Proceedings, 10th World Congress of Genetics Applied to Livestock Production

Genome-Wide Association Study on Body Weight Reveals Major Loci on OAR6 in Australian Merino Sheep

H.A. Al-Mamun*, S.A. Clark*, P. Kwan* and C. Gondro*
*University of New England, Armidale, Australia

ABSTRACT: Body weight (BW) is an important trait for meat production in sheep. In order to clarify genes and chromosomal regions that might be associated with body weight in sheep, a genome-wide association study using Illumina 50K Ovine SNP chip on a population of 1,743 Australian Merino sheep was carried out. A total of 39 SNPs were found to be associated with BW at Bonferroni-corrected genome-wide significance of 1%. One region on OAR6, between 40.3 Mb and 42.9 Mb, included 13 significant SNPs that were identified as being associated with the trait. The most significant SNP (OAR6_41936490.1) has 2.11 kg of allele substitution effect, while the second most significant SNP (s17946.1) has 2.13 kg of effect on body weight. These correspond to 24.33% and 24.57%, respectively, of the phenotypic standard deviations for BW in Australian Merino sheep.
Keywords: Genome-Wide Association Study; Body Weight, Sheep**.**

Introduction

In sheep, body weight (BW) is considered an important economic trait for meat production. GWAS are being broadly applied to discover candidate genes for many quantitative traits not only in cattle but also in many other species (Lango Allen et al. (2010); Yang et al. (2010); Gu et al. (2011); Schneider et al. (2012); Signer et al. (2012)). Several GWAS studying growth and meat production traits in cattle have been published in recent years (Maltecca et al. (2009); McClure et al. (2010); Cole et al. (2011)). For bovine carcass weight, studies have revealed major QTLs on chromosomes 6 and 14 (Mizoshita et al. (2005); Setoguchi et al. (2009); Nishimura et al. (2012)). Mizoshita et al. (2005) identified a 1.1 Mb QTL region on the bovine chromosome 14, while Setoguchi et al. (2009) discovered a 591 Mb QTL region on chromosome 6, both related to bovine carcass weight. Karim et al. (2011) mapped a QTL having a major effect on bovine stature to a 780-kb region on the bovine chromosome 14. Nishimura et al. (2012) identified three major QTLs, namely BTA6, BTA8, and BTA14 for carcass weight in Japanese black cattle. Together, these 3 QTLs explained approximately one-third of the genetic variance. The QTL region on BTA14 contains the genes PLAG1, CHCHD7, PENK, MOS, LYN and TGS1 which have been found previously having association with stature in both cattle and human (Gudbjartsson et al. (2008); Weedon et al. (2008); Pausch et al. (2011); Pryce et al. (2011)). The same QTL region on BTA14 was also found to affect birth weight and size in zebu cattle (*Bos primigenius indicus*) (Utsunomiya et al. (2013)). In brown Swiss cattle, 74 genome-wide significant SNPs have been found to be associated with one or more production traits, including fertility, confirmation, udder health and workability in BTA6, BTA11, BTA24 and BTA25 (Guo et al. (2012)). Lee et al. (2013) has identified a major QTL region for carcass weight on BTA14 in the Korean Cattle (Hanwoo) that explains at least 10% of the genetic variation of the carcass weight in these cattle.

Although many GWAS have identified important candidate genes in different species, relatively few QTL studies have been reported in sheep. At the time of writing, only 789 sheep QTLs have been curated in the Sheep QTLdb (Hu et al. (2013)) within which the number of QTLs for meat production is very small. Most of the reported QTL studies on sheep were based on sparse microsatellite markers that exhibited relatively wide QTL confidence intervals, making it very difficult to identify specific genes that affect the targeted quantitative traits. We have come across only a few complete GWAS that focused on growth and meat production traits in sheep (Riggio et al. (2013); Zhang et al. (2013)). Among them, GWAS and fine mapping of QTL for body weight were performed on the sheep chromosome 21 only (Jonas et al. (2010)).

