Copyright Statement

The digital copy of this dissertation is protected by the Copyright Act 1994 (New Zealand).

This dissertation may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- you will use the copy only for the purposes of research or private study
- you will recognise the author's right to be identified as the author of the dissertation and due acknowledgement will be made to the author where appropriate
- you will obtain the author's permission before publishing any material from the dissertation.
Analysis of the Keratin Associated

Protein Gene KRTAP6-1 in New Zealand Wiltshire Sheep

A Dissertation
submitted in partial fulfilment
of the requirements for the Degree of
Bachelor of Agricultural Science with Honors

at
Lincoln University
by
Emma Wilding

Lincoln University
2016
Abstract of a Dissertation submitted in partial fulfilment of the requirements for the Degree of Bachelor of Agricultural Science with Honors.

Abstract

Analysis of the Keratin Associated Protein Gene KRTAP6-1 in New Zealand Wiltshire Sheep

by

Emma Wilding

Keratin associated proteins (KAPs) and keratins (Ks) are key elements in the structure of wool fibre and determine the phenotypes seen in sheep. The abundance of the KAP gene family (KRTAP-n) varies in wool fibre. Variation within the ovine KAP gene KRTAP6-1 was screened for using PCR-SSCP analysis. One hundred and eight New Zealand Wiltshire sheep from a North Canterbury stud were sampled for DNA, and additional to five previously reported variants of KRTAP6-1, A – E, two new variants, F and G, were found. The sampled Wiltshire sheep had five KRTAP6-1 gene variants, A, B, D, F and G, but variants C and E were not found. Of the five gene variants identified, B was the most common at a frequency of 56%, followed by the new variants F and G at 16% and 12% respectively. A and D had frequencies of 11% and 4% respectively. The KRTAP-n genes coding for wool fibre traits are multi-locic and highly polymorphic making selective breeding for desirable traits difficult. By gaining a better understanding of the genes involved in wool fibre properties more consistent fibre production may be achieved. Improved knowledge may offer us clues as to what underpins the fundamental differences between Wiltshire sheep and other breeds.

Keywords: KAP6-1, KRTAP6-1 gene, Wiltshire, variation, wool shedding, wool trait, sheep.
Acknowledgements

This work was financially supported by Lincoln University department of Animal Science, and the Lincoln University Gene-Marker Laboratory and Professor Jonathan Hickford supervised the research undertaken. S. Byun is thanked for her extensive technical assistance.
# Table of Contents

Abstract ............................................................................................................ ii
Acknowledgements .......................................................................................... iii
Table of Contents ............................................................................................ iv
List of Tables .................................................................................................... v
List of Figures .................................................................................................. v

Chapter 1 Review of Literature ........................................................................ 1
  1.1 Objectives of study .................................................................................... 7

Chapter 2 Materials and Methods ..................................................................... 8
  2.1 Sheep blood samples ................................................................................ 8
  2.2 PCR primers and PCR amplification ......................................................... 8

Chapter 3 Results ............................................................................................... 10
  3.1 Variants of KRTAP6-1 ............................................................................. 10
  3.2 Alignment of KRTAP6-1 nucleotide sequence and variants ....................... 13
  3.3 Alignment of KAP6-1 amino acid sequence and variants ......................... 15

Chapter 4 Discussion ......................................................................................... 16
  4.1 Variation within KRTAP6-1 .................................................................... 16
  4.2 KRTAP6-1 variant C ................................................................................ 16
  4.3 New variants of KRTAP6-1 ..................................................................... 17
  4.4 Specificity of SSCP gel conditions ......................................................... 18
  4.5 Limitations of PCR-SSCP analysis .......................................................... 18
  4.6 Potential further research ....................................................................... 20
  4.7 Conclusion ............................................................................................... 21

References ....................................................................................................... 22
List of Tables

Table 3.1  Genotype frequencies within the sampled Wiltshire sheep population..........................................................12

List of Figures

Figure 1.1  Electron microscopy image of wool fibre from a cross-bred animal demonstrating the extent of variation between individual wool fibres (J. Hickford, personal communication November 7th, 2016). ..............................................................................................................2

Figure 3.1  Seven distinct banding patterns of KRTAP6-1 variants represented in both homozygous and heterozygous forms. 10.5% acrylamide:bisacrylamide (37.5:1: Bio-Rad) gel, 17°C, 260V, 24 hours. .........................................................................................................................11

Figure 3.2  Seven distinct banding patterns of KRTAP6-1 variants represented in both homozygous and heterozygous forms in 10.5% acrylamide:bisacrylamide (37.5:1: Bio-Rad) gel with 1.5% glycerol, 17°C, 260V, 24 hours.........................................................................................................................12

Figure 3.3  Nucleotide sequence of KRTAP6-1 variants (Note: KAP6-1 in image is referring to KRTAP6-1) .................................................................................................................14

Figure 3.4  Amino acid sequence of KAP6-1 variants........................................................................................................15
Chapter 1

Review of Literature

Wool is a natural fibre that is in global demand. It possesses natural qualities such as fire resistance, having antistatic properties, being a good insulator and when woven or knitted, and it maintains breathability. Uses for wool range from making high-quality clothing such as suits and underwear, through to the manufacture of carpet and rugs. The value of wool is based on various fibre properties such as mean fibre diameter (MFD), fibre diameter standard deviation (FDSD), the coefficient of variation of fibre diameter (CVFD), prickle factor (PF), wool staple length, strength and colour. Finer fibre is ideal for next-to-skin use, such as undergarments, suits or baby apparel, whereas coarse fibre is suited for carpets where it will withstand tread wear.

