

RESEARCH ARTICLE

Variations in T Cell Transcription Factor Sequence and Expression Associated with Resistance to the Sheep Nematode *Teladorsagia circumcincta*

Hazel Wilkie, Anton Gossner, Stephen Bishop[†], John Hopkins*

The Roslin Institute & R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, United Kingdom

[†] Deceased.

* john.hopkins@ed.ac.uk



OPEN ACCESS

Citation: Wilkie H, Gossner A, Bishop S, Hopkins J (2016) Variations in T Cell Transcription Factor Sequence and Expression Associated with Resistance to the Sheep Nematode *Teladorsagia circumcincta*. PLoS ONE 11(2): e0149644. doi:10.1371/journal.pone.0149644

Editor: Maxwell John Scott, North Carolina State University, UNITED STATES

Received: August 5, 2015

Accepted: February 3, 2016

Published: February 18, 2016

Copyright: © 2016 Wilkie et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This project was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Grant to The Roslin Institute and by a BBSRC Industrial CASE scholarship to JH and SB, for HW, sponsored by Sheep Improved Genetics Ltd and AHDB Beef and Lamb (Grant BB/J012653/1). The initial work was funded by the Natural Environment Research Council (NERC) through its 'Post-Genomics and Proteomics' programme (grant NE/

Abstract

This study used selected lambs that varied in their resistance to the gastrointestinal parasite *Teladorsagia circumcincta*. Infection over 12 weeks identified susceptible (high adult worm count, AWC; high fecal egg count, FEC; low body weight, BW; low IgA) and resistant sheep (no/low AWC and FEC, high BW and high IgA). Resistance is mediated largely by a Th2 response and IgA and IgE antibodies, and is a heritable characteristic. The polarization of T cells and the development of appropriate immune responses is controlled by the master regulators, T-bet (*TBX21*), GATA-3 (*GATA3*), ROR γ t (*RORC2*) and ROR α (*RORA*); and several inflammatory diseases of humans and mice are associated with allelic or transcript variants of these transcription factors. This study tested the hypothesis that resistance of sheep to *T. circumcincta* is associated with variations in the structure, sequence or expression levels of individual master regulator transcripts. We have identified and sequenced one variant of sheep *TBX21*, two variants of *GATA3* and *RORC2* and five variants of *RORA* from lymph node mRNA. Relative RT-qPCR analysis showed that *TBX21*, *GATA3* and *RORC2* were not significantly differentially-expressed between the nine most resistant (AWC, 0; FEC, 0) and the nine most susceptible sheep (AWC, mean 6078; FEC, mean 350). Absolute RT-qPCR on all 45 animals identified *RORVv5* as being significantly differentially-expressed ($p = 0.038$) between resistant, intermediate and susceptible groups; *RORCv2* was not differentially-expressed ($p = 0.77$). Spearman's rank analysis showed that *RORAv5* transcript copy number was significantly negatively correlated with parameters of susceptibility, AWC and FEC; and was positively correlated with BW. *RORCv2* was not correlated with AWC, FEC or BW but was significantly negatively correlated with IgA antibody levels. This study identifies the full length *RORA* variant (*RORAv5*) as important in controlling the protective immune response to *T. circumcincta* infection in sheep.

D000645/1). No funding was obtained from Sheep Improved Genetics Ltd (SIG). Their support was essential to obtain the BBSRC CASE Scholarship, but all the funds came from either BBSRC or AHND Animal Health Beef and Lamb (originally called Eblex). SIG had no influence on the design of the experiments and have no rights to the data.

Competing Interests: This project was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Grant to The Roslin Institute and by a BBSRC Industrial CASE scholarship to JH and SB, for HW, supported by Sheep Improved Genetics Ltd (no funding was obtained from SIG Ltd) and AHDB Beef and Lamb (Grant BB/J012653/1). This does not alter the authors' adherence to PLOS ONE policies on sharing data. The initial work was funded by the Natural Environment Research Council (NERC) through its 'Post-Genomics and Proteomics' programme (grant NE/D000645/1).

Introduction

The abomasal strongylid *Teladorsagia circumcincta* is a major cause of sheep parasitic gastroenteritis [1, 2]. The most susceptible animals are weaned lambs [3], but many eventually suppress larval development and egg production [4] through the development of IgA and IgE anti-parasite antibodies [5–7]. The ability to control infection is a heritable characteristic and most flocks consist of animals with a range of susceptibilities. Indeed, IgA levels and fecal egg counts (FEC) have been used as selectable markers for resistance [4, 8, 9] and antigens that promote the production of abomasal IgA antibodies have been identified as potential vaccine candidates [10].

The candidate gene approach for the identification of molecular markers for selection aims to evaluate the relationship between phenotype and a variation in a gene [11]. Several studies have analysed abomasal mucosa to identify genes associated with resistance to *T. circumcincta* or the related parasite *Haemonchus contortus* [12–14]. However, the induction of the immune response to these parasites occurs in the draining abomasal (gastric) lymph node (ALN) and the events within that node are likely to determine the quality and quantity of the response that occurs within the mucosa and the consequent clinical outcome.

This current study exploited parasite-naïve Blackface lambs with diversity in their predicted genetic resistance to *T. circumcincta* [15]. Both immunological [16] and microarray analyses [17] of ALN linked Th2 responses to high IgA levels, low FEC and resistance, and also showed that Th1/Th17 T cell activation in susceptible sheep resulted in granulomatous inflammation and low antibody levels that failed to control infection (high AWC and FEC). In mouse and human gastrointestinal nematode infections, resistance is determined by Th2 activation [18] associated with a balanced Th1/Th2/Treg response [19]. Uncontrolled Th1 and/or Th17 activation leads to clinical disease [20]. Consequently, the clinical outcome of infection is mediated by differential T cell activation.

The multiple effector functions of CD4 T cells are achieved by the differentiation of multi-potential precursors into distinct polarized subsets, which is largely regulated by the master regulators T-bet, GATA-3 and ROR γ t; transcription factors that transactivate the genes and mediate the subset-specific functions [21]. T-bet (*TBX21*) promotes Th1 differentiation by transactivating *IFNG* and increasing IFN γ production, and by repressing Th2 activation [22, 23]. GATA-3 (*GATA3*) regulates Th2 development [24] by direct binding to enhancers and/or promoters of the Th2 cytokine gene cassette (*IL4*, *IL5* and *IL13*) and inducing their expression; it also inhibits Th1 development [25]. Th17 cells express neither T-bet nor GATA-3 [26]; antigen-activated naïve T cells in the presence of TGF β and IL-6 differentiate into IL-17A producing T cells regulated by ROR γ t (*RORC2*) [27]. Maximum production of IL-17A also requires the related transcription factor ROR α (*RORA*) [28]. ROR α is also critical for the development and function of the nuocyte [29] subset of innate lymphoid cells 2 (ILC2), which are crucial in the induction of Th2 cells and the development of anti-helminth immunity [30] in mice. Tregs have also been shown to play a role in anti-helminth immunity in mice [19], however previous work in *T. circumcincta*-infected sheep [16] has shown that there is no differential expression of the Treg transcription factor *FOXP3* in resistant and susceptible sheep.

