Pure

Scotland's Rural College

Genome-wide association studies of parasite resistance, productivity and immunology traits in meat in Scottish Blackface sheep

Pacheco, A.; Banos, G.; Lambe, N.; McLaren, A.; McNeilly, T. N.; Conington, J.

Published in:
Animal

DOI:
10.1016/j.animal.2023.101069

First published: 22/01/2024

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available

Link to publication

Citation for pulished version (APA):
Pacheco, A., Banos, G., Lambe, N., McLaren, A., McNeilly, T. N., & Conington, J. (2024). Genome-wide association studies of parasite resistance, productivity and immunology traits in meat in Scottish Blackface sheep. *Animal*, Article 101069. Advance online publication. https://doi.org/10.1016/j.animal.2023.101069

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. Feb. 2024

Animal xxx (xxxx) xxx

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Genome-wide association studies of parasite resistance, productivity and immunology traits in Scottish Blackface sheep

A. Pacheco [a,*], G. Banos [a], N. Lambe [a], A. McLaren [a], T.N. McNeilly [b], J. Conington [a]

[a] Scotland's Rural College, Roslin Institute Building, Easter Bush, Midlothian EH25 9RG, United Kingdom
[b] Moredun Research Institute, Pentlands Science Park, Bush Loan, Midlothian EH26 0PZ, United Kingdom

ARTICLE INFO

Article history:
Received 5 August 2023
Revised 18 December 2023
Accepted 21 December 2023
Available online xxxx

Keywords:
Breeding
Gastrointestinal parasites
Genomic analysis
Immunity
Sheep

ABSTRACT

Gastrointestinal parasitism represents a global problem for grazing ruminants, which can be addressed sustainably by breeding animals to be more resistant against infection by parasites. The aim of this study was to assess the genetic architecture underlying traits associated with gastrointestinal parasite resistance, immunological profile and production in meat sheep, and identify and characterise candidate genes affecting these traits. Data on gastrointestinal parasite infection (faecal egg counts for Strongyles (**FEC_S**) and *Nematodirus* (**FEC_N**) and faecal oocyst counts for *Coccidia*, along with faecal soiling scores (**DAG**), characterised by the accumulation of faeces around the perineum) and production (live weight (**LWT**)) were gathered from a flock Scottish Blackface lambs at three and four months of age. Data on the immune profile were also collected from a subset of these lambs at two and five months of age. Immune traits included the production of Interferon-γ (**IFN-γ**), Interleukin (**IL**)**-4** and **IL-10** following stimulation of whole blood with pokeweed mitogen (**PWM**) or antigen from the gastric parasite *Teladorsagia circumcincta* (**T-ci**), and serum levels of *T. circumcincta*-specific immunoglobulin A (**IgA**). Animals were genotyped with genome-wide DNA arrays, and a total of 1 766 animals and 45 827 Single Nucleotide Polymorphisms (**SNPs**) were retained following quality control and imputation. Genome-wide association studies were performed for 24 traits. The effects of individual markers with significant effects were estimated, and the genotypic effect solutions were used to estimate additive and dominance effects, and the proportion of additive genetic variance attributed to each SNP locus. A total of 15 SNPs were associated at least at a suggestive level with FEC_S, FEC_N, DAG, IgA, PWM-induced IFN-γ and IL-4, and T-ci-induced IL-10. This study uncovered 52 genes closely related to immune function in proximity to these SNPs. A number of genes encoding C-type lectins and killer cell lectin-like family members were close to a SNP associated with FEC_N, while several genes encoding IL-1 cytokine family members were found to be associated with IgA. Potential candidate genes belonging to or in close proximity with the Major Histocompatibility Complex (**MHC**) were revealed, including Homeostatic Iron Regulator and butyrophilin coding genes associated with IFN-$\gamma_{(PWM)}$, and IL-17 coding genes associated with IgA. Due to the importance of the MHC in the control of immune responses, these genes may play an important role in resistance to parasitic infections. Our results reveal a largely complex and polygenic genetic profile of the studied traits in this Scottish Blackface sheep population.
© 2024 The Authors. Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Implications

Parasitological traits are widely used to select animals for increased resistance. Moreover, adaptive immunity can potentially be used in tandem with more commonly used traits related to resistance. Here, genome-wide association studies were performed on faecal counts of nematode and coccidian parasites along with cytokine and immunoglobulin A production, revealing notable signals in genomic regions harbouring genes with immunological significance, including genes belonging to or in close proximity to the Major Histocompatibility Complex. If enough variation underpins these regions, phenotypes can potentially respond to selection. If validated, the information presented can potentially be included in breeding programmes aimed at selection for increased resistance in sheep.

