Genetic factors controlling wool shedding in a composite Easycare sheep flock

Genetic factors controlling wool shedding in a composite Easycare sheep flock

O. Matika, S. C. Bishop, R. Pong-Wong, V. Riggio and D. J. Headon

The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland, UK

Corresponding author:

Oswald Matika E-mail: oswald.matika@roslin.ed.ac.uk

Tel: +44 (0) 131 651 9100

Fax: +44 (0) 131 651 9105

Summary

Historically, sheep have been selectively bred for desirable traits including wool characteristics. However, recent moves towards extensive farming and reduced farm labour have seen a renewed interest in easy care breeds. The aim of this study was to quantify the underlying genetic architecture of wool shedding in an Easycare flock.

Wool shedding scores were collected from 565 pedigreed commercial Easycare sheep from 2002 to 2010. The wool scoring system was based on a 10 point (0-9) scale, with score zero for animals retaining full fleece and nine for those completely shedding.

DNA was sampled from 200 animals of which 48 with extreme phenotypes were genotyped using 50k SNP chip. Three genetic analyses were performed: heritability analysis, complex segregation analysis to test for a major gene hypothesis and a genome wide association study (GWAS) to map regions in the genome affecting the trait. Phenotypes were treated as a continuous or binary and categories.

High estimates of heritability (0.80 when treated as a continuous, 0.65-0.75 as binary and 0.75 as categories) for shedding were obtained from linear mixed model analyses. Complex segregation analysis gave similar estimates (0.80±0.06) to those above with additional evidence for a major gene with dominance effects. Mixed model association analyses identified four significant ($p<0.05$) SNPs. Further analyses of these four SNPs in all 200 animals revealed that one of the SNP displayed dominance effects similar to those obtained from the complex segregation analyses.

In summary, we found strong genetic control for wool shedding, demonstrated the possibility of a single putative dominant gene controlling this trait and identified four SNPs that may be in partial linkage disequilibrium with gene(s) controlling shedding.

Keywords: Genetic control, Sheep, Wool shedding, GWAS, Segregation analysis.

Introduction

During the process of domestication wild animals have undergone selective breeding to alter their growth, reproduction, behaviour and morphology. In sheep one of the most notable morphological differences between most domestic breeds and their wild ancestor, the mouflon, is the growth of a non-shedding woolly fleece. This fleece has been produced by selection for continuous year-round hair growth, rather than a seasonally intermittent hair growth with conspicuous moulting of the winter coat in the spring.

Physiological hair shedding is a result of cyclic bouts of hair follicle growth activity, programmed destruction and dormancy. The intermittent nature of the growth phases means that hair fibres do not grow continuously, but instead newly grown fibres displace the hairs remaining from the previous bout of growth, the latter being shed from the skin. The hair length at a particular site on the body, or on a given species, depends largely on the duration of the follicles' active growth phase (Stenn & Paus 2001). The hair cycle is modulated by several endocrine inputs (Randall 2007; Chen & Chuong 2012), with seasonal changes in hormone profiles playing a major role in

species with strongly seasonal life cycles. In mammals, such seasonal timing relies on interactions between the hypothalamus, pituitary, pineal and thyroid glands, together with the input of environmental cues such as day length. This results in the production of hormones such as melatonin, prolactin and thyroid hormone with profound effects on a range of target organs (Hazlerigg & Wagner 2006). The coupling of hair growth cycles with season permits the growth of a winter coat that is shed in the spring to allow for greater insulation specifically in the coldest part of the year, and also enables seasonal changes in colouration as hairs only acquire pigment while actively growing. In non-shedding sheep, this potentially valuable wool requires annual shearing to maintain the health of the animal and this represents a cost to producers who, based on the recently prevailing market value of wool, often focus on the meat qualities of their flock.

Recent trends towards more extensive farming and reduced farm labour have seen a renewed interest in low cost and low maintenance breeds of sheep. Consequently, breeding for easier-to-manage sheep, as a solution to counteract the effects of reduced labour available on farms, has been proposed as one of the selection objectives in current extensive farming systems in the UK (Conington *et al.* 2010). One attribute contributing to such breeds is natural wool shedding, as this trait when fully expressed negates the requirement for annual shearing. Other advantages conferred by wool shedding include a reduced susceptibility to blowfly strike (Litherland *et al.* 1992) and likely an increased ability to withstand heat stress in summer, these attributes probably representing the selective advantage of seasonal moulting in wild species.

Predisposition to shedding varies greatly between breeds and is under some degree of genetic control. Historically, amongst farmed breeds in the UK wool shedding has been observed and studied in the Wiltshire Horn, in studies comparing this breed to non-shedding breeds such as Blackface and Tasmanian Merino sheep (Slee 1959; Slee & Carter 1961). The gene variants which lead to shedding are presumably the ancestral forms of the genes since, in addition to being observed in the mouflon, this phenotype has predominantly been observed in feral sheep populations (Rudge 1983; Orwin & Whitaker 1984; Van Vuren & Bakker 2009), wild Soay sheep (Boyd *et al.* 1964; Ryder 1971; Doney *et al.* 1974), and other breeds such as the Blackhead Persian sheep (Duerden & Boyd 1930). It is also seen in composites such as Easycare, Dorper and Katahdin (Pollott 2011), often as a result of a deliberate breeding policy. Whilst environmental factors undoubtedly contribute to fleece shedding or retention, recent evidence suggests that wool shedding is a complex trait which is influenced by a dominant gene of major effect (Pollott 2011).