In addition to controlling T cell differentiation, all four transcription factors contribute to the pathogenesis of chronic inflammatory diseases. T-bet plays a role in the abnormal expression of Th1 cytokines in human Crohn's disease [31] and GATA-3 is prominent in the development of ulcerative colitis [32]. Furthermore, *GATA3* gene variants and deletion mutants have been linked to a number of other inflammatory pathologies, including asthma and IgE-mediated allergy [33, 34]. The major function of Th17 cells is in the development of inflammatory reactions, and many inflammatory diseases have been ascribed to increased Th17 activity

[35]. Consequently, ROR γ t and ROR α have also been linked to abnormal inflammation [36, 37].

In this study we characterise the different transcript variants of the four master regulators expressed in sheep and then compare the expression of these individual variants in animals of defined resistance status. Finally we quantify expression levels of variants to enable their correlation with quantitative phenotypes of resistance to *T. circumcincta*.

Materials and Methods

Animals and experimental design

Female Blackface lambs (10–13 weeks old), from a flock previously used for quantitative genetic and QTL analyses [38], were housed in worm-free conditions. Ten lambs were sham-infected controls and 45 lambs were infected with ~2300 infective L3 *T. circumcincta* larvae three times a week for 12 weeks. At post mortem, two days after the last infection, the abomasal AWC ranged from 0 to 11300 and FEC from 0–950 eggs per g (S1 Table). The lambs selected for analysis were chosen to maximize the power of detecting differential expression. Consequently, animals were ranked (1–45) according to their infection level [15]. Full details of the animals, animal husbandry, infection protocols, quantitative phenotypes and population genetic analyses have been described previously [15, 16]. Animal experiments were approved by University of Edinburgh Ethical Review Committee and conducted under an Animals (Scientific Procedures) Act 1986 Project Licence.

Animals were housed in a large open barn in three pens; infected animals were in two adjacent pens of approximately 180 m² each and the uninfected lambs were in a pen of approximately 50 m². All animals were bedded on clean straw and had *ad libitum* access to hay and water supplemented twice daily with Maize Lamb Pellets (16.0% protein; Carrs Billington, Carlisle, UK). Animals were examined at least daily. All lambs were vaccinated with Heptavac P Plus at 5 and 6 weeks and Scabivax (both MSD Animal Health, UK) at 6 weeks. Most lambs were administered 2 ml Hexasol (Norbrook Pharmaceuticals, UK) to treat respiratory infections. Infected animals showing mild symptoms of visceral pain associated with parasite infection were treated, under veterinary instructions, with 2 ml Finadyne (MSD Animal Health). Lambs were killed by intravenous administration of Euthetal (Merial Animal Health, UK).

Sample collection and total RNA isolation

Abomasal (gastric) lymph nodes (ALN) were removed immediately post mortem and stored at –80°C in RNAlater (Ambion, UK). Total RNA was isolated using the Ribopure Kit (Ambion) according to the manufacturers' instructions. Contaminating DNA was removed by On-column PureLink[®] DNase I treatment (Ambion). The quantity, quality and integrity of the RNA samples were determined using a NanoDrop ND-1000 spectrophotometer and Agilent 2200 TapeStation system; all had an RNA Integrity Number of > 7.5.

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 primers used to amplify overlapping sections of genes are shown in S2 Table. 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 BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, UK). The GeneRacer™ 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 (S2 Table) were designed 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.

RT-qPCR quantification of transcript variants

Primers for quantitative real-time RT-PCR (RT-qPCR) were designed and optimized for all gene variants (S2 Table); primer specificity was confirmed by sequencing. The *RORC2* full length RT-qPCR primers did not distinguish between *RORC2* and *RORC1*, however *RORC1* is not expressed in lymph node but is expressed in liver (S1 Fig). RT-qPCR was performed in 15 µl volumes containing 7.5 µl FastStart Universal SYBR Green Master (Rox) 2x concentration (Roche), 2 µl template cDNA, 0.25–1.0 µl primers (S2 Table) and nuclease-free water. Reactions were prepared using a CAS-1200™ robot and performed on a Rotor-Gene 6 (Qiagen). Amplification (conditions in S2C Table) was followed by dissociation curve analysis. Relative expression levels were quantified in cDNA in triplicate from two separate RT-qPCR reactions for each of the nine most resistant (ranked 1–9) and the nine most susceptible sheep (ranked 37–45), and duplicate no-template controls were included in all runs. Optimized RT-qPCR assays had an efficiency between 95–105% and R² value of 0.98–0.99.

Relative gene expression levels were calculated in GenEx 5 Standard Programme (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 of *RORA*v2 and v5 in all 45 infected animals, a standard curve of linearized plasmid DNA was used with a dynamic range that spanned at least five orders of magnitude. For each point on the standard curve, copy numbers were calculated from Cq values using the following formula:

$$\text{molecules per ng} = ([1 \times 10^{-9}] / (M \text{ g/mol})) \times [6.03 \times 10^{-23} \text{ molecules/mol}]$$

M = size of plasmid 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 for each individual sample. The normalization factor was calculated using the method described by Vandesompele [39], 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.

Statistical analysis

Relative gene expression levels were analyzed in GenEx using an unpaired, 2-tailed t-test to determine the difference between groups. Graph Pad Prism 6 for Windows (Graph Pad

Software, USA) was used for statistical analysis of the absolute expression data. The data were grouped into resistant, intermediate and susceptible ($n = 15$ per group) and a one-way ANOVA was performed to determine overall significance and Tukey's multiple comparisons test (within ANOVA) was used to determine significance between groups. The correlations between transcript levels and quantitative phenotypes were analyzed with Spearman's correlation coefficient (r_s). P-values ≤ 0.05 were considered statistically significant.

