Student Final Report No. 61110033 **October 2016**

Identifying and Exploiting the Molecular Basis of Resistance to Gastro-intestinal Nematodes in Sheep

October 2016

Student Final Report No. 61110033

Identifying and Exploiting the Molecular Basis of Resistance to Gastro-intestinal Nematodes in Sheep

Hazel E. Wilkie

The Roslin Institute, The University of Edinburgh, Edinburgh, EH25 9RG

Supervisor: Professor John Hopkins

This is the final report of a PhD project (AHDB Beef and Lamb Ref: 61110033) that ran from October 2012 to September 2016. The work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and AHDB Beef and Lamb.

CONTENTS

1. Abstract

Teladorsagia circumcincta is a common species of parasitic worm that infects sheep in the UK. The parasite causes an estimated £84 million per annum in production loss to UK farmers, through loss of appetite and growth impairment in young stock. Animals are treated regularly with anthelmintic (drench) throughout spring and summer to reduce their worm burden. However, the overuse of these drugs has led to the evolution of drug-resistant parasites which are extremely difficult to kill. This has led to research into other ways to manage the infection, with the aim to limit the development of drug resistance whilst maintaining flock health.

Sheep naturally acquire resistant immunity to the worms with repeated exposure to small numbers of the parasite. This immunity is also heritable, so lambs can be bred to naturally develop resistance to the parasite. The aims of this project were; 1) understand how sheep respond to worm infection, 2) understand and identify genes associated with the response and 3) identify variation within those genes which may contribute to resistance.

A number of immune response genes were sequenced within this project which led to the identification of several mutations within these genes. Quantifying the expression levels of the genes in resistant and susceptible lambs illustrated associations between the genes and the development of resistance. Therefore, the genes investigated in this project have the potential to aid genetic selection of worm-resistant sheep in the future.

2. Introduction

One of the most common gastro-intestinal parasites of sheep in the UK is the nematode *Teladorsagia circumcincta*, which infects the sheep's fourth stomach (the abomasum) (Dunn, 1969). Young lambs are highly susceptible and become infected soon after weaning. In time, lambs develop an immune response against the worms (Stear *et al*., 1999). This can either lead to clearance of the infection, termed resistance measured by reduced faecal egg count (FEC), reduced adult worm count (AWC), increased anti-parasite antibody (IgA) production and/or increased body weight (BW); or susceptibility, measured by high FEC, high AWC, low IgA and low BW (Beraldi *et al.,* 2008).

The production of protective antibodies is regulated by the differential polarization of T cells, and studies in mice have indicated the critical role of T helper type 2 (Th2) cells in the control of worm infection and the development of resistance (Allen and Maizels, 2011). Inflammatory responses linked with the activation of Th1 and Th17 cells lead to tissue damage and exacerbated disease, resulting in the susceptibility traits (Ahern *et al*., 2008).

The development of polarized T helper cell subsets is under the control of cytokine receptors (which bind to signalling (cytokine) molecules) and transcription factors (which are activated by cytokine receptor signalling and switch on other genes, initiating downstream pathways) (Murphy and Weaver, 2016).

Th2 cells express the receptor for the cytokine interleukin-25 (IL25); the receptor is made up of two subunits – *IL17RA* and *IL17RB*. Signalling through this receptor activates the transcription factor *GATA3* (Murphy and Reiner, 2002). Th17 cells express the IL23 receptor, made up of the *IL23R* and *IL12RB1* subunits. Signalling through this receptor activate two transcription factors: *RORC2* and *RORA* (Gaffen, 2009). Th1 cells express the receptor for IL12, which is made up of the *IL12RB1* and *IL12RB2* subunits. When IL12 binds to its receptor the transcription factor *TBX21* is activated (Murphy and Weaver, 2016).

Alternative splicing is a common mechanism by which multiple variants of the same gene are created (Breitbart *et al*., 1987). This can introduce non-productive mutations (where the variant is useless and immediately broken down by the cell) as well as beneficial mutations (such as different expression of different variants in different tissues).

The hypothesis of this project was that genes involved in differential T helper cell activation could be key to the development of resistance and susceptibility. Furthermore, it was believed that mutations within these cytokine receptor and transcription factor genes may cause inhibition of certain responses (e.g. Th17) leading to activation of the competitive response (e.g. Th2).

To investigate this hypothesis, experimentally infected Blackface ewe lambs were utilised which had been bred to have a range of susceptibilities to *T. circumcincta* (Beraldi *et al*., 2008). The transcription factors (*GATA3, RORA, RORC2* and *TBX21*) were sequenced in the abomasal lymph node (ALN; the draining lymph node and site of immune response initiation in this infection) and splice variants were identified. Assays were designed to quantify the expression of these genes and variants in the ALN and the abomasal mucosa (AM; the site of infection). The cytokine receptor subunits (*IL17RA, IL17RB, IL23R* and *IL12RB1*) were sequenced by Nicol *et al*. (2016). Quantification assays were also designed by Nicol *et al*. (2016); these were then optimised for use in the current study. Genes that were significantly differentially expressed between the most resistant and susceptible lambs were then analysed further for correlations with resistant traits (FEC, AWC, IgA and BW), indicating genes with potential as selection markers of resistant lambs. These results were published by Wilkie *et al*. (2016a) and Wilkie *et al*. (2016b).

3. Materials and methods

3.1. Animals and experimental design

Female Blackface lambs were ~13 weeks old were housed in worm-free conditions; 45 lambs were infected with ~2300 infective L3 *T. circumcincta* larvae three times a week for 12 weeks, and 10 were sham-infected controls. At post mortem the AWC ranged from 0 to 11,300 and FEC from 0 – 950 eggs per g (Table 1) and the animals were ranked (1 - 45) according to their infection level. All details of animals and animal husbandry, infection protocols, phenotypes and population genetic analyses have been previously described (Beraldi *et al*., 2008). Animal experiments were approved by University of Edinburgh Ethical Review Committee and conducted under an Animals (Scientific Procedures) Act 1986 Project Licence.

Table 1 Phenotypic markers collected at post mortem from artificially infected Blackface lambs. (Beraldi *et al.*, 2008).

Sheep	Infection rank[1]	AWC[2]	FEC[3]	BW (kg)	IgA[4]
92	1	0	0	39	1.195
100	2	0	0	38	0.63
21	3	0	0	37.5	0.798
20	4	0	0	37	0.633
58	5	0	0	37	0.077
50	6	0	0	37	0.373
110	7	0	0	36.5	0.384
54	8	0	0	36	1.695
116	9	0	0	36	1.066
25c	10	80	0	36	0.856
155	11	100	0	35	0.126
52	12	100	0	35	0.547
34	13	100	25	35	0.154
184	14	200	0	35	0.782
123	15	300	0	34	0.706
10	16	400	0	34	0.706
193	17	420	75	34	0.232
102	18	200	25	34	0.232
40	19	600	0	32.5	0.596
12	20	900	0	32	0.804
125	21	800	50	31	0.703
62	22	1200	0	30	0.142
172	23	1700	175	30	0.21
181	24	1200	25	30	0.161
165	25	2400	0	30	0.732
19	26	2300	175	29	1.51
8	27	2400	475	29	0.539
138	28	2400	75	29	0.183
48	29	2600	100	29	0.259
30	30	3100	125	28	0.468
82	31	3300	175	28	0.245
190	32	2900	225	28	0.219
178	33	3800	100	27.5	0.073
59	34	3900	250	27.5	0.84
191	35	4200	275	27	0.06
65	36	4700	150	27	0.033
60	37	5400	75	27	0.697
119	38	5300	250	26	0.151
131	39	4000	125	26	0.035
28	40	6000	200	26	0.126
114	41	6000	200	26	0.451
38	42	5200	525	25	0.155
173	43	6200	200	22	0.047
183	44	5300	950	20	0.209
109	45	11300	625	15	0.141

[1]Ranked on basis of both adult worm count (AWC) and FEC, [2]AWC; total number adult worms in abomasal contents post mortem, [3]Faecal egg counts per g faeces post mortem, [4]Relative levels of serum anti-*T. circumcincta* L3 IgA.

3.2. Sample collection and RNA isolation

Abomasal lymph node (ALN) and mucosa (AM) was removed immediately post mortem and stored at –80˚C in RNAlater (Ambion, UK). Total RNA was isolated from ~ 20mg tissue using the Ribopure Kit (Ambion) according to the manufacturers' instructions, and genomic DNA was removed by on-column PureLink® DNase I treatment (Ambion). RNA quantity, quality and integrity was assessed by a NanoDrop ND-1000 spectrophotometer and Agilent 2200 TapeStation system; all samples had an RNA Integrity Number of >7.5.

