Analysis of single nucleotide polymorphisms variation associated with important economic and computed tomography measured traits in Texel sheep

Citation for published version:

Digital Object Identifier (DOI):
10.1017/S1751731117002488

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Animal

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Analysis of single nucleotide polymorphisms variation associated with important economic and computed tomography measured traits in Texel sheep

D. Garza Hernandez 1,a, S. Mucha1,4, G. Banos1,2, K. Kaseja1, K. Moore1, N. Lambe1, J. Yates3, L. Bunger1

1Animal and Veterinary Sciences, Scotland’s Rural College, Easter Bush, Midlothian EH25 9RG, Scotland, United Kingdom
2Roslin Institute, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, Scotland, United Kingdom
3British Texel Sheep Society, National Agricultural Centre, Stoneleigh Park, Kenilworth, Warwickshire, CV8 2LG
4Poznan University of Life Sciences, Department of Genetics and Animal Breeding, Wolynska 33, 60-637 Poznan, Poland
5Present address: Universidad Autónoma de Nuevo León (UANL), Pedro de Alba S/N, Ciudad Universitaria, San Nicolás de los Garza, N.L. México 66451.

Corresponding author: Lutz Bunger Email: Lutz.Bunger@sruc.ac.uk

Analysis of SNP variation in Texel Sheep
Abstract

Sheep are an important part of the global agricultural economy. Growth and meat production traits are significant economic traits in sheep. The Texel breed is the most popular terminal sire breed in the UK, mainly selected for muscle growth and lean carcasses. This is a study based on a genome-wide association approach that investigates the links between some economically important traits, including Computed Tomography (CT) measurements, and molecular polymorphisms in UK Texel sheep. Our main aim was to identify Single Nucleotide Polymorphisms (SNP) associated with growth, carcass, health and welfare traits of the Texel sheep breed. This study used data from 384 Texel rams. Data comprised 10 traits, including 2 CT measured traits. The phenotypic data were placed in four categories: growth traits, carcass traits, health traits and welfare traits. De-regressed estimated breeding values (EBV) for these traits together with sire genotypes derived with the Ovine 50K SNP array of Illumina were jointly analysed in a genome wide association analysis. Eight novel chromosome-wise significant associations were found for carcass, growth, health and welfare traits. Three significant markers were intronic variants and the remainder intergenic variants. This study is a first step to search for genomic regions controlling CT based productivity traits related to body and carcass composition in a terminal sire sheep breed using a 50K SNP genome-wide array. Results are important for the further development of strategies to identify causal variants associated with CT measures and other commercial traits in sheep. Independent studies are needed to confirm these results and identify candidate genes for the studied traits.

Keywords: Sheep, Texel, CT, Associated, GWAS.
Implications

Sheep are an important part of the global agricultural economy. To the best of our knowledge GWAS for CT based productivity traits, for a UK terminal sire breed, has not been widely researched. The main aim of this work was to exploit improved genotypic tools, specifically the Illumina OvineSNP50 chip, allowing a simultaneous genotyping for up to 54,241 SNPs to identify those SNPs associated with growth, carcass composition, health and welfare traits of Texel sheep using de-regressed estimated breeding values of rams.
Introduction

Sheep are an important part of the global agricultural economy. They are particularly well adapted to convert short herbage to meat, milk and wool and they are very important to meet global needs for food security for an increasing population around the world (Hopkins and Lobley, 2009).

Currently the Texel breed is the most popular terminal sire breed in the UK accounting for 30% of all purebred rams used for crosses to maternal sheep breeds (Pollott, 2014) and is mainly selected for muscle growth and lean carcasses (Hopkins and Lobley, 2009).

There are only a few methods to predict body composition in live sheep. Over the last few decades mainly ultrasound technologies had been used on farm animals for evaluation of carcass composition (Silva, 2016). However, computed tomography (CT), a non-invasive imaging technology, can accurately measure carcass traits in vivo such as muscle and fat (Bünger et al., 2011), muscularity (Jones et al., 2002) and tissue weights (Macfarlane et al., 2006). Additionally, it has been evidenced the potential of CT scanning to improve eating quality and tissue distribution of sheep meats (Macfarlane et al., 2009). As CT scanning is however more expensive than ultrasound, a two-step-procedure is recommended. Only the best 15-20% of selection candidate ram lambs measured by ultrasound would be subsequently CT scanned (Lewis, 2004).