Introduction

The advancements in the fields of molecular biology and genomics have held the promise that information at the DNA level

* Corresponding author.
E-mail address: Antonio.Pacheco@sruc.ac.uk (A. Pacheco).

https://doi.org/10.1016/j.animal.2023.101069
1751-7311/© 2024 The Authors. Published by Elsevier B.V. on behalf of The Animal Consortium.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article as: A. Pacheco, G. Banos, N. Lambe et al., Genome-wide association studies of parasite resistance, productivity and immunology traits in Scottish Blackface sheep, Animal, https://doi.org/10.1016/j.animal.2023.101069

would contribute to greater genetic improvement of farmed animals (Meuwissen et al., 2016). Further advancements in next-generation sequencing technology have allowed for *de novo* sequencing of livestock, including sheep, opening the opportunity to create high-density Single Nucleotide Polymorphism (**SNP**) chips. The development of the *Illumina OvineSNP50 BeadChip* microarray was made possible as part of the *International Sheep Genomics Consortium* (**ISGC**) (Rupp et al., 2016). The practicality of using genomic selection in small ruminants has been assessed internationally, including Australia, New Zealand, France, and UK (Rupp et al., 2016). Widespread use of DNA information has resulted in several important breakthroughs, such as the development of genomic selection methodologies and the discovery of many thousands of SNP markers associated with multiple animal traits of interest to farmers, breeders, and consumers (Meuwissen et al., 2016). Genomic selection is particularly useful for polygenic, lowly heritable animal traits with complex genetic architecture and selective breeding may bring ethical benefits by reducing the number of animals exposed to disease (Rupp et al., 2016).

Genome-wide association studies (**GWASs**) have facilitated the identification of loci associated with important animal traits, based on the joint analysis of animal phenotypes and genome-wide SNP genotypes. Furthermore, immunogenetic studies have been crucial in identifying numerous genes involved in shaping the immune repertoire (Acevedo-Whitehouse and Cunningham, 2006). Nevertheless, the great variety of studied parasite species and sheep breeds has contributed to a general lack of agreement among studies (Atlija et al., 2014), which supports the hypothesis that the traits studied may represent different aspects of host-parasite interaction during infection (Gutiérrez-Gil et al., 2009). Discrepancies in the scientific literature may be also attributed to the complex nature of these traits and differences in trait definitions between studies. Gene ontology (**GO**) resources can help narrow down the search for candidate genes (Brown et al., 2013) and even though the annotation of the sheep genome is good, some regions still remain to be annotated, making the identification of candidates in all regions containing significant SNP difficult (Mucha et al., 2015).

The aims of the present study were to (i) assess the genetic architecture underlying lamb traits associated with parasitic infection (faecal egg and oocyst counts and an animal soiling score), immunological profile (cytokines and parasite-specific immunoglobulin A) and production (represented by live BW), and to (ii) identify and characterise candidate genes affecting these traits.

Material and methods

Phenotypic data

Scottish Blackface lambs reared at a Scotland's Rural College experimental hill farm in the Pentland hills, Midlothian, Scotland were used for this study. The flock was managed under typical hill conditions and consisted of different genetic selection lines: Selection (**S**) and Control (**C**) lines as described by Conington et al., (2006) and, (since 2011), a Faecal Egg Count (**F**) line, which includes animals selected for parasitic resistance based on their breeding values for faecal count scores. Animals in lines S and C were selected using a multi-trait index that combines maternal and lamb performance traits (Conington et al., 2001), and were selected for high- and average-performing animals, respectively. Lambs from each of these three genetic lines were used in the study. A fourth group included lambs born from a selection of ewes (~40 per year) from across the three genetic lines that were mated to bought-in rams linking the flock with the Scottish Blackface (**SBF**) sheep industry breed improvement programme (**L**).

Parasitic infection and production data used for the present study were collected individually on lambs born from 2016 to 2018. Parasitic infection data included faecal counts for Strongyles ($\mathbf{FEC_S}$), *Nematodirus* ($\mathbf{FEC_N}$), *Coccidiosis* (**FOC**), and a 5-point faecal soiling score (0 = good, 4 = poor; **DAG**). Production data consisted of live BW (**LWT**) records.

Additionally, blood samples were collected to assess the animal's immunological profiles. Three distinct cytokines were analysed: Interferon-gamma (**IFN-γ**), Interleukin (**IL**)**-4** and **IL-10**. Each of these cytokines was subjected to two distinct and separate stimulation protocols. Whole blood stimulation assays were used to characterise the adaptive immune response traits of this flock in response to pokeweed mitogen (**PWM**), a mitogenic lectin capable of stimulating B and T lymphocytes irrespective of antigenic specificity, and somatic antigen of the common gastrointestinal nematode *Teladorsagia circumcincta* (**T-ci**), to activate parasite-specific lymphocytes. Cytokine phenotypes will henceforth be designated as IFN-$\gamma_{(PWM)}$, IL-$4_{(PWM)}$ and IL-$10_{(PWM)}$ to describe PWM stimulation, and IFN-$\gamma_{(T\text{-}ci)}$, IL-$4_{(T\text{-}ci)}$ and IL-$10_{(T\text{-}ci)}$ to describe a T-ci stimulation. Additionally, the levels of T. circumcincta-specific immunoglobulin (**Ig**)**A** in serum were also measured.

Each set of traits was recorded twice. Chronologically, immunological traits were first recorded at an average age of 53 days, followed by parasitic infection and production traits at an average of 92 days of age. These sets of data are part of the 1st recording occasion. The 2nd recording occasion included measurements of parasitic infection and production traits at an average of 126 days of age, followed by the collection of immunological traits at an average age of 157 days. The order and structure of trait recording are justified by logistical priorities due to the complexity of collecting blood samples and faeces within shorter time intervals without putting the lambs through unnecessary stress.