3.3. Cloning and sequencing of transcription factors

cDNA was synthesised from 1 µg RNA using SuperScript™ II RT with RNaseOUT (Invitrogen, UK) and oligo-dT(15) primer (Promega, UK). The predicted sequences of the full length sheep genes were obtained by NCBI-BLAST of the bovine sequences against the Oar v3.1 sheep genome assembly (http://www.livestockgenomics.csiro.au/sheep/oar3.1.php/). Primers were selected using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and reanalysed using Net Primer (http://www.premierbiosoft.com/netprimer/). The sequencing primers have been published by Wilkie *et al*. (2016a) and Nicol *et al*. (2016). Each primer set was used in RT-PCR using FastStart Taq (Roche, UK) as per manufacturers' instructions. PCR products were fractionated by agarose gel electrophoresis, visualized by gel red /UV transillumination, purified using MinElute PCR Purification Kit (Qiagen), ligated into pGEM-T Easy vector (Promega) and transformed into JM109 High Efficiency Competent Cells (Promega). Colonies that contained the inserted vector and target sequence were incubated overnight at 37°C in LB broth with 50 µg/ml ampicillin. Purified plasmid was extracted from the culture using QIAprep Spin MiniPrep Kit (Qiagen) following the manufacturer's protocol. Three clones from six resistant and six susceptible lambs for each insert were sequenced using T7 and SP6 primers, with BigDye1Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, UK). The GeneRacerTM Kit for full-length, RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends (Invitrogen) was used to sequence the 5' and 3' untranslated region (UTR) of the transcription factors. Primers and nested primers were designed (Wilkie *et al*., 2016a) and the protocol was followed as detailed by the manufacturer, using SuperScript™ III protocol for reverse transcription reaction and Platinum(R) Taq DNA Polymerase High Fidelity protocol for the amplification of the final product, which were cloned as described above.

3.4. RT-qPCR quantification of transcript variants

cDNA was synthesised from 1.0 µg RNA using SuperScript™ II RT with RNaseOUT (Invitrogen, UK) and oligo-dT(15) primer (Promega, UK) in 20 µl final volume. Quantification primers for the transcription factor and cytokine receptor variants have been published by Wilkie *et al*. (2016a) and Nicol *et al*. (2016). Primers were designed to overlap exon/exon boundaries; all amplicons were sequenced to ensure specificity. Each reaction contained 7.5 µl FastStart Universal SYBR Green Master (Rox) 2x concentrated master mix (Roche), 2 µl template cDNA (diluted 1/10), 0.25 – 1.0 µl of each primer at 10 mM and nuclease-free water to a final volume of 15 µl. Reactions were prepared using a CAS-1200™ robot and performed on a Rotor-Gene Q (Qiagen). Amplification was followed by dissociation curve analysis. PCR optimization was performed on cDNA from a pool of samples. Not all variants could be quantified in the AM; the signals for *GATA3v1, RORAv2, v3, v4* and *v5*; and *IL23Rv1, v2, v3, v4* and *v5; IL12RB1v1, v2* and *v4*; and *IL17RBv1* were too low for accurate quantification as they were detected only after > 30 cycles PCR, outside the linear part of the standard curve.

Relative expression levels were measured in duplicate from two separate RT reactions for each of the nine most resistant and the nine most susceptible lambs, with duplicate no-template controls included in all runs. Optimized RT-qPCR assays had an efficiency >95% and R^2 value of >0.98. Absolute copy numbers were also calculated from duplicate samples from two separate RT reactions for all lambs.

Relative transcript levels were calculated in GenEx 5 (MultiD Analyses AB, Sweden) using the comparative 2-(ΔΔ Cq) method and normalized to the geometric mean of *GAPDH* and *SDHA*; fold changes were calculated from ΔCq values using GenEx. To calculate copy number in all 45 infected and the 10 uninfected sheep, a standard curve of linearized plasmid was used with a dynamic range of at least five orders of magnitude. For each point on the standard curve, copy numbers were calculated from Cq values:

Molecules per ng = ([1x 10-9]/(M g/mol)] x [6.03x10-23 molecules/mol])

M= plasmid size x 660g/mol per bp.

The expression levels were normalized by dividing the copy number derived from the standard curve, by the calculated normalization factor (Vandesompele *et al*., 2002) for each sample, using the geometric mean of *GAPDH* and *SDHA*. RT replicates were averaged per animal and multiplied by the dilution factor (x100) to calculate the copy number per µg of total RNA.

3.5. Statistical analysis

Relative transcript levels were analysed in GenEx using an unpaired, 2-tailed t-test to determine the difference between groups. Graph Pad Prism 6.07 for Windows (Graph Pad Software, USA) was used for statistical analysis of the copy number expression data. The data were grouped into resistant, intermediate, susceptible (n =15 per group) and uninfected control (n =10). One-way

ANOVA was performed to determine overall significance and Tukey's multiple comparisons test within ANOVA was used to determine significance between groups. Correlations between transcript levels and quantitative phenotypes were analysed with 2-tailed Spearman's correlation coefficient (r_s); P-values ≤ 0.05 were considered statistically significant.

4. Results

4.1. Identification of transcription variants of sheep transcription factors (Wilkie *et al*. 2016a)

Cloning and sequencing identified two transcription variants of *GATA3* and *RORC2* and five variants of *RORA*. *GATA3* is encoded on the plus strand of chromosome 13 (Oar v3.1; NC_019464). The difference between full length *GATA3* (LN848231) and *GATA3v1* (LN848232) is the deletion of a single codon (g.806_808delGAA) at the 5' end of exon 3 (Chr13: 12,082,935–12,082,937) which represents the deletion of a glutamic acid (E) at position 260 within the protein (Figure 1). *RORC2* is encoded on the minus strand of chromosome 1 (NC_019468) and, in comparison to the full length gene (LN848233), *RORC2v1* (LN848234) has a 36 bp deletion (g.1237_1272del) at the 3' end of exon 7 (Chr1: 100,653,158–100,653,123; Figure 2), which represents the deletion of 12 amino acids (GKYGGVELFRAL) at position 359–370 of the protein. Analysis of ovine *TBX21* identified only one sequence, identical to XM_004012818.2. *RORA* is encoded on the plus strand of chromosome 7 (NC_019464) where five transcript variants (LN848235 –LN848239) were identified in the region spanning Chr7: 46,097,028 – 46,731,648 (Figure 3). All five variants are identical from position 46 of *RORAv1* (Chr7: 46,691,081), which represents the 5' end of exon 6. The five variants have variable usage of exons 1 to 6 and consequently 5' UTRs (untranslated regions) of different lengths. *RORAv1, v2, v3* and *v5* have unique translation start sites, while *RORAv2* and *v4* have the same translation start site and encode identical derived protein sequences of 387 amino acids. *RORAv5* encodes the largest protein of 513 amino acids.