This study builds on the observation of Pollott (2011), and aimed to quantify the underlying genetic architecture of wool shedding, using an Easycare sheep flock bred specifically to comprise individuals which shed their fleece. We performed our study using both quantitative genetic analyses, to confirm the inheritance of wool shedding and to suggest a putative mode of inheritance for this characteristic, and DNA analyses to identify regions of the genome containing loci influencing the predisposition of the fleece to shed.

Materials and methods

Data description

Wool shedding score data from 565 sheep were collected from 2002 to 2010 in a commercial flock of Easycare sheep maintained at latitude 55.57° N. Pedigree records were also available for these sheep. This flock had a history of Lleyn, Meatlinc and Blackface rams being mated to an Easycare ewe flock. Animals were scored by the farmer for wool shedding in their second year of life, in June every year, with the scoring system based on a 10 point (0-9) scale. Animals which retained all their wool were scored zero and those which completely shed their fleece were scored nine. A summary of the available data and pedigree information is given in Table 1 and a full description of the scoring system is shown in Table 2, along with the number of animals scored in each category. This scoring system was developed by the farmer and photos of animals representative of the 10 scores are presented in the supplementary material.

A total of 200 phenotyped sheep were DNA sampled, using a non-invasive commercial nasal collection kit (Performagene Livestock, supplied by DNA Genotek, Kanata, Canada). Of these 200 sheep, 48 animals of extreme phenotypes (7 and 13 with non-shedding scores of 0 and 1, respectively, and 28 with score 9) were genotyped using the Illumina® Ovine 50SNP BeadChip (50k SNP chip). Quality control (QC) measures applied to the genotype data included cut-off rates of 0.05 for minor allele frequency and 0.80 for call rates. Deviations from Hardy-Weinberg equilibrium were not considered as an appropriate method for excluding SNPs since the population was a composite of several breeds. After QC, 45,133 SNPs were available for further analyses.

Statistical analyses

Three genetic analyses were performed: (i) a heritability analysis, (ii) a segregation analysis to determine if the data provided evidence for the segregation of a putative major gene affecting wool shedding and (iii) a genome wide association study (GWAS) to map regions in the genome affecting the trait.

Heritability estimation. For heritability estimation, the data were analysed using ASReml software (Gilmour *et al.* 2009), fitting linear mixed models using restricted maximum likelihood (REML). Phenotypes were treated as a continuous trait (WSCORE), or coded as 0,1 (BIN1 defined as wool scores 0,1 *vs*. the rest, BIN2 defined as wool scores 0,1,2 *vs.* the rest) and as categories (WSCAT where wool scores 0,1 were non-shedders, scores 2-7 were medium shedders and 8,9 were complete shedders). Year (2002-2010), sex (males and females), type of birth (singles, twins, triplets) and age of dam (1 to 6+ years) were fitted as fixed effects. These fixed effects were used in all subsequent analyses. Random effects fitted included animal and permanent environmental effects due to dam and litter. The following models were fitted:

$$\mathbf{y}=\mathbf{X\beta} + \mathbf{Z}_1\mathbf{u} + \boldsymbol{e} \qquad (1)$$

$$\mathbf{y}=\mathbf{X\beta} + \mathbf{Z}_1\mathbf{u} + \mathbf{Z}_2\mathbf{l} + \mathbf{Z}_3\mathbf{pe} + \boldsymbol{e} \qquad (2)$$

where **y** is the vector of phenotypes with **β, u, l, pe** and ***e*** being vectors of fixed effects, additive polygenic, litter effects, permanent environment due to the dam and residuals, respectively, with incidence matrices **X** , $\mathbf{Z}_1$, $\mathbf{Z}_2$ **and** $\mathbf{Z}_3$**.** The terms **u**, **l**, **pe** and **e** were assumed to be normally distributed: $N(0, A\sigma^2_a)$, $N(0, I\sigma^2_l)$, $N(0, A\sigma^2_{pe})$,

and $N(0, I\sigma^2_e)$, respectively.

The data were further explored deleting either the litter or permanent environment terms from model (2). All traits were analysed as continuous and in addition, the binary data were analysed using sire and sire plus dam models fitting a logit link function.

$$\mathbf{y=X\beta + Zs + Zd + } \boldsymbol{e} \qquad \textbf{(3}\boldsymbol{)}$$

where **y** is again the vector of phenotypes with **β, s , d** and ***e*** being vectors of fixed effects, additive sire, additive dam and residuals, respectively, with incidence matrices **X and Z.** The terms **s**, **d** and **e** were assumed to be normally distributed: $N(0, A\sigma^2_s)$, $N(0, A\sigma^2_d)$, and $N(0, I\sigma^2_e)$, respectively.

It was not possible to account for breed-of-origin effects since this information was not available on many animals.