In this paper, we report the results of a genome wide scan for SNPs that might be associated with body weight (BW) in sheep by using the data on 1,743 Australian Merino sheep. The objective of this study was to identify the set of significant SNPs associated with body weight, and to explore and predict the genomic regions around these SNPs. The list of shortlisted genomic regions can then be used as the basis for further fine mapping analyses, with the hope of eventually determining the causal mutations that are associated with body weight in sheep.

Materials and Methods

Phenotypic data. The data used in this study consisted of phenotypic records from 1,743 Merino sheep from the Australian sheep CRC information nucleus flock. There were 1,088 males and 693 females and the animals were sired by 111 sires. The maximum number of sheep per sire was 46 and the minimum number was 1. The trait used for this study was post weaning body weight (PWWT) which is an early life body weight measurement. The average age of the animals during the measurements of the phenotypes was 287.5 days with the minimum 148 days and the maximum 431 days.

Genotyping and quality control. All animals were genotyped using the Illumina 50k ovine SNP chip (Illumina Inc., San Diego, CA, USA), which includes 54 977 SNP. A number of quality control measures were

applied to all SNP as follows: SNP were removed if they had a call rate of less than 95%, a GC score of less than 0.6, a minor allele frequency of less than 0.01, a SNP heterozygosity greater than 3s.d. from the mean (mean heterozygosity, 0.374; s.d., 0.129). Data was also removed if a SNP was out of Hardy-Weinberg equilibrium (a P-value cut-off of $1 \cdot 10^{-15}$), had no genome location or was in greater than 0.99 LD with another SNP on the chip. After these quality control measures were applied, 48,640 SNP were used. Missing genotypes were imputed using fast PHASE (Scheet and Stephens, 2006). After executing quality control, the number of SNPs on each chromosome varied between 912 on OAR21 and 6,016 on OAR1, while the adjacent distance ranged from 46.52 kb on OAR9 to 60.72 kb on OAR21.

Statistical Analysis. A genome wide association analysis was performed using a linear mixed effects model (using ASREML (Gilmour et al., 2009)). The following fixed effects were fitted in the model: sex, birth type, rearing type, age of dam, contemporary group (birth year * birth month * site * management group) and age-at-trait recording. To account for family effects sire was also fitted as a random effect. The additive allele substitution effect of each SNP(j) was calculated by fitting the following mixed model:

$$y = Xb + Zg + \varepsilon \qquad (1)$$

where, y is the vector of the individual body weights, X is a design matrix relating the fixed effects (as described above), including each SNP(*j*) to each animal and b is a vector of fixed effects (including SNP effects). Z is a matrix allocating records to sire groups, g is a vector of sire effects for all animals and ε is the residual.

The percentage of the genetic variance explained by each significant SNP was calculated according to the formula:

$$\%V_{gi} = 100 \times 2p_i q_i a_i^2 / \sigma_g^2 \qquad (2)$$

where, p_i and q_i are the allele frequencies for the *i-th* SNP, a_i is the estimated additive effect of the *i-th* SNP on BW and σ_g^2 is the REML estimate of the polygenic variance for BW in sheep (obtained from ASReml). We used the GCTA software (Yang et al., 2011) to partition the genetic variance onto different chromosomes.

Results and Discussion

The GWAS identified one major chromosome region, ranging from 40.3 Mb to 42.9 Mb on chromosome 6, which is associated with body weight in Merino sheep (Figure 1). The quantile-quantile (Q-Q) plot with P<0.01 for body weight revealed a large deviation from the distribution under the null hypothesis, indicating that strong associations were obtained (Figure 2). Significant Bonferroni-corrected genome-wide associations ($P<1.15 \times 10^{-11}$) were detected for 13 SNPs for carcass weight on chromosome 6 (Figure 1). Among the five most significant SNPs, four are located on the OAR6 chromosome. The most significant SNP was OAR6_41936490.1 ($P = 2.36 \times 10^{-16}$) located in the 41.9 Mb region on OAR6 (Figure 1). The other 3 SNPs on this chromosome were s17946.1 ($P = 7.97 \times 10^{-14}$) in the 41.3 Mb region, OAR6_41877997.1 ($P = 2.49 \times 10^{-12}$) in the 41.8 Mb region, OAR6_41003295.1 ($P = 2.4 \times 10^{-11}$) in the 41 Mb region, respectively. One SNP among the five most significant SNPs is OAR3_128968872.1 ($P = 2.16 \times 11^{-16}$) in the 12.8 Mb region on OAR3 chromosome. We investigated LD (r^2 value) structure in the region from 40.3 Mb to 42.9 Mb on OAR6 which has 13 significant SNPs. We found three LD blocks with LD coefficients ranging from 0.1 to 0.99 within each LD blocks.

