*Proc. Assoc. Advmt. Breed. Genet. 21: 69-72*

# IDENTIFICATION OF LOCI ASSOCIATED WITH PARASITE RESISTANCE IN AUSTRALIAN SHEEP

**M. Al Kalaldeh, J.H.J. van der Werf and C. Gondro**

University of New England, School of Environmental and Rural Science, Armidale 2350, Australia

## SUMMARY

This study aimed to identify loci underlying variation in parasite resistance, as measured by worm egg count (WEC), in a large multi-breed sheep population using genome-wide association studies (GWAS) and regional heritability mapping (RHM) approaches. A total of 7153 animals with both genotype data and WEC phenotypes were included in this analysis. Strong evidence of association was observed on chromosome 2 by both approaches. However, RHM had a greater power to identify loci than GWAS analysis. RHM identified an additional region at the genome-wide significance level on chromosome 6. This region was also previously found to be associated with mastitis resistance and facial eczema susceptibility in sheep, indicating that some pleiotropic effects are possibly affecting a wide range of sheep diseases. Three other regions on chromosome 1, 3 and 24 reached the suggestive threshold. However, the regions accounted for a small proportion of genetic variance ($h_g^2 < 0.01$). It seems that parasite resistance is a complex disease with a large number of genes involved in the mechanism of resistance.

## INTRODUCTION

Gastrointestinal nematode infections are one of the most important health problems affecting sheep and other grazing ruminants in Australia and worldwide. Selection for parasite resistance has been suggested as a viable method for parasite control (Roeber *et al.*, 2013). Most breeding programs for parasite resistance are based on phenotypic indicators, particularly worm egg counts (WEC) in faeces, but trait measurement is unattractive, costly and time consuming. Therefore, it would be very useful to select directly for parasite resistance. To date, several quantitative trait loci (QTL) mapping studies have been conducted for parasite resistance in sheep (e.g. Dominik *et al.*, 2010 and Marshall *et al.*, 2009). However, little overall consensus has emerged from these studies. This may be due to the physiological complexity of parasite resistance, and the fact that these studies are very diverse, involving a variety of analytical approaches, experimental designs, parasite species and sheep breeds. Further, genome-wide association studies (GWAS) for complex diseases, such as parasite resistance, have generally failed to explain the majority of genetic variation influencing the trait (Kemper *et al*. 2011). The objective of this study was to identify loci underling variation in parasite resistance in a multi-breed sheep population.

## MATERIALS AND METHODS

**Animals.** Parasite resistance trait, as measured by WEC, was investigated in a multi-breed sheep population from the Sheep Cooperative Research Centre information nucleus flock (INF). A total of 7,539 animals with both genotype data and WEC phenotypes were included in this analysis. Various breeds were represented in the population (Table 1) but with a significant proportion of Merino sheep, and only this breed had a substantial proportion of purebred animals. The remaining breeds were mainly represented by their crosses with Merino (van der Werf *et al*. 2010).

**Genotypes.** Animals were genotyped using the 50k Ovine marker panel (Illumina Inc., SanDiego, CA, USA). SNPs were removed if they had a minor allele frequency (MAF) < 1%, an Illumina Gentrain score (GC) less than 0.6, a call rate less than 95%, or not in Hardy-Weinberg equilibrium. Furthermore, positions of SNPs were obtained from the latest sheep genome

*Ovis_aries_v3.1*, and any SNP with unknown position was removed. After applying these quality measures, 7,539 animals and 48198 SNPs were retained.