Table 1 summarises data used for the present study. Since data pertained to two distinct phases in the lambs' growth phase, measurements in each recording occasion were considered as different traits in the ensuing analyses. This was confirmed with preliminary bivariate analyses between the two recording occasions resulting in correlations that were significantly lower than unity. Thus, a total of 24 animal phenotypes were eventually analysed.

Genotypic data

DNA samples for genotyping assays were collected through nasal swabs. DNA was then quantified and genotyped using three arrays including the Illumina OvineSNP15k BeadChip, the Illumina OvineSNP50k BeadChip and the Illumina OvineSNP HD Beadchip (ISGC; https://www.sheephapmap.org) with densities of 15, 50 and 600 thousand SNPs, respectively. Table 2 shows the number of animals genotyped per year of study and array density. Overall, 1 006 animals were genotyped with a low-density SNP chip (15 k), 778 animals with a 50 k SNP chip and 17 using a high-density SNP chip (HD), totalling 1 801 animal genotypes.

Prior to imputation, genotypic data were subjected to quality control checks. Therefore, an SNP was removed if its call rate fell below 90% or the minor allele frequency (**MAF**) was less than 5%. As such, a total of 1 716 and 4 504 SNPs were removed for failing the SNP call rate and MAF thresholds, respectively. Additionally, individual sample-level call rate was also used as a quality control check. Animal genotypes with a call rate below 90% were excluded from the analyses, thereby removing 35 samples. SNP data were also filtered to remove SNPs located on the sex chromosomes. The software PLINK v1.9 (Purcell et al., 2007) was used in all quality control checks.

Table 1
Number of Scottish Blackface sheep lambs for which parasitic infection, production and immunology phenotypes were recorded, and their respective age at the time of data collection.

Traits	Phenotypes	Recording occasion no. 1		Recording occasion no. 2	
		No. Animals	Average age (days)	No. Animals	Average age (days)
Parasitic infection					
	Parasitology				
	FEC_S	1 561	92	1 536	126
	FEC_N	1 561	92	1 536	126
	FOC	1 561	92	1 536	126
	DAG	1 561	92	1 536	126
Production	LWT	1 561	92	1 536	126
Immunology					
	Cytokines				
	IFN-$\gamma_{(PWM)}$	972	53	1 068	157
	IL-$4_{(PWM)}$	972	53	1 068	157
	IL-$10_{(PWM)}$	972	53	1 068	157
	IFN-$\gamma_{(T\text{-}ci)}$	972	53	1 068	157
	IL-$4_{(T\text{-}ci)}$	972	53	1 068	157
	IL-$10_{(T\text{-}ci)}$	972	53	1 068	157
	IgA	949	53	1 045	157

FEC_S = *Strongyles* faecal egg count; FEC_N = *Nematodirus* faecal egg count; FOC = *Coccidia* oocyst count; DAG = dag scores; LWT = live weight; IFN-γ = Interferon-gamma; IL-4 = Interleukin-4; IL-10 = Interleukin-10; IgA = Immunoglobulin A. PWM = Pokeweed mitogen stimulant; T-ci = *T. circumcincta*-specific antigen stimulant.

Table 2
Number of Scottish Blackface sheep lambs genotyped yearly and respective SNP densities.

Year	Animals	No. SNP	No
2016	681	15 k	370
		50 k	311
2017	481	15 k	314
		50 k	150
		HD	17
2018	639	15 k	322
		50 k	317
		Total	1 801

SNP = single nucleotide polymorphism; HD = high-density SNP chip.

Following the quality control steps described above, animal genotypes were imputed to a subset of the most informative SNPs from the 50 k array panel (Ovine SNP50 chip). Imputation was executed with the software FindHap v3 (VanRaden et al., 2011), and SNP positions were based on the Oar_v3.1 version of the sheep genome assembly. After these quality control and imputation steps, 45 827 SNPs and 1 766 animals were retained for analysis. After imputation, a genomic relationship matrix was constructed based on the first method described by VanRaden (2008). The genotypes used in this study were previously quality-assured and imputed for a different project involving the same population of Scottish Blackface sheep.

Population structure

Principal component analysis was first undertaken on the genotypic data to investigate possible population structure using the genomic relationship matrix between animals to identify the principal components explaining variation among individual samples and reveal potential population stratification. Possible stratification was examined across multiple factors, including genetic line, sex, year, and grazing location, in order to determine the need for possible consideration in the ensuing GWAS analyses. The principal components were analysed through the Eigen-decomposition of the genomic relationship matrix to check if there is any recorded covariate that could influence population structure at a genetic level. Relevant analyses were conducted with GEMMA v0.84.1 (Zhou and Stephens, 2012), and the resulting plots were visualised using R software v3.5.1 (R Core Team, 2021).

Genome-wide association studies

The data described above were analysed to determine associations between genetic variants and trait phenotypes. The following model was applied separately for each of the 24 traits listed in Table 1:

$$y = \mu + Xb + Zu + e \quad (1)$$

where y is the animal record (trait phenotype), μ is the overall mean, b and u correspond to the vector of fixed and random effects, respectively, X and Z are the corresponding design matrices, and e is the vector of residual effects. Table 3 summarises the fixed effects pertaining to each trait. A stepwise backward elimination approach was implemented whereby fixed effect significance was determined in preliminary analyses with the model (1). Fixed effects with P-values <0.05 were included in the model for each trait. Additionally, the first three principal components were included in the model (1) as covariates to account for population stratification revealed through principal component analysis. Fixed effects include year of birth, sex of the lamb, genetic line of the lamb, location of the lamb at different stages of development (location when lambs are at the mid-point between birth and weaning and location of the lambs at the time of weaning), age of the lamb at the time of recording, birth-rearing rank (single or twins) and age of the dam (in years) at the time of parturition.