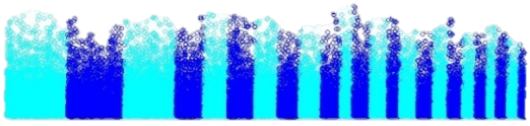

Figure 1. Manhattan plot of GWAS for body weight traits of 1,743 sheep

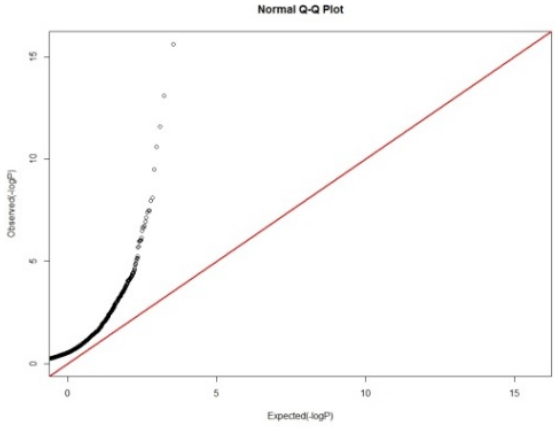

Figure 2. Q-Q plot for observed and expected -log(p-value)

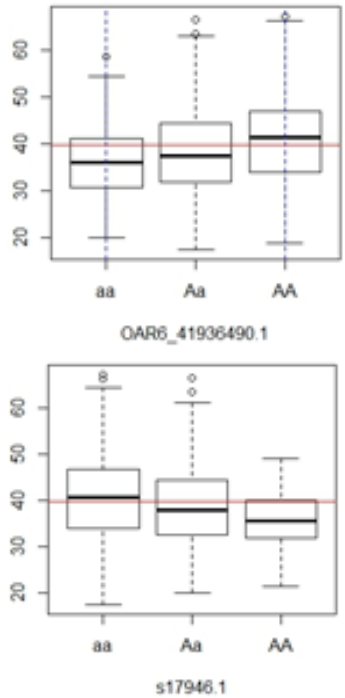

Figure 3: Effect of the two most significantly associated SNPs on body weight

The nearest gene to the most significant SNP (OAR6_41936490) is LOC101103153, which is similar to the ribosomal protein L10a gene in human. The genomic region between 40.3 Mb and 42.9 Mb on chromosome 6 encompasses the genes KCNIP4, TRNS-GGA, LOC101104829, TRNAW-CCA, GPR125, LOC101103153, and GBA3. The region from 36 - 42 Mb in OAR6 of Scottish Blackface lambs has been reported to contain several SNPs that are associated with body weight in different weeks. (Riggio et al. (2013)). Recently, the orthologous KCNIP4-GPR125 locus and its surrounding 8.6 Mb region (71.6 - 80.2 Mb on GGA4) in chicken was reported to be associated with body weight and average daily weight gain (ADG) from 6 weeks to 12 weeks (Gu et al. (2011)). In our study, we did not find any functional relationship between the orthologous genes (from 40.3 Mb to 42.9 Mb in OAR6) and body weight/height traits in either human or in cattle. We then investigated a broader region between 37.1 MB and 41.9 Mb (a 4.8 Mb interval) on chromosome 6. The region from 37.1 Mb to 37.5 Mb comprises genes LAP3, NCAPG and LCORL, which has been identified as a locus associated with height and weight related traits in many species. The NCAPG-LCORL locus corresponds to a height locus and associated with bovine stature and adult human height (Pryce et al. (2011)). In human, the LCORL gene encodes a transcription factor that appears to function in spermatogenesis. Polymorphisms in this gene are associated with measures of skeletal frame size and adult height. In cattle, a non-synonymous variant in the NCAPG gene has been found as a potential causative variant for body frame size: height, length and width at puberty. In horse, the NCAPG-LCORL locus is located on ECA 3 and a QTL located shortly upstream of the LCORL gene was reported as associated with the height at withers (Signer et al. (2012)).