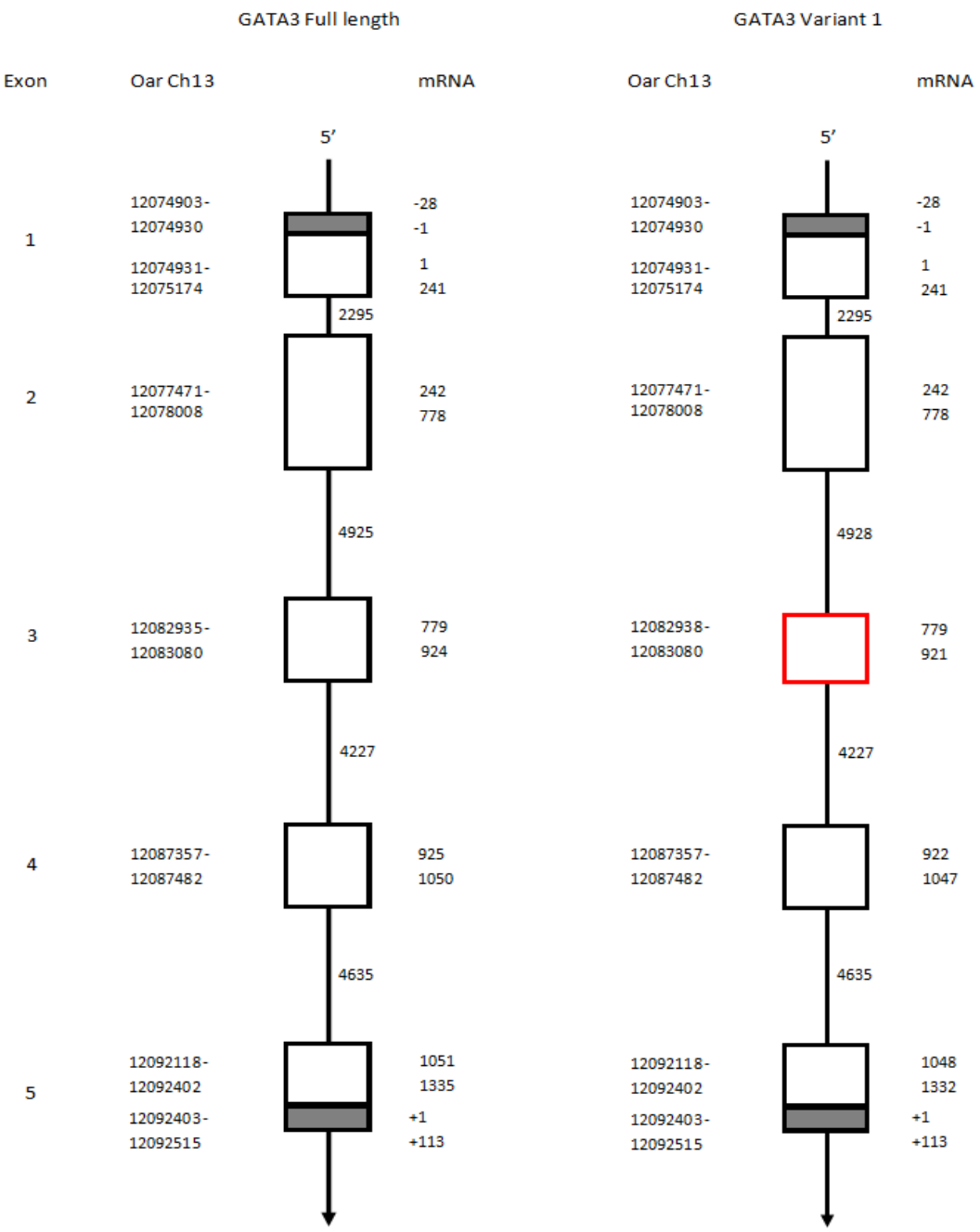


Figure 1 *GATA3* full length and variant exon maps.

Untranslated regions (UTRs) are shaded grey, coding regions are white, arrows represent strand direction, numbers between boxes are intron lengths, numbers in the mRNA column represent the location of the coding region within exons, the red box represents the variant exon. Chromosome locations are based on Oar v3.1.

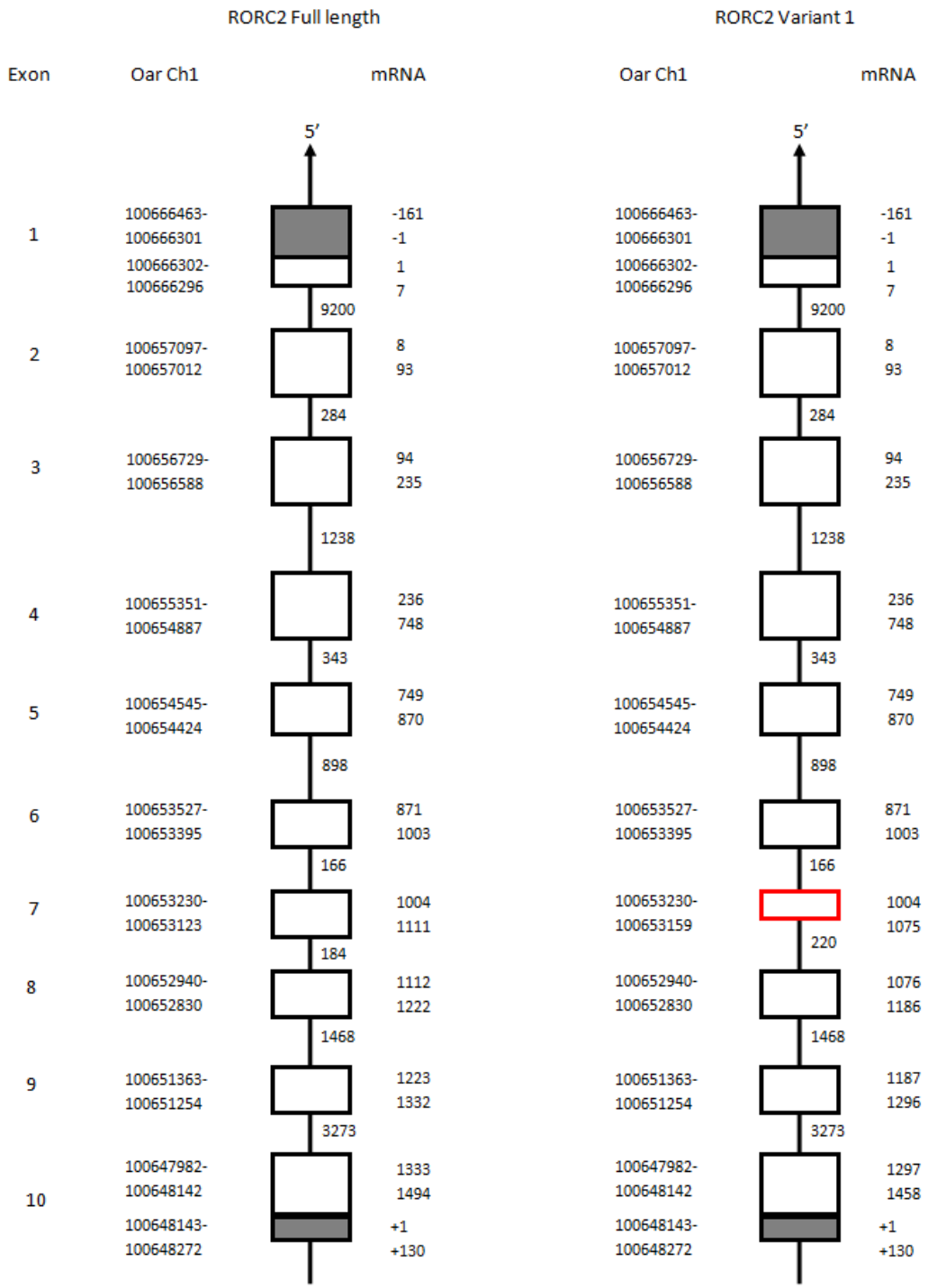


Figure 2 *RORC2* full length and variant exon maps.

UTRs are shaded grey, coding regions are white, arrows represent strand direction, numbers between boxes are intron lengths, numbers in the mRNA column represent the location of the coding region within exons, the red box represents the variant exon. Chromosome locations are based on Oar v3.1.