The GEMMA software (v0.94.1) (Zhou and Stephens, 2012) was used to first compute the genomic relationship matrix between animals and then conduct GWAS with the model (1). Significance thresholds were obtained and applied to GWAS results after Bonferroni correction for multiple testing as $-\log_{10}(0.05/N)$ and $-\log_{10}(1/N)$ for genome-wide and suggestive significance, respectively, where N is the total number of SNPs (N = 45 827). After Bonferroni correction, the genome-wide significance threshold ($P \leq 0.05$) was set at $P = 1.09 \times 10^{-6}$ which corresponds to $-\log10(P) = 5.96$ while the suggestive significance threshold level (accounting for one false positive per genome scan) was set at $P = 2.18 \times 10^{-5}$ corresponding to $-\log10(P) = 4.66$.

Quantile-quantile (**Q-Q**) plots were used to verify whether the distribution of the observed $-\log10(P)$ values deviated from the expected exponential distribution under the null hypothesis of no genetic association and no LD between SNPs. Individual SNP marker associations with the studied traits were visualised by plotting the resulting $-\log10(P)$ values in Manhattan plots. The geno-

Table 3
Statistically significant fixed effects included for each trait model analysed in Scottish Blackface sheep lambs.

Traits	Fixed effects
Recording occasion no. 1	
FEC_S	Year, sex, line, brrnk, age, three principal components
FEC_N	Year, sex, line, mkgraz, age, three principal components
FOC	Year, sex, line, mkgraz, age, three principal components
DAG	Year, sex, line, mkgraz, three principal components
LWT	Year, sex, line, mkgraz, brrnk, age, three principal components
IFN-$\gamma_{(PWM)}$	Year, sex, line, mkgraz, age, three principal components
IL-$4_{(PWM)}$	Year, sex, line, mkgraz, brrnk, three principal components
IL-$10_{(PWM)}$	Year, sex, line, mkgraz, three principal components
IFN-$\gamma_{(T\text{-}ci)}$	Year, sex, line, mkgraz, brrnk, age, three principal components
IL-$4_{(T\text{-}ci)}$	Year, sex, line, mkgraz, three principal components
IL-$10_{(T\text{-}ci)}$	Year, sex, line, mkgraz, dage, three principal components
IgA	Year, sex, line, mkgraz, age, dage, three principal components
Recording occasion no. 2	
FEC_S	Year, sex, line, mkgraz, wngraz, brrnk, three principal components
FEC_N	Year, sex, line, mkgraz, wngraz, three principal components
FOC	Year, sex, line, mkgraz, wngraz, age, three principal components
DAG	Year, sex, line, mkgraz, wngraz, dage, three principal components
LWT	Year, sex, line, mkgraz, wngraz, brrnk, dage, three principal components
IFN-$\gamma_{(PWM)}$	Year, line, mkgraz, age, three principal components
IL-$4_{(PWM)}$	Year, sex, line, mkgraz, age, three principal components
IL-$10_{(PWM)}$	Year, line, mkgraz, brrnk, age, three principal components
IFN-$\gamma_{(T\text{-}ci)}$	Sex, line, brrnk, age, three principal components
IL-$4_{(T\text{-}ci)}$	Year, sex, line, brrnk, three principal components
IL-$10_{(T\text{-}ci)}$	Year, line, wngraz, age, three principal components
IgA	Year, sex, line, wngraz, three principal components

Fixed effects key: year = year of birth, sex = sex of the lamb, line = genetic line of the lamb, mkgraz = location of the lamb at the mid-point between birth and weaning, wngraz = location of the lamb at weaning, age = age of the lamb at recording, brrnk = birth-rearing rank and dage = age of the dam at the time of parturition. FEC_S = *Strongyles* faecal egg count; FEC_N = *Nematodirus* faecal egg count; FOC = *Coccidia* oocyst count; DAG = dag scores; LWT = live weight; IFN-γ = Interferon-gamma; IL-4 = Interleukin-4; IL-10 = Interleukin-10; IgA = parasite-specific Immunoglobulin A. PWM = Pokeweed mitogen stimulant; T-ci = *T. circumcincta*-specific antigen stimulant.

mic inflation factor λ was calculated to determine possible estimate inflation based on the method described by Amin et al., (2007). Factor λ checks for any systematic deviations observed from expected *P*-values that could result from the remaining, unaccounted-for population substructure.

The magnitude of the effects of individual markers with significant effects identified in GWAS was estimated using model (1) with the addition of the fixed effect of the corresponding SNP locus genotype. These analyses were conducted with the ASReml v3.0 software (Gilmour et al., 2009). The genotypic effect solutions were used to estimate the additive (a) and dominance (d) effects, and the proportion of additive genetic variance (PV_A) due to each SNP locus as follows:

$$a = (AA - BB)/2$$

$$d = AB - ((AA + BB)/2)$$

$$PV_A = (2pq(a + d(q - p))^2)/V_A$$

where *AA*, *BB* and *AB* correspond to the solutions of the respective genotypic effect levels, p and q correspond to the allelic frequencies of *A* and *B* at the SNP locus and V_A correspond to the additive genetic variance of the trait.

