Short communication

A candidate gene approach to study nematode resistance traits in naturally infected sheep

Hazel Wilkie, Valentina Riggio, Oswald Matika, Louise Nicol, Kathryn A. Watt, Rona Sinclair, Alexandra M. Sparks, Daniel H. Nussey, Josephine M. Pemberton, Ross D. Houston, John Hopkins

Keywords:
Teladorsagia circumcincta
SNPs
IL23R
T cells
Parasite resistance
Sheep

1. Introduction

Young lambs are particularly susceptible to parasitic nematode infection, which is a major economic and welfare burden to sheep production (Wright, 2013). With repeated exposure, most lambs develop some level of protective immunity against the parasites (McRae et al., 2015). The use of genetic markers in selective breeding for resistant animals (currently selected based on lower faecal egg count (FEC), higher antibody (IgA) levels and higher weight compared to susceptible lambs; Berald et al., 2008) would allow for improved accuracy of selection and reduced need for routine phenotyping.

While genome wide association study (GWA) is a powerful method for detecting loci associated with the trait of interest, it requires a very large sample size to achieve adequate statistical power to detect loci with small to medium effects (Korte and Farlow, 2013). An alternative is the candidate gene approach in which the effects of polymorphisms within genes of relevance to the trait of interest are investigated (Brown et al., 2013). In the case of nematode resistance, candidate genes include those that are involved in the immune response to nematode infection.

Gene expression studies in sheep infected with Teladorsagia circumcincta have indicated that increased expression of inflammatory T helper type 1 (Th1) and/or Th17-associated genes correlate with nematode susceptibility (Gossner et al., 2012) while increased expression of Th2-associated genes correlate with resistance (Wilkie et al., 2016). During previous sequencing work (Nicol et al., 2016; Wilkie et al., 2016a) we identified potential SNPs within TBX21 (expressed by Th1 cells), IL23R and RORC2 (expressed by Th17 cells). In the present study, these SNPs were sequenced and then tested for associations with nematode resistance and associated traits in naturally infected domestic Scottish Blackface and free-living Soay ewe lambs. The two aims of the current study were (i) to confirm SNPs in the nematode resistance candidate genes TBX21, IL23R and RORC2 in different sheep breeds, and (ii) to test for associations between these SNPs and traits related to nematode infection.
resistance in naturally infected populations.

2. Materials and methods

2.1. SNP identification and validation in the experimental Scottish Blackface lambs

Numerous potential SNPs were identified in the T helper cell transcription factor and cytokine receptor genes during previous expression and cDNA sequencing experiments (Nicol et al., 2016; Wilkie et al., 2016a). Twelve experimental Scottish Blackface lambs (described in full by Beraldi et al., 2008) were used to confirm the sequence of all potential SNPs. Genomic DNA was extracted from 20 mg abomasal lymph node tissue, preserved in RNAlater (Ambion, UK), using the Wizard® SV Genomic DNA Purification System (Promega, UK) and was quantified by NanoDrop ND-1000 spectrophotometry; 200 ng was used as PCR template. FastStart Taq (Roche, UK) was used as per manufacturer’s instructions using the primers in Table S1. Six clones per SNP per lamb were sequenced as described previously (Wilkie et al., 2016a). A total of twelve SNPs in TBX21, RORC2 and IL23R were thus confirmed using this method. After sequence confirmation it transpired that all twelve SNPs had already been identified in the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/), therefore the SNP accession numbers (Table S1) were extracted.

2.2. Naturally infected Scottish Blackface population

DNA samples from Scottish Blackface ewe lambs (born between 2001 and 2003) previously analysed by Riggio et al. (2013), were used for SNP association analysis. Lambs were born on pasture and were continuously exposed to natural mixed nematode infection (including a number of genera such as Teladorsagia, Trichostrongylus, Oesophagostomum, Chabertia, Bunostomum, Cooperia, Nematodirus and Haemonchus). All lambs were sequentially sampled at 16, 20 and 24 weeks for strongyle (excluding Nematodirus) FEC and body weight; in an attempt therefore only 116 lambs had weight data at this time point in the current study. Serum anti- T. circumcincta IgA levels were measured from blood samples collected when lambs were 24 weeks old.