Complex segregation analysis. Secondly, data were explored using complex segregation analyses (Walling *et al.* 2002), implemented using a Gibbs sampler, to formally investigate the major gene hypothesis. For this Bayesian segregation analysis, data were fitted as a continuous variable, the genetic component was partitioned into a major gene and a polygenic component, and fixed effects were fitted as described above. Both additive and dominance effects were estimated for the putative major gene, and the full model was as follows:

$$\mathbf{y=X\beta + Zu + Zg + Zd + } \boldsymbol{e} \qquad \textbf{(4}\boldsymbol{)}$$

where **y** is the vector of phenotypes with **β, u, g, d** and ***e*** being vectors of fixed

effects, additive polygenic, major gene effect, major gene dominance effect and residuals, respectively, with incidence matrices **X** and **Z.** The terms **u**, **g, d** and **e** were assumed to be normally distributed: $N(0, A\sigma^2_a)$, $N(0, A\sigma^2_g)$, $N(0, A\sigma^2_d)$ and $N(0, I\sigma^2_e)$, respectively. Reduced models lacking either the dominance term or both the additive and dominance terms were also fitted, and these models were used to test for evidence of an additive effect and a dominance effect, in turn.

Flat priors were used for both fixed effects and variance components. Parameters were drawn from the posterior conditional distributions. A chain of 15,020,000 cycles was run for each trait, with a burn-in of 20,000 rounds, keeping every 50th sample for inference of posterior features (i.e. 300,000 iterations were used for inference). Bayes Factors as a summary of evidence provided by the data in favour of one proposed mode of inheritance *vs.* another were used to distinguish between different models; the test statistic being twice the difference between the natural logarithm of the Bayes Factor in the contrasted models (Kass & Raftery 1995). Values greater than 10 were considered as very strong evidence against accepting the model supporting the null hypothesis.

Genome-wide association analysis. The GWAS was performed using the GenABEL package (Aulchenko *et al.* 2007) in R environment (http://www.r-project.org), with data fitted as a continuous variable. The first step consisted of fitting mixed models with both fixed and polygenic effects to each trait, the latter to account for genetic relationships amongst the 48 animals, fitting the same fixed effects as above. Covariances amongst the polygenic effects were fitted using the marker-based genomic relationship matrix. Thus, this step was equivalent to fitting model (1),

except that the A-matrix was replaced by the G-matrix. Secondly, association was tested using an mmscore function (Chen & Abecasis 2007) on the residuals, which have been corrected for familial relatedness using the genomic relationship matrix, and thus should be independent of pedigree or prior selection. After Bonferroni correction, the genome-wide ($p<0.05$) and the suggestive (i.e., one false positive per genome scan) significance threshold were $p<1.11x10^{-6}$ and $2.22\ x10^{-5}$, respectively. Although only two of the SNPs crossed these thresholds, the top 20 SNPs identified with p-values greater than $4x10^{-4}$ were then tested in full mixed model analyses (model 1) fitting the same fixed effects as described above, polygenic effects and SNP as a fixed effect. Since only 48 animals were genotyped using the 50k chip, potentially giving poor corrections for fixed effects, the association analysis was conducted using the data from all 565 animals to better correct for the fixed effects. Consequently, the A matrix was fitted rather than the G matrix, and an additional fixed effect category of 'genotyped or not' was fitted to account for animals without genotype information.

Following the association analyses, four SNPs out of the 20 tested were significant using $p<0.05$ nominal threshold, including the 2 SNPs significant from the GWAS. Therefore these four SNPs were considered most likely to be in linkage disequilibrium (LD) with the causative mutation and were assessed in all 200 animals for which DNA was available. We observed that these four SNPs were located on the same chromosomes as the known Texel and Lleyn mutations at myostatin locus (c.*1232G > A on OAR2, (Clop *et al.* 2006)) and the GDF9 locus on OAR5 (Hanrahan *et al.* 2004). Since the Easycare flock under study had a known history of

Meatlinc (and hence Texel) and Lleyn introgressions, these two mutations were also genotyped in the 200 DNA sampled animals. Association analyses were then performed in ASReml on the six SNPs fitting the SNP as fixed effects (model 1, fitting the A matrix). These analyses also enabled us to estimate the additive and dominance effects of each SNP, which was not possible in the original 48 genotyped animals as they were chosen due to their extreme phenotypes. The phenotypes explored were those of wool scores treated as: a) a continuous variable (WSCORE), b) as three categories (WSCAT) and c) as a binary trait (BIN1 and BIN2). The results of these analyses were decomposed into genetic effects and variance components as follows. Defining AA, BB and AB to be the predicted trait values for each genotype class, p and q to be the SNP allele frequencies and VA to be the total additive genetic variance of the trait obtained when no SNP effects are included in the model, genetic effects were then calculated as follows: additive effect, a = (AA - BB)/2; dominance effect, d = AB - [(AA + BB)/2]; and proportion of genetic variance due to SNP = $[2pq\ (a + d(q - p))^2]/VA$.

Results

Heritability estimation

All linear mixed model analyses showed wool shedding to be highly heritable (Table 3). When wool shedding score was analysed as a continuous trait all heritabilities were close to 0.8, and the effects of litter and permanent environmental effects due to the dam were both negligible. Similarly high heritability estimates were obtained when wool shedding was treated both as binary (BIN1 = 0.65±0.08, BIN2 =

0.79±0.08) and as categories (WSCAT = 0.75±0.08). Furthermore, the heritabilities estimated from the sire and dam model using the logit link function were also very high (BIN1 sire-based heritability = 1.17±0.41, dam-based heritability = 0.70±0.27, BIN2 sire-based heritability = 1.09±0.37, dam-based heritability = 0.68±0.25) with a total additive heritability of 0.93±0.17 for BIN1 and 0.89±0.16 for BIN2 (Supplementary Table 1).