Markets for wool are very specific and the manufacturers and consumers requirements determine how farmers set their breeding objectives for wool production. Typically small, high-value niche markets for superfine wool (low MDF) exist at one end of the spectrum, while prices are lower for larger markets for strong and coarse (high MFD) fibre. Consistency in fibre properties both within and between fleeces is desirable, as it improves wool quality and processing performance. It allows the wool to better compete with synthetics and other man-made fibres. Improvement in fibre consistency would raise the value of a fleece, but the ability to achieve this improvement is hampered by a limited understanding of wool genes and proteins. The keratin associated proteins (KAPs) are a key structural component of wool fibre, and exploring how they differ between sheep breeds will help us to better understand their functions and interactions. This understanding should also allow us to breed for and produce better and more consistent wool.

The New Zealand Wiltshire sheep is a large white-faced rare breed and exhibits fundamental differences to more common and conventional New Zealand sheep breeds, such as the Romney and Merino. It is an exceptionally old breed, usually farmed for prime-lamb meat production and not fibre production. They are a persistent breed and economically desirable due to their high fecundity, committed mothering ability and production of fast-growing lambs. Some are horned, however conventionally polled sheep are more desirable as it makes for a safer and more manageable animal. Wiltshires have short white hair on their heads, bellies, legs, and around the perineal region, however there are varying degrees of coverage that exist. Underneath the tail is bare skin, removing the need for docking and significantly reducing the risk of dag formation and fly strike. Wiltshires have a natural photoperiod-induced wool shedding ability (Craven, Ashby, Scobie & Nixon, 2007). Typically, this shedding occurs in the spring, but there is a large amount of variation in shedding
ability and timing. Inconsistency in the timing of shedding makes their wool difficult to harvest, but this is production wastage is offset by the removal of the requirement to pay for shearing. Wiltshire sheep typically produce wool with a mean fibre diameter between 30-33 microns. In fibre production, the breeding goals for Wiltshire sheep are different to other breeds, with emphasis placed on the timing and consistency of shedding.

Figure 1.1  Electron microscopy image of wool fibre from a cross-bred animal demonstrating the extent of variation between individual wool fibres (J. Hickford, personal communication November 7th, 2016).

There is large variation in both Wiltshire fibre properties and shedding ability, and both between individuals, and between sites on the body, illustrated in figure 1.1. This is partly the result of heritable differences and genetic interactions, but also largely influenced by external factors such as nutrient intake, hormone levels, pregnancy, and climate influences. The Wiltshires shedding ability sets it apart from breeds conventionally farmed for their wool, such as the Merino, but the wool follicle growth cycle it experiences is very similar to that of conventional breeds (Craven et al., 2007). The physiological mechanisms within the wool growth cycle that occur are consistent across breeds,
with three key phases in this seasonal cycle; anagen, catagen and telogen. Anagen is the phase of active mitotic growth during the summer and spring. The catagen phase is then signalled by a photoperiod response (Parry, Pearson, Morris & Still, 1991), and then regression of the follicle occurs where the blood supply is cut off and no new cells are added to the fibre. The telogen phase can be described as dormancy, when the follicle rests and the attached fibre becomes a dead hair ready to be shed.

There is variation again in the extent and length of each follicle phase a sheep will go through, Wiltshires can be viewed at one extreme end of the follicle cycle continuum. A less extreme example is the Merino; a breed which does not completely lose the fibre produced, but does go through summer high growth phases and winter dormancy (Sumner, Clarke, Pearson & Speedy, 1994). Other sheep breeds have less pronounced wool growth phases.

Parry et al. (1991) examined primary and secondary follicle activity in twelve New Zealand Wiltshire sheep, and found mean peak follicle activity (approximately 90% active) was between January and April which decreased to 50% between June and July. Secondary follicle activity was consistently higher than in primary follicles. The telogen phase lasted approximately two months, during the winter and anagen activity induced shedding at the end of winter. Shedding was observed as a gradual process, beginning on the neck, tail and underside, and spreading up the flanks to the topside of the sheep. Although all sheep exhibited a similar trend in timing there was a large amount of variation between the rate of activity and timing of phases, despite sheep being kept under identical controlled conditions. Both Parry et al. (1991) and Craven et al. (2007) highlight the range of differences between individuals observed.

Wool is composed of keratin and other proteins, which are chemically inert and insoluble. The keratins are not only found in wool, but also in hair, fingernails, hooves and horns. The physical properties of a wool fibre are determined by the structural components; and variation in these components can explain some of the wool fibre variation both within and between sheep breeds. Wool fibre is a complex structure. There are a huge number of polymorphic genes involved in fibre production, many containing multiple alleles. The interaction of these genes determines the phenotype expressed within the continuums of traits such as MFD, FDSD, staple length, crimp and strength (Wang, Zhang, Yang, Wang, Rong, Pei, W., ... & Wang, 2014).