Results

Identification of transcription variants of sheep transcription factors

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 (S2 Fig). *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), which represents the deletion of 12 amino acids (GKYGGVELFRAL) at position 359–370 (S3 Fig) of ROR γ t. 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. All five variants are identical from position 46 of *RORAv1* (Chr7: 46,691,081), which represents the 5' end of exon 6 (S4 Fig). The five variants have variable usage of exons 1 to 6 and consequently 5' UTRs of different lengths (S5 Fig). *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 (Fig 1). The derived protein sequences of all five variants are identical 3' from amino acid 83 of ROR α v1.

RT-qPCR quantification of transcription factor expression

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 (S1 Table). 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 1 shows the relative expression (fold change) of each transcription factor and variant between the resistant and susceptible groups and shows that all but *RORCv5* were equally expressed in the ALN of both groups. *RORAv5* was 1.57 fold higher in the resistant animals but $p = 0.08$. *RORCv3* could not be quantified because expression levels were too low for accurate measurement.

Correlation of transcription factor levels and quantitative phenotypes

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; *RORAv2* was also quantified as the control *RORA* variant. *RORAv2* expression in all 45 animals

RORαv1	-----MMYFVIAAMKAQIEII	16
RORαv2	-----MMYFVIAAMKAQIEII	0
RORαv3	-----MPV-----SSEGTERRRHDPGISVTKKTHTSQIEII	31
RORαv5	MYGVFRQAGFLWMGSASGRPGEKTKTSVSCSSLLQLLTESHSTLGISVTKKTHTSQIEII	60
RORαv1	<u>PCKICGDKSSGIHYGVITCEGCKGFFRRSQSNATYSCPRQKNCLIDRTSRNRCQHCRLO</u>	76
RORαv2	-----PCKICGDKSSGIHYGVITCEGCKGFFRRSQSNATYSCPRQKNCLIDRTSRNRCQHCRLO	0
RORαv3	<u>PCKICGDKSSGIHYGVITCEGCKGFFRRSQSNATYSCPRQKNCLIDRASRNRCQHCRLO</u>	91
RORαv5	<u>PCKICGDKSSGIHYGVITCEGCKGFFRRSQSNATYSCPRQKNCLIDRTSRNRCQHCRLO</u>	120
RORαv1	KCLAVGMS RD AVK FG RM S KK QR DSL YAEVQK HRM QQ Q RD H QQ Q PGEAEPLTPTYSISAN	136
RORαv2	-----MSRD AVK FG RM S KK QR DSL YAEVQK HRM QQ Q RD H QQ Q PGEAEPLTPTYSISAN	54
RORαv3	KCLAVGMS RD AVK FG RM S KK QR DSL YAEVQK HRM QQ Q RD H QQ Q PGEAEPLTPTYSISAN	151
RORαv5	KCLAVGMS RD AVK FG RM S KK QR DSL YAEVQK HRM QQ Q RD H QQ Q PGEAEPLTPTYSISAN	180

Fig 1. Ovis aries RORA transcript variants. Derived NH₂-protein sequences of RORαv1, v2, v3 and v5. Bold is the zinc-finger DNA binding domain; underlined are B domains. The derived protein sequences of RORαv2 and RORαv4 are identical. Sequences are identical after amino acid 1 (M) of RORαv2.

doi:10.1371/journal.pone.0149644.g001

was 162938 ± 14799 (transcript copy number per µg RNA ± SD), in comparison to 108806 ± 11211 (p = 0.0045) for RORCv5.

Comparison of RORAv2 and RORAv5 expression in 15 most resistant (rank 1–15, mean AWC 59 and FEC 1.7), 15 intermediate (rank 16–30, AWC 1508 and FEC 87) and the 15 most susceptible (rank 31–45, AWC 5167 and FEC 288) sheep (Fig 2) shows that RORAv2 is expressed equally in all three infected groups (164411 ± 80305, 148781 ± 139319 and 175622 ± 68746), with one-way ANOVA p = 0.77. In contrast RORAv5 expression levels were significantly different between the infected groups (p = 0.038); mean levels in the resistant sheep were 148439 ± 91397 in comparison with the intermediate (93161 ± 44477, p ≤ 0.05) and the susceptible group (84818 ± 69870, p ≤ 0.05).

Expression levels of RORAv2 and RORAv5 in the uninfected controls were 14336 ± 12927 and 11308 ± 6571 respectively and not significantly different (p = 0.52). RORAv2 was significantly up-regulated 9–11 fold in all three infected groups (p < 0.0003) in comparison with the uninfected controls, with no significant discrimination between resistant, intermediate and susceptible animals. In contrast RORAv5 was significantly increased 13 fold (p < 0.0001) in the resistant sheep but only 7 (p ≤ 0.005) and 6 (p ≤ 0.005) fold in the intermediate and susceptible sheep, in comparison to the controls.

Table 1. Relative quantification of transcription factor transcripts in ALN.

Gene	Fold change R vs. S	p value
TBX21	-1.10	0.61
GATA3	1.06	0.62
GATA3v1	-1.12	0.47
RORC2	1.18	0.42
RORC2v1	1.12	0.54
RORAv1	-1.13	0.79
RORAv2	-1.07	0.83
RORAv4*	-1.04	0.88
RORAv5	1.57	0.08

* RORAv3 signal was too low for accurate measurement.