Gene and gene ontology annotation in associated regions

The presence of candidate genes 500 kb upstream or downstream of the significant SNPs for each of the studied traits at either genome-wide or suggestive levels was investigated. Protein-coding genes that were found within the candidate regions considered were retrieved from the Ensemble Genes 91 database, which is based on the Oar v3.1 ovine reference genome. This was achieved using the BioMart tool (Kinsella et al., 2011).

The classification of genes in accordance with biological function was performed by using the Database for Annotation, Visualisation and Integrated Discovery (DAVID) v6.8 tool (Huang et al., 2009). Gene ontology (GO) terms (The Gene Ontology Consortium, 2021) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways (Kanehisa et al., 2014) were identified using the same tool.

Results

Principal component analysis

This analysis revealed population stratification attributed to the different genetic lines (Supplementary Fig. S1) that constitute the population studied. The first, second and third principal components explain 23.9, 18.9 and 9.6% of variance, respectively. The genetic lines appear to be distinct, illustrating the effect of directional selection. The flock is split into S, C and F selection lines. A fourth group comprising fewer animals (L line), produced from mating a number of ewes from each of the three lines each year to rams from other recorded SBF sheep flocks, are distributed throughout the plot, although the majority seem to be genetically closer to the F and C selection lines.

Genome-wide association studies results

Manhattan plots of GWAS results and the corresponding Q-Q plots of observed *P*-values against expected *P*-values for each trait are shown in Figs. 1–3. Q-Q plots are used as a form of validation of the results displayed in the Manhattan plots. Results show that the genomic relationship matrix and the inclusion of the three first principal components were sufficient to account for nearly all population structures in this study, as reflected by the inflation factor (λ) being near unity. The average inflation factor was 1.000 ± 0.009, ranging from 0.9802 to 1.012 for IFN-$\gamma_{(PWM)}$ (2nd recording occasion) and IgA (2nd recording occasion), respectively. Inflation factors reflect that the necessary adjustments have already been made.

The study revealed several SNPs in genomic regions that can potentially be associated with the traits studied (Table 4). A total of 15 SNPs were associated at least at a suggestive level with FEC_S, FEC_N, DAG, IFN-$\gamma_{(PWM)}$, IL-$4_{(PWM)}$, IL-$10_{(T\text{-}ci)}$ and IgA (Table 4). Of these 15 SNPs, only one was significant at a genome-wide level, after multiple trait adjustment (IgA at the 2nd recording occasion). Each SNP association was unique to a specific trait, and no single SNP affected multiple traits. No significant SNPs associated with FOC, LWT, IFN-$\gamma_{(T\text{-}ci)}$, IL-$4_{(T\text{-}ci)}$ and IL-$10_{(PWM)}$ were found.

The magnitude of significant SNP markers was manifested in additive and dominant genetic effects calculated for each one, along with the proportion of total genetic variance explained by each locus (Table 4). Most SNPs had a significant additive effect on the corresponding trait (OAR7_53177511.1, OAR3_220227474.1, s15560.1,