The wool fibre outer layer is called the cuticle, a scaled protective layer with a waxy coating to repel water, while retaining the ability to absorb moisture. The cuticle impacts the wool’s durability and its ability to felt and shrink (Koehn, Plowman, Morton, & Dyer, 2015). Underneath is the cortex containing long, overlapping cells surrounded by a cell membrane complex made up of vulnerable proteins and waxy lipids held together by relatively weak covalent bonds. These cells, called ortho-
cortical and para-cortical cells, differ chemically. They are distributed more uniformly in fine fibre than in coarse fibre, and when moisture is absorbed they expand in different ways. This determines fibre crimp, which is also directly related to fibre diameter, hence why finer fibre crimps more readily than coarse. Coiled macrofibrils exist within the cortical cells, made up of very strong microfibrils embedded within a protein matrix. The microfibrils are composed of helical spring-like coils of protein chains, which are supported by hydrogen bonds and disulphide bonds prevent the extension of the tiny coils, while maintaining the flexibility of the wool.

Historically the protein components of wool fibre were classified into two sub categories; high and low sulphur proteins. Subsequently, these were re-classified as α-keratins (Ks) and keratin-associated proteins (KAPs), respectively. The α-keratins are a strong coiled amino acid sequence, with reoccurring heptamer repeats (Fraser, MacRae, Parry & Suzuki, 1986). They are intercoiled, held together predominantly by ionic bonds (although hydrophobic and hydrogen bonds also contribute) into keratin intermediate filaments (KIFs), which are set within a KAP matrix.

Ks and KAPs contribute up to 90% of the weight in a wool fibre (Koehn et al., 2015). Keratin associated proteins (KAPs) are a crucial element in the fibre structure (Reviewed by Gong, Zhou, Forrest, Li, Wang, Dyer, Luo & Hickford, 2016). They provide the matrix in which keratin intermediate filaments lie and are attached by covalent bonds. In effect the coiled KIFs interact through disulphide bonds with matrix proteins to build macro fibrils (Rogers, 2003). Disulphide bonds are a key element of wool structure, as they stabilise the structure of the KIFs in the KAP matrix. Cystine residues are involved in the formation of these disulphide bonds both within and between KAPs and keratins (Wang, Parry, Jones, Idler, Marekov & Steinert, 2000). These KIF macro fibrils help to determine the properties of the cortical cells, thus contributing to wool fibre properties. While KAPs do not have a large impact on KIF structure, they are believed to impact the co-ordination of KIFs into larger sized arrangements.

The chemical structure of KAPs varies considerably. They are rich in sulphur-containing amino acids; the sulphur content of wool is important as sulphur attracts water molecules, giving wool its water absorbent properties. The chemical structure of the KAPs is mostly distinguished by the presence of either cysteine, or glycine-tyrosine, or both. The KAPs can be categorized into three groups; the ultra-high sulphur content KAPs, KAPs 4, 5, 9 and 17, containing >30 mol% cysteine; the high sulphur content KAPs, KAPs 1–3, 10–16 and 23, 24 and 26.1, containing <30 mol% cysteine and the high glycine-tyrosine (HGT) family, containing KAP6-1, (Zhou, Gong, Li, Luo & Hickford, 2015), alongside KAPs 6–8 and 18–22 (Koehn et al., 2015).

High sulphur content KAPs exist predominantly in the protein matrix, interacting with intermediate filaments (Dowling, Jones & Speakman, 1990), while the low sulphur (LS) KAPs and those with high
garcene-tyrosine content, are more concentrated in the orthocortex (Zhou et al., 2015), and in the microfibrils (Dowling, Jones, & Speakman, 1990). The occurrence of LS KAPs in cortical cells means they are likely to impact the traits that cortical cells govern, such as crimp, fibre diameter and prickle factor (Zhou et al., 2015). Even with similar cuticle structure, clear differences exist between sheep possessing different numbers of different KAPs (Koehn et al., 2015), and it can therefore be assumed that KAPS contribute to sheep breed differences.

The multi-genic KAP6 family (KAP-n) contains KAPs with low sulphur and high glycine-tyrosine content (Gong, Zhou & Hickford, 2011). They are expressed in the orthocortex (Gong et al., 2016; Zhou et al., 2015) along with genes from the KAP8 family. The KAP-n family contains variation within the majority of its known genes; KRTAP6-1, 6-2, 6-3, 6-4 and 6-5. KRTAP6-1, 6-2 and 6-3 were recently assigned within this family, and KRTAP6-1 was found by Gong, Zhou and Hickford (2011) to have five gene variants, A – E, with variation in both coding sequence and sequence length. Zhou et al. (2015) found that of the three variants A, B and C, C was unique in that it contained a 57 base pair deletion, causing a loss of 19 amino acids. The genotype was found to be associated with changes in wool fibre properties such as an increase in MFD, FDS, CVFD, and prickle factor (PF). While the C variant of KRTAP6-1 was associated with fibre traits, it cannot be confidently ascertained that it is the factor that controls these traits as the associations may just reflect correlated traits.