Complex segregation analysis

Complex segregation analysis fitting a polygenic model, gave the same heritability (0.80±0.06) as that obtained from the REML analyses of wool shedding as a continuous trait (Table 4). However, from inspection of the Bayes Factors there was strong evidence that the major gene model gave a better fit to the data that the polygenic model. The test statistic ((-2Log$_e$(Bayes Factor $_{\text{model2-model1}}$)) was 433 comparing the major gene (additive effect) model with the polygenic model, and 135 comparing the major gene with dominance model against the major gene (additive effect) model. Thus, the segregation analysis revealed significant evidence for a locus with a major effect on fleece shedding, with the mode of inheritance of this putative locus suggesting that shedding is likely to be dominant. The frequency of the putative wool shedding allele was 0.37. However, this locus does not explain all of the genetic variation in wool shedding and the heritability of fleece shedding, after accounting for the putative major gene effect, was 0.35.

Wool shedding distributions are presented in Figure 1, showing the expected frequency of each putative genotype from the segregation analyses and a histogram for residuals obtain for wool score after accounting for fixed effects. The residuals

show a bimodal distribution, which is captured by the posterior frequency distributions for the inferred genotype classes.

The phenotypes and inferred genotypes of the 48 SNP genotyped animals are shown in Table 5, with 'F' representing the 'Fixed' (i.e. non-shedding) allele and 'S' representing the 'Shedding' allele. The frequency for the "shedding" allele from the inferred genotypes was 0.39 in this subset of animals.

Association Analyses

A Manhattan plot of the GWAS results is presented in Figure 2. The single highly significant SNP on OAR3 was discounted on account of having a very low minor allele frequency which yielded a pattern of relationships between the genotype and phenotype that was inconsistent with the population structure. Of the 45 successfully genotyped animals (i.e., three animals had missing genotype at this locus), 42 were of the 'CC' genotype and were distributed across all phenotypes. Four of the 20 top SNPs identified by GWAS (GenABEL), two on OAR2 ($P < 4.5x10^{-6}$ and $9.1x10^{-5}$, respectively) and two on OAR5 ($8.0x10^{-5}$ and $1.0x10^{-4}$, respectively) were significant ($p<0.05$) when tested using mixed model association analyses. The two significant SNPs on OAR2 are moderately linked (Figure 3), being separated by 13 SNPs (0.44Mb apart), with the shown r^2 values reflecting moderate LD among the SNPs spanning the interval between SNP1 and SNP2. The SNPs on OAR5 were further apart at 59.41Mb and unlinked ($r^2=0.03$).

When tested on the population of 200 animals, all four SNPs that had been identified in the SNP chip dataset (48 animal) showed significant ($p<0.05$) associations with wool shedding (Table 6). However, there was no association ($p>0.05$) between the

myostatin locus SNP and any of the wool shedding phenotypes, and the GDF9 SNP was monomorphic with the wild type allele fixed in this population. Genotype means, additive and dominance effects and QTL heritabilities are shown in Table 7. The most significant SNP, SNP2, displayed dominance ($p<0.05$) for increased shedding, in agreement with the results from the segregation analyses. This SNP also displayed the largest QTL heritability, apparently explaining more than 50% of the genetic variance when wool shedding was analysed as a binary trait. It should be noted, however, that the estimated additive and dominance effect sizes are smaller than those obtained from the segregation analyses, which is suggestive of incomplete LD between this SNP and the causative mutation.

Discussion

The high estimates of heritability in our study are consistent with previous field observations that a high proportion of Wiltshire Horn F1 crosses shed their wool (Slee 1959; Slee & Carter 1961, 1962). More recently, Pollott (2011) reported heritability estimates of 0.54 in lambs and 0.26 in animals of all ages in one flock using Easycare, Wiltshire Horn, Katahdin and Dorper shedding animals. The estimated heritability in the current study was somewhat higher, at ca. 0.8. However, since this flock was recently established (2002) and had a history of introgressions, not being able to account for breed in these analyses may have biased the genetic estimates upwards.

The complex segregation analysis not only confirmed the high genetic estimates, but also suggested that shedding was largely, but by no means entirely, controlled by a major gene with dominance effects. As we regard seasonal fleece shedding as the

ancestral trait, the dominant nature of this locus suggests that the derived trait, i.e. fleece retention, is caused by a recessive mutation. This result is in agreement with that drawn by Pollott (2011), *viz*. that the likely mode of inheritance was autosomal dominance after examining Mendelian ratios between shedders and non-shedders, derived from Wiltshire Horn F1. A plot of the residuals, after removing fixed effects, in our data (Figure 1) also displays a bimodal distribution which is consistent with a major gene hypothesis. This is more evident when we overlay the genotypic probabilities obtained using the Bayesian segregation analysis (Figure 1).