**Table 1. Proportions of different breeds in the population**

| **Breed** | **BL** | **COR** | **COOP** | **EF** | **WD** | **PD** | **TEX** | **AF** | **PS** | **MER** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Proportion (%)** | 11.1 | 0.8 | 10 | 0.7 | 0.4 | 1.8 | 2.3 | 2 | 1.1 | 69.8 |

Border Leicester: **BL**, Corriedale: **COR**, Coopworth: **COOP**, East Friesian: **EF**, White Dorper:**WD**, Poll Dorset: **PD**, Texel: **TEX**, Australian Finnsheep: **AF**, Prime Samm: **PS**, Merino:**MER**

**Data analysis.** Genome-wide association studies (GWAS) and regional heritability mapping (RHM) approaches were performed using ASReml-R (Butler *et al.*, 2009). GWAS was performed using the GRAMMAR approach (Aulchenko *et al.*, 2007). In the first step, we fitted the following animal model to the data:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_1 q} + \boldsymbol{Z_2 a} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is a vector of cube root transformed WEC records, $\boldsymbol{X}$ is a design matrix of fixed effects, $\boldsymbol{b}$ is a vector of fixed effects, $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are design matrices of random effects, $\boldsymbol{q}$ is a vector of random breed effects, $\boldsymbol{a}$ is a vector of random genetic effects, and $\boldsymbol{e}$ is the vector of residuals. The following distributions were assumed: $\boldsymbol{q} \sim N(0, I\sigma_q^2)$, $\boldsymbol{a} \sim N(0, A\sigma_a^2)$ and $\boldsymbol{e} \sim N(0, I\sigma_e^2)$, where $\boldsymbol{A}$ is the numerator relationship matrix (NRM) calculated from deep pedigree records, $\boldsymbol{\sigma_a^2}$ is the additive genetic variance explained by pedigree, $\boldsymbol{\sigma_q^2}$ is the variance of breed effects, and $\boldsymbol{\sigma_e^2}$ is the residual variance. The fixed effects were sex, rearing type × birth type, contemporary group (flock site × group of management × year of birth), age of animal at WEC recording and its quadratic polynomial. Second, residuals obtained from the animal model were treated as corrected phenotypes for a single- SNP regression:

$$\hat{\boldsymbol{y}} = \boldsymbol{1\mu} + \boldsymbol{Wa} + \boldsymbol{e}$$

where $\hat{\boldsymbol{y}}$ is a vector of adjusted phenotypes, $\boldsymbol{\mu}$ is the overall mean, $\boldsymbol{W}$ is a vector of a single SNP`s genotype for each of the animals, $\boldsymbol{a}$ is the effect size of the SNP, and $\boldsymbol{e}$ is the vector of residuals. The second approach was RHM, in which each chromosome was divided into windows of predefined number of SNPs, and the variance attributable to each window was calculated. In this analysis, two window sizes were used, 100-SNP and 50-SNP windows. The following model was fitted to the data:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_1 q} + \boldsymbol{Z_2 a} + \boldsymbol{Z_3 g} + \boldsymbol{e}$$

where the terms are as described in the animal model, and $\boldsymbol{g}$ is the regional genomic effect estimated from SNPs within each window. $\boldsymbol{g}$ was assumed to be distributed as $N(0, G\sigma_g^2)$, where $G$ is the regional genomic relationship matrix built from SNPs within each window, and $\sigma_g^2$ is the regional genomic variance. Phenotypic variance, $\sigma_p^2$, was then given by $\sigma_q^2 + \sigma_a^2 + \sigma_g^2 + \sigma_e^2$. The whole heritability was calculated as $h_a^2 = \sigma_a^2/\sigma_p^2$, whereas the regional heritability was calculated as $h_g^2 = \sigma_g^2/\sigma_p^2$ . Significance thresholds of GWAS and RHM were determined using the Bonferroni correction (significance threshold = α / N, where N is the number of tests) at the genome-wide (α =0.05) and suggestive (α=1) levels.