Over time, selective breeding has successfully resulted in improvements being made in production animals, but breeding for wool production is a slow process (Purvis & Franklin, 2005), due to the huge complexity in genes determining the fibre properties. The KAPs are an essential structural component of wool fibre, and gene interactions and the occurrence of multiple alleles in their genes determines wool traits. In this study, variation in ovine KRTAP6-1 was screened for to determine if differences in wool fibre qualities between Wiltshire sheep and other non-shedding New Zealand based breeds, such as Merinos, can be identified. The expression of KAPs begins within the wool follicle. It has been proposed that the KAP6 family is limited to the orthocortex, with much higher expression in wool fibre than in human hair.

The breeding process for wool traits can become very complex, like many other phenotypic traits there are multiple genes that code for varying forms of a phenotype. The wool structure is determined by many genes that code for K proteins and KAPs, influencing the wool staple length, strength, diameter, color etc. (McKenzie, Abbott, Zhou, Fang, Merrick, Forrest, Sedcole & Hickford, 2010).

Over 100 KAPs and K genes have been identified (Gong, Zhou, Hodge, Dyer & Hickford, 2015), and multigenic systems like make wool traits difficult to breed for as interacting genes all contributing in their individual way may create 'noise' impacting a single phenotype. This complexity means that
when animals are artificially selected for a desirable wool trait, although some genetic gain may be made, it is likely that some desirable genes will not be passed on, particularly if they are located distally on a chromosome or on separate chromosomes. Additional to this each gene has multiple alleles, sometimes there are only two alleles for a gene however some traits are under the influence of multiple alleles meaning there are more than two forms of the same gene.

Breeding for wool-related traits was historically to satisfy a market demand for a quality end product destined for a specific use. Sought after wool fibre is fine, strong and consistent, however in the case of breeding Wiltshire sheep, wool is not considered a production trait. Shed wool has the potential to be harvested, however currently the economic for doing this is low.

The key aim for a Wiltshire sheep breeder fleece is to obtain sheep that have a ‘clean’ shed with timing consistent to that of the rest of the flock. While this may come across as wastage of a commodity that could be sold, it means farmers can focus on production traits with the potential to generate higher returns than the harvest of wool; sheep serve a dual purpose, they are farmed for both their wool and meat. Fecundity, lamb growth rates and meat production are worth far more in New Zealand and international markets than wool is, unless the fibre is of very high quality and low MFD. Wiltshires also present the opportunity to sell to a niche market as a low maintenance breed of stock, suited for small lifestyle blocks, hobbyist and older people who are not capable an annual shear of their flock. Often the cost of shearing a small number of sheep compared to the returns of the fibre generates a loss, and wool is better left to be shed at no cost.

Over time selective breeding has shifted fibre growth away from a seasonal shedding cycle in order to preserve wool as a harvestable commodity, to meet a market demand. Interest is focused on the fundamental differences in the physical structure of the fibre, driven by gene expression, rather than seasonal shedding which is driven by hormonal responses. A better understanding of the genes involved in wool growth and variation in wool properties may allow more consistent fibre production, and in this respect analysis of key wool genes in Wiltshire sheep may offer us some clues as to what underpins the difference between this breed that only seasonally produces wool of low value and other non-shedding breeds that consistently produce higher value wool.
1.1 Objectives of study

The primary objective of this study was to screen for ovine genetic variation in the keratin associated protein gene KRTAP6-1. The New Zealand Wiltshire was chosen for analysis due to the removed nature of its breeding and genetics from conventional breeds, such as the Merino. Any variation seen in the sampled Wiltshires could then be compared to that seen in previous studies, such as Zhou et al. (2015). Improving the understanding of this gene may allow for increased accuracy when breeding for wool-related traits in the future, alongside giving the chance of identifying the driving factor for wool shedding in Wiltshire sheep.
Chapter 2
Materials and Methods

2.1 Sheep blood samples

One hundred and five polled New Zealand Wiltshire ewes, and three rams, were sampled to determine whether variation in ovine KRTAP6-1 is consistent with that seen previously in New Zealand Merinos. The stud is North Canterbury based and breeding ewes are selected on the basis of clean and consistent shedding ability, and all ewes are ear tagged to sire.

Blood samples were collected from these sheep, directly onto blotting paper, by crushing the tip (5mm) of the ear until the second drop of blood appeared. The animal’s corresponding tag number was recorded alongside its blood sample. Samples were then allowed to dry at room temperature for five days and then stored in folders.

DNA for analysis was then collected from the blotting paper using 1.2 mm punches, and purified using the two-step procedure described by Zhou et al. (2006).

2.2 PCR primers and PCR amplification

PCR primers (5’-TCTACCGAGAACAACCTC-3’ and 5’ AGGCAAGTCTTTAGTAGGAC-3’) (with nucleotide co-ordinates relative to the KRTAP6-1 sequence Genbank M95719) were designed based on published ovine KRTAP6 sequences (Fratini et al., 1993; Gong et al., 2011, as cited in Zhou et al., 2015) and the Oar_v3.1 chromosome 1 sequence (NC_019458), to specifically amplify the entire coding sequence of intronless KRTAP6-1, located on chromosome 1. PCR amplification was carried out in a 15 μl reaction consisting of purified DNA sample from blotting paper, 9.69 μl of purified water, 1.5 μl of 10xQ, 1.5 μl of 5Q, 1.5 mM of MgCl2, 0.150 μM of dNTP (Eppendorf, Hamburg, Germany), 0.25 μM of primer, 0.5 U of Taq DNA Polymerase (Qiagen, Hilden, Germany) and 1 x reaction buffer supplied. Thermal cycling began with 2 minutes at 94°C, followed by 35 cycles of 30 seconds at 60°C, 30 seconds at 72°C and a final extension of 5 minutes at 72°C. Amplification took place in Lincoln University Gene Marker laboratory’s S1000 thermal cyclers (Bio-Rad, Hercules, CA, USA).