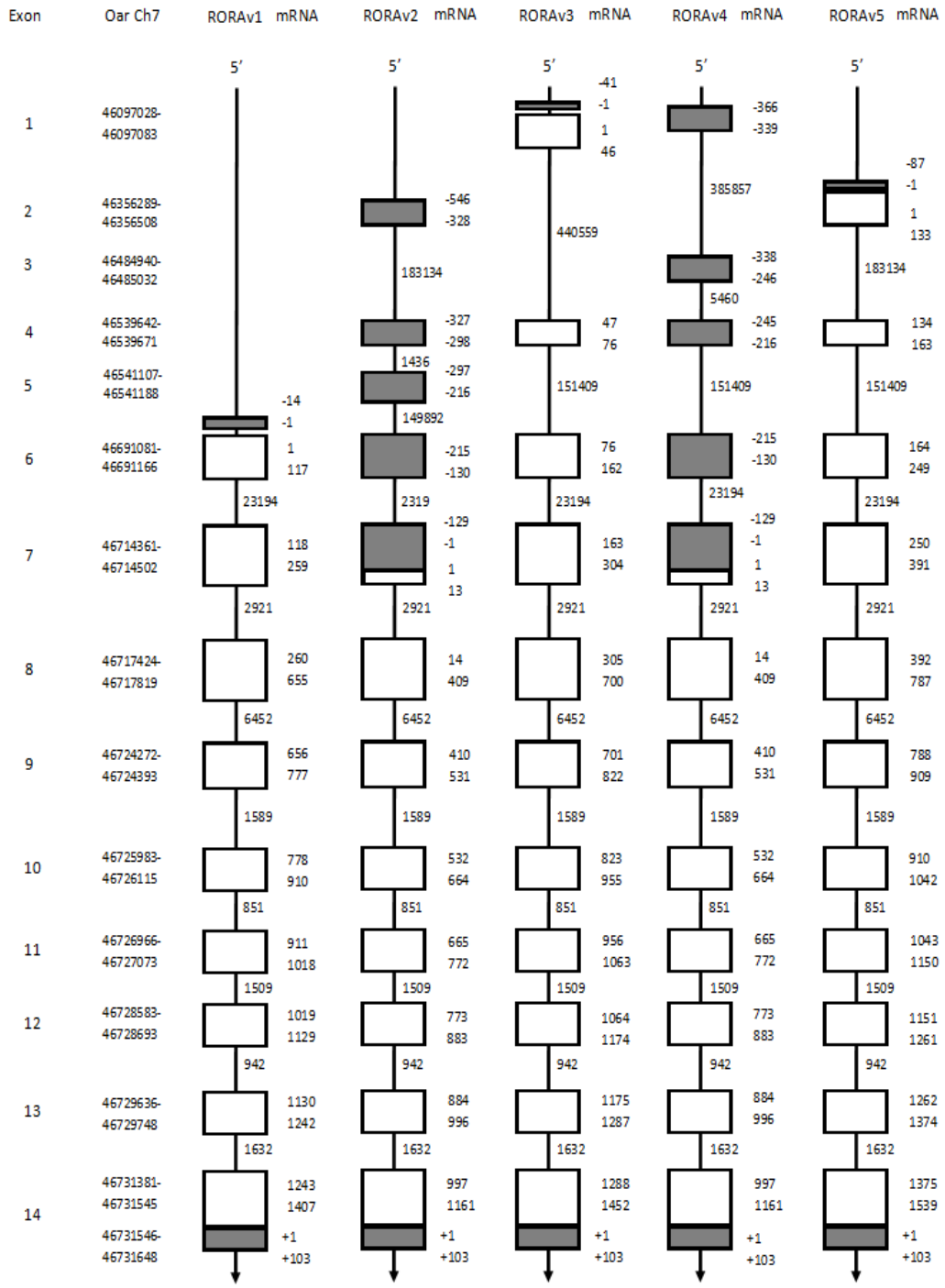


Figure 3 *RORA* variant exon maps.

UTRs are shaded grey, coding regions are white, arrows represent strand direction, numbers between boxes are intron lengths, numbers in the mRNA column represent the location of the coding region within exons. Chromosome locations are based on Oar v3.1.

4.2. RT-qPCR quantification of transcription factor expression (Wilkie *et al*., 2016a and 2016b)

Relative RT-qPCR assays were developed for *TBX21*, for the two variants of *GATA3* and *RORC2* and for five variants of *RORA*. These were used to compare the nine most resistant (R) and the nine most susceptible (S) lambs (Table 1). The resistant group (rank 1–9) had no detectable AWC or FEC, mean body weight (BW) of 37 kg and high IgA antibody levels (mean 0.76 relative units). The susceptible lambs (rank 37–45) were those with the highest AWC (mean 6078, maximum 11300), high FEC (mean 350, maximum 950), low BW (mean 23.7 kg) and low IgA (mean 0.22 units). Table 2 shows the relative expression (fold change) of each transcription factor and variant between the resistant and susceptible groups and shows that all but *RORAv5* were equally expressed in the ALN of both groups. *RORAv5* was 1.57 fold higher in the resistant animals but p = 0.08. *RORAv3* could not be quantified because expression levels were too low for accurate measurement.

In the abomasal mucosa, *GATA3* was significantly increased 2.06 fold (p = 0.0002) in the resistant animals and *RORC2v1* (-1.30 fold, p = 0.03) was significantly decreased in the resistant animals compared to susceptible lambs (Table 2). Full length *RORC2* was also significantly decreased in the resistant group (-1.41 fold, p = 0.01). *TBX21* and *RORAv1* were not significantly differentially-expressed.

Table 2 Relative transcription factor expression in the ALN and AM.

Gene	ALN		Abomasal Mucosa	
	Fold Change (R vs. S)	**p value**	**Fold Change (R vs. S)**	**p value**
GATA3	1.06	0.62	**2.06**	**0.0002**
GATA3v1	-1.12	0.47	*	
TBX21	-1.10	0.61	1.24	0.19
RORA (Total)	-1.10	0.61	*	
RORAv1	-1.13	0.79	-1.37	0.25
RORAv2	-1.07	0.83	*	
RORAv3	*†		*	
RORAv4	-1.04	0.88	*	
RORAv5	1.57	0.08	*	
RORC2	1.18	0.42	**-1.41**	**0.01**
RORC2v1	1.12	0.54	**-1.30**	**0.03**

* Signals obtained after 30 PCR cycles were too low to quantify.

† Positive control confirmed assay was optimised.

Bold: significant P values.

4.3. Correlation of transcription factor levels in ALN and quantitative phenotypes (Wilkie *et al*., 2016a)

Absolute quantitative analysis was performed on *RORAv5* as it was the only transcription factor variant that showed evidence of differential expression between the R and S groups in the ALN; *RORAv2* was also quantified as the control *RORA* variant.

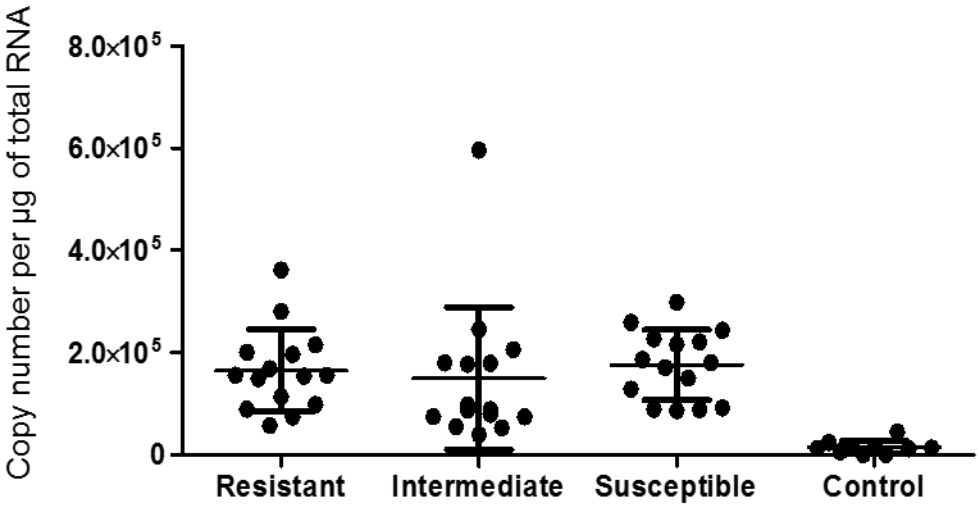
ANOVA confirmed the expression of *RORAv2* was equal in all three infected groups (mean copy number per µg RNA ± SD; resistant 164411 ± 80305, intermediate 148781 ± 139319, susceptible 175622 ± 68746), p value 0.77. Mean expression level of *RORAv2* in the uninfected controls was 14336 ± 12927, which was a significant reduction compared to the resistant (p value < 0.01), intermediate (p value < 0.01) and susceptible (p value < 0.001) groups (Figure 4A).

RORAv5 mean expression levels were 148439 ± 91397 in resistant, 93161 ± 44477 in intermediate, 84818 ± 69870 in susceptible and 11308 ± 6571 in control animals, (p value < 0.0001; Figure 4B). Tukey's multiple comparison post-hoc test determined a significant difference between expression in the resistant and susceptible groups (p value < 0.05). The uninfected control group mean expression level was also significantly lower than the resistant (p value < 0.0001), intermediate (p value < 0.01) and susceptible (p value < 0.01) groups.

Spearman Rank correlation analysis found a significant negative correlation between *RORAv2* and IgA levels (r_s -0.30, p value < 0.05) but there were no significant correlations between this transcript and AWC, FEC or BW (Figure 5).

RORAv5 significantly negatively correlated with AWC (r_s -0.53, p value < 0.05) and FEC (r_s -0.55, p value < 0.05) and positively correlated to BW (r_s 0.53, p value < 0.05) but there was no significant association with IgA (r_s 0.22, p value > 0.05; Figure 5).

A. *RORAv2*



B. *RORAv5*

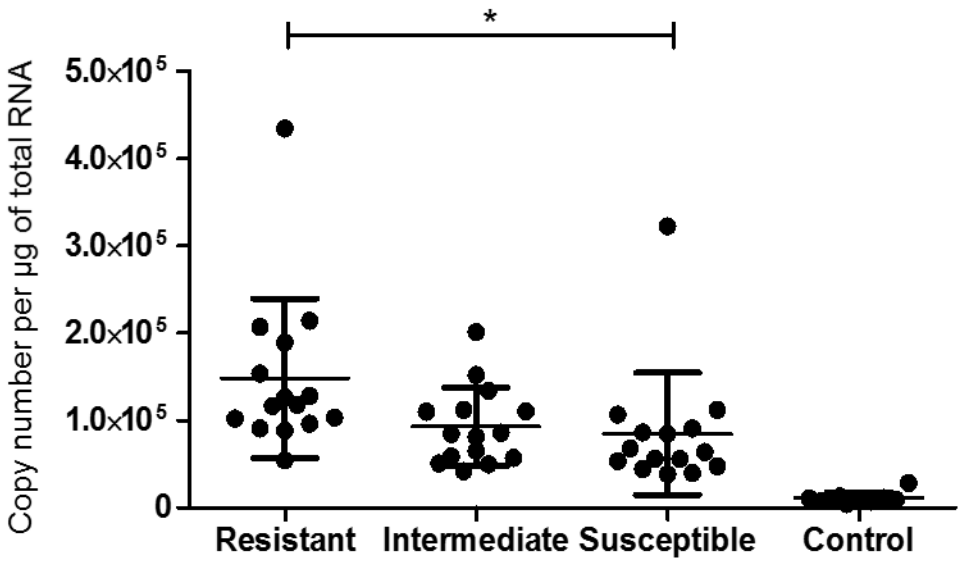


Figure 4 *RORAv2* and *RORAv5* expression levels in the ALN of 45 infected lambs and 10 uninfected controls (Wilkie *et al.*, 2016a). * p value < 0.05.