In our analysis, we calculated the proportion of additive genetic variance explained by each SNP which has adjusted *p*-value <= 0.01, detected in single-marker regression analysis. The most significant SNP explained a large percentage (7.22%) of the total additive genetic variance. We also calculated the chromosome wise percentage of the explained genetic variance (estimated using GCTA software) of body weight trait for the 1,743 sheep and found that chromosome 6 explains the largest proportion of additive genetic variance.

The allele substitution effects for each of the significant SNPs were calculated by using linear regression analysis. The most significant SNP (*OAR6_41936490.1; P=2.36×10^{-16}*) has 2.11 kg of allele substitution effect, while the second most significant SNP (s17946.1*; P=7.97×10^{-14}*) has an effect of 2.13 kg for carcass weight. They correspond to 24.82% and 29.45%, respectively, of the phenotypic standard deviation for body weight in Australian Merino sheep (Figure 3).

Conclusion

In summary, our GWAS identified 39 SNPs with genome-wide significance and we found a major QTL region on OAR6 for body weight in sheep. Our results include both the previously reported QTL regions and some new QTL regions for body weight traits. These findings are anticipated to facilitate the discovery of causative variants for body weight in the future and could inform marker-assisted selection.

References

Cole, J. B., G. R. Wiggans, et al. (2011). BMC genomics 12: 408.

Gilmour,A. R., Gogel B. J., et al. (2009). ASReml User Guide Release 3.0.

Gu, X., C. Feng, et al. (2011). PloS one 6(7.

Gudbjartsson, D. F., G. B. Walters, et al. (2008). Nature genetics 40(5): 609-615.

Guo, J. Z., H. Jorjani, et al. (2012). BMC genetics

Hu, Z. L., C. A. Park, et al. (2013). Nucleic Acids Research 41(Database issue): D871-879.

Jonas, E., P. C. Thomson, et al. (2010). 9th WCGALP, Leipzig, Germany.

Karim, L., H. Takeda, et al. (2011). Nat Genet 43(5): 405-413.

Lango Allen, H., K. Estrada, et al. (2010). Nature 467(7317): 832-838.

Lee, S. H., et al. (2013). PloS one 8(10.

Maltecca, C., K. A. Weigel, et al. (2009). Animal genetics 40(1): 27-34.

McClure, M. C., N. S. Morsci, et al. (2010). Animal genetics 41(6): 597-607.

Mizoshita, K., A. Takano, et al. (2005). Mammalian genome 16(7): 532-537.

Nishimura, S., T. Watanabe, et al. (2012). BMC genetics 13: 40.

Pausch, H., K. Flisikowski, et al. (2011). Genetics 187(1): 289-297.

Pryce, J. E., B. Hayes, et al. (2011). Genetics 187(3): 981-984.

Purcell, S., B. Neale, et al. (2007). American journal of human genetics 81(3): 559-575.

Riggio, V., O. Matika, et al. (2013). Heredity 110(5): 420-429.

Scheet, P. and Stephens, M. (2006). American Journal of Human Genetics 78: 629-644.

Schneider, J. F., L. A. Rempel, et al. (2012). Journal of animal science 90(10): 3353-3367.

Setoguchi, K., M. Furuta, et al. (2009). BMC genetics 10: 43.

Signer-Hasler, H et al. (2012). PloS one 7(5.

Utsunomiya, Y. T., A. S. do Carmo, et al. (2013). BMC genetics 14: 52.

Weedon, M. N., H. Lango, et al. (2008). Nature genetics 40(5): 575-583.

Yang, J. A., B. Benyamin, et al. (2010). Nature genetics 42(7): 565-U131.

Yang J., Manolio T.A., et al. (2011). Nat Genet 43: 519–525.

Zhang, L., J. Liu, et al. (2013). PloS one 8(6).