Amplicons were first visualised using electrophoresis in 1% agarose gels using 19 TBE buffer (89 mM Tris, 89 mM boric acid, 2mM Na2EDTA) containing 200ng/ml of ethidium bromide.

PCR amplicons were screened for sequence variation using single strand conformational polymorphism (SSCP) analysis. PCR extracted DNA first had 100 μl of loading dye added (98%
formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene–cyanol), and was then
denatured at 100 °C for 5 minutes, followed by rapid cooling on wet ice. Aliquots of 7 μl were loaded
into acrylamide:bisacrylamide (37.5:1) (Bio-Rad) gel wells. SSCP analysis of amplicons was optimised
under two different sets of conditions. Firstly 10.5% acrylamide:bisacrylamide (37.5:1: Bio-Rad) gels
at 260V for 24 hours, at 17°C in 0.5 x TBE buffer. Secondly, 10.5% acrylamide:bisacrylamide (37.5:1:
Bio-Rad) at 260V for 24 hours at 18°C with 1.5% glycerol added. Electrophoresis was performed using
Protean II xi cells (Bio-Rad).

Gels were silver-stained according to the method of Byun et al. (2009) as cited in Zhou et al., (2015).
Amplicons of previously genotyped sheep containing specific KRTAP6-1 variants were included in the
SSCP gels as reference standards. Heterogenous amplicons produced complex SSCP banding
patterns. Based on the comparison of reference standards to Wiltshire sheep DNA in the banding
patterns seen, variation present between samples was determined.

Single bands of interest were recovered directly from the SSCP gel as a gel slice. This was macerated
and the DNA was eluted into 50 μl TE buffer by incubating at 70°C for 20 min. 1 micro litre of this
eluted solution was used as a template for further PCR amplification using the original primers. The
PCR amplicons were directly sequenced at the Lincoln University DNA Sequencing facility. Sequence
alignments, translations and comparisons were carried out using DNAMAN (version 5.2.10, Lynnon,
BioSoft, Vaderaui, Canada).


Chapter 3

Results

3.1 Variants of *KRTAP6-1*

Seven independent variants of *KRTAP6-1* were identified using PCR-SSCP analysis. For five of these, polymorphism has been previously identified in studies by Zhou et al. (2015) including allele variants A, B, C, D and E. Variants E and C were not seen in the sampled population. The banding patterns seen in the SSCP gel (figure 3.1) show animals carrying homologous and heterogenous forms of the five variants. When further analysed it was revealed that these five variants each have unique elements to their gene sequence. Differentiating variants in the SSCP gel is most clear when viewing the lower banding pattern, however variants B and D, B and E, F and G exhibit exceptional similarity. The addition of 1.5% glycerol to a repeat of the initial gel caused the amplified DNA samples to move faster through the gel matrix, and extended the spaces between DNA bands, causing the upper bands to be further defined. This increased clarity helps to determine the difference between variants exhibiting high similarity.

Of the five variants that were seen, the respective frequencies were 56%, 16%, 12%, 11% and 4% for B, F, G, A and D. The most common variant was homologous BB at 32% (table 3.1).
Figure 3.1 Seven distinct banding patterns of KRTAP6-1 variants represented in both homozygous and heterozygous forms. 10.5% acrylamide:bisacrylamide (37.5:1: Bio-Rad) gel, 17°C, 260V, 24 hours.
Figure 3.2 Seven distinct banding patterns of KRTAP6-1 variants represented in both homozygous and heterozygous forms in 10.5% acrylamide:bisacrylamide (37.5.1: Bio-Rad) gel with 1.5% glycerol, 17°C, 260V, 24 hours.

Table 3.1 Genotype frequencies within the sampled Wiltshire sheep population.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Appearances</th>
<th>% Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>1</td>
<td>0.9</td>
</tr>
<tr>
<td>BB</td>
<td>34</td>
<td>31.8</td>
</tr>
<tr>
<td>AB</td>
<td>15</td>
<td>14.0</td>
</tr>
<tr>
<td>BD</td>
<td>4</td>
<td>3.7</td>
</tr>
<tr>
<td>FD</td>
<td>2</td>
<td>1.9</td>
</tr>
<tr>
<td>DG</td>
<td>2</td>
<td>1.9</td>
</tr>
<tr>
<td>AG</td>
<td>4</td>
<td>3.7</td>
</tr>
<tr>
<td>BG</td>
<td>15</td>
<td>14.0</td>
</tr>
<tr>
<td>FG</td>
<td>5</td>
<td>4.7</td>
</tr>
<tr>
<td>FF</td>
<td>3</td>
<td>2.8</td>
</tr>
<tr>
<td>AF</td>
<td>3</td>
<td>2.8</td>
</tr>
<tr>
<td>BF</td>
<td>19</td>
<td>17.8</td>
</tr>
</tbody>
</table>
3.2 Alignment of *KRTAP6-1* nucleotide sequence and variants

Seen in figure 3.3 the variants that arise in *KRTAP6-1* are predominantly single nucleotide polymorphisms (SNPs), with the exception of a 52 base-pair deletion occurring in variant C (figure 3.3) which results in the loss of 19 amino acids.