doi:10.1371/journal.pone.0149644.t001

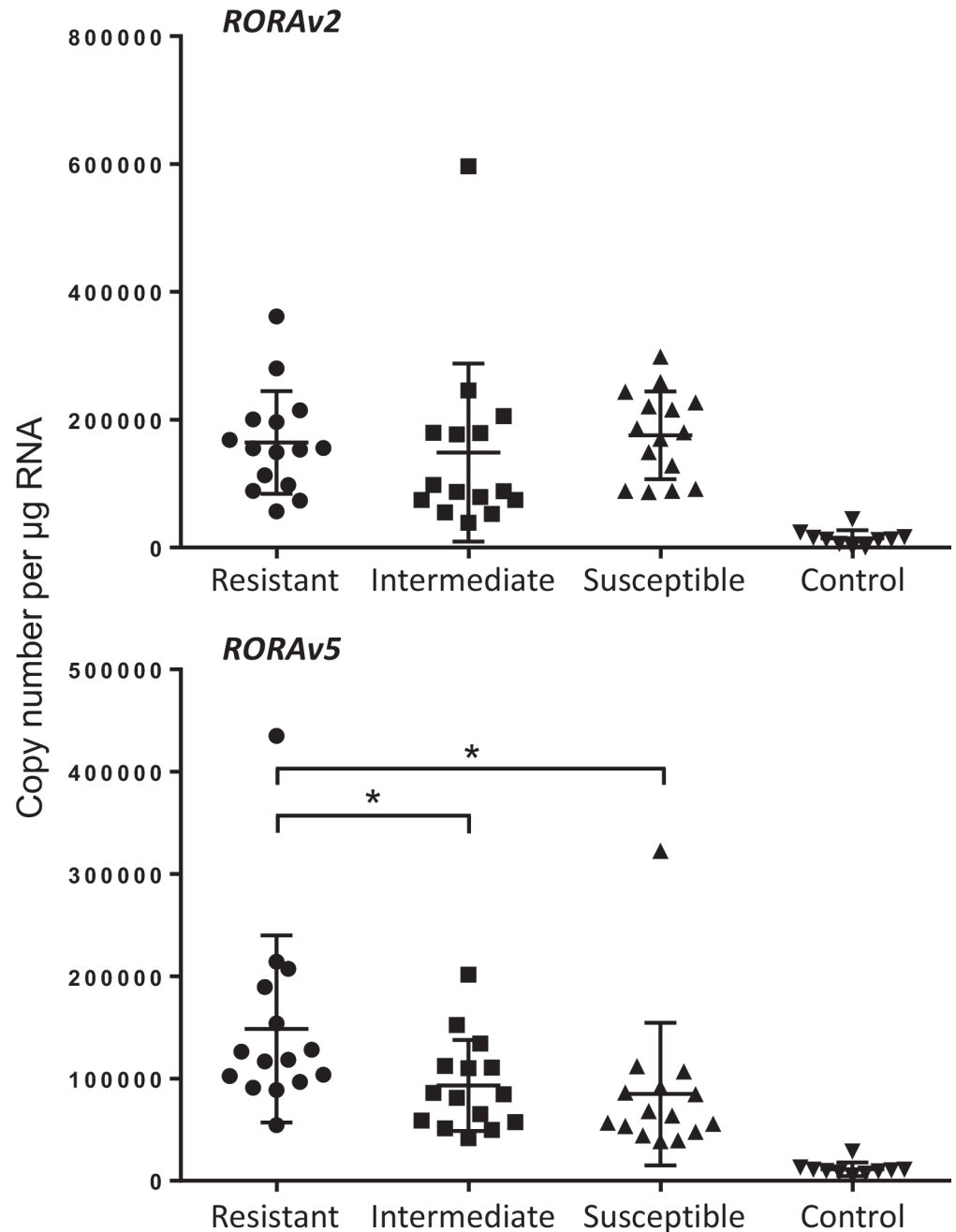


Fig 2. Expression of *RORAv2* and *RORAv5* in the ALN of *T. circumcincta* infected sheep. Copy number per µg total RNA in rank 1–15 resistant sheep (AWC 0–300, mean 59; FEC 0–25, mean 1.7); rank 16–30 intermediate (AWC 200–3100, mean 1508; FEC 0–475, mean 87) and rank 31–45 susceptible (AWC 2900–11300, mean 5167; FEC 75–950, mean 288) sheep. Error bars are means ± SD. One-way ANOVA for *RORAv2* $p = 0.77$, infected animals only ($p < 0.0003$ with controls); *RORAv5* $p = 0.038$, infected animals only ($p < 0.0001$ with controls); * $p \leq 0.05$ (Tukey's multiple comparison test within ANOVA). $P \leq 0.05$ for all three infected groups vs. uninfected controls for both *RORAv2* and *RORAv5*.

doi:10.1371/journal.pone.0149644.g002

Spearman's rank analysis was used to quantify the correlation of expression levels of *RORAv2* and *RORAv5* in relation to the quantitative phenotypes, AWC, FEC, BW and IgA antibody levels (Fig 3). *RORAv2* expression levels were not significantly correlated with AWC,