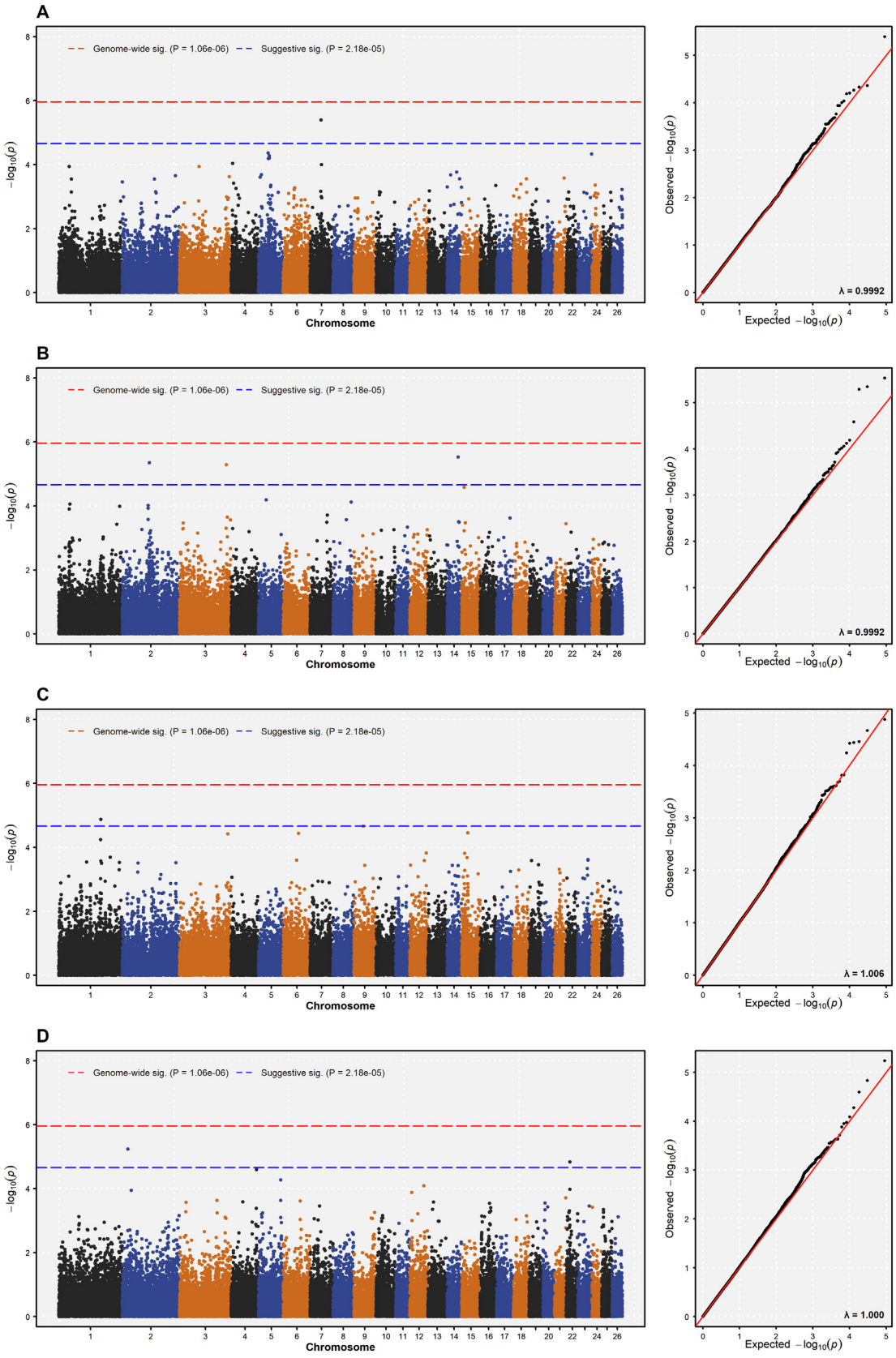

Fig. 1. Manhattan plots displaying genome-wide association study (GWAS) results (*P*-values) and corresponding quantile-quantile (Q-Q) plots (observed *P*-values against expected *P*-values) for (A) FEC_S (1st recording occasion), (B) and (C) FEC_N (1st and 2nd recording occasion, respectively), and D) DAG phenotypes in Scottish Blackface sheep lambs. Red and blue dashed lines represent the genome-wide (*P*-value = 1.06×10^{-6}) and suggestive significance thresholds (*P*-value = 2.18×10^{-5}), respectively. FEC_S = Strongyles faecal egg counts; FEC_N = *Nematodirus* faecal egg counts; DAG = dag scores.