Variants A and C are identical, with the exception of C’s 57 base pair deletion. G is also very similar to A and C, with one SNP, where the base change is from adenine to guanine, causing an amino acid substitution. Variant E contains a silent SNP from guanine to cytosine. Variants B and F each contain three SNPs, none of which are at coinciding locations, and all of which are silent.
Figure 3.3 Nucleotide sequence of KRTAP6-1 variants (Note: KAP6-1 in image is referring to KRTAP6-1)
3.3 Alignment of KAP6-1 amino acid sequence and variants

The amino acids with the potential to be transcribed from each of the seven nucleotide base variants of KRTAP6-1 are illustrated on the template in figure 3.4 below. The majority of the SNPs in KRTAP6-1 are silent, however variant G displays a difference.

In variant G at the 26th amino acid (109th nucleotide seen in the given KRTAP6-1 sequence) there is a putative amino acid change from the wild type; illustrated in figure 3.4. A base substitution from adenine to guanine causes an amino acid substitution. The wild type of the gene at the point where the substitution occurs codes for the amino acid tyrosine (Y), nitrogenous bases TAT, which converts to UAU in an RNA transcript. Base A is substituted for G, changing the codon to TGT, which converts to UGU in an RNA transcript, coding for amino acid cysteine (C). This amino acid substitution in variant G occurs proximate to where the 57 base pair deletion in variant C occurs.

Figure 3.4  Amino acid sequence of KAP6-1 variants.

<table>
<thead>
<tr>
<th>MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</th>
<th>Majority</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 20 30 40</td>
<td></td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-A</td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-B</td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-C</td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-D</td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-E</td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-F</td>
</tr>
<tr>
<td>41 MCGYGGNYGGGLCGSYGGGLCGGYGSGFRRLLC</td>
<td>KAP6-1-G</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</th>
<th>Majority</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 60 70 80</td>
<td></td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-A</td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-B</td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-C</td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-D</td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-E</td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-F</td>
</tr>
<tr>
<td>167 CGYGYGSRSLCGSGYGYGSGSRSLCGSGYGYGSGF</td>
<td>KAP6-1-G</td>
</tr>
</tbody>
</table>

Decoration 'Decoration #:1': Box residues that match the Consensus exactly.


Chapter 4

Discussion

4.1 Variation within KRTAP6-1

Variation in the KRTAP6 family was recognised by Gong, Zhou and Hickford (2011) in Merino, Romney, Coopworth and cross breed sheep, and further investigated within KRTAP6-1 by Zhou et al. (2015) in Southdown x Merino cross lambs. In this study one hundred and eight Wiltshires were DNA sampled and genotyped using PCR-SSCP analysis. For one animal no genotype was recorded, as the DNA did not successfully display a banding pattern during multiple rounds of SSCP analysis. In the Wiltshires successfully genotyped, variation was identified within KAP6-1; variants A, B and C, previously identified by Zhou et al. (2015), and variants D and E. Two variants, F and G, had not been seen before in sampled sheep populations but were found to exist in this particular population; there are now seven known variants of KRTAP6-1. Surprisingly, the gene variant C was not seen in the sampled Wiltshires, however C tends to occur less frequently in populations previously tested, compared to A and B. The variation profile seen within the sampled flock of Wiltshires can be considered as representative of this individual Wiltshire population, and not the New Zealand Wiltshire population, as the sample size is not extensive enough to assume a breed-wide representation.

In this study gene testing using the primers targeting KRTAP6-1 began with a broad sample of fifty conventional breeds of sheep. The DNA screened for variation was randomly selected from stores of DNA at the Lincoln University Gene Marker Laboratory, previously submitted from properties around New Zealand for commercial testing of specific genetic traits. Breeds sampled were Merino, Romney, Suffolk, Corriedale, Texel, Perendale, Poll Dorset, Dorset Down, Dorper and Coopworth. Five samples per breed were screened for variation. Variants A-E were found to be present in other breeds using the PCR-SSCP analysis conditions specified here. However, the two KRTAP6-1 variants F and G, that not been previously identified, were also not seen in any of the conventional sheep breeds screened for variation, suggesting that these variants are unique to Wiltshires.

4.2 KRTAP6-1 variant C

The KRTAP6-1 variant C did not occur within the sampled Wiltshire flock, but was identified by Zhou et al. (2015) in sampled Southdown Merino cross lambs. Zhou et al. (2015) found that the presence of variant C was associated with increased wool MFD, FDSD, CVFD and prickle factor. Individuals carrying variant C had coarser and less consistent wool. Seen in the SSCP gel in figure 3.1 there is a large margin of difference between the lower banding pattern of variant C and other gene variants.
This large difference is most likely due to the 57 base pair deletion that occurs (figure 3.3), which probably results in smaller, and less, DNA fragments. As a result, these fragments move faster through the SSCP gel matrix when electrophoresis is carried out, hence the variant C’s band is much lower on the gel than other samples not containing the 57 base pair deletion.