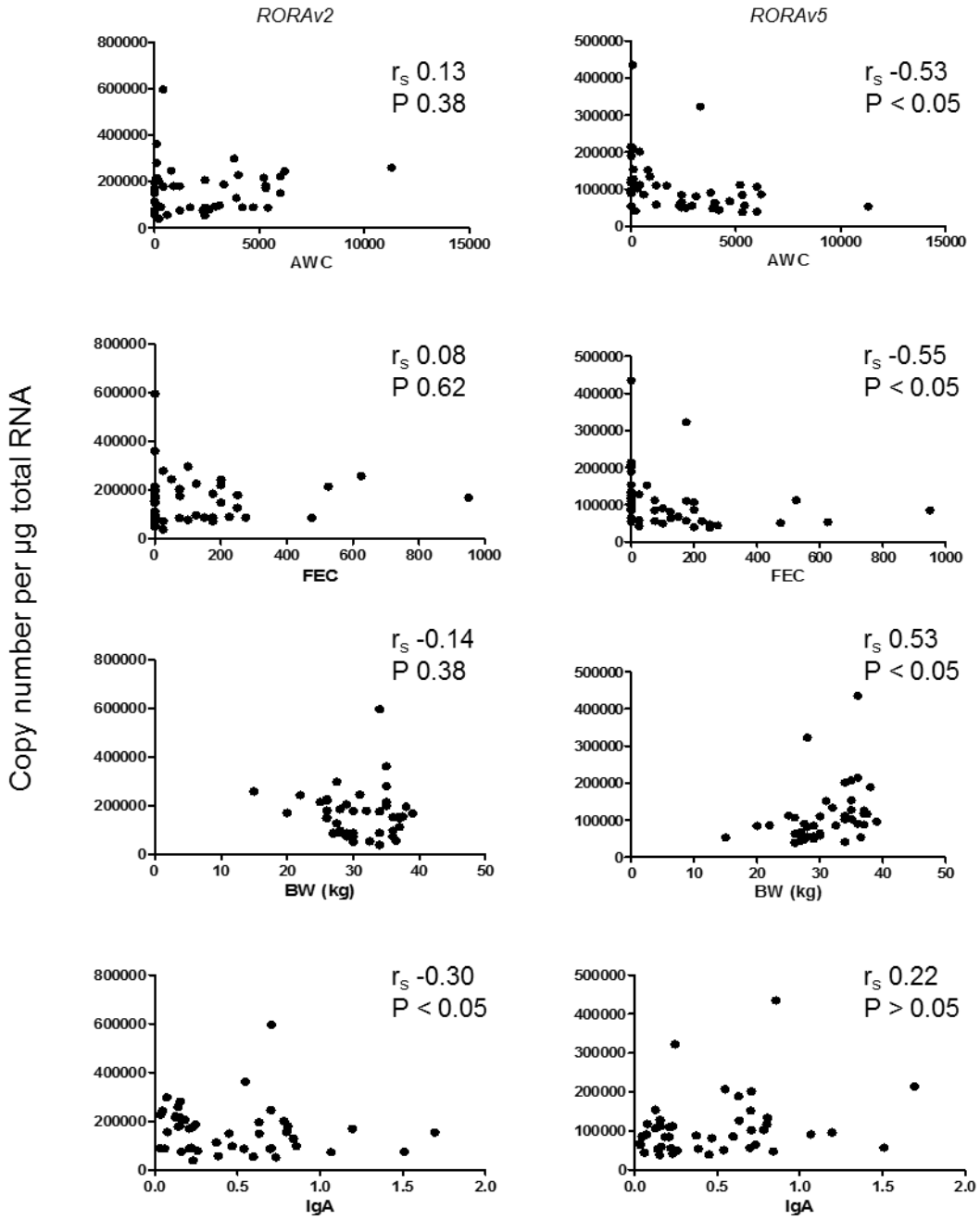


Figure 5 Correlation analysis of the phenotypic parameters and *RORAv2* and *RORAv5* in ALN (Wilkie *et al.*, 2016a).

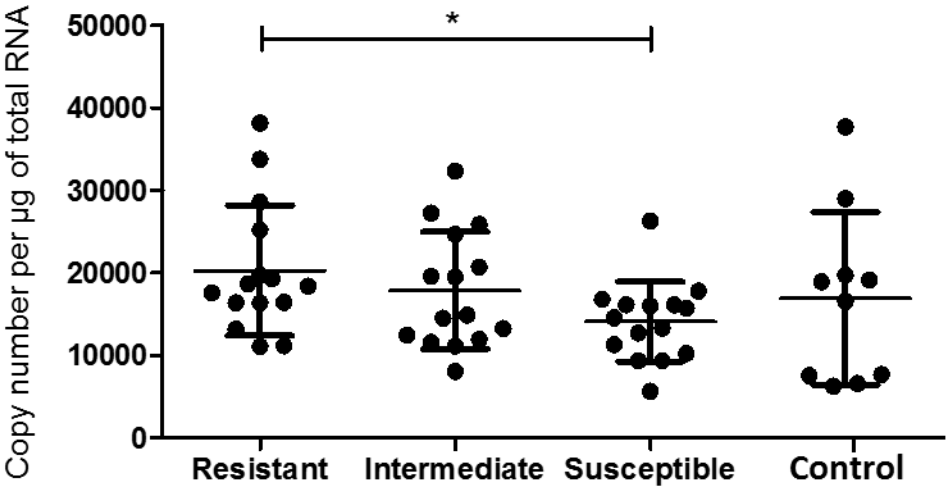
AWC: adult worm count in the abomasum; FEC; faecal egg count (eggs per gram faeces); BW: body weight in Kg; IgA: relative *T. circumcincta* L3 IgA antibody level; r_S: Spearman Rank correlation coefficient; P: p value.

4.4. Correlation of transcription factor levels in AM and quantitative phenotypes (Wilkie *et al*., 2016b)

ANOVA determined that the mean expression level of *GATA3* (resistant 20291 ± 7904, intermediate 17869 ± 7116, susceptible 14102 ± 4831, control 16929 ± 10479) did not differ significantly overall, p value 0.17. However, Tukey's post-hoc multiple comparison test identified a significant difference between the resistant and susceptible groups (p value 0.015; Figure 6A). In the control group, *RORC2v1* mean levels (48352 ± 20764) were significantly higher ($p < 0.0001$) than the resistant (18919 ± 6998), intermediate (17630 ± 5482) and susceptible (21420 ± 6757) groups (Figure 6B).

Spearman's rank correlation analysis found that *GATA3* was significantly negatively correlated with AWC (r_s -0.42, p value 0.004) and significantly positively correlated with BW (r_s 0.44, p value 0.003) and IgA (r_s 0.32, p value 0.03) but there was no significant correlation with FEC (Figure 7). Correlation analysis identified no significant correlations between *RORC2v1* and any of the quantified phenotypes (Figure 7).

A. *GATA3*



B. *RORC2v1*

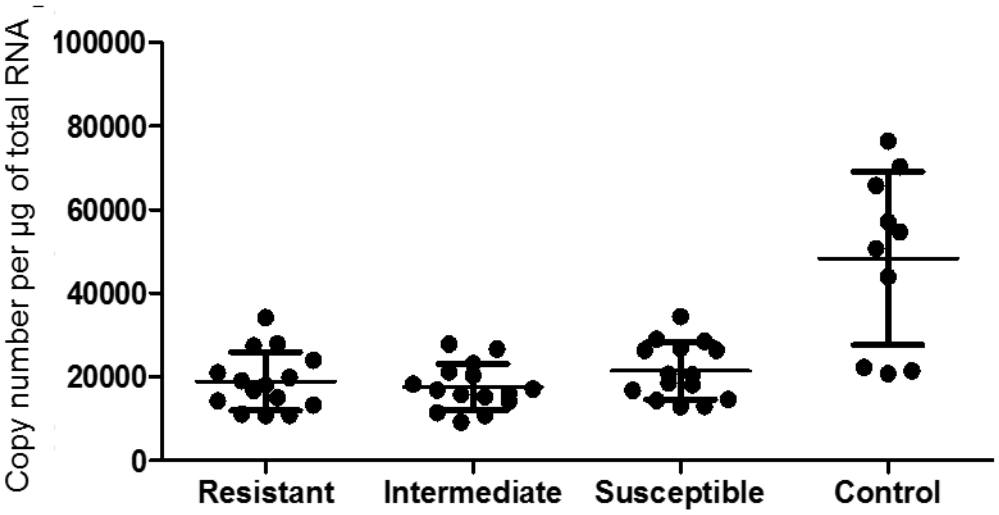


Figure 6 *GATA3* and *RORC2v1* expression levels in the AM of 45 infected lambs and 10 uninfected controls (Wilkie *et al.*, 2016b). * p value < 0.05.