Sheep genetics studies

Breeders focus sheep selection on production traits, including carcass composition and growth traits but also integrate other traits such as meat quality, disease resistance, lambing ease and survival (Bünger et al., 2011). According to the animal QTL database
there are currently (06/2017) 1,515 sheep QTLs curated in the animal QTL database (Hu et al., 2013) representing 222 different sheep traits, reported in 126 publications. However, one of the main limitations of unscrambling the genetic architecture underlying production traits in sheep has been the relative lack of information on the sheep genome in addition to the lack of accurate phenotypic data obtained (Zhang et al., 2013).

Currently, knowledge of the major genes or QTL associated with carcass composition and growth traits in sheep is limited (Zhang et al., 2013). Walling et al. (2004) pioneered the first accounts of QTL studies for growth and carcass conformation traits in domesticated sheep covering several genomic regions, which led to characterization of the Texel muscling QTL (TM-QTL).

With the advent of genome-wide panels of single nucleotide polymorphisms (SNPs) and using the approach of a genome-wide association study (GWAS), it has become possible to identify and localize QTLs for complex traits in many livestock species (Georges, 2007). However, to date, only a small number of GWASs in sheep have been conducted because of either limited information available for the sheep genome and funding. These studies have been mainly focused on sheep growth, ultrasound-measured meat traits and body composition traits (Cavanagh et al., 2010, Zhang et al., 2013, Bolormaa et al., 2016, Matika et al., 2016)

Moreover, GWAS with high accuracy CT measured body composition traits are still very rare in the literature. Donaldson et al. (2014) used spine characteristics measured from X-ray computed tomography (CT) scans in order to investigate if there were any subsequent associations between TM-QTL inheritance and underlying spine characteristics (Donaldson et al., 2014). Also, Cavanagh et al. (2010) performed a QTL
mapping study in sheep based on in vivo obtained CT data providing predictions for 13 traits describing major fat depots, lean muscle, bone, body proportions and body weight; they identified 3 highly significant, 15 significant, and 11 suggestive QTL on eleven chromosomes. But, no tissue-specific QTL were identified. Furthermore, Matika et al. (2016) conducted recently a genome-wide association study (GWAS) for carcass composition phenotypes, including bone, fat and muscle components, which were captured using CT. The GWAS analyses revealed multiple SNPs and quantitative trait loci (QTL) that were associated with effects on carcass composition traits and were significant at the genome-wide level.

In this study we performed a genome wide association study to identify those SNPs associated with growth, carcass composition, health and welfare traits, including 2 CT measured phenotypes, of Texel sheep using de-regressed EBVs of rams.
Material and Methods

Traits and phenotypes

A total of 384 Texel rams descended from 252 sires and 351 dams were analysed for 10 productivity traits including 2 CT measured traits. These rams represent a group of well-monitored animals as only a proportion (10-20%) of the initial selection candidates will be put forward to CT scanning based on ultrasound results.

The phenotypic data were provided by the Signet Sheep breeder Service and comprised EBVs progeny test derived for: birth weight (BW), eight week body weight (EWW) and scan weight (SW), which is the live weight at US scanning at about 21 weeks of age. These were considered as growth traits. As carcass traits were used US measured fat depth (FD) and muscle depth (MD) which are obtained by US-scanning at the at the third lumbar vertebra at 90 degrees to the backbone. The CT measured carcass traits: fat weight (FW), CT lean weight (LW) and the muscularity score (MU), a measure of carcass shape (Bünger et al., 2011), were also included. Details on the CT measured traits have been reported earlier (Bünger et al., 2011). Faecal egg count (FEC) as a measure of worm egg count in sample from lambs at 21 weeks of age, and, Lambing ease (LE) as a direct assessment of the ease with which ram progeny will be born.