The further SNP association analysis (with 200 animals) allowed us to discount the hypothesis that the regions on OAR2 and OAR5 found using GWAS on 48 animals could have been breed effects due to population admixture arising from either the Texel or Lleyn ancestry. In addition, because the 200 animals comprised the whole spectrum of shedding scores, rather than being selected extremes, we were able to estimate dominance effects for each locus. Dominance was most pronounced for the most significant SNP, i.e. SNP2, and at this SNP the evidence for dominance was strongest when the trait was expressed as a categorical trait, i.e. fleece shed *vs*. retained.

The quantitative nature of the shedding trait is regionally restricted in the skin such that unshed wool in animals with intermediate shedding scores is not distributed across the body, but instead has a strong tendency to persist on the back and hindquarters. This region phenomenon may arise from the initiation of shedding typically occurring on the belly and propagating across the skin as a wave, as documented in other animal species (Plikus & Chuong 2008). The range of shedding

phenotypes that we observed in our population by the time of scoring in mid-summer (see supplementary material) most likely illustrates the action of mutations that either modulate the timing of initiation of shedding on the belly, or that influence the speed of propagation of the shedding wave towards the back.

A hypothesis that is consistent with these results is that one of the two regions identified is in partial LD with the causative mutation which triggers fleece shedding and the other is more likely to modify the rate and extent of shedding. This is consistent with the finding from the segregation analyses that fleece shedding remained a heritable trait even after accounting for the effect of a major segregating locus. Further, from our marker analyses, SNP3 and SNP4 were generally significant after fitting SNP2, however the level of significance of these two SNPs tended to be greater when shedding was analysed as a continuous or three-category variable than when shedding was analysed as a binary trait (results not shown). It should be noted that the SNPs identified in this study do not on their own fully explain all the genetic variance for wool shedding and this may be an indication that they are in partial LD with the causative mutation or that full expression of the shedding trait involves other loci that this study did not have the power to detect.

Conclusions

We found high estimates of heritability for wool shedding in Easycare sheep, confirmed by analyses of SNP markers on a subset of animals, with complex segregation analyses suggesting that shedding is partially explained by a putative single dominant gene. GWAS results have identified two regions which may contain the putative causal mutation. Furthermore, an extended SNP analyses provided

estimates of the effects of these two previously identified regions and revealed dominance effects consistent with the segregation analysis. Although our results contribute to the understanding of wool genetics, it should be noted that this study was conducted with a limited number of animals in one Easycare flock.

Acknowledgements

The authors would like to thank Ann, Tom and Sandy Welsh for access to data and DNA, Sam Boon (SIGNET) for the pedigree information. Funding was provided by a SPARK award from Biosciences KTN and BBSRC.

References

Aulchenko Y.S., Ripke S., Isaacs A. & Van Duijn C.M. (2007) GenABEL: an R library for genorne-wide association analysis. *Bioinformatics* **23**, 1294-6.

Boyd J.M., Doney J.M., Gunn R.G. & Jewel P.A. (1964) The Soay sheep of the island of Hirta, St Kilda. A study of a feral population. *Proceedings of the Zoological Society of London* **142**, 129-63.

Chen C.-C. & Chuong C.M. (2012) Multi-layered environmental regulation on the homeostasis of stem cells: The saga of hair growth and alopecia. *Journal of Dermatological Science* **66**, 3-11.

Chen W.M. & Abecasis G.R. (2007) Family-based association tests for genomewide association scans. *American Journal of Human Genetics* **81**, 913-26.

Clop A., Marcq F., Takeda H., Pirottin D., Tordoir X., Bibe B., Bouix J., Caiment F., Elsen J.M., Eychenne F., Larzul C., Laville E., Meish F., Milenkovic D., Tobin J., Charlier C. & Georges M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics* **38**, 813-8.

Conington J., Collins J. & Dwyer C. (2010) Selection for easier managed sheep. *Animal Welfare* **19**, 83-92.

Doney J.M., Ryder M.L., Gunn R.G. & Grubb P. (1974) *Colour, conformation, affinities, fleece and patterns of inheritance of the Soay sheep*. The Athlone Press of the University of London, London.