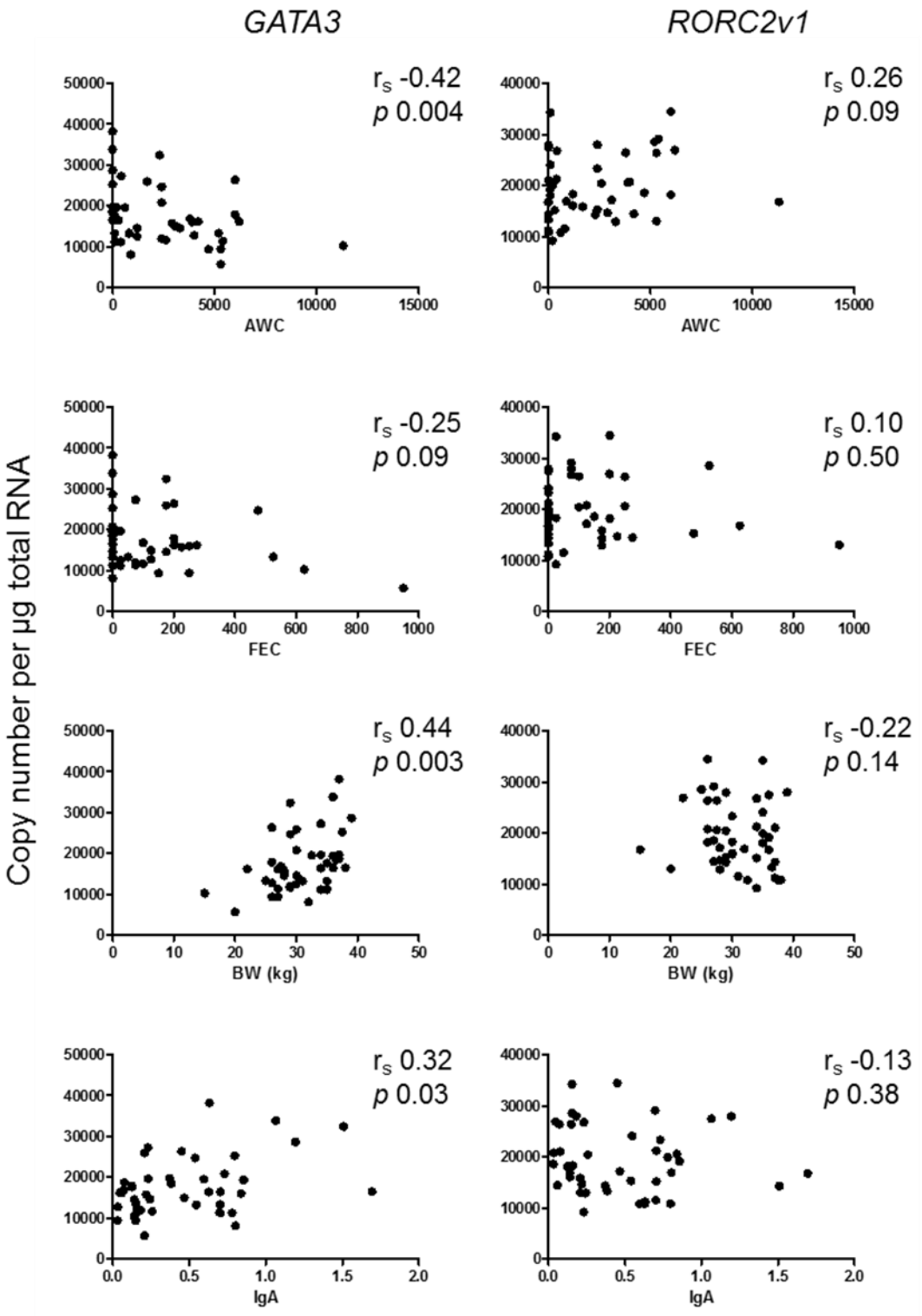


Figure 7 Correlation analysis of the phenotypic parameters and G*ATA3* and *RORC2v1* in AM (Wilkie *et al.*, 2016b).
AWC: adult worm count in the abomasum; FEC; faecal egg count (eggs per gram faeces); BW: body weight in Kg; IgA: relative *T. circumcincta* L3 IgA antibody level; r_s: Spearman Rank correlation coefficient; *p*: p value.

4.5. RT-qPCR quantification of cytokine receptor expression (Wilkie *et al*., 2016b)

Table 3 shows the relative expression (resistant vs. susceptible, fold change and P-value) of each quantifiable variant in the ALN and AM. In the ALN, *IL17RB* was significantly increased 1.62 fold (p = 0.03) in the resistant animals and *IL23R* approached significance (-2.76 fold, p = 0.08) compared to susceptible lambs. In the AM, only *IL17RBv2* was significantly differentially expressed (2.10 fold, p = 0.01). All other subunit genes were not significantly differentially-expressed or not quantifiable.

Table 3 Relative cytokine receptor expression in the ALN and AM.

Gene	ALN		Abomasal Mucosa	
	Fold Change (R vs. S)	**P-value**	**Fold Change (R vs. S)**	**P-value**
IL17RA	-1.19	0.30	1.12	0.39
IL17RB	**1.62**	**0.03**	1.81	0.08
IL17RBv1	1.15	0.62	*	
IL17RBv2	*		**2.10**	**0.01**
IL17RBv3	*		-1.09	0.82
IL23R	-2.76	0.08	*	
IL12RB1	-1.08	0.61	1.42	0.11
IL12RB1v3	-1.27	0.23	1.36	0.15
IL12RB1v4	-1.04	0.89	*	

* Signals obtained after 30 PCR cycles were too low to quantify. *IL23R* variants 1-5 and *IL12RB1* variants 1-2 were too low to quantify in either tissue, for which a positive control confirmed the assays were optimised. **Bold**: significant P values.

4.6. Correlation of cytokine receptor levels in AM and quantitative phenotypes (Wilkie *et al*., 2016b)

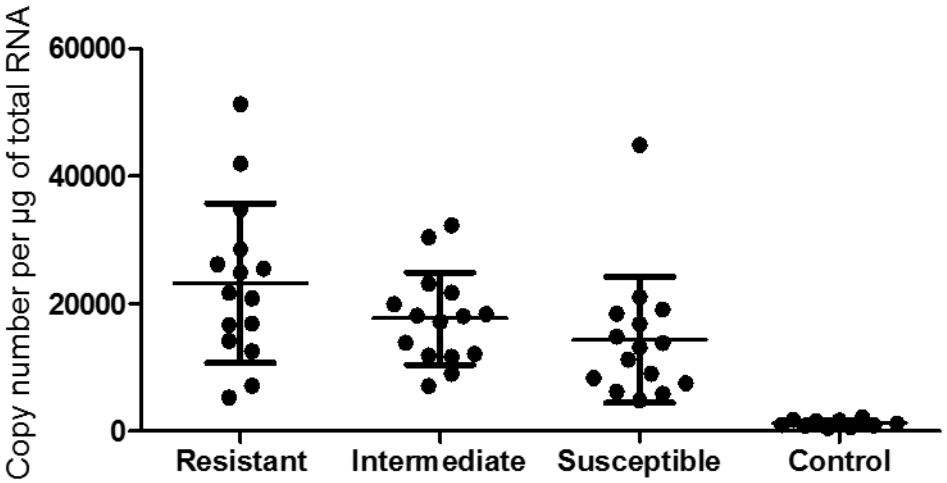
In the abomasal mucosa, *IL17RB* was expressed in equal levels in all infected groups (resistant 23207 ± 12487, intermediate 17653 ± 7217, susceptible 14338 ± 9907), but was significantly lower in the control group (1234 ± 538), compared to resistant (p value < 0.0001), intermediate (p value < 0.0001) and susceptible (p value < 0.001) groups (Figure 8A).

Mean levels of *IL17RBv2* were not significantly different between any group (resistant 1182 ± 826, intermediate 987 ± 579, susceptible 652 ± 544, control 1006 ± 625; Figure 8B), p value 0.18.

The correlation analysis found a significant negative correlation between *IL17RB* and AWC (r_s -0.37, p value 0.01) and a significant positive correlation with BW (r_s 0.40, p value 0.006; Figure 9). There was no significant correlation between *IL17RB* and FEC (r_s -0.26, p value 0.08) or IgA (r_s 0.26, p value 0.09).