GWAS accuracy can also be affected by systematic environmental effects. De-regressed EBVs are an alternative to raw phenotypic measurements, because they represent aggregate phenotypes adjusted for systematic environmental effect. The phenotypic data used therefore consisted of de-regressed estimated breeding values (EBVs) of standard commercial traits.
Statistical model for de-regressed breeding values

The official Texel EBVs were used, those breeding values were derived from the following model:

\[ y = Xb + Za + e, \]

where \( y \) is the vector of phenotypic observations for one of the analysed traits, \( b \) is the vector of fixed effects with design matrix \( X \) (relating observations to fixed effects), which varied depending on the trait, \( a \) is the vector of random animal effects, with design matrix \( Z \) (relating observations to random effects) and \( e \) is the vector of random residuals. The list of effects is summarized in Supplementary Table S1.

Random effects are assumed to be normally distributed with zero means and the following covariance structure:

\[ \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_e^2 \end{bmatrix} \]

where \( A \) is the pedigree-based relationship matrix, \( \sigma_a^2 \) is the genetic variance, and \( \sigma_e^2 \) is the residual variance.

The software package MIX99 was used for de-regression (Lidauer M, 2011), using a full animal pedigree with effective offspring contributions (EOC) as weighting factors. The de-regression procedure was based on the method published by Jairath et al. (1998), involving solving the mixed model equations with a full pedigree to obtain the right-hand side or de-regressed EBVs. Thus DRPs represent daughters averages adjusted for fixed effects and contributions from parents and relatives in the pedigree (Jairath et al., 1998).
EOC were calculated as:

\[ EOC_i = \frac{rel_i \cdot kdau}{1 - rel_i} \]

\[ kdau = \frac{4 - h^2}{h^2} \]

where \( rel_i \) is the reliability of EBV for animal \( i \) and \( h^2 \) is the heritability of one of the analysed traits.

The use of effective daughter or progeny contribution as a weighting factor is used to avoid biases in sire variances (Fikse and Banos, 2001). The EOC provides a measure of the precision of the daughter information used to compute the de-regressed EBV of the animal as the estimates of reliability used in the computation accounts for factors such as contemporary group (CG) structure for the ram’s daughters, the correlation between observations on the same daughter and the reliability of the performance of the daughters’ dams.

A Shapiro and Wilk’s W-statistic test, conducted using the R-package (R Core Team, 2013) was used to test data distribution for normality (Royston, 1995). Traits not normally distributed were rank transformed to a normal distribution for their use in subsequent analysis. This rank-transformation method has been reported to give a consistent performance in identifying causal polymorphisms with a slight increase in false positive rate (Goh et al., 2009). This method was used because according to Goh et al. (2009) for small sample size or genetic effects, the improvement in sensitivity for rank transformation outweighs the slight increase in false positive rate.

\section*{Genotyping}
All rams were genotyped with the ovine 50k SNP chip (54,241 SNPs across the genome with an average of 20.4 SNPs per Mb) by AgResearch. The order of the SNPs was based on the Ovis_aries_4.0 assembly released by the International Sheep Genomics Consortium (Jiang et al., 2014).

Quality control (QC) was performed with the GenABEL R package by considering genotypes of all rams (Aulchenko et al., 2007). The QC excluded 1,564 SNPs with call rates lower than 95%, 3,891 SNPs with minor allele frequencies less than 1%, 98 X-linked SNPs that were likely to be autosomal (cut off odds > 1000) and 777 SNPs not in Hardy-Weinberg equilibrium (p-value <1x10e-5). The call rate per individual was always higher than 90% so no animal was removed from the analysis. After applying these quality control criteria 48,433 SNPs (89%) located on 26 autosomes and on the X chromosome were used in the subsequent analyses.

**Statistical Model for GWAS**

A Multidimensional Scaling Analysis (MDS) was performed first to evaluate the genetic structure of the population. For each trait, SNP effects were then tested, by a single marker regression, with a mixed animal model including the genomic kinship matrix (identity by state) between the genotyped animals, adjusted for allele frequencies. Kinship was computed based on the method proposed by Astle and Balding (2009), using GenABEL, to control for population structure or polygenic effect (Astle and Balding, 2009). The following model was used:

\[ y = X\beta + Zu + e \]

where \( y \) is the vector of de-regressed EBV of rams, \( \beta \) is a vector of coefficients for the SNP effects, \( u \) is the vector of random animal effects, \( e \) is the vector of random residual effects, and \( X \) and \( Z \) are incidence matrices relating observations to fixed and random
animal effects, respectively. Random animal effect followed a normal distribution MVN(0, $G\sigma_u^2$) where $G$ is the genomic kinship matrix and $\sigma_u^2$ is the polygenic variance; and the random residual effects of the model was assumed to be MVN(0, $I\sigma_e^2$), where $\sigma_e^2$ is the residual variance and $I$ is an identity matrix. Each trait was analysed separately and all analyses were run with GenABEL.

