucleotide substitution in the coding region was synonymous, and the other was in intron 3. Two unique variants were observed for the exon 7 region amplified and these were named *a'* and *b'*. They contained a SNP c.*20G/A (Fig. 1), with the nucleotide substitution occurring in a 3′ untranslated region.

The frequencies of *KRT31* promoter variants *A* to *C* in the Southdown × Merino-cross sheep were 56, 29 and 15%, respectively. Five different genotypes were observed with the following genotypes having frequencies over 5%: *AA* (28%); *AB* (30%); *AC* (18%); *BB* (8%) and *BC* (7%). The remaining genotype, *CC*, was only observed in 15 sheep.

Associations between variation in the KRT31 *promoter and wool traits*

No associations were detected with *A* and *B* in the GLMMs, but the presence of *C* was found to be associated with an increase in GFW, CFW and MSL (Table 2). These effects persisted in the multi-variant GLMMs. The phenotypic change was 5.9%, 7.2% and 4.8%, respectively, for GFW, CFW and MSL.

With the five genotypes (*AA*, *AB*, *AC*, *BB* and *BC*) that occurred at a frequency >5%, an effect of genotype on GFW, CFW and MSL was detected. Genotype *BC* sheep produced wool of higher GFW, CFW and MSL than *AA*, *AB*, *AC* and *BB* sheep (Table 3). Although the overall genotypic effect on GFW was significant ($P = 0.015$), the pairwise comparison revealed there were no significant pairwise differences in GFW between sheep of differing genotype (Table 3).

Discussion

Although many of the SNPs detected in *KRT31* occurred in introns and would not change the putative amino acid chain produced, variation in the introns of genes can influence gene expression by changing the activity of enhancer elements, or by influencing the production and splicing of primary transcripts (Fong and Zhou, 2001; Furger *et al.*, 2002; Kwek *et al.*, 2002). Further analysis of the effect of these SNPs could, therefore, involve precise analysis of the expression level of different variants in sheep of different genotypes.

Three promoter sequence variants were identified and defined by two SNPs (c.−864C/G, and c.−847A/G). Variation in promoters could affect transcription factor binding sites or enhancer elements, which potentially affects the expression of the gene. Keratin promoter regions have proven invaluable for targeting transgene expression to specific compartments within epithelia (Liu *et al.*, 2003). For instance, the promoter region of *KRT15* was found to target hair follicle bulge cells with specificity, and the *KRT5* and *KRT14* promoters drive expression of transgenes to the relatively undifferentiated basal cell layer of the epidermis and hair follicle (Liu *et al.*, 2003). It could, therefore, be claimed

https://doi.org/10.1017/S0021859619000406 Published online by Cambridge University Press

with some confidence that *KRT31* promoter variation might affect gene expression, at least until proven otherwise.

The exon 1, exon 3 and exon 7 sequence variants were identified by two (c.55A/C and c.348+31A/G), two (c.522C/T and c.588 +14C/T) and one (c.*20G/A) SNPs, respectively. These SNPs were either synonymous SNPs in the coding region, or located in intron regions, or the 3′ untranslated region. This would not change the putative amino acid sequence but could affect the rate of transcription. The occurrence of SNPs in intron regions can cause inactivation of mRNA splice donor site, which can ultimately result in the formation of premature stop codons or exon skipping, hence yielding a shorter mRNA (Kimchi-Sarfaty *et al.*, 2007). For example, a transition from G to A at the fifth position of intron-32 of the dystrophin gene causes a splicing error leading to transcript termination (Tran *et al.*, 2005).

Despite there being no evidence of variation in the amino acid sequence of K31, variation in the gene's promoter was found to be associated with a number of wool traits, including GFW, CFW and MSL. The presence of *C* in a sheep's genotype was associated with an average additional 200 g of CFW. This effect seemed to be consistent with the findings for MSL, suggesting the increase in GFW and CFW may have come about because of increased fibre growth, as opposed to an increase in fibre diameter traits. This effect was similar to the findings reported by Gong *et al.* (2015) with the KAP1-2 gene. This may, however, simply reflect the observation made by Gong *et al.* (2015) that MSL, CFW and GFW are moderately positively correlated traits.