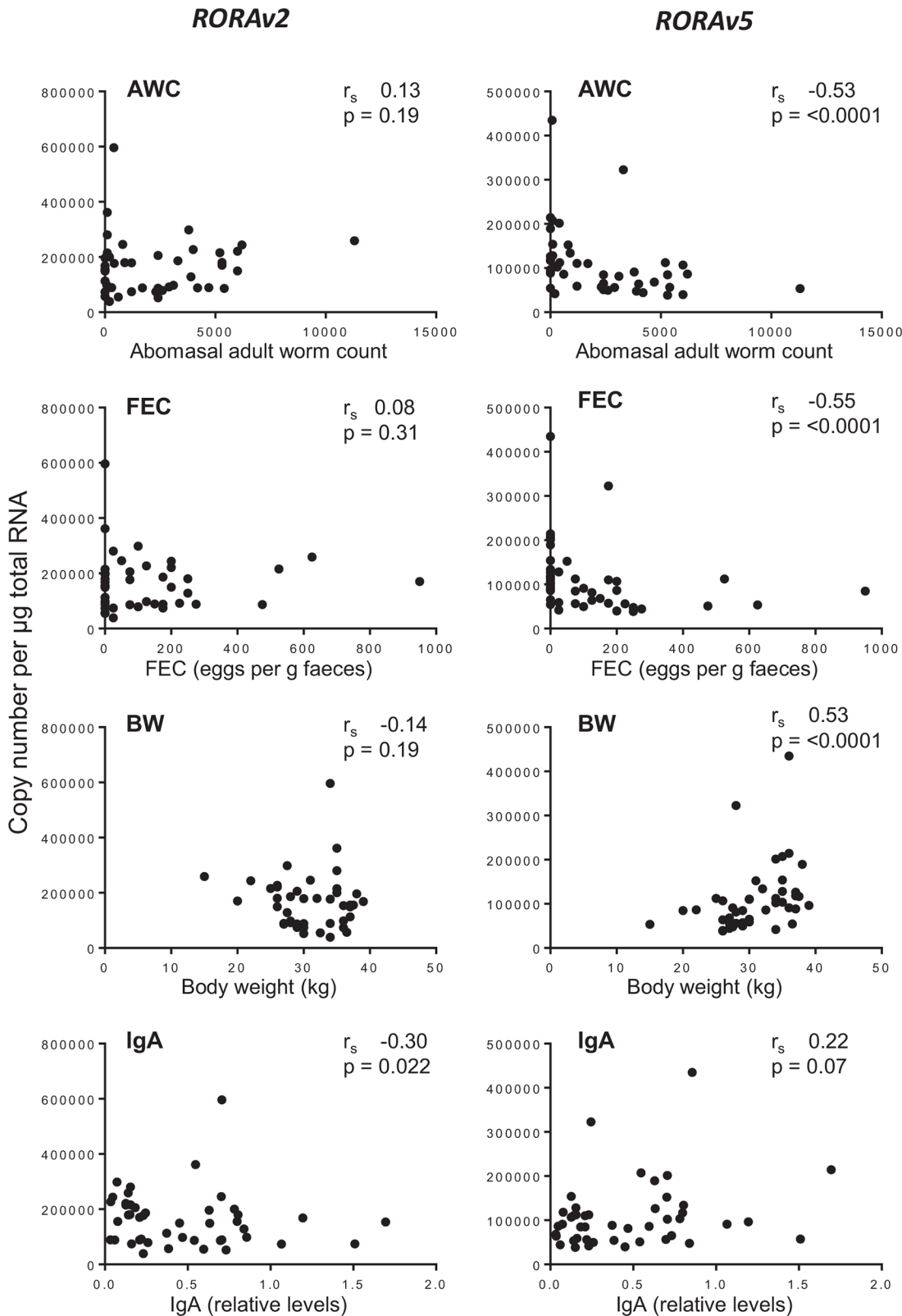


Fig 3. Correlation analysis of the phenotypic parameters and RORA transcript variants. Correlation of AWC, FEC, BW and IgA with *RORAv2* and *RORAv5* copy number per μg total RNA in ALN of *T. circumcincta* infected sheep. r_s —Spearman's rank correlation coefficient.

doi:10.1371/journal.pone.0149644.g003

FEC or BW but were significantly negatively correlated with IgA (r_s -0.30, p = 0.022). In contrast *RORCv5* was significantly negatively correlated with AWC (r_s -0.53, p = <0.0001) and FEC (r_s -0.55, p = <0.0001), and significantly positively correlated with BW (r_s 0.53, p = <0.0001); but was not significantly correlated with IgA antibody (r_s 0.22, p = 0.07).

Discussion

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. Initial characterization of mRNA from abomasal lymph node failed to find any splice variants of sheep *TBX21*, identified two transcript variants each of *GATA3* and *RORC2*, and five transcript variants of *RORA*. *TBX21* has been unambiguously identified in relatively few species, including humans, cattle and mice, and in all these species only a single gene has been characterized.

The sequence difference between the two ovine *GATA3* variants is also found in the two variants of human *GATA3* with no reported phenotypic consequence. The deleted aa260 is not within either of the two conserved zinc-finger domains (aa264–286 and aa318–342) nor the functionally-critical DNA-binding YxKxHxxxRP motif (YYKLHNINRP) at aa345–354 [40] although it is close to the NH₂-zinc finger and could possibly affect protein structure. Only one *Bos taurus* *GATA3* sequence has been identified (NM_001076804.1) that encodes an identical protein (isoform X2) to full length sheep *GATA3*. There are also four predicted *Bos* transcript variants, X1, X2 and X3 encode isoform X2, and variant X4 encodes a protein (isoform X1) identical to sheep *GATA3v1*.

Humans express two transcript variants of *RORC*; transcript variant 1 encodes full length *RORC* (isoform a or ROR γ) of 518 amino acids with major functions in regulating embryo development as well as testicular and liver functions [41]. Transcript variant 2 (isoform b or ROR γ t) is 485 amino acids and has a major role in regulating Th17 T cells [27]. The two sheep *RORC2* variants are homologous to human transcript variant 2, and the 12 amino acid deletion in *RORC2v1* is within the predicted ligand binding domain [42]. One major *RORC2* variant also exists in humans; this has a deletion of exons 5–8, also within the ligand binding domain, and functions to suppress Th17 function by the repression of *IL17A* and *IL21* gene transcription [43].

Previous studies, using RT-qPCR of cytokine transcripts and whole genome microarrays, have shown that a strong Th2 response is associated with resistance and Th1/Th17 activation linked to susceptibility [16, 17]. However, comparison of resistant and susceptible animals showed that there was no differential expression of any of the three master regulators (or variants) that control T cell polarization, *TBX21*, *GATA3* or *RORC2*. A possible explanation for this is that the lymph nodes studied had been responding to *T. circumcincta* infection for 12 weeks and that the regulatory events controlling T cell polarization occurred much earlier in infection; during initial activation of T helper cells.

The identification of five *RORA* transcript variants in sheep is not an unusual finding, as several species express multiple *RORA* variants, including four in humans. All four human isoforms have identical zinc-finger DNA and ligand-binding domains but differ in their NH₂-terminal sequences (A/B domains [44]) and it seems that these differences are associated with cellular expression levels [45] and determine functions [46, 47]. Indeed in mice, NH₂-terminal truncated ROR α isoforms are linked to cerebellar cell atrophy and the ‘staggerer’ phenotype [48]. The five sheep *RORA* variants also encode four distinct proteins. The derived protein sequences of *RORAv2* and *RORAv4* are identical; they possess the conserved ligand-binding domain but lack A/B domains and the DNA-binding domain. The *RORAv1* protein (ROR α v1) contains both DNA- and ligand-binding domains and a 17 amino acid B domain immediately

upstream of the DNA-binding domain, which is identical to that of human Rora isoform d (NP_599024.1). ROR α 3 and ROR α 5 possess both DNA- and ligand-binding domains and identical B domains, which is the same sequence as that of human isoform a, and distinct from that of ROR α 1. ROR α 3 possesses an upstream A domain of 13 amino acids and ROR α 5 has an A domain of 44 amino acids, different to any human isoform.

Of the four measurable variants of *RORA* only *RORA*v5 showed consistent differential expression in the resistant vs susceptible comparison and quantitative levels of *RORA*v5 were highly significantly correlated with quantitative parameters of resistance. This is likely to be related to the length and identity of the A/B domains. Human Rora isoforms a and b have distinct A/B domains and display different binding specificities. The A/B domains of isoform a mediate almost complete nuclear localization [44] making it a strong transcription factor that activates a broad range of genes. In contrast the A/B domains of isoforms b, c and d mediate cytoplasmic localization and consequently they are weaker transcription factors with a narrower range of targets [46].

This does not fully explain the link between sheep *RORA*v5 and parasite resistance. Analysis of the sheep immune response to *T. circumcincta* shows that resistance is correlated with the development of a Th2 response, including the production of IL-13 and up-regulated transcription of genes associated with wound healing [17]. In humans a major function of Rora isoform a (most similar to ROR α 5) is the inhibition of NF- κ B associated inflammation [37], and it is clear from previous studies in sheep that resistant animals are those with repressed inflammatory gene expression [16, 17]. ROR α is also an essential transcription factor for nuocyte development in mice; these mediate Th2 polarization [29], optimize tissue repair [49] and consequently have a critical role in parasite expulsion. However, if the link between ROR α v5 and parasite resistance is due to this ILC2 subset in sheep it does not involve the GATA-3⁺ ILC2 cells [49, 50] as this transcription factor is not differentially-expressed in parasite-resistant sheep. It is also known that ROR α is an important accessory transcription factor for the development and function of Th17 cells [28], but the master regulator of these cells is ROR γ t which, like GATA-3, is not differentially-expressed; furthermore Th17 activation seems to be more associated with susceptibility than resistance [16].