## RESULTS AND DISCUSION

The most significant GWAS results were observed on chromosome 2. Two SNPs, OAR2_119123707.1 and OAR2_119557086.1, were significantly associated with parasite resistance at the genome-wide and suggestive levels, respectively (Figure1). These results were also confirmed with RHM using 100 SNP window size (Figure 2). Both of GWAS and RHM

analyses generally agreed when there was a strong evidence of association (e.g.: chromosome 2). In this study, however, RHM detected more genomic regions significantly associated with parasite resistance therefore suggesting the method has greater power than GWAS analysis. For example, RHM using 100 SNP window size identified a significant region on chromosome 6 that was below the suggestive level by GWAS analysis. Furthermore, a region on chromosome 24 reaching the suggestive threshold was detected only with RHM. However, all significant regions identified by RHM explained a small proportion of WEC variation (RHM ranged from 0.0036 to 0.01), indicating that parasite resistance is a largely polygenic trait with a large number of loci involved in conferring resistance. Nagamine *et al.* (2012) showed that RHM captured more of the genetic variation than a single-SNP GWAS approach, especially when associated SNPs have very small effects to be declared significant at the genome wide level.

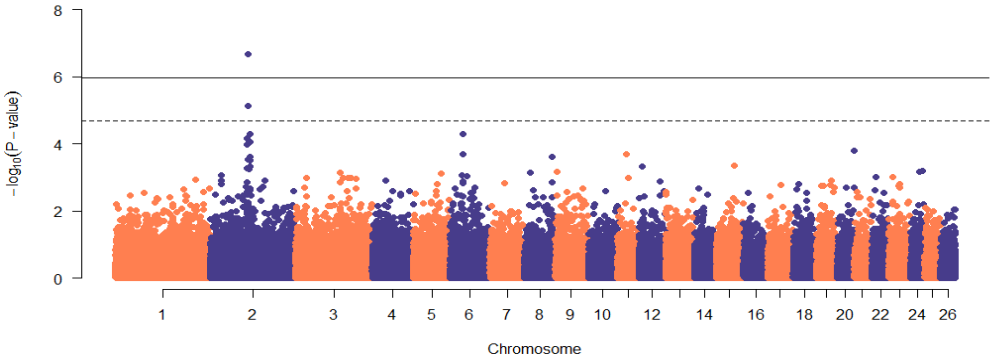


**Figure1. Manhattan plot of GWAS results.The solid line represents the genome-wide significance threshold (α = 0.05) and the dashed line represents the suggestive threshold (α = 1).**

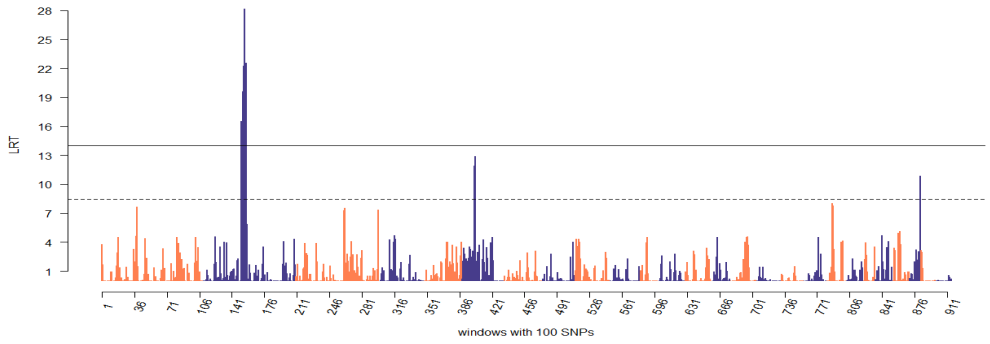


**Figure2. Regional heritability mapping (RHM) across the genome. The solid line represents the genome-wide significance threshold (α = 0.05) and the dashed line represents the suggestive threshold (α = 1).**