Duerden J.E. & Boyd E. (1930) The blackhead Persian. A primitively-coated fat-rumped sheep. *Union So Africa Dept Agric Bull* **82**, 1-46.
Gilmour A.R., Gogel B.J., Cullis B.R. & Thompson R. (2009) ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.
Hanrahan J.P., Gregan S.M., Mulsant P., Mullen M., Davis G.H., Powell R. & Galloway S.M. (2004) Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (Ovis aries). *Biology of Reproduction* **70**, 900-9.
Hazlerigg D.G. & Wagner G.C. (2006) Seasonal photoperiodism in vertebrates: from coincidence to amplitude. *Trends in Endocrinology and Metabolism* **17**, 83-91.
Kass R.E. & Raftery A.E. (1995) Bayes Factors. *Journal of the American Statistical Association* **90**, 773-95.
Litherland A.J., Sorenson E., Niezen J. & Bishop D. (1992) A pilot evaluation of shedding sheep breeds compared with non-shedding breeds for susceptibility to nematodes and flystrike. *Proceedings of the New Zealand Society of Animal Production, Vol 52 1992* **52**, 233-5.
Orwin D.F.G. & Whitaker A.H. (1984) Feral sheep (Ovis-aries L) of Arapawa Island, Marlborough sounds, and a comparison of their wool characteristics with those of four other feral flocks in New Zealand. *New Zealand Journal of Zoology* **11**, 201-24.
Plikus M.V. & Chuong C.M. (2008) Complex hair cycle domain patterns and regenerative hair waves in living rodents. *Journal of Investigative Dermatology* **128**, 1071-80.
Pollott G.E. (2011) A suggested mode of inheritance for wool shedding in sheep. *Journal of Animal Science* **89**, 2316-25.
Randall V.A. (2007) Hormonal regulation of hair follicles exhibits a biological paradox. *Seminars in Cell & Developmental Biology* **18**, 274-85.
Rudge M.R. (1983) A reserve for feral sheep on Pitt Island, Chatham group, New-Zealand. *New Zealand Journal of Zoology* **10**, 349-63.
Ryder M.L. (1971) Wool growth cycles in Soay sheep. *Journal of Agricultural Science* **76**, 183-&.
Slee J. (1959) Fleece shedding, staple length and fleece weight in experimental Wiltshire Horn-Scottish Blackface sheep crosses. *Journal Agriculural Science (Camb.)* **53**, 209-23.
Slee J. & Carter H.B. (1961) A comparative study of fleece growth in Tasmanian Fine Merino and Wiltshire Horn ewes. *Journal of Agricultural Science* **57**, 11-&.
Slee J. & Carter H.B. (1962) Fibre shedding and fibre-follicle relationships in the fleeces of Wiltshire Horn x Scottish Blackface sheep crosses. *Jour Agric Sci* **58**, 309-26.
Stenn K.S. & Paus R. (2001) Controls of hair follicle cycling. *Physiological Reviews* **81**, 449-94.
Van Vuren D.H. & Bakker V.J. (2009) Rapid morphological change in an insular population of feral sheep. *Journal of Zoology* **277**, 221-31.
Walling G.A., Bishop S.C., Pong-Wong R., Gittus G., Russel A.J.F. & Rhind S.M. (2002) Detection of a major gene for litter size in Thoka Cheviot sheep using Bayesian segregation analyses. *Animal Science* **75**, 339-47.

Table 1. Data structure

No. of wool records		565
Pedigree Structure:	Total no. animals	1474
	Sires	49
	Sires of sires	22
	Dams of sires	23
	Dams	351
	Sires of dams	39
	Dams of dams	155

Table 2. Description of wool score

Description	Wool score	No. of Animals
Full fleece	0	30
Fleece loss opens the neck and tail area	1	64
Fleece loss widen around the neck and/or tail area	2	28
The remaining fleece resembles a large waist coat	3	34
The remaining fleece resembles waist coat	4	21
The remaining fleece resembles small waist coat	5	13
A band of fleece remains on the back	6	32
Fleece tufts remain on gigot	7	30
Fleece tufts remain on shoulder or/and tail	8	39
Clean	9	274

Table 3. Heritability estimates of wool shedding as a continuous trait (WSCORE) fitting an animal model plus permanent environmental effects due to the dam (pe) and litter, wool score as binary trait ([¥]BIN1 and [‡]BIN2) and wool score as categories ([†]WSCAT)

	Model						
	WSCORE				BIN1	BIN2	WSCAT
Parameter	Animal	Animal,pe	Animal,litter	Animal,pe,litter	Animal	Animal	Animal
σ^2_{ani}	7.56	8.02	7.64	8.02	0.08	0.13	0.37
σ^2_{pe}		0.48		0.48			
σ^2_{lit}			0.32	0			
σ^2_{resid}	1.84	1.2	1.51	1.2	0.05	0.04	0.12
σ^2_{phen}	9.42	9.7	9.46	9.7	0.13	0.17	0.50
s.e.	0.72	0.78	0.73	0.78	0.01	0.01	0.04
h^2_{direct}	0.80	0.83	0.81	0.83	0.65	0.79	0.75
s.e.	0.06	0.06	0.06	0.06	0.08	0.08	0.07
pe^2		0.05		0.05			
s.e.		0.03		0.03			
litter			0.03	0			
s.e.			0.06	0			
logL	-833.2	-831.98	-831.98	-831.98			

Where WSCORE (0-9 wool scores), [¥]BIN1 (wool scores 0,1 *vs.* the rest), [‡]BIN2 (wool scores 0-2 *vs.* the rest) and [†]WSCAT (wool scores 0,1 represent non-shedders, scores 2-7 medium shedders and scores 8,9 complete shedders)

Table 4. Summary of complex segregation analyses estimates when fitting polygenic effects and a major gene assuming additive and dominance effects

Model	Estimates	s.e.
1) Polygenic		
Variances/heritability estimates		
$\sigma^2_{residual}$	1.90	0.50
$\sigma^2_{additive}$	7.63	1.02
h^2	0.80	0.06
2) Major gene (Additive)		
Additive effect size	5.88	0.17
Variances/heritability estimates		
$\sigma^2_{residual}$	1.37	0.18
$\sigma^2_{additive}$	0.57	0.21
h^2	0.29	0.10
3) Major gene (Dominance)		
Additive effect size	3.11	0.12
Dominance effect size	2.90	0.14
Variances/heritability estimates		
$\sigma^2_{residual}$	0.99	0.12
$\sigma^2_{additive}$	0.55	0.16
h^2	0.35	0.09