Spearman rank correlation analysis found a significant negative correlation between *IL17RBv2* and AWC (r_s -0.38, p value 0.01) and FEC (r_s -0.40, p value 0.006), and a significant positive correlation with BW (r_s 0.42, p value 0.004) and IgA (r_s 0.41, p value 0.005), illustrated in Figure 9.

A. *IL17RB*



B. *IL17RBv2*

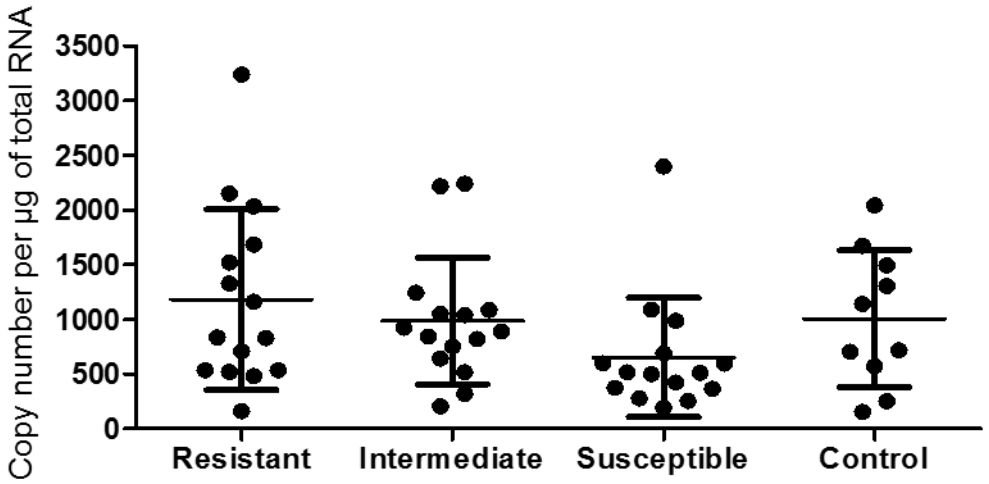


Figure 8 *IL17RB* and *IL17RBv2* expression levels in the AM of 45 infected lambs and 10 uninfected controls (Wilkie *et al.*, 2016b). * p value < 0.05.

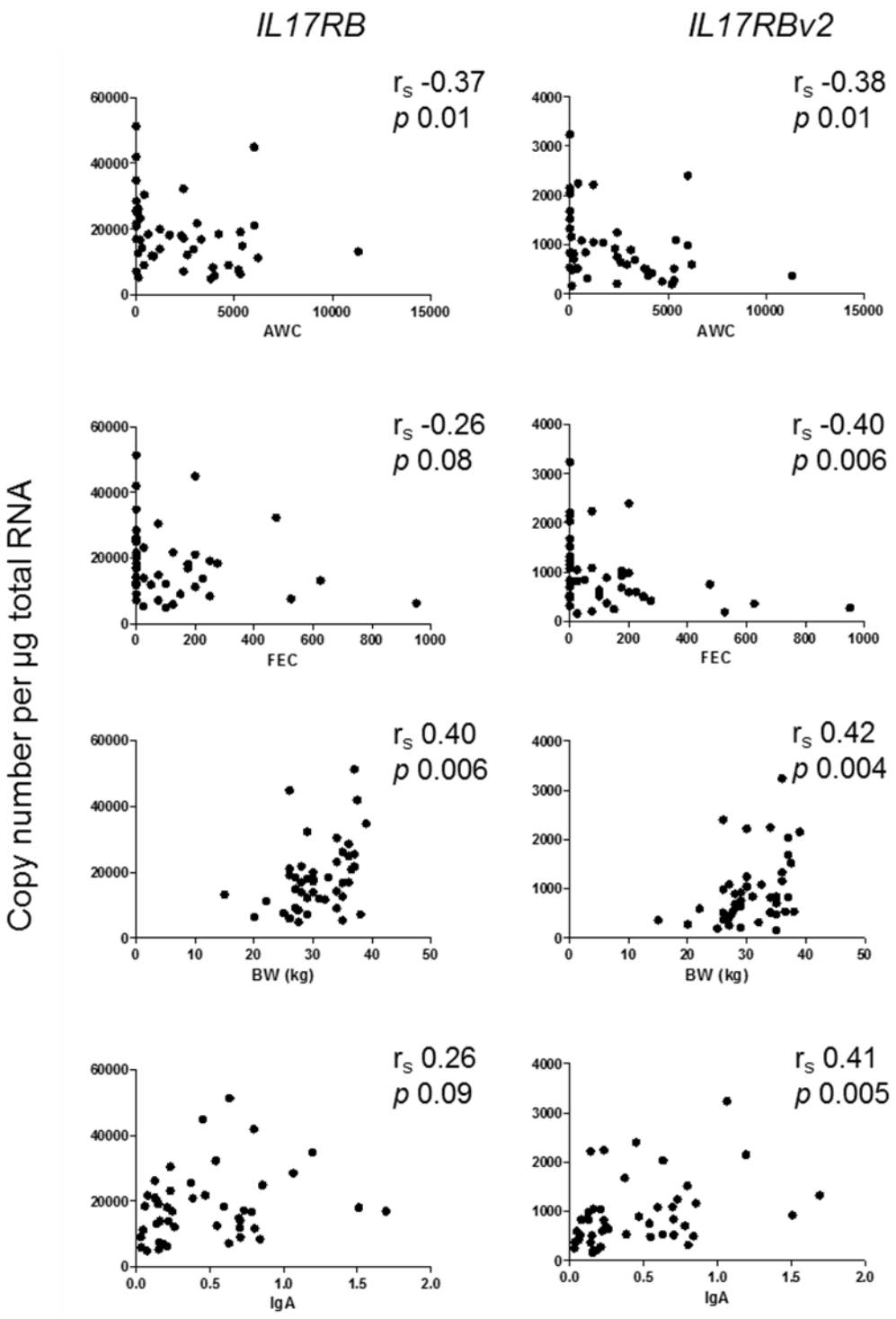


Figure 9 Correlation analysis of the phenotypic parameters and *IL17RB* and *IL17RBv2* in AM (Wilkie *et al.*, 2016b). AWC: adult worm count in the abomasum; FEC; faecal egg count (eggs per gram faeces); BW: body weight in Kg; IgA: relative *T. circumcincta* L3 IgA antibody level; r_s: Spearman Rank correlation coefficient; *p*: p value.

5. Discussion

Previous studies on the immunological basis of resistance to *T. circumcincta* using this model have highlighted the role of Th2 cells in resistance and Th1/Th17 cell activation in susceptibility (Pemberton *et al*., 2011; Gossner *et al*., 2012). This study represents the first systematic attempt to assess the role of the four major immune system transcription factors in an ovine infectious disease context.

In this study two transcript variants of *GATA3* were detected, the DNA-binding domains of which share 100% peptide sequence identity with the DNA-binding domains of known human transcript variants. It is therefore predicted that the ovine transcripts will have similarly located functional domains. Both the human and ovine truncated variant have amino acid E260 deletion, however, there are no known functional effects of this mutation in humans and E260 is situated outside of any functional domains (Gaynor *et al*., 2009). Within the abomasal lymph node, there was no significant difference between the resistant and susceptible groups in *GATA3* or *GATA3v1* expression. This result therefore suggests that at this late time point in chronic infection, there are similar levels of Th2 cells in the ALN of both resistant and susceptible lambs (Wilkie *et al*., 2016a). Within the abomasal mucosa, *GATA3v1* levels were not detectable, however, *GATA3* was significantly increased 2.10 fold in the resistant compared to susceptible lambs (Table 2). This is the first indication of a difference in expression of the shorter variant compared to the wild type gene, which potentially indicates a functional effect of the variant, although the mechanism of action is not clear. Correlation analysis in the mucosa found significant negative correlations between *GATA3* and AWC, and significant positive correlations with BW and IgA (Wilkie *et al*., 2016b). Therefore, not only is this gene increased in resistant lambs, it is also associated with the phenotypes of resistance. This is evidence of a robust association between *GATA3* expression and the maintenance of resistance in chronic worm infection.

There were no splice variants of sheep *TBX21* identified (Wilkie *et al*., 2016a). *TBX21* has been unambiguously sequenced in relatively few species, including humans, cattle and mice, and in all these species only a single gene has been characterized. Despite being detectable in the ALN and AM, there was no significant difference in *TBX21* expression between the resistant and susceptible lambs. It is concluded that Th1 cells are not increased in the susceptible lambs relative to the resistant lambs at this time point. However, it is possible Th1 cells play a role in the early induction of the inflammatory response to nematodes, but this can not be investigated in the current model.