This procedure consisted of two steps: firstly it estimated the polygenic and residual variance, not accounting for marker effects and fitting the genomic kinship matrix in the model. Secondly, these estimated variance components were used to estimate all the marker effects (fitting in the model the genotypes and the previously estimated residuals). The $j$-th marker was fitted in the single-marker-based linear mixed model without removing the $j$-th marker from the $G$ matrix. Evidence has shown analytically that, if variance components are kept constant, the estimation of the regression of phenotype on $m$ markers is invariant with respect to whether or not the marker(s) tested for association is(are) included when constructing the $G$ matrix (Gianola et al., 2016).

Significance of the results was tested at genome-wise and chromosome-wise levels, including a strict Bonferroni correction for multiple-testing, corresponding to $1 \times 10^{-6}$ and $3.5 \times 10^{-5}$, respectively.

In order to address possible population stratification problems, the inflation in the test statistic was monitored with factor lambda, which does not depend on allele frequencies (Aulchenko et al., 2007). The allele effects estimated by GenABEEL refer to the least frequent allele in the population and are expressed in trait phenotypic standard deviation (STD) units. Genes located on or around the identified SNPs were examined using the ENSEMBL database and the Ovis_aries_3.1 and 4.0 assembly released by the International Sheep Genomics Consortium (Jiang et al., 2014). And
finally JBrowse was used to identify previously associated QTLs in the tagged regions (Skinner et al., 2009).
Results

Descriptive statistics
For the 10 analysed traits (de-regressed EBVs) the means and standard deviations are shown in Table 1. The normal distributions of the 10 traits were tested with the Shapiro-Wilk’s test (Table 1). For EWW, FD, FW, FEC and LE traits the null hypothesis of following a normal distribution was rejected according to a p value ≤ 0.1, which has been previously suggested as an acceptable threshold for this type of analysis (Royston, 1995). These records were rank-transformed to a normal distribution for their use in the subsequent analyses.

Genome Wide Association Analysis
A multidimensional scaling analysis using the GenABEL package showed that no genetic stratification was present in this population. Also, the average inflation factor (λ) was $1.008 \pm 0.007$, with a maximum value of 1.021 for FEC and a minimum of 1 for FD, FW and MU. Therefore, the population structure is not expected to affect the results of GWAS in the present study.

No genome-wise significant associations were found between any SNP and trait. However, 8 chromosome-wise significant SNPs were found for EWW, FD, MD, LW, FEC, and LE (Figure 1). These SNPs were located on chromosomes 3, 4, 6, 11, 16 and 17, respectively (Table 2). None of the associated SNPs found had been previously associated with any trait in sheep.

The proportion of total variance explained by each SNP was obtained by first scanning using the score test and then reevaluating best hits, individually, using Maximum
Likelihood with significant SNP allelic effect fitted as covariate. The variance explained for chromosome wise significant SNP associated with EWW, FD, LW, MD and FEC were 0.029, 0.061, 0.062, 0.060 and 0.051, respectively. And for LE, each significant marker explained a variance of 0.006, 0.038 and 0.013.
Discussion

Until very recently, limited information on the sheep genome and lack of phenotypic data for many important traits have resulted in only a few studies on SNPs associated with production and welfare traits in sheep (Zhang et al., 2013). It has been suggested that the use of more precise phenotypes derived from CT measures will lead to more accurate phenotypes for genetic analyses (Cavanagh et al., 2010).

To date, only a small number of GWAS in sheep have been conducted, those have been mainly focused on sheep growth, ultrasound-measured meat traits and body composition traits (Cavanagh et al., 2010, Zhang et al., 2013, Bolormaa et al., 2016, Matika et al., 2016). Moreover, genetic analyses with high accuracy CT-measured body composition traits are still very rare in the literature (Walling et al., 2004, Donaldson et al., 2014, Bolormaa et al., 2016, Matika et al., 2016).