2.3. Naturally infected soay population

DNA samples from feral Soay ewe lambs (born between 2002 and 2015) were also used for SNP association analysis; lambs with the highest and lowest FEC from each year were selected. In August, when lambs were between 16 and 20 weeks of age, around 60% of the resident population of Hirta, St Kilda, was rounded up in temporary corral traps; body weight was measured, faecal and blood samples were taken (Clutton-Brock and Pemberton, 2004). The Soay sheep were exposed to a mixture of gastro-intestinal nematodes including the genera Teladorsagia, Trichostrongylus, Bunostomum, Chabertia, Trichuris, Capillaria and Nematodirus.

2.4. Phenotypic analysis of naturally infected lambs

For all lambs in both populations, strongyle (excluding Nematodirus) FEC was obtained using a modified McMaster Method and is reported as eggs per g faeces (epg) (Gulland and Fox, 1992; Riggio et al., 2013). Plasma fractions of the whole blood sample were used to determine serum anti- T. circumcincta IgA levels measured by ELISA (as described in Riggio et al., 2013 and Nussey et al., 2014) and expressed as: (sample OD – background OD)/ (positive control OD – background OD) on each ELISA plate.

2.5. SNP genotyping allele frequencies and LD analysis

Genotyping was conducted by LGC Genomics using KASP™ technology (http://www.lgcgroup.com/products/kasp-genotyping-chemistry/). Genomic DNA was provided at 5 ng/μl per sample. Of the twelve SNPs identified, three did not pass quality control and were not analysed further.

For the nine SNPs genotyped in the two populations, Hardy Weinberg Equilibrium (HWE) was calculated and all SNPs were found to be in HWE (data not shown). A Chi Square test with 2° of freedom confirmed differences in genotype frequency between the populations (Table S2) using Graph Pad Prism v5 (Graph Pad software, USA).

The R² between each SNP pair was estimated using Plink v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/) (Purcell et al., 2007) to explore linkage disequilibrium (LD) within the Blackface and Soay populations.

2.6. Statistical analyses for SNP associations

A generalised linear model (GLM) in SAS 9.4 (2015) was used to explore the significance of the fixed effects. Significant effects were then fitted in a linear mixed model in ASReml (Gilmour et al., 2015) to estimate genetic parameters. The response variables were FEC, IgA and weight (at various ages in the Blackface lambs). The fixed effects fitted for the Blackface population were: early management group, lamb birth year, dam age at sampling (years) and litter size. The fixed effects fitted for the Soay population were: lamb birth year, lamb age at sampling (weeks), dam age at lambing (years) and litter size. In both cases animal was fitted as a random effect.

The animal models were fitted as follows:

\[ y = Xb + Zu + e \]

Where y is a vector of observations on the specific trait; b is a vector of non-genetic fixed effects; u is direct additive genetic effects; X and Z are corresponding incidence matrices and e is the residual. The normal distributions assumed for the random effect were as follows: animal \~ N(0, Aαα), e \~ N(0, \sigma^2e), where A is the numerator relationship matrix, and \sigma^2 is an identity matrix of the order equal to the number of records. The relationship matrix was constructed using pedigree records of 960 Blackface sheep and 730 Soay sheep. A sire model (fitted as above but with sire as the random effect) was then used in the Blackface population to confirm associations. The SNPs were fitted as fixed effects in the association analyses.

Multiple testing was accounted for using a Bonferroni correction threshold of \( P \leq 0.002 \). This was calculated as \( P = \frac{0.05}{\alpha + (\beta + \gamma)} \); where \( \alpha \) corresponds to the (5) independent LD blocks of SNPs; \( \beta \) represents the (3) independent traits tested in the Soay lambs; \( \gamma \) corresponds to the (3) independent traits tested in the Blackface lambs (weight at different time points were correlated therefore were treated as one trait for correction purposes).

SNP effects were estimated as:

Additive effect, \( a = (AA - aa)/2 \);

Dominance effect, \( d = Aa - [(AA + aa)/2] \);

Percentage of additive variance (VA) due to the SNP = \([2pq (a + d (p - q))]/VA\)

where AA, Aa and aa were the predicted trait values for each genotype class; p and q were the allelic frequencies at the SNP locus. Animals with missing data were excluded from relevant analyses; the total number of lambs included in each analysis is detailed in Table S3. Statistical power was calculated using GWApower software (Feng et al., 2011).
Table 1
Validation of SNP locations and effects in ovine IL23R, RORC2 and TBX21.