Analysis of *RORA*v2 expression showed no correlation with AWC, FEC or BW and a low but significant negative correlation with IgA, i.e. a positive correlation with susceptibility. ROR α v2 possesses neither the A/B nor DNA-binding domains, but does have the ligand-binding domain. Both *RORA*v2 and *RORA*v5 are expressed at high levels only after infection. however *RORA*v5 showed a 13 fold up-regulation in resistant animals, but only a 7 and 6 fold increase in the intermediate and susceptible groups. In contrast, the infection-associated increase in *RORA*v2 expression was similar in all three groups. It is possible that ROR α v2, which is found at higher levels than ROR α v5 in all infected lambs, could compete with ROR α v5 for ligand binding and inhibit its function in those animals with lower levels of *RORA*v5, with consequent pathological effects in susceptible lambs.

In conclusion, we have identified transcript variants of the master regulators of T cell immunity in sheep. There are two variants of *GATA3* and *RORC2*, five variants of *RORA* and a single *TBX21* transcript. Analysis of expression of each variant in sheep with defined resistance status to the gastrointestinal parasite *T. circumcincta* showed that only *RORA*v5 is differentially expressed in resistant compared to susceptible sheep, with expression levels that correlate with the phenotypic parameters of resistance. The largest variant of *RORA* (*RORA*v5) seems to be important in controlling the protective immune response to *T. circumcincta* infection in sheep.

Supporting Information

S1 Fig. Agarose gel electrophoresis of sheep *RORC1* and *RORC2*. RT-PCR products using: (A) *RORC1* primers, Lanes 1 and 2 mRNA from liver; Lane 3 mRNA from liver (no RT control); Lane 4 mRNA from lymph node; (B) *RORC2* primers, Lane 1 mRNA from lymph node (no RT control); Lanes 2 and 3 mRNA from lymph node.

(PDF)

S2 Fig. *Ovis aries* *GATA3* transcript variants. (A) 5' nucleotide sequences (LN848231 and LN848232). (B) derived NH₂-protein sequences (CRI68167.1 and CRI68168.1).

(PDF)

S3 Fig. *Ovis aries* *RORC2* transcript variants. (A) 5' nucleotide (LN848233 and LN848234). (B) derived NH₂-protein sequences (CRI68169.1 and CRI68170.1).

(PDF)

S4 Fig. *Ovis aries* *RORA* nucleotide sequences. 5' nucleotide sequences of *RORA* transcript variants (LN848235, LN848236, LN848237, LN848238, LN848239).

(PDF)

S5 Fig. *Ovis aries* *RORA* transcript variants. *RORAv1*, *v2*, *v3*, *v4* and *v5* intron and exon structures, mapped to Oarv3.1 genome assembly.

(PDF)

S1 Table. Quantitative phenotypic data and normalized copy numbers of *RORAv2* and *RORAv5*, in ALN of Scottish Blackface lambs persistently infected with *T. circumcincta*.

(PDF)

S2 Table. (A) Primer sets for amplification of sheep transcription factors. (B) Primer sequences for 5' and 3' RACE. (C) Primer sets for RT-qPCR with optimized parameters and volumes.

(PDF)

Acknowledgments

We thank Professor Josephine Pemberton (Institute of Evolutionary Biology, University of Edinburgh) and Dr Dario Beraldi (now at CRUK, Li Ka Shing Centre, Cambridge) for original collaboration; Joan Docherty (The Marshall Building) for animal husbandry and collection of weight data and blood samples; Emily Clark (Roslin Institute) for ovine liver RNA.

Author Contributions

Conceived and designed the experiments: JH SB AG. Performed the experiments: HW JH. Analyzed the data: HW AG JH. Contributed reagents/materials/analysis tools: JH AG. Wrote the paper: JH AG HW. Played a major role in obtaining funding: SB.

References

1. Urquhart GM, Armour J, Duncan JL, Dunn AM, Jennings FW. *Veterinary Parasitology*: Avon: Longman Scientific and Technical; 1987 1987.
2. Brunson RV. The economic impact of nematode infection in sheep: implications for future research and control. In: Heath ACG, editor. *The economic importance of parasites of livestock in New Zealand*. Miscellaneous Publication No. 1: New Zealand Society for Parasitology; 1988. p. 4–16.
3. Stear MJ, Strain S, Bishop SC. Mechanisms underlying resistance to nematode infection. *Int J Parasitol.* 1999; 29(1):51–6. PMID: [10048819](https://pubmed.ncbi.nlm.nih.gov/10048819/)