The main aim of the present study was to identify SNPs associated with traits currently in the selection index for a UK Terminal sire breed (Texel Sheep), including CT based productivity traits. In the UK, CT scanning has been used in sheep breeding programs since 2000. However, as CT scanning is more expensive than ultrasound, a two-step-procedure is recommended. Only the best 15-20% of selection candidate ram lambs measured by ultrasound are usually subsequently CT scanned (Lewis, 2004, Bünger et al., 2011).

A total of 384 Texel rams were analysed for 10 productivity traits including 2 CT measured traits. It should be noted that the dataset used in the present study was limited in its size, largely due to the restricted availability of CT-measured rams, due to CT costs. However, because this study analysed a small group of preselected animals...
we acknowledged that the power to detect genome wide significant associations was diminished.

*Genome Wide Association Analysis*

In the current study no genome-wide significant association for any of the analysed traits was found. However, 8 chromosome-wise significant SNPs were found for: EWW, FD, MD, LW, FEC and LE. These SNPs were located on chromosomes 3, 4, 6, 11, 16 and 17, and were found to be either intronic or intergenic variants. None of the significant SNPs had been previously associated with any trait in sheep. However, chromosomes 11 and 16 have been previously tagged by SNPs associated with muscle, body and carcass weight (Cavanagh *et al.*, 2010).

We identified as candidate genes, those which were either directly tagged by a significant SNP (intronic variant) or those located within genomic regions of 30 kb up and downstream of an associated marker (Bolormaa *et al.*, 2016). However, due to the current relatively poor status of the ovine genome annotation, little information regarding the function of the tagged genes was obtained.

Regions tagged for EWW and LE have not been previously associated with any significant growth or welfare traits. However, two identified markers for LE, on chromosomes 6 and 17 (OAR6_108683365.1 and OAR17_11963200.1), belong to suggestive QTLs previously associated with parasite resistance (Beh *et al.*, 2002, Marshall *et al.*, 2009). Former studies have reported a low to moderate genetic correlation between lambing ease and birth weight (Brown, 2007), while a moderate genetic correlation between birth weight and parasite resistance has been suggested (Verbeek *et al.*, 2011). However, more information would be needed to estimate the genetic correlation between parasite resistance and welfare traits such as LE.
The region tagged by OAR16_20147789.1, significantly associated with FD, is an intronic variant of the NDUFAF2 gene, which encodes a NADH dehydrogenase (ubiquinone) complex I, assembly factor 2, a molecular chaperone for mitochondrial complex I assembly. OAR16_20147789.1 is located in a QTL region, which has been previously associated with final body weight, percent lean and subcutaneous fat area (Cavanagh et al., 2010).

SNP s26074.1 was found to be significantly associated with LW. This SNP, is an intergenic variant, which is located in a QTL region formerly associated with body and carcass weight (Cavanagh et al., 2010).

The region identified by SNP OAR11_12972551.1, was significantly associated with MD. This SNP is an intronic variant of the ACACA gene. ACACA encodes an acetyl-CoA carboxylase alpha, which is considered as a key enzyme of fatty acid synthesis in the mammary gland by catalysing the first step of fatty acid synthesis in mammalian cytosol. This gene has been described as a candidate gene for fat content in sheep, due to an observed significant association with variation in milk fat content, and change of fat composition in several sheep breeds (Bolormaa et al., 2016). Moreover, OAR11_12972551.1 is located in QTL regions associated with body weight (Raadsma et al., 2009), fat synthesis (Bolormaa et al., 2016), internal fat amount and hot carcass weight (Cavanagh et al., 2010).

Thus, results of significant associations with carcass traits provide evidence of a possible effect on FD, LW and MD by QTLs previously reported by Raadsma et al. (2009), Cavanagh et al. (2010) and Bolormaa et al. (2016).