| SNP Accession No. | Chr (strand) | Genome position | Gene | Exon | cDNA position | Protein position | Effect |
|-------------------|--------------|-----------------|------|------|---------------|------------------|--------|
| rs5935336447      | 1 (+)        | 42,469,152      | IL23R| 3    | 1292C > T    | P110S            | Missense|
| rs408363839       | 1 (+)        | 42,499,077      |     | 7    | 1823A > G    | N287D            | Missense|
| rs426358915       | 1 (+)        | 42,512,431      |     | 8    | 1934G > A    | V324M            | Missense|
| rs405079951       | 1 (+)        | 42,512,460      |     | 8    | 1963A > T    | K333N            | Missense|
| rs413830024       | 1 (+)        | 42,487,890      | Intron 4-S | 492-115A > G | Intron variant |        |
| rs415145731       | 1 (+)        | 42,487,903      | Intron 4-S | 492-102C > T | Intron variant |        |
| rs159639535       | 1 (-)        | 100,654,917     | RORC2| 5    | 880G > C     | E294Q            | Missense|
| rs30822388        | 1 (-)        | 100,653,186     |     | 8    | 1219G > A    | A404T            | Missense|
| rs428174832       | 1 (-)        | 100,648,086     |     | 3’ UTR | *25T > C    | Downstream gene variant|
| rs415026575       | 1 (-)        | 100,648,002     |     | 3’ UTR | *109A > G    | Downstream gene variant|
| rs426434073       | 11 (-)       | 38,270,681      | TBX21| 3’ UTR | *861A > G    | Downstream gene variant|
| rs411294999       | 11 (-)       | 38,270,671      |     | 3’ UTR | *871A > G    | Downstream gene variant|

1 SNP accession numbers from dbSNP database.
2 Genome position is nucleotide number in Oar v4.0.
3 cDNA position is relative to the ‘A’ of the start codon denoted as position 1.
4 Protein position indicates the amino acid residue affected. SNPs in **bold** were not sequenced in the naturally infected lambs. Chr: Chromosome.

3. Results

3.1. SNP confirmation in Blackface lambs

A total of twelve SNPs in IL23R, RORC2 and TBX21 were verified by sequencing genomic DNA in experimental Blackface lambs (Table 1). Six SNPs were identified in IL23R on Chr 1 plus strand (NC_019458.2). Three SNPs (p.P1105, c.492-115A > G and c.492-102C > T) did not pass LGC Genomics quality control and were not analysed further (bold in Table 1). IL23R p.N287D was located in exon 7 within the extra-cellular domain. IL23R p.V324M and IL23R p.K333N were in exon 8, between the ligand-binding motif and the transmembrane domain (Table 1). Four SNPs were identified in RORC2, on Chr 1 minus strand (NC_019458.2); RORC2 p.E294Q was located in exon 5, between the predicted ligand and DNA-binding domains. RORC2 p.A404T was situated in exon 8, within the ligand-binding domain. RORC2 c.*25T > C and RORC2 c.*109A > G were situated 25 and 109 nucleotides downstream of the stop codon respectively (Table 1). TBX21 is located on ovine Chr 11 minus strand (NC_019468.2); TBX21 c.*861A > G and TBX21 c.*871A > G, both were located downstream of the stop codon within the predicted 3’ UTR (Table 1).

3.2. Phenotype analysis of naturally infected populations

FEC data were natural log transformed (In(epg + 15)) and IgA data were cube-root transformed for normal distributions, weight data were approximately normally distributed without transformation (Figs. S1 and S2). At all time points, mean FEC of the Blackface ewe lambs was lower than mean FEC in the Soay population (Fig. S3A) and mean Blackface body weight was greater than Soay mean body weight (Fig. S3B). Litter size was larger in the Blackface compared to the Soay population (Fig. S3C) and dams were older at lambing in the Soay compared to the Blackface (Fig. S3D).

3.3. SNP frequencies and LD analysis in naturally infected populations

Within each population, all SNPs were in HWE. The loci RORC2 c.*25T > C, RORC2 c.*109A > G and TBX21 c.*861A > G were fixed in the Soay population (Table S2). Among those loci that segregated in both populations, there was a significant difference (P < 0.05) in allele frequencies at IL23R p.N287D, IL23R p.V324M, RORC2 p.E294Q and TBX21 c.*871A > G (Table S2).

In the Blackface population, RORC2 c.*25T > C and RORC2 c.*109A > G were in complete LD, whilst p.N287D and p.K333N in the IL23R gene had the second highest R2 value of 0.55 for this population (Fig. S4A). IL23R p.N287D and p.K333N were in complete LD in the Soay population, while RORC2 p.E294Q and p.A404T had the second highest R2 value of 0.61 in this population (Fig. S4B). Lower R2 values were observed for other SNP pairs within genes. SNPs in different genes were not in LD.