4. Stear MJ, Strain S, Bishop SC. How lambs control infection with *Ostertagia circumcincta*. *Vet Immunol Immunopathol.* 1999; 72(1–2):213–8. PMID: [10614511](#)
5. Smith WD, Jackson F, Jackson E, Graham R, Williams J, Willadsen SM, et al. Transfer of immunity to *Ostertagia circumcincta* and IgA memory between identical sheep by lymphocytes collected from gastric lymph. *Res Vet Sci.* 1986; 41(3):300–6. PMID: [3809720](#)
6. Halliday AM, McAllister HC, Smith WD. Kinetics of the local immune response in the gastric lymph of lambs after primary and challenge infection with *Teladorsagia circumcincta*. *Parasite Immunol.* 2010; 32(2):81–90. doi: [10.1111/j.1365-3024.2009.01173.x](#) PMID: [20070822](#)
7. Murphy L, Eckersall PD, Bishop SC, Pettit JJ, Huntley JF, Burchmore R, et al. Genetic variation among lambs in peripheral IgE activity against the larval stages of *Teladorsagia circumcincta*. *Parasitology.* 2010; 137(8):1249–60. doi: [10.1017/S0031182010000028](#) PMID: [20233490](#)
8. Shaw R, Morris C, Wheeler M, Tate M, Sutherland I. Salivary IgA: A suitable measure of immunity to gastrointestinal nematodes in sheep. *Vet Parasitol.* 2012; 186(1–2):109–17. doi: [10.1016/j.vetpar.2011.11.051](#) PMID: [22153121](#)
9. Sayers G, Sweeney T. Gastrointestinal nematode infection in sheep—a review of the alternatives to anthelmintics in parasite control. *Anim Health Res Rev.* 2005; 6(2):159–71. PMID: [16583780](#)
10. Nisbet AJ, Knox DP, McNair CM, Meikle LI, Smith SK, Wildblood LA, et al. Immune recognition of the surface associated antigen, Tc-SAA-1, from infective larvae of *Teladorsagia circumcincta*. *Parasite Immunol.* 2009; 31(1):32–40. doi: [10.1111/j.1365-3024.2008.01070.x](#) PMID: [19121081](#).
11. Pemberton JM, Beraldi D, Craig BH, Hopkins J. Digital gene expression analysis of gastrointestinal helminth resistance in Scottish blackface lambs. *Mol Ecol.* 2011; 20(5):910–9. doi: [10.1111/j.1365-294X.2010.04992.x](#) PMID: [21324010](#)
12. Diez-Tascon C, Keane OM, Wilson T, Zadissa A, Hyndman DL, Baird DB, et al. Microarray analysis of selection lines from outbred populations to identify genes involved with nematode parasite resistance in sheep. *Physiol Genomics.* 2005; 21(1):59–69. doi: [10.1152/physiolgenomics.00257.2004](#) PMID: [15623564](#)
13. Rowe A, Gondro C, Emery D, Sangster N. Genomic analyses of *Haemonchus contortus* infection in sheep: abomasal fistulation and two *Haemonchus* strains do not substantially confound host gene expression in microarrays. *Vet Parasitol.* 2008; 154(1–2):71–81. doi: [10.1016/j.vetpar.2008.02.014](#) PMID: [18387746](#)
14. Knight PA, Griffith SE, Pemberton AD, Pate JM, Guarneri L, Anderson K, et al. Novel gene expression responses in the ovine abomasal mucosa to infection with the gastric nematode *Teladorsagia circumcincta*. *Vet Res.* 2011; 42(1):78.
15. Beraldi D, Craig BH, Bishop SC, Hopkins J, Pemberton JM. Phenotypic analysis of host-parasite interactions in lambs infected with *Teladorsagia circumcincta*. *Int J Parasitol.* 2008; 38:1567–77. doi: [10.1016/j.ijpara.2008.04.011](#) PMID: [18589425](#)
16. Gossner AG, Venturina VM, Shaw DJ, Pemberton JM, Hopkins J. Relationship between susceptibility of Blackface sheep to *Teladorsagia circumcincta* infection and an inflammatory mucosal T cell response. *Vet Res.* 2012; 43(1):26. doi: [10.1186/1297-9716-43-26](#)
17. Gossner A, Wilkie H, Joshi A, Hopkins J. Exploring the abomasal lymph node transcriptome for genes associated with resistance to the sheep nematode *Teladorsagia circumcincta*. *Vet Res.* 2013; 44(1):68. doi: [10.1186/1297-9716-44-68](#)
18. Anthony RM, Rutitzky LI, Urban JF, Staderker MJ, Gause WC. Protective immune mechanisms in helminth infection. *Nat Rev Immunol.* 2007; 7(12):975–87. PMID: [18007680](#)
19. Maizels RM, Yazdanbakhsh M. Immune regulation by helminth parasites: cellular and molecular mechanisms. *Nat Rev Immunol.* 2003; 3(9):733–44. PMID: [12949497](#)
20. Maizels RM, Pearce EJ, Artis D, Yazdanbakhsh M, Wynn TA. Regulation of pathogenesis and immunity in helminth infections. *J Exp Med.* 2009; 206(10):2059–66. doi: [10.1084/jem.20091903](#) PMID: [19770272](#)
21. Zhu J, Yamane H, Paul WE. Differentiation of effector CD4 T cell populations. *Annu Rev Immunol.* 2010; 28:445–89. doi: [10.1146/annurev-immunol-030409-101212](#) PMID: [20192806](#)
22. Lazarevic V, Glimcher LH, Lord GM, Yu F, Sharma S, Edwards J, et al. T-bet: a bridge between innate and adaptive immunity. *Nat Rev Immunol.* 2013; 13(11):777–89. doi: [10.1038/nri3536](#) PMID: [24113868](#)
23. Hwang ES, Szabo SJ, Schwartzberg PL, Glimcher LH. T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. *Science.* 2005; 307(5708):430–3. doi: [10.2307/3840063](#) PMID: [15662016](#)
24. Zheng W, Flavell RA. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell.* 1997; 89(4):587–96. PMID: [9160750](#).