Finally, SNP s30868.1 associated with FEC, is an intronic variant of the ZNF227 gene, which encodes a zinc finger protein 227, probably involved in transcriptional regulation.
This gene is a paralogue of the ZNF229 gene, which has been previously associated with tuberculosis susceptibility in African human populations (Thye et al., 2010). Also, s30868.1 tags a QTL region formerly reported to be associated with Immunoglobulin A level, an antibody that plays a crucial role in the immune function (Atlija et al., 2016). This suggests that there might be a worm resistance QTL on chromosome 4.

A large number of QTLs have been identified for traits related to parasite resistance in sheep (Beh et al., 2002, Marshall et al., 2009, Atlija et al., 2016) suggesting that those traits are not determined by individual genes acting alone but rather by complex multigene interactions. Thus, further identification of SNPs in strong LD with the casual variants, could contribute to the implementation of these results in breeding schemes for the Texel breed population.

The proportion of total variance explained by the significant SNPs was low, which is in agreement with Hayes and Goddard (2010), who explained that a small number of markers with validated associations would explain a small portion of the genetic variance in complex traits (Hayes and Goddard, 2010). This suggests that if alleles of large effect were present in our data, those would be in such a low frequency that they individually could only explain a small proportion of the variance.

Further improvement in sheep GWAS could be achieved by increasing the sample size and using the new ovine 700K HD chip, which has a much denser distribution of SNPs across the genome and thus should have higher LD with the potential QTLs controlling the traits of interest.

The present study found 8 chromosome-wise significant SNPs for 6 traits among them a CT measured trait (LW). Tagged regions on chromosomes 4, for worm resistance (FEC), 11 and 16, for carcass traits (MD, LW and FD), are consistent with other
studies, where QTL regions have been found for Immunoglobulin A level and meat and carcass traits, respectively. Whereas regions tagged on chromosomes 3, 6 and 17 for LE and EWW can be considered novel. Among the tagged genes ZNF227, ACACA and NDUFAF2 were found. Hence, these genes could be considered as candidate genes for future research to further dissect the genomic architecture of the traits.

Conclusions

This study is one of very few studies using CT-derived carcass traits and other productivity traits already integrated in the selection index for terminal sire sheep breeds. It revealed some significant associations between genomic markers and important traits in sheep production. Further fine mapping the regions around these markers could lead to the identification of causative genes and better molecular predictors of CT based carcass composition, which might help to decrease phenotyping costs in the longer term. Results may also be integrated and inform genomic selection approaches and future SNP chip designs. The result may also guide similar studies in the other important Terminal Sire Breeds in the UK and beyond.

Acknowledgements

The research leading to this publication was supported by funds of the Scottish government (RERAD). DGH was supported by The Mexican National Council for Science and Technology. The opinions expressed and arguments employed in this paper are the sole responsibility of the authors and do not necessarily reflect those of RERAD.
This work acknowledges the support of AHDB Beef and Lamb, HCC and QMS in funding the use of the CT scanner and their annual contribution towards the provision of performance recording services delivered by Signet Breeding Services.

Many thanks to the British Texel Sheep society for their long-term support of the CT-scanning activities, for providing the tissue samples for subsequent DNA extraction and permitting the use of the genotypic information. Thanks go also to John McEwan and AgResearch in New Zealand for the genotyping.