Seen in figure 3.4, this loss of bases also results in a loss of 19 amino acids, which is likely the reason behind the coarser wool traits seen in animals carrying variant C. The 57 base pair deletion in variant C causes a loss of four aromatic, and two basic, amino acid residues (Zhou et al., 2015) which contribute to the binding properties of KAP6-1 and IF proteins. The binding strength is weakened, causing coarser and less consistent fibre in carriers of variant C observed by Zhou et al. (2015).

While the lack of variant C suggests a breed difference between the sampled Wiltshires and other sampled animals, Zhou et al. (2015) identified the occurrence of variant C to be very low, at 5.8%. This indicates that variant C is a less common form of KRTAP6-1 and the sample size used here may simply not be large enough to account for C. It cannot be assumed that Wiltshires do not carry variant C.

The 57 base pair deletion in variant C may affect the structure and/or expression of the KAP6-1 protein and/or its interaction with IF proteins. This would consequently have an effect on wool fibre diameter associated traits. However, it was also speculated by Zhou et al. (2015) that the association between coarser fibre traits and variant C could be due to gene linkage on chromosome 1, or maybe other KAP6, 7 or 8 genes. This could mean that KRTAP6-1 is not the driver behind the observed traits, but merely the link between phenotype and other genetic traits, if it is linked at all. To confirm this theory further research should be undertaken into a wider variety of breeds.

### 4.3 New variants of KRTAP6-1

In the Wiltshire population tested, the frequency of variants F and G was high, 16% and 12% respectively, considering they have never been seen in previously screened populations. F and G were the second and third most seen variants, following variant B. This is consistent with the results achieved by Zhou et al. (2015) who found variant B to have the highest frequency at 49% in a sampled population.

In variant G at the 26\textsuperscript{th} amino acid there is a putative amino acid change from the wild type (illustrated in figure 3.4). A base substitution from adenine to guanine causes an amino acid substitution. The wild type of the gene at the point where the substitution occurs codes for the amino acid tyrosine (Y), nitrogenous bases TAT, which converts to UAU in an RNA transcript. Base A is substituted for G, changing the codon to TGT, converting to UGU in an RNA transcript, which codes for amino acid cysteine (C). It is at this point unknown what fibre related trait the change from Y to C
change could impact, if any, however there is a likelihood it will create phenotypic change, even if the change is exceptionally small and not obvious upon visual inspection.

The substituted amino acid cysteine is a crucial element for wool growth and usually the first limiting amino acid for fibre synthesis. It is involved in the interaction of KIFs with matrix proteins, through the formation of disulphide bonds, both within and between KAPs and Ks (Wang, Parry, Jones, Idler, Marekov & Steinert, 2000). The substitution of cysteine seen might have an effect on the structure and/or the expression of the KAP6-1 protein and its interaction with IF proteins, hence the potential to impact MFD related traits in variant G carriers.

Because the amino acid substitution seen in variant G occurs at the same base as the deletion in variant C begins, is it is likely that the phenotypical consequences will be similar.

To verify the theory that variant G has an impact on phenotypical traits, further testing of the association between carriers of G and the differences of MFD related traits should be carried out. Additional to further Wiltshire sampling, the Dorper is a candidate for further testing, as it also exhibits a seasonal fleece shed, and it can be assumed that similar genes may be found.

4.4 Specificity of SSCP gel conditions

Breed screening discussed in section 4.1, found no other sheep breeds carrying KRTAP6-1 variants F and G, suggesting that these variants could be unique to the Wiltshire. However, only five samples from each breed (50 altogether) were screened, and it is likely that if variants F and G did exist in other breeds they have been missed due to too small a sample size. Wider screening has the potential to reveal more carriers of F and G.

The SSCP conditions used to identify these seemingly unique variants are very specific, gels were run under two different sets of conditions to gain a clear enough image to confidently separate gene variants from one another; Firstly 10.5% acrylamide:bisacrylamide (37.5:1: Bio-Rad) gels were run at 260V for 24 hours, at 17°C in 0.5 x TBE buffer. Secondly, 10.5% acrylamide:bisacrylamide (37.5:1: Bio-Rad) were run at 260V for 24 hours at 18°C with 1.5% glycerol added. The similarity of the banding patterns of variants B and D, B and E, F and G is high, and there is a possibility that these new variants have occurred in the previous testing, but were not recognised due to insufficient conditions to visibly isolate them.

4.5 Limitations of PCR-SSCP analysis

Prior to SSCP analysis, DNA and primer viability were tested using agarose gels. These gels are far smaller, taking less than 25 samples, compared to an SSCP which can run 50 at a time. They are also less specific than SSCP gels, however they are quick carry out, saving the time and effort of running
inviable SSCP gels. Once DNA and primer viability was confirmed, agarose gels were not again required.