25. Zhu JF, Yamane H, Cote-Sierra J, Guo LY, Paul WE. GATA-3 promotes Th2 responses through three different mechanisms: induction of Th2 cytokine production, selective growth of Th2 cells and inhibition of Th1 cell-specific factors. *Cell Research*. 2006; 16(1):3–10. PMID: [16467870](#)
26. Weaver CT, Hatton RD, Mangan PR, Harrington LE. IL-17 family cytokines and the expanding diversity of effector T cell lineages. *Annu Rev Immunol*. 2007; 25:821–52. PMID: [17201677](#)
27. Ivanov II, McKenzie BS, Zhou L, Tadokoro CE, Lepelley A, Lafaille JJ, et al. The orphan nuclear receptor ROR γ directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell*. 2006; 126(6):1121–33. PMID: [16990136](#).
28. Yang XO, Pappu BP, Nurieva R, Akimzhanov A, Kang HS, Chung Y, et al. T helper 17 lineage differentiation is programmed by orphan nuclear receptors ROR α and ROR γ . *Immunity*. 2008; 28(1):29–39. doi: <http://dx.doi.org/10.1016/j.immuni.2007.11.016> PMID: [18164222](#)
29. Wong SH, Walker JA, Jolin HE, Drynan LF, Hams E, Camelo A, et al. Transcription factor ROR alpha is critical for nuocyte development. *Nat Immunol*. 2012; 13(3):229–U43. doi: [10.1038/ni.2208](#) PMID: [22267218](#)
30. Neill DR, Wong SH, Bellosi A, Flynn RJ, Daly M, Langford TK, et al. Nuocytes represent a new innate effector leukocyte that mediates type-2 immunity. *Nature*. 2010; 464(7293):1367–70. doi: [10.1038/nature08900](#) PMID: [20200518](#)
31. Lazarevic V, Glimcher LH. T-bet in disease. *Nat Immunol*. 2011; 12(7):597–606. doi: [10.1038/ni.2059](#) PMID: [21685955](#)
32. Ohtani K, Ohtsuka Y, Ikuse T, Baba Y, Yamakawa Y, Aoyagi Y, et al. Increased mucosal expression of GATA-3 and STAT-4 in pediatric ulcerative colitis. *Pediatrics International*. 2010; 52(4):584–9. doi: [10.1111/j.1442-200X.2009.03019.x](#) PMID: [20030749](#)
33. Huebner M, Kim D-Y, Ewart S, Karmaus W, Sadeghnejad A, Arshad SH. Patterns of GATA3 and IL13 gene polymorphisms associated with childhood rhinitis and atopy in a birth cohort. *J Allergy Clin Immunol*. 2008; 121(2):408–14. doi: [10.1016/j.jaci.2007.09.020](#) PubMed PMID: [174547148](#).
34. Pykalainen M, Kinos R, Valkonen S, Rydman P, Kilpelainen M, Laitinen LA, et al. Association analysis of common variants of STAT6, GATA3, and STAT4 to asthma and high serum IgE phenotypes. *J Allergy Clin Immunol*. 2005; 115(1):80–7. PMID: [15637551](#)
35. Kanai T, Mikami Y, Sujino T, Hisamatsu T, Hibi T. ROR γ -dependent IL-17A-producing cells in the pathogenesis of intestinal inflammation. *Mucosal Immunol*. 2012; 5(3):240–7. doi: [10.1038/mi.2012.6](#) PMID: [74280514](#).
36. Matsuoka K, Inoue N, Sato T, Okamoto S, Hisamatsu T, Kishi Y, et al. T-bet upregulation and subsequent interleukin 12 stimulation are essential for induction of Th1 mediated immunopathology in Crohn's disease. *Gut*. 2004; 53(9):1303–8. PMID: [15306590](#)
37. Delerive P, Monte D, Dubois G, Trottein F, Fruchart-Najib J, Mariani J, et al. The orphan nuclear receptor ROR alpha is a negative regulator of the inflammatory response. *EMBO Reports*. 2001; 2(1):42–8. PMID: [11252722](#)
38. Davies G, Stear MJ, Benothman M, Abuagob O, Kerr A, Mitchell S, et al. Quantitative trait loci associated with parasitic infection in Scottish blackface sheep. *Heredity*. 2006; 96(3):252–8. PMID: [16391549](#)
39. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002; 3:RESEARCH0034. PMID: [12184808](#)
40. Shinnakasu R, Yamashita M, Shinoda K, Endo Y, Hosokawa H, Hasegawa A, et al. Critical YxKxHxxxRP motif in the C-terminal region of GATA3 for its DNA binding and function. *J Immunol*. 2006; 177(9):5801–10. doi: [10.4049/jimmunol.177.9.5801](#) PMID: [17056504](#)
41. Lee CH, Wei L-N, Copeland NG, Gilbert DJ, Jenkins NA. Genomic structure, promoter identification, and chromosomal mapping of a mouse nuclear orphan receptor expressed in embryos and adult testes. *Genomics*. 1995; 30(1):46–52. doi: [10.1006/geno.1995.0007](#) PMID: [edsecd.273458](#).
42. Medvedev A, Chistokhina A, Hirose T, Jetten AM. Genomic structure and chromosomal mapping of the nuclear orphan receptor ROR gamma (RORC) gene. *Genomics*. 1997; 46(1):93–102. PMID: [9403063](#)
43. Rauen T, Juang YT, Hedrich CM, Kis-Toth K, Tsokos GC. A novel isoform of the orphan receptor ROR γ suppresses IL-17 production in human T cells. *Genes Immun*. 2012; 13(4):346–50. doi: [10.1038/gene.2011.85](#) PMID: [76244293](#).
44. Aschrafi A, Meindl N, Firla B, Brandes RP, Steinhilber D. Intracellular localization of ROR α is isoform and cell line-dependent. *Biochimica et Biophysica Acta (BBA)—Mol Cell Res*. 2006; 1763(8):805–14. doi: <http://dx.doi.org/10.1016/j.bbamcr.2006.05.006>.
45. Besnard S, Heymes C, Merval R, Rodriguez M, Galizzi J-P, Boutin JA, et al. Expression and regulation of the nuclear receptor ROR α in human vascular cells. *FEBS Letters*. 2002; 511(1–3):36–40. doi: [http://dx.doi.org/10.1016/S0014-5793\(01\)03275-6](http://dx.doi.org/10.1016/S0014-5793(01)03275-6) PMID: [11821045](#)

46. Giguère V, Tini M, Flock G, Ong E, Evans RM, Otulakowski G. Isoform-specific amino-terminal domains dictate DNA-binding properties of ROR alpha, a novel family of orphan hormone nuclear receptors. *Genes Dev.* 1994; 8(5):538–53. doi: [10.1101/gad.8.5.538](https://doi.org/10.1101/gad.8.5.538) PMID: [7926749](https://pubmed.ncbi.nlm.nih.gov/7926749/)
47. Migita H, Satozawa N, Lin J-H, Morser J, Kawai K. ROR α 1 and ROR α 4 suppress TNF- α -induced VCAM-1 and ICAM-1 expression in human endothelial cells. *FEBS Letters.* 2004; 557(1–3):269–74. [http://dx.doi.org/10.1016/S0014-5793\(03\)01502-3](http://dx.doi.org/10.1016/S0014-5793(03)01502-3). PMID: [14741380](https://pubmed.ncbi.nlm.nih.gov/14741380/)
48. Steinmayr M, Andre E, Conquet F, Rondi-Reig L, Delhaye-Bouchaud N, Auclair N, et al. Staggerer phenotype in retinoid-related orphan receptor alpha-deficient mice. *Proc Natl Acad Sci USA.* 95(7):3960–5. PMID: [9520475](https://pubmed.ncbi.nlm.nih.gov/9520475/)
49. Tait Wojno ED, Artis D. Innate lymphoid cells: Balancing immunity, inflammation, and tissue repair in the intestine. *Cell Host Microbe.* 2012; 12(4):445–57. doi: <http://dx.doi.org/10.1016/j.chom.2012.10.003> PMID: [23084914](https://pubmed.ncbi.nlm.nih.gov/23084914/)
50. Yagi R, Zhong C, Northrup DL, Yu F, Bouladoux N, Spencer S, et al. The transcription factor GATA3 is critical for the development of all IL-7R α -expressing innate lymphoid cells. *Immunity.* 2014; 40(3):378–88. doi: <http://dx.doi.org/10.1016/j.immuni.2014.01.012> PMID: [24631153](https://pubmed.ncbi.nlm.nih.gov/24631153/)