References


### Table 1 Descriptive statistics for the de-regressed EBVs of the analysed traits.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Unit</th>
<th>Acronym</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth Traits</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth Weight</td>
<td>kg</td>
<td>BW</td>
<td>0.48</td>
<td>0.81</td>
<td>-2.19</td>
<td>2.89</td>
<td>0.88</td>
</tr>
<tr>
<td>Eight Week Weight</td>
<td>kg</td>
<td>EWW</td>
<td>3.24</td>
<td>11.30</td>
<td>-27.01</td>
<td>43.26</td>
<td><strong>0.10</strong></td>
</tr>
<tr>
<td>Scan Weight</td>
<td>kg</td>
<td>SW</td>
<td>7.17</td>
<td>7.60</td>
<td>-14.69</td>
<td>35.22</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Carcass Traits</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat Depth</td>
<td>mm</td>
<td>FD</td>
<td>-0.08</td>
<td>1.74</td>
<td>-6.1</td>
<td>5.78</td>
<td><strong>0.07</strong></td>
</tr>
<tr>
<td>Muscle Depth</td>
<td>mm</td>
<td>MD</td>
<td>1.73</td>
<td>3.42</td>
<td>-8.64</td>
<td>12.4</td>
<td>0.16</td>
</tr>
<tr>
<td>Fat Weight</td>
<td>kg</td>
<td>FW</td>
<td>0.79</td>
<td>1.75</td>
<td>-4.05</td>
<td>6.50</td>
<td><strong>0.10</strong></td>
</tr>
<tr>
<td>Lean Weight</td>
<td>kg</td>
<td>LW</td>
<td>2.17</td>
<td>2.01</td>
<td>-3.53</td>
<td>8.70</td>
<td>0.74</td>
</tr>
<tr>
<td>Muscularity</td>
<td>Ratio</td>
<td>MU</td>
<td>3.3</td>
<td>5.85</td>
<td>-12.94</td>
<td>18.14</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>Health Trait</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecal Egg Count</td>
<td>Log values</td>
<td>FEC</td>
<td>0.12</td>
<td>0.58</td>
<td>-2.72</td>
<td>4.77</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Welfare Trait</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambing Ease</td>
<td>Score units (1-6)</td>
<td>LE</td>
<td>0.05</td>
<td>11.98</td>
<td>-70.11</td>
<td>24.83</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

SD = Phenotypic standard deviation, 384 tested individuals, Significant p values, for Shapiro and Wilk’s W-statistic test, (p ≤ 0.1) in bold. Fat and Lean weights were measured by CT (as described by Bunger et al. (2011))
Table 2 Chromosome-wide significant SNPs associated with important economic traits and size of estimated effects.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position OAR v3.1 / OAR v4.0</th>
<th>Allele Effect</th>
<th>SD</th>
<th>P-value</th>
<th>Trait</th>
<th>Nearest Gene (Code)</th>
<th>Nearest Gene (Name)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAR17_22884911.1</td>
<td>17</td>
<td>20425356 / 20428283</td>
<td>-</td>
<td>0.388</td>
<td>0.09</td>
<td>EWW</td>
<td>PCDH18 [454.22]</td>
<td>Protocadherin 18</td>
</tr>
<tr>
<td>OAR16_20147789.1</td>
<td>16</td>
<td>18368560 / 18365229</td>
<td>-</td>
<td>0.439</td>
<td>0.10</td>
<td>FD</td>
<td>NDUFAF2</td>
<td>Ubiquinone oxidoreductase complex assembly factor 2</td>
</tr>
<tr>
<td>s26074.1</td>
<td>11</td>
<td>8271088 / 8261942</td>
<td>0.673</td>
<td>0.15</td>
<td>2.6E-05</td>
<td>LW</td>
<td>CUEDC1 [37.38]</td>
<td>CUE domain containing 1</td>
</tr>
<tr>
<td>OAR11_12972551.1</td>
<td>11</td>
<td>13110133 / 13079564</td>
<td>-</td>
<td>1.115</td>
<td>0.25</td>
<td>MD</td>
<td>ACACA</td>
<td>Acetyl-CoA carboxylase alpha</td>
</tr>
<tr>
<td>s30868.1</td>
<td>4</td>
<td>56089343 / 56074079</td>
<td>-</td>
<td>0.336</td>
<td>0.07</td>
<td>FEC</td>
<td>ZNF227</td>
<td>Zinc finger protein 227</td>
</tr>
<tr>
<td>OAR6_108683365.1</td>
<td>6</td>
<td>98702734 / 98597850</td>
<td>0.341</td>
<td>0.07</td>
<td>6.8E-06</td>
<td>LE</td>
<td>NKX6 [193.99]</td>
<td>NK6 homeobox 1</td>
</tr>
<tr>
<td>Chr (Chromosome); Allele effect (deviations from the mean); SD (standard deviation) of the allele effect; P-value for the significance of the association; Units for FEC and LE on the transformed scale; SNPs located within known ovine genes are highlighted in bold; the nearest genes were identified using the ENSEMBL Genome Browser; the number in brackets is the distance from SNP to the nearest gene.</td>
<td>Figure Captions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Figure 1: Manhattan plots for EWW, FD, LW, MD, FEC and LE traits, blue line refers to the genome-wise threshold and the red line to the chromosome-wise significance threshold.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>