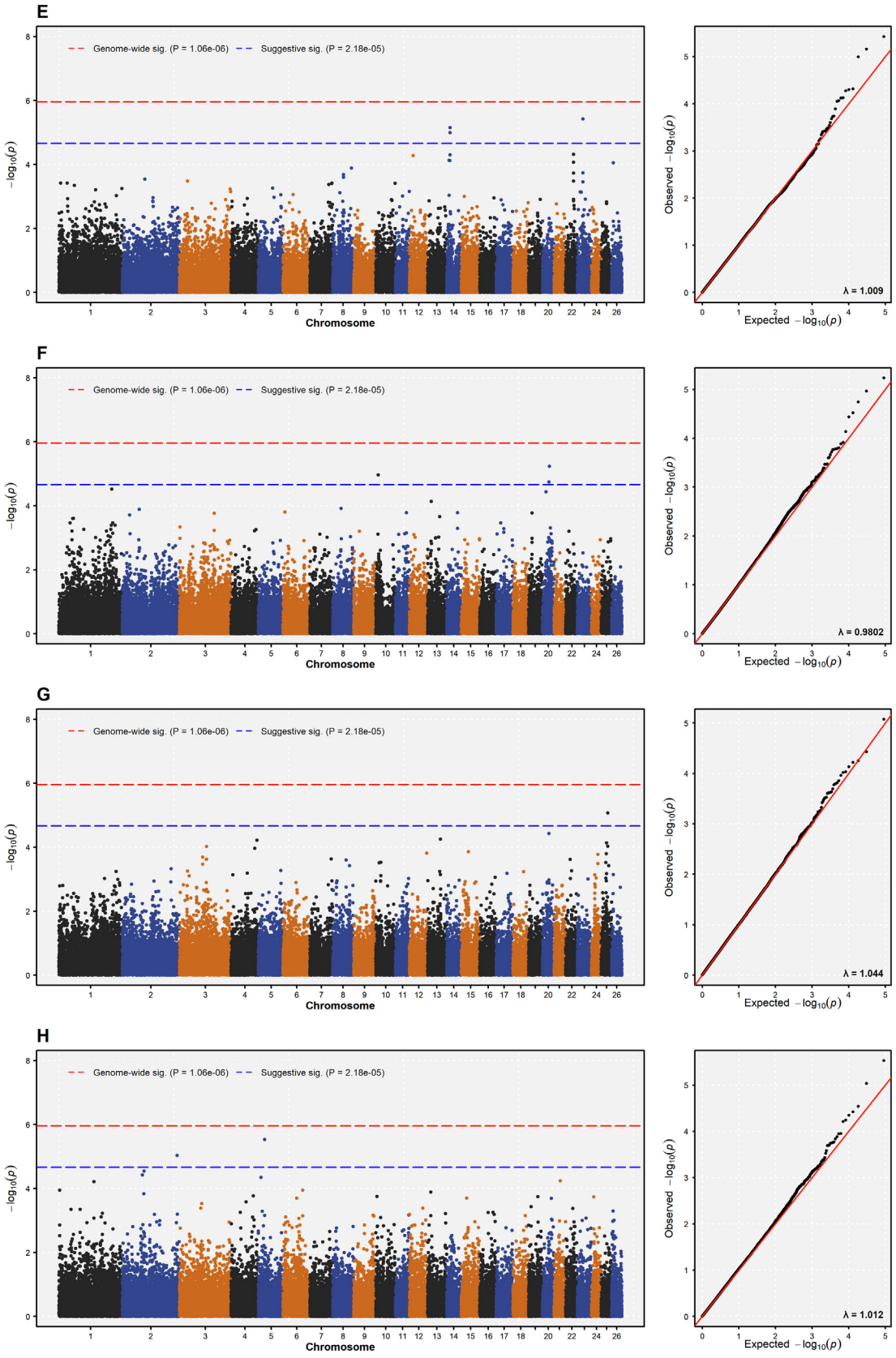

Fig. 2. Manhattan plots displaying genome-wide association study (GWAS) results (*P*-values) and corresponding quantile–quantile (Q-Q) plots (observed *P*-values against expected *P*-values) for (E) IL-$4_{(PWM)}$ (1st recording occasion), (F) IFN-$\gamma_{(PWM)}$, (G) IL-$4_{(PWM)}$ and (H) IL-$10_{(T.ci)}$ (2nd recording occasion) phenotypes in Scottish Blackface sheep lambs. Red and blue dashed lines represent the genome-wide (*P*-value = 1.06×10^{-6}) and suggestive significance thresholds (*P*-value = 2.18×10^{-5}), respectively. IFN-γ = Interferon-gamma; IL-4 = Interleukin-4; IL-10 = Interleukin-10; PWM = Pokeweed mitogen stimulant; T.ci = Larval (L4) *T. circumcincta*-specific antigen stimulant.

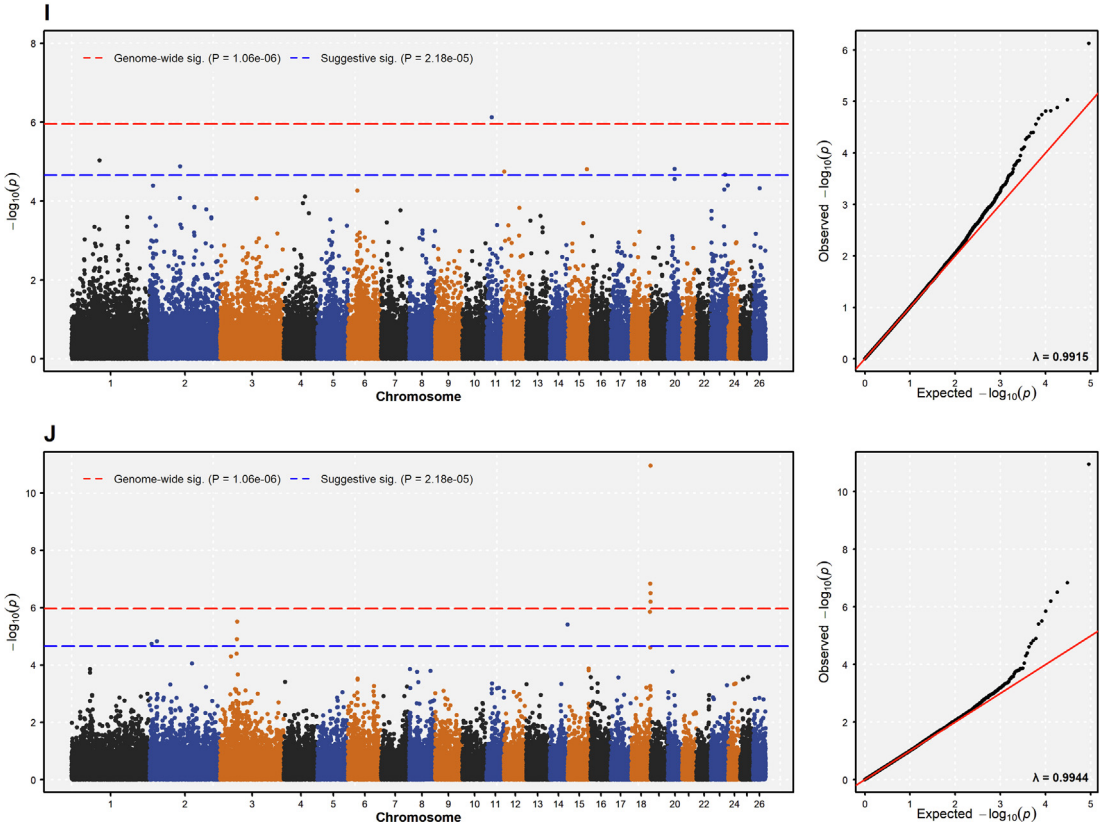

Fig. 3. Manhattan plots displaying genome-wide association study (GWAS) results (*P*-values) and corresponding quantile–quantile (Q-Q) plots (observed *P*-values against expected *P*-values) for (I) and (J) IgA (1st and 2nd recording occasions, respectively) phenotypes in Scottish Blackface sheep lambs. Red and blue dashed lines represent the genome-wide (*P*-value = 1.06×10^{-6}) and suggestive significance thresholds (*P*-value = 2.18×10^{-5}), respectively. IgA = Immunoglobulin A.