PCR-SSCP analysis of DNA is an effective and efficient method for analysis of genotypes, particularly when large sample numbers are involved. Up to 25 samples can be loaded in a single gel, and two gels are loaded into a buffer for electrophoresis at a time. Extracting and amplifying DNA through PCR is the most laborious element of the typing process, taking one day to carry out, once samples are loaded into SSCP gels after denaturation and the addition of loading dye, a 24-hour wait is required before gel staining. In this study there was a trial period that occurred to optimise the gel conditions, approximately 25 gels were run before the ideal conditions to isolate the variants of KRTAP6-1 were decided upon. Condition optimisation included altering the percentage of the acrylamide: bisacrylamide gels, and extending the time taken for electrophoresis, to achieve a clearer lower banding pattern, but not for so long that they ran off the gel; a complication encountered by variant C. Two sets of bands were visualised, upper and lower as seen in figure 3.1. Once the lower banding pattern was optimised the addition of glycerol was tested. Glycerol tended to increase the clarity of the upper banding pattern, which helped to determine the difference between exceptionally similar variants B and D, B and E and F and G.

Samples tended to be clearest with fresh blood samples that had soaked through the blotting paper at collection, and with a larger volume of amplified DNA loaded into gel well (7μl, versus 5 μl). Once amplified, DNA samples lasted 2-3 weeks in the chiller before deterioration, and re-extraction was required.

For one animal in the sampled population a genotype was not recorded, due to unsuccessful visualising of the banding pattern in the stained SSCP gel. DNA was extracted multiple times to rule out the possibility of contamination or human error during DNA extraction and PCR, and it was concluded that the extracted DNA was inviable for either PCR, or SSCP visualisation. This could be explained by the DNA sample on the blotting paper not being strong enough. Non-specific binding of primers could also have the same effect, cutting DNA in locations other than the binding site it was designed for. This is unlikely, as primers were specifically designed to amplify the entire intronless region of KRTAP6-1, however there is a small possibility that alternative identical DNA sequences exist in the genome and were targeted by the primers. This would result in extra fragments of DNA being amplified, which could cause interference with the desired sequence, resulting in the absence of bands, or non-specific DNA bands appearing.

Another possibility is that this sheep carries a variant of KRTAP6-1, with one or more nitrogenous base changes within the primer binding site. This polymorphism in a homozygous animal could prevent primer attachment and hence no PCR product would be seen. To determine if this is the case
the primer binding sites of the KRTAP6-1 used in this study would need to be amplified using alternate primers on the DNA of the animal in question.

There is the potential for other variants of KRTAP6-1 to exist, which are not being amplified using the current primers used, which raises the question of whether all the variants of KRTAP6-1 within the successfully genotyped population are being captured. Heterozygous animals may have one variant form of KRTAP6-1 amplified, and another form silenced due to non-specific binding of primers. They would appear to be homozygous for the one variant seen on SSCP gels, but in reality are heterozygous, with one unamplified allele.

If a silent polymorphism in the primer binding region occurred that still allowed the primers to attach, the sequence would still be amplified, and the PCR product would use the primers as a template for amplification, rather than the sequence itself. This would result in a viable PCR product that appeared identical to other gene sequences, and the inconsistency in the gene promoter region would be missed.

4.6 Potential further research

As far as we know the new variants of KRTAP6-1, F and G, are confined to the Wiltshires in the sampled population, however there is the possibility that they exist in other Wiltshires or alternate breeds but have not been identified yet. To determine if this is the case Wiltshires from a separate lineage than those tested here should be screened for the presence of F and G. Furthermore, the Dorper is a candidate for being a carrier of these new variants, due to their high level of relatedness to the Wiltshire and similar wool-related characteristics.

Although KAP6-1 is known to be primarily expressed in the orthocortex of wool fibre, relative content in primary follicles compared to secondary follicles is still unknown. Comparing follicles and their relative shedding ability could highlight significant differences in fibre properties and bring to light new breeding objectives for Wiltshire breeders.

Investigating other genes with similar functions to KRTAP6-1 (i.e. orthocortex expressed) could bring to light more information on the impact of KAPs and Ks on fibre traits. Determining what KAP/s contribute to the spring wool break seen in Wiltshires would open the possibility of breeding for more consistency in a shedding flock. Because of the incomplete nature of the shed in some animals and the variation of shedding ability within a flock it is presumed that shedding is under the control of multiple polymorphic genes. Additionally, animals are observed to shed more readily when in good health and at higher live weights, compared to animals under nutritional stress. This is a hormonal response, and to identify the genetic drivers behind fleece shedding a hormonal inconsistencies need to be removed.
The Dorper is a candidate breed for further screening for similarities in variants within KRTAP6-1. Due to its relatedness to the Wiltshire it is likely that similar genotypes will be seen, along with the potential for new variants to arise.

4.7 Conclusion

There is a possibility that the newly identified variants F and G of KRTAP6-1 could contribute to Wiltshire shedding ability, particularly as thus far they are the only breed known to carry these variants. Some shedding Wiltshires sampled, did not carry these variants, and while F and G most likely influence wool traits, it is very unlikely that variants F and G are major contributors to the wool shedding ability. The vast and polymorphic nature of not only KRTAP6-1, but all KAP and K proteins make it exceptionally difficult to pinpoint specific genes that determine desirable wool traits.

Over time, selective breeding has successfully changed and improved wool traits to meet a specific market requirement, and undoubtedly advances in fibre production will continue to be made. Improving the understanding of the genes that drive fibre production traits may allow for further improvement in fibre production and bring us closer to understanding the reasons behind major breed differences, such as variation in MFD and shedding ability. Improved understanding will help to hasten the development of economically viable breeds in New Zealand, and internationally.
References