The possibility exists that the effects observed for *KRT31* may be due to its linkage to other *KRTs* and *KRTAPs* on the same chromosome. Seventeen other hair *KRTs* and *KRTAPs* have been identified on sheep chromosome 11 near *KRT31* (Oar_v4.1 reference assembly), including *KRT32, KRT33a, KRT33b, KRT34 – KRT36, KRT38 – KRT40*, *KRTAP3-3*, *KRTAP3-2*, *KRTAP1-1*, *KRTAP1-2*, *KRTAP1-3*, *KRTAP1-4*, *KRTAP4-3* and *KRTAP4-1*. These *KRTs* and *KRTAPs* are clustered, potentially variable and expressed in the wool fibre. Their proximity to each other might mean that it would be difficult to isolate and illustrate the independent effects of individual *KRTs* and *KRTAPs*.

Clean fleece weight and MFD are two important traits in wool production and selection for either CFW or MFD may lead to a change in the other trait. However, in the current study, while there was an association with variation in GFW, CFW and MSL, there was no association observed with MFD. This suggests that selecting for improvements in GFW and CFW may be possible, without having an adverse effect on MFD. This is consistent with the conclusions of Gong *et al.* (2015), and supported by the Trangie QPLU$ project (CSIRO, Australia), which has revealed that both CFW and MFD can be improved concurrently using genetic selection (Mortimer *et al.*, 2006). It also would require further testing in more sheep of different breeds, genders and ages.

Author ORCIDs. H. Zhou, 0000-0001-7620-3160
J. Wang, 0000-0002-9811-0085

Financial support. The current work was in part supported by funding from the Foundation for Research Science and Technology (FRST) (C10X0710), Basic Research Creative Groups of Gansu Province (18JR3RA190) and the Lincoln University Gene-Maker Laboratory. We thank the support of the AGMARDT Postdoctoral Fellowship to Hua Gong, and the New Zealand Guardian Trust for the Vernon Willey Trust Fellowship to HZ.

Conflict of interest. None.

Ethical standards. Not applicable.

References

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

Fong YW and Zhou Q (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929–933.

Furger A, O'Sullivan JM, Binnie A, Lee BA and Proudfoot NJ (2002) Promoter proximal splice sites enhance transcription. *Genes & Development* **16**, 2792–2799.

Gong H, Zhou H, Yu Z, Dyer J, Plowman JE and Hickford JGH (2011) Identification of the ovine keratin-associated protein KAP1-2 gene (*KRTAP*1-2). *Experimental Dermatology* **20**, 815–819.

Gong H, Zhou H, Hodge S, Dyer JM and Hickford JGH (2015) Association of wool traits with variation in the ovine KAP1-2 gene in Merino cross lambs. *Small Ruminant Research* **124**, 24–29.

Gong H, Zhou H, Forrest RHJ, Li S, Wang J, Dyer JM, Luo Y and Hickford JGH (2016) Wool keratin-associated protein genes in sheep - a review. *Genes* **7**, article number 24. doi: 10.3390/genes7060024.

Hediger R, Ansari HA and Stranzinger GF (1991) Chromosome banding and gene localizations support extensive conservation of chromosome structure between cattle and sheep. *Cytogenetics & Cell Genetics* **57**, 127–134.

Heid HW, Werner E and Franke WW (1986) The complement of native α-keratin polypeptides of hair-forming cells: a subset of eight polypeptides that differ from epithelial cytokeratins. *Differentiation* **32**, 101–119.

Heid HW, Moll I and Franke WW (1988) Patterns of expression of trichocytic and epithelial cytokeratins in mammalian tissues. I. Human and bovine hair follicles. *Differentiation* **37**, 137–157.