3.4. SNP associations in naturally infected Scottish Blackface population

A statistical power calculation based on the sample size of the current study (n = 230) and assuming an R2 between marker and causative variant of 0.3 indicated that the study had a power of 0.82 to detect QTL explaining 5% of the phenotypic variation (h2 > 0.05).

There were no significant (P ≤ 0.002) associations between the SNPs and the resistance traits in the Blackface lambs (Table S3). However, a nominally significant association (p = 0.007 and p = 0.004 fitting animal and sire model, respectively) was identified for IL23R p.V324M and body weight at 20 weeks in this population (Table S3). The estimated additive genetic effect was 1.60 ± 0.51 kg (p = 0.002) with no significant dominance effects (0.60 ± 0.77 kg, p = 0.14). The percentage of additive variance (VA) due to the SNP was 6.41%. The heritability estimate for body weight at 20 weeks had a large point estimate but was not different from zero (0.72 ± 0.48), however, the predicted genotypic effects for weight at 20 weeks were significant (p = 0.01), with higher mean (± SEM) body weight in GG homozygotes (32.5 ± 1.1 kg) compared to AA homozygotes (29.3 ± 1.0 kg; Fig. 1).

3.5. SNP associations in naturally infected Soay population

There were so significant associations between any of the SNPs and body weight at 20 weeks based on IL23R p.V324 M allele in Blackface lambs. G allele encodes V; A allele encodes M. *P value = 0.01. Error bars are SEM.
phenotypes in the Soay population (Table S3). The association (P = 0.071) between IgA and RORC2 p.A404T (Table S3) was the closest to the threshold but was not analysed further.

4. Discussion

This study identified SNPs within ovine TBX21, RORC2 and IL23R genes in an experimentally challenged Scottish Blackface population. The SNPs were primarily situated within or in close proximity to functional domains of the genes, and it was hypothesised that they could result in altered gene function. This has been previously reported in human IL23R (Duerr et al., 2006) in which a SNP within IL23R was protective against inflammatory bowel disease (Di Meglio et al., 2011). Therefore, mutations within inflammation-associated genes can protect against inflammatory pathways potentially leading to increased Th2 and IgA responses (Langrish et al., 2004). The ovine SNPs within IL23R, TBX21 and RORC2 were therefore analysed for associations with FEC, IgA and body weight in two different naturally infected lamb populations. An advantage of the current study was the sequencing of the SNPs prior to association analysis to confirm that these SNPs represent real polymorphisms (Table 1) not predicted mutations. This study used two unrelated and phenotypically distinct (Fig. S2) sheep breeds indicating that the SNPs identified were not breed-specific. Functional analysis is required to confirm if these SNPs affect gene expression or protein structure and function as is hypothesised.

No significant associations were found between any of the SNPs and the traits considered in either of the populations. This is most likely due primarily to the small sample size and the polygenic nature of the traits (Sweeney et al., 2016).

In conclusion, this study identified and confirmed segregating SNP markers in genes of biological relevance to host immune response, and potentially nematode resistance. The nominally significant association between IL23R p.V324 M and body weight under infection indicates that this SNP may be worthy of further investigation in additional, larger sheep populations to test its association with nematode resistance.

Conflict of interests

The authors declare that they have no conflicts of interest.

Acknowledgements

The authors wish to thank Dr Tom McNeilly, Moredun Research Institute, for providing the L3 antigen used in the Soay antibody ELISA tests. This project was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Grants to The Roslin Institute (BB/J004235/1 and BB/J004324/1). HW was a BBSRC Industrial CASE scholar, sponsored by Sheep Improved Genetics Ltd and financially supported by BBSRC and AHDB Beef and Lamb (Grant BB/J012653/1). The Soay sheep project data was funded by NERC and ERC grants awarded to JMP and others, the Soay sheep immune data was funded by a BBSRC David Phillips fellowship awarded to DHN. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetpar.2017.06.010.