Table 5. Number of genotyped animals classified by phenotype and inferred genotype from the segregation analysis

	Wool Shedding Score			
Inferred genotype	**0**	**1**	**9**	Total
FF	7	12	0	19
FS	0	1	20	21
SS	0	0	8	8
Total:	7	13	28	48

Where 'F' represents the putative 'Fixed' (i.e. non-shedding) allele and 'S' represents the putative 'Shedding' allele. $p = f(F) = 0.615$, $q = f(S) = 0.385$

Table 6. P-values and allele frequencies (q) for SNPs fitted in models where wool shedding was treated as continuous (WSCORE), binary ([¥]BIN1, [†]BIN2) and categories ([‡]WSCAT) for 200 animals

SNP	Myo SNP[*]	SNP1	SNP2	SNP3	SNP4
WSCORE	0.371	0.032	0.004	0.042	0.003
BIN1	0.100	0.002	<.001	0.046	0.031
BIN2	0.579	0.082	<.001	0.048	0.009
WSCAT	0.112	0.013	0.005	0.036	<.001
allele freq (q)	0.099	0.323	0.314	0.046	0.259

Where WSCORE (0-9 wool scores), [¥]BIN1 (wool scores 0,1 *vs.* the rest), [†]BIN2 (wool scores 0-2 *vs.* the rest) and [‡]WSCAT (wool scores 0,1 represent non-shedders, scores 2-7 medium shedders and scores 8,9 complete shedders). Myo SNP[*] is the myostatin locus (c.*1232G > A) Clop A *et al*. 2006 *Nature Genetics* **38**, 813-8.

Table 7. Summary of SNP association analysis results giving genotype means, additive and dominance (dom) effects and their standard errors (±s.e.) for all 200 genotyped animals, using full mixed model analyses

	WSCORE				BIN1				BIN2				WSCAT			
Genotype	SNP1	SNP2	SNP3	SNP4	SNP1	SNP2	SNP3	SNP4	SNP1	SNP2	SNP3	SNP4	SNP1	SNP2	SNP3	SNP4
0	7.56	7.92	7.84	8.33	0.016	0.032	0.047	0.011	0.065	0.040	0.059	0.013	2.70	2.71	2.69	2.82
(±s.e.)	0.69	0.62	0.61	0.63	0.076	0.068	0.069	0.071	0.092	0.082	0.082	0.084	0.16	0.14	0.14	0.14
1	7.30	7.58	6.32	7.00	0.105	0.058	0.215	0.131	0.128	0.083	0.257	0.147	2.57	2.65	2.33	2.47
(±s.e.)	0.74	0.67	0.92	0.65	0.082	0.074	0.104	0.075	0.099	0.089	0.123	0.088	0.17	0.15	0.21	0.15
2	5.72	5.44	*	5.98	0.277	0.392	*	0.186	0.272	0.424	*	0.356	2.22	2.16	*	2.28
(±s.e.)	0.94	0.90	*	1.08	0.100	0.100	*	0.126	0.123	0.120	*	0.148	0.21	0.21	*	0.25
Additive	0.921	1.238	1.527	1.176	0.131	0.180	0.168	0.087	0.103	0.192	0.197	0.171	0.237	0.277	0.358	0.269
(±s.e.)	0.355	0.375	0.740	0.487	0.037	0.042	0.083	0.056	0.046	0.050	0.099	0.066	0.079	0.086	0.168	0.110
Dom	0.655	0.899	*	0.155	0.042	0.154	*	0.032	0.040	0.149	*	0.037	0.105	0.213	*	0.074
(±s.e.)	0.448	0.424	*	0.539	0.051	0.049	*	0.064	0.062	0.059	*	0.075	0.102	0.098	*	0.123
h^2_{qtl}	0.06	0.12	0.02	0.06	0.09	0.24	0.02	0.04	0.04	0.19	0.05	0.06	0.07	0.12	0.03	0.08
VA_{prop}	0.08	0.15	0.03	0.09	0.27	0.72	0.07	0.12	0.12	0.51	0.10	0.17	0.11	0.19	0.03	0.12

Where WSCORE (0-9 wool scores), BIN1 (wool scores 0,1 *vs.* the rest), BIN2 (wool scores 0-2 *vs.* the rest) and WSCAT (wool scores 0,1 represent non-shedders, scores 2-7 medium shedders and scores 8,9 complete shedders). Genotype categories 0 and 2 represent homozygotes and 1 heterozygotes

h^2_{qtl} is given by $[2pq(a + d(q-p))^2/VP]$ where p and q are allele frequencies with a and d representing additive and dominance effects, respectively, and VP is the total phenotypic variance.

VA_{prop} is given by $[2pq(a + d(q-p))^2/VA]$ where VA is the total additive genetic variance. VP and VA are estimated from a model ignoring SNP effects.