The two ovine *RORC2* variants identified are homologous to human *RORC2*. The truncated protein predicted to be encoded by *RORC2v1* has a 12 amino acid deletion in the predicted ligand-binding

domain (Wilkie *et al*., 2016a). There is a known variant of human *RORC2* in which exons 5 – 8 (within the hinge region of the ligand-binding domain) are deleted, causing suppression of IL17 production (Rauen *et al*., 2012). It was therefore hypothesised that the ovine variant might also affect the function of Th17 cells and may be associated with resistance rather than susceptibility. However, there was no significant difference between *RORC2* and *RORC2v1* in the resistant and susceptible groups in the ALN (Wilkie *et al*., 2016a). This indicates there is little difference between the levels of Th17 cell activation in the ALN of resistant and susceptible lambs in late-stage chronic worm infection. In the abomasal mucosa, both *RORC2* and *RORC2v1* were significantly increased in the susceptible lambs relative to the resistant lambs, thereby indicating the ovine variant does not appear to have altered function (Wilkie *et al*., 2016b) unlike the human variant.

The identification of five variants of *RORA* within the ovine lymph node (Wilkie *et al*., 2016a) highlights the hitherto unexplored polymorphic capability of *RORA* within the immune system. The variants differ in their 5' UTR sequence and translational start site, particularly in their A/B domain sequence which is required for nuclear localisation (Aschrafi *et al*., 2006). It is predicted that all ovine variants encode a ligand-binding domain but that ovine *RORAv2* and *RORAv4* do not encode the DNA-binding domain or the A/B domain. Consequently, it is hypothesised that ovine *RORAv2* and *RORAv4* remain in the cell cytoplasm and act as competitive inhibitors of the other variants; in essence they could sequester the activating ligand, reducing downstream signalling within the cell (Wilkie *et al*., 2016a). Since *RORAv5* was correlated with resistance and *RORAv2* was correlated with susceptibility in the ALN, a model is hypothesised in which the ratio of *RORAv2*:*RORAv5* expression is a critical factor in cell phenotype outcome, such that resistant lambs have high enough *RORAv5* expression to overcome the inhibition of *RORAv2*, while in susceptible lambs, *RORAv5* activation and downstream signalling is successfully inhibited by *RORAv2*. *RORAv5* could therefore be used as a marker of resistant lambs in the future.

Quantitation analysis of the cytokine receptor subunits sequenced by Nicol *et al*. (2016) identified only *IL17RB* and its variant *IL17RBv2*, as significantly differentially expressed between resistant and susceptible lambs. Both these transcripts were robustly correlated with the traits of resistance (Figure 9), indicating this gene as a potential marker of resistant lambs in the future.

GATA3 is the critical transcription factor of Th2 polarization and *IL17RB* is a major component of the IL-25 receptor, expressed by Th2 cells. The results for these two transcripts imply that Th2 cells, at the site of parasite infection (AM) and immune response initiation (ALN), may play a role in the maintenance of the resistance phenotype.

Data using mouse models of gastrointestinal nematode infections have shown that components of the excretory-secretory products of infecting helminths can suppress the immune response and

can activate regulatory cells to enhance the suppression (Grainger *et al*., 2010). These factors can lead to the inhibition of Th1 and the promotion of Th2 development. It is possible that some of the variation in the expression of the T cell genes in this study have been influenced by parasite secretions, for example RORC2v1 expression which was significantly lower in infected lambs compared to uninfected lambs (Figure 7). *T. circumcincta* secreted products have also been shown to promote Foxp3-expression (a regulatory cell marker) *in vitro* (McNeilly *et al*., 2013).

This study investigated the abomasal lymph node and mucosal expression of variants of the transcription factors and cytokine receptors associated with differential T cell activation. Measurements of expression of each variant, in selected lambs with well-defined phenotypes of resistance to the abomasal parasite *T. circumcincta*, identified *RORAv5*, *GATA3* and *IL17RB* levels were positively correlated with resistance and negatively correlated with susceptibility. This indicates that Th2-polarized T cells may play an important role in the maintenance of resistance and suggests these genes could be useful markers of worm-resistant sheep in the future.

6. Industry messages

The overall goal of this project was to identify candidate genes associated with resistance that could be used as selection markers for future breeding programmes. Using an experimentally-infected Blackface lamb model, genes with characterised immune response functions in mice were investigated in sheep. This confirmed that these genes have robust associations with disease outcome in worm-infected lambs.

Current genetic techniques (such as SNP chips) use predominantly predicted mutations from cattle to interrogate the ovine genome for disease-associated genes. The genes, and mutations with them, discovered in this project, are real ovine mutations and represent novel markers that could be used for the genetic selection of worm-resistant sheep in the future.

7. References

Ahern, P. P., Izcue, A., Maloy, K. J. and Powrie, F. 2008. The interleukin-23 axis in intestinal inflammation. Immunological Reviews 226: 147-159.

Allen, J. E. and Maizels, R. M. 2011. Diversity and dialogue in immunity to helminths. Nature Reviews Immunology 11(6): 375-388.

Aschrafi, A., Meindl, N., Firla, B., Brandes, R. P. and Steinhilber, D. 2006. Intracellular localization of RORalpha is isoform and cell line-dependent. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1763(8): 805-814.

Beraldi, D., Craig, B. H., Bishop, S. C., Hopkins, J. and Pemberton, J. M. 2008. Phenotypic analysis of host-parasite interactions in lambs infected with *Teladorsagia circumcincta*. International Journal for Parasitology 38(13): 1567-1577.

Breitbart, R. E., Andreadis, A. and Nadal-Ginard, B. 1987. Alternative splicing: a ubiquitous mechanism for the generation of multiple protein isoforms from single genes. Annual Reviews of Biochemistry 56: 467-495.

Dunn, A. M. 1969. Veterinary Helminthology. Willian Heinemann Medical Books Ltd.

Gaffen, S. L. 2009. Structure and signalling in the IL-17 receptor family. Nature Reviews Immunology 9(8): 556-567.

Gaynor, K. U., Grigorieva, I. V., Nesbit, M. A., Cranston, T., Gomes, T., Gortner, L. and Thakker, R. V. 2009. A missense *GATA3* mutation, Thr272Ile, causes the hypoparathyroidism, deafness, and renal dysplasia syndrome. The Journal of Clinical Endocrinology and Metabolism 94(10): 3897-3904.

Gossner, A. G., Venturina, V. M., Shaw, D. J., Pemberton, J. M. and Hopkins, J. 2012. Relationships between susceptibility of Blackface sheep to *Teladorsagia circumcincta* infection and an inflammatory mucosal T cell response. Veterinary Research 43(26).

Grainger, J. R., Smith, K. A., Hewitson, J. P., McSorley, H. J., Harcus, Y., Filbey, K. J., Finney, C. A., Greenwood, E. J., Knox, D. P., Wilson, M. S., Belkaid, Y., Rudensky, A. Y. and Maizels, R. M. 2010. Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-beta pathway. Journal of Experimental Medicine 207(11): 2331-2341.

McNeilly, T. N., Rocchi, M., Bartley, Y., Brown, J. K., Frew, D., Longhi, C., McLean, L., McIntyre, J., Nisbet, A. J., Wattegedera, S., Huntley, J. F. and Matthews, J. B. 2013. Suppression of ovine lymphocyte activation by *Teladorsagia circumcincta* larval excretory-secretory products. Veterinary Research 44(70).

Murphy, K. and Weaver, C. 2016. Janeway's Immunobiology. 9th Edition. Garland Science.

Murphy, K. M. and Reiner, S. L. 2002. The lineage decisions of helper T cells. Nature Reviews Immunology 2: 933-944.

Nicol, L., Gossner, A., Watkins, C., Chianini, F., Dalziel, R. and 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. Veterinary Research 47(1): 27.

Pemberton, J. M., Beraldi, D., Craig, B. H. and Hopkins, J. 2011. Digital gene expression analysis of gastrointestinal helminth resistance in Scottish blackface lambs. Molecular Ecology 20(5): 910-919.

Rauen, T., Juang, Y. T., Hedrich, C. M., Kis-Toth, K. and Tsokos, G. C. 2012. A novel isoform of the orphan receptor RORgammat suppresses IL-17 production in human T cells. Genes and Immunity 13(4): 346-350.

Stear, M. J., Strain, S. and Bishop, S. C. 1999. Mechanisms underlying resistance to nematode infection. International Journal for Parasitology 29: 51-56.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., van Roy, N., De Paepe, A. and Speleman, F. 2002. Accurate normalisation of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology 3(7).

Wilkie, H., Gossner, A., Bishop, S. and 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. and Hopkins, J. 2016b. Mucosal expression of T cell gene variants is associated with differential resistance to *Teladorsagia circumcincta*. PLoS ONE: submitted.