References

Beraldi, D., Craig, B.H., Bishop, S.C., Hopkins, J., Pemberton, J.M., 2008. Phenotypic analysis of host-parasite interactions in lambs infected with Teladorsagia circumcincta. Int. J. Parasitol. 38, 1567–1577.
Brown, E.A., Pilkington, J.G., Nussey, D.H., Watt, K.A., Hayward, A.D., Tucker, R., Graham, A.L., Paterson, S., Beraldi, D., Pemberton, J.M., Slate, J., 2013. Detecting genes for variation in parasite burden and immunological traits in a wild population: testing the candidate gene approach. Mol. Ecol. 22, 757–773.
Clutton-Brock, T.H., Pemberton, J.M., 2004. Soay Sheep: Dynamics and Selection in an Island Population. Cambridge University Press.
Davies, G., Stear, M.J., Benoisthan, M., Abazogb, O., Kerr, A., Mitchell, S., Bishop, S.C., 2006. Quantitative trait loci associated with parasitic infection in Scottish blackface sheep. Heredity 96, 252–258.
Di Meglio, P., Di Cesare, A., Lagger, U., Chua, C.C., Napolitano, L., Villanova, F., Toni, I., Capon, F., Trembath, R.C., Peris, K., Nestle, F.O., 2011. The IL23R R381Q gene variant protects against immune-mediated diseases by impairing IL23-induced Th17 effector responses in humans. PLoS One 6, e2024.
Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhardt, A.H., Abraham, C., Regueiro, M., Griffiths, A., Dassopoulos, T., Bitton, A., Yang, H., Targan, A., Datta, L.W., Koneru, E.O., Schumm, L.P., Lee, A.T., Gregersen, P.K., Barrand, M.M., Rotter, J.L., Nicolae, D.L., Cho, J.H., 2006. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314, 1461–1463.
Feng, S., Wang, S., Chen, C., Lan, L., 2011. GWAPower: a statistical power calculation software for genome-wide association studies with quantitative traits. BMC Genet. 12, 12.
Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J., Thompson, R., 2015. ASReml User Guide Release 4.1 Structural Specification. VSN International Ltd.
Gossner, A.G., Venturina, V.M., Shaw, D.J., Pemberton, J.M., Hopkins, J., 2012. Relationships between susceptibility of Blackface sheep to Teladorsagia circumcincta infection and an inflammatory mucosal T cell response. Vet. Res. 43, 26.
Gullden, F.M.D., Fox, M., 1992. Epidemiology of nematode infections of Soay sheep (Ovis aries L.) on St. Kilda. Parasitology 105, 481–492.
Korte, A., Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9, 29.
Langrish, C.L., McKenzie, B.S., Wilson, N.J., De Waal Malefyt, R., Kastelein, R.A., Cua, D.J., 2004. IL-12- and IL-23: master regulators of innate and adaptive immunity. Immunol. Rev. 202, 96–105.
McRae, K.M., Stear, M.J., Good, B., Keane, O.M., 2015. The host immune response to gastrointestinal nematode infection in sheep. Parasite Immunol. 37, 605–613.
Nicol, L., Gossner, A., Watkins, C., Chianini, F., Dalziel, R., Hopkins, J., 2016. Variations in IL-23 and IL-25 receptor gene structure, sequence and expression associated with the two disease forms of sheep paratuberculosis. Vet. Res. 47, 27.
Nussey, D.H., Watt, K.A., Clark, A., Pilkington, J.G., Pemberton, J.M., Graham, A.L., McNeilly, T.N., 2014. Multivariate immune defences and fitness in the wild: complex but ecologically important associations among plasma antibodies, health and survival. Proc. R. Soc. B 281, 201432931.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, M., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
Riggio, V., Matsika, O., Pong-Wong, R., Stear, M.J., Bishop, S.C., 2013. Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. Heredity 110, 420–431.
Sweeney, T., Hanrahan, J.P., Ryan, M.T., Good, B., 2016. Immunogenomics of gastrointestinal nematode infection in ruminants – breeding for resistance to produce food sustainably and safely. Parasite Immunol. 38, 569–586.
Wilkie, H., Gossner, A., Bishop, S., Hopkins, J., 2016a. Variations in T cell transcription factor sequence and expression associated with resistance to the sheep nematode Teladorsagia circumcincta. PLoS One 11, 2.
Wilkie, H., Nicol, L., Gossner, A., Hopkins, J., 2016b. Mucosal expression of T cell gene variants is associated with differential resistance to Teladorsagia circumcincta. PLoS One 11, 12.
Wright, N., 2013. Economic Impact of Health and Welfare Issues in Beef Cattle and Sheep in England. ADAS UK Ltd.. http://beefandlamb.ahdb.org.uk/wp/wp-content/uploads/2013/2004/Economic-Impact-of-Health-Welfare-Final-Rpt-170413.pdf.