Itenge-Mweza TO, Forrest RH, McKenzie GW, Hogan A, Abbott J, Amoafo O and Hickford JGH (2007) Polymorphism of the KAP1.1, KAP1.3 and K33 genes in Merino sheep. *Molecular and Cellular Probes* **21**, 338–342.

Itenge TO, Hickford JGH, Forrest RH, McKenzie GW and Frampton CM (2010) Association of variation in the ovine KAP1.1, KAP1.3 and K33 genes with wool traits. *International Journal of Sheep and Wool Science* **58**, 1–19.

Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV and Gottesman MM (2007) A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**, 525–528.

Kwek KY, Murphy S, Furger A, Thomas B, O'Gorman W, Kimura H, Proudfoot NJ and Akoulitchev A (2002) U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nature Structural Biology* **9**, 800–805.

Liu Y, Lyle S, Yang Z and Cotsarelis G (2003) Keratin 15 promoter targets putative epithelial stem cells in the hair follicle bulge. *Journal of Investigative Dermatology* **121**, 963–968.

Lynch MH, O'Guin WM, Hardy C, Mak L and Sun TT (1986) Acidic and basic hair/nail ("hard") keratins: their colocalization in upper cortical and cuticle cells of the human hair follicle and their relationship to "soft" keratins. *Journal of Cell Biology* **103**, 2593–2606.

Mortimer S, Taylor P and Atkins K (2006) The Trangie QPLU$ selection lines: responses in clean fleece weight and fibre diameter on completion of ten rounds of selection. In Pope CE (ed.) *Trangie QPLU$ Merinos – Open Day 2006. Proceedings of the Trangie QPLU$ Open Day*. Orange, Australia: NSW DPI, pp. 7–11.

Parsons YM, Piper LR and Cooper DW (1994) Linkage relationships between keratin-associated protein (KRTAP) genes and growth hormone in sheep. *Genomics* **20**, 500–502.

Popescu C and Höcker H (2007) Hair - the most sophisticated biological composite material. *Chemical Society Reviews* **36**, 1282–1291.

Rogers GR, Hickford JGH and Bickerstaffe R (1993) Mspi RFLP in the gene for a type I intermediate filament wool keratin. *Animal Genetics* **24**, 218.

Rogers GR, Hickford JGH and Bickerstaffe R (1994) A potential QTL for wool strength located on ovine chromosome 11. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* **21**, 291–294.

Roldan DL, Dodero AM, Bidinost F, Taddeo HR, Allain D, Poli MA and Elsen JM (2010) Merino sheep: a further look at quantitative trait loci for

https://doi.org/10.1017/S0021859619000406 Published online by Cambridge University Press

wool production. *Animal: An International Journal of Animal Bioscience* **4**, 1330–1340.

Sumner RMW, Forrest RHJ, Zhou H, Henderson HV and Hickford JGH (2013) Association of the KRT33A (formerly KRT1.2) gene with live-weight and wool characteristics in yearing Perendale sheep. *Proceedings of the New Zealand Society of Animal Production* **73**, 158–164.

Tran HTT, Takeshima Y, Surono A, Yagi M, Wada H and Matsuo M (2005) A G-to-A transition at the fifth position of intron-32 of the dystrophin gene inactivates a splice-donor site both in vivo and in vitro. *Molecular Genetics & Metabolism* **85**, 213–219.

Yu Z, Gordon SW, Nixon AJ, Bawden CS, Rogers MA, Wildermoth JE, Maqbool NJ and Pearson AJ (2009) Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles. *Differentiation* **77**, 307–316.

Zhou H, Hickford JGH and Fang Q (2006) A two-step procedure for extracting genomic DNA from dried blood spots on filter paper for polymerase chain reaction amplification. *Analytical Biochemistry* **354**, 159–161.

https://doi.org/10.1017/S0021859619000406 Published online by Cambridge University Press