* Not estimable as there were no animals homozygous for the minor allele

Figure 1. Posterior distributions of putative genotypes from the complex segregation analyses and histogram of residuals for wool shedding, after correcting for fixed effects

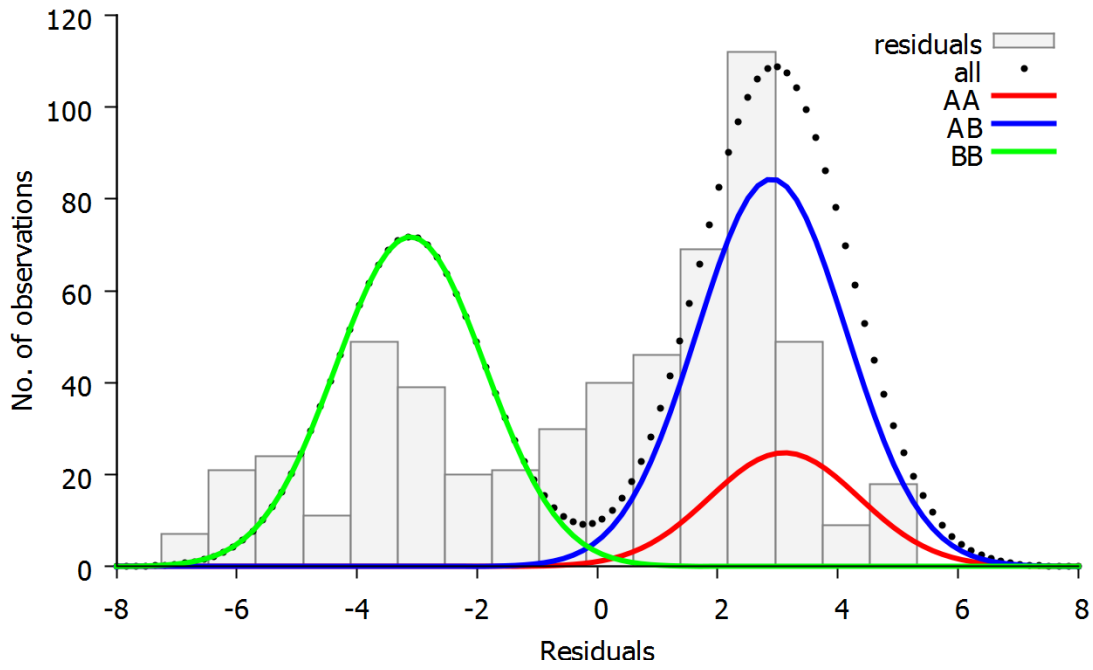

Figure 2. Manhattan plot displaying the GWAS results (p-values corrected for the genomic inflation factor λ) for wool score

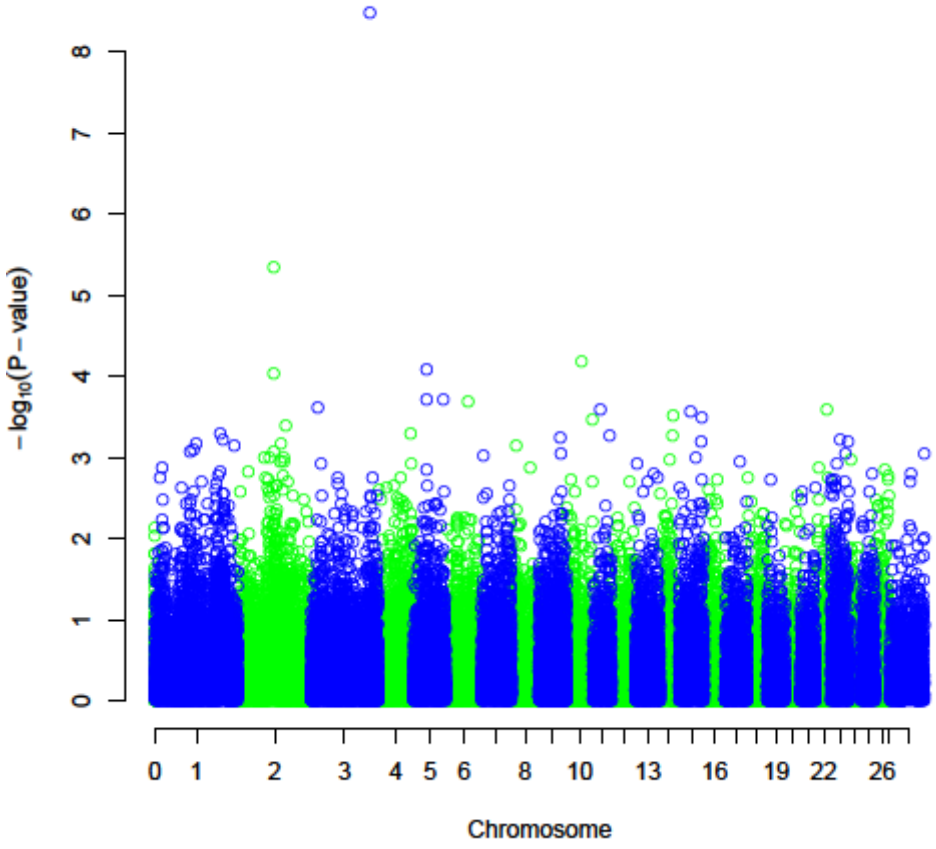

Figure 3: Linkage disequilibrium values (r^2 values, expressed as %) between SNPs in the interval between SNP1 (represented by number 1) and SNP2 (represented as number 13) on OAR 2