Significant regions with 100 SNP window size as well as those below the suggestive level were also analysed with a 50 SNP window size. The 50 SNP window size analysis confirmed the significant regions on chromosome 2 and 6, but did not confirm the region on chromosome 24. From all the region below the suggestive level, only two regions on chromosome 1 and 3 were significant. The results for RHM using 50-SNP window size are given in Table 2. Comparison

with other studies showed that significant region in chromosome 2 was contained within previously identified QTLs for parasite resistance (Hu *et al.,* 2013). This region has also been found to be associated with mastitis resistance in sheep (Jonas *et al.* 2011). Candidate genes in this region include: DEAD box polypeptide 60 (DDX60) and annexin A10 (ANXA10), which their expression found to be involved with immune response. Significant region in chromosome 6 has recently been identified by Riggio *et al.,* (2013) for parasite resistance using the 50K-SNP array. Potential candidate genes in this region include: polycystin-2 (PKD2) and ATP binding cassette G member 2 (ABCG2), which have been reported as being under selection in a study of large number of breeds (Kijas et al., 2012). Although, ABCG2 has been investigated as a candidate gene for facial eczema in sheep (Duncan *et al.*, 2007).

**Table 2: Summary of significant regions for RHM using 50 SNP window size analysis**

| OAR | Window start | Window finish | LRT | $h_g^2$ | Candidate genes |
|---|---|---|---|---|---|
| 2 | 105083320 | 107564404 | 10.07 | 0.0043 | PALLD, DDX60, ANXA10 |
| 2 | 106585530 | 108470142 | 16.12 | 0.0045 | |
| 2 | 107564404 | 109633672 | 16.13 | 0.0048 | |
| 2 | 109633672 | 113113775 | 15.59 | 0.0051 | |
| 2 | 110827578 | 114955024 | 16.69 | 0.0051 | |
| 2 | 113113775 | 116350674 | 18.42 | 0.0084 | |
| 6 | 34614727 | 38019817 | 16.97 | 0.0083 | PKD2, ABCG2, SP1 |
| 6 | 36522166 | 39035619 | 10.70 | 0.0054 | |
| 1 | 92157812 | 94722198 | 10.14 | 0.0048 | CD58, CD2, CD101, IGSF3, VTCN1, FAM46C |
| 3 | 129451837 | 1311779166 | 9.23 | 0.0047 | SOCS2 |

## CONCLUSION

This study has been successful at identifying QTLs for parasite resistance in a large multi-breed sheep population. The most significant regions were detected on chromosome 2 and 6. Four other regions on chromosome 1, 3 and 24 reached the suggestive threshold. These results also showed that there are a number of common genes that are underlying resistance to a wide range of parasite species. Furthermore, some of these common genes are possibly pleiotropic with other sheep diseases.

## REFERENCES

Aulchenko Y. S., de Koning D-J., Haley C. (2007) *Genetics* **177**:577-585.
Butler, D.G., *et al.* (2007) ASReml-R reference manual. QLD Dept. of Primary Industries.
Dominik, S., Hunt, P.W., McNally, J et al. (2010) *Parasitology* **137**: 1275-1282.
Duncan, E., *et al.* (2007). *Animal genetics*, **38,** 126-131.
Hu Z-L., Park C. A., Wu X-L., Reecy J. M. (2013), *Nucleic Acids Res* **41**:D871–D879.
Jonas, E., *et al.* (2011) *Genetics Selection Evolution.* **43,** 22.
Kemper, K. E., Emery, D. L., Bishop, S. C. et al. (2011) *Genetics research* **93**:203-219.
Kijas, J. W., *et al.* (2012) *PLoS biology*, **10,** e1001258.
Marshall, K., Maddox, J.F., Lee, S.H., et al. (2009) *Animal Genetics*. **40**:262-272.
Nagamine, Y., Pong-Wong, R., Navarro, P. et al. (2012) *PLoS ONE* **7**:e46501.
Roeber, F., Jex, A. R., Gasser, R.B. (2013) *Parasites Vectors* **6**:153.
Van der Werf J.H.J.*et al.* (2010) *Anim.Prod.Sci.* **50:** 998.
Riggio, V., Matika, O., Pong-Wong, R. et al. (2013). *Heredity.* **110**:420-429.