Table 4
SNP information: SNPs presented show the strongest association with FEC_S, FEC_N, DAG, IFN-$\gamma_{(PWM)}$, IL-$4_{(PWM)}$, IL-$10_{(T\text{-}ci)}$ and IgA in Scottish Blackface sheep lambs.

Traits	OAR	SNP	Position (bp)	*P*-value	Beta (SE)	a (SE)	*P*-value	d (SE)	*P*-value
FEC_S[1]	7	OAR7_53177511.1	48116542	8.91×10^{-6}	−0.247 (0.054)	−0.1132 (0.0404)	0.0053	0.0530 (0.0653)	0.4174
FEC_N[1]	3	OAR3_220227474.1	204319907	5.31×10^{-6}	−0.691 (0.147)	−0.3707 (0.0887)	0.0000	−0.0495 (0.1013)	0.6257
	14	s15560.1	47964362	3.11×10^{-6}	0.425 (0.093)	−1.0407 (0.2881)	0.0003	0.4092 (0.3016)	0.1754
DAG[1]	22	OAR22_20459522.1	16596984	1.16×10^{-5}	−0.052 (0.012)	0.0615 (0.0785)	0.4338	0.0333 (0.0898)	0.7109
IL-$4_{(PWM)}$[1]	14	OAR14_14894096.1	14691342	8.17×10^{-6}	−0.134 (0.03)	−0.1132 (0.404)	0.0053	0.0279 (0.0452)	0.5381
IgA[1]	1	OAR1_106406464.1	99129941	1.06×10^{-5}	0.017 (0.01)	−0.0170 (0.0037)	0.0000	−0.0096 (0.0040)	0.0170
	20	s67322.1	24425338	1.75×10^{-5}	0.021 (0.00)	0.0079 (0.0069)	0.2095	0.0213 (0.0068)	0.0018
FEC_N[2]	1	OAR1_197073884.1	182705596	1.52×10^{-5}	0.444 (0.102)	0.4446 (0.1060)	0.0000	−0.1301 (0.1203)	0.2797
IFN-$\gamma_{(PWM)}$[2]	20	OAR20_33522555.1	30534971	6.91×10^{-6}	0.140 (0.031)	−0.0924 (0.0377)	0.0144	0.0936 (0.0407)	0.0220
IL-$4_{(PWM)}$[2]	25	OAR25_30051833.1	28730885	9.09×10^{-6}	0.191 (0.043)	−0.2636 (0.0620)	0.0000	−0.0648 (0.0647)	0.3167
IL-$10_{(T\text{-}ci)}$[2]	2	OAR2_252448674.1	239031909	1.06×10^{-5}	0.118 (0.025)	−0.0890 (0.0219)	0.0001	−0.0214 (0.0253)	0.3997
IgA[2]	2	OAR2_26992982.1	26058151	1.67×10^{-5}	0.053 (0.012)	0.1353 (0.0573)	0.0186	−0.0184 (0.0952)	0.8472
	3	s67484.1	59634347	3.48×10^{-6}	0.087 (0.016)	−0.0783 (0.0304)	0.0102	0.0165 (0.1140)	0.6059
	14	s52366.1	62552503	4.51×10^{-6}	0.131 (0.028)	−0.1946 (0.1132)	0.0861	−0.0817 (0.1140)	0.4740
	18	s08970.1	68450168	4.49×10^{-11}	−0.079 (0.012)	−0.0899 (0.0119)	0.0000	0.0271 (0.0135)	0.0451

[1] Recording occasion no. 1.
[2] Recording occasion no. 2. FEC_S = *Strongyles* faecal egg count; FEC_N = *Nematodirus* faecal egg count; DAG = dag scores; IFN-γ = Interferon-gamma; IL-4 = Interleukin-4; IL-10 = Interleukin-10; IgA = parasite-specific Immunoglobulin A. PWM = Pokeweed mitogen stimulant; T-ci = *T. circumcincta*-specific antigen stimulant. OAR = sheep chromosome; SNP = single nucleotide polymorphism; bp = base-pair position; a = additive effect; d = dominance effect.

OAR14_14894096.1, OAR1_106406464.1, OAR1_197073884.1, s67484.1, OAR20_33522555.1, OAR25_30051833.1, OAR2_252448674.1, OAR2_26992982.1 and s08970.1), and only four had a significant dominance effect (OAR1_106406464.1, s67322.1, OAR20_33522555.1 and s08970.1). Boxplots in Fig. 4 plot phenotypes adjusted for fixed effects against genotypes. With the exception of SNP OAR1_106406464.1, the plots fail to show the dominance effect. Individual genes neighbouring significant SNPs mentioned above are summarised in Tables 4, 5 and Supplementary Table S1, and gene function is described in Supplementary Table S2. The results confirmed that 52 genes neighbouring relevant SNPs are involved in immune function in this study.

Parasitic infection traits

The Cyclin B2 (*CCNB2*) and the Ring finger protein 111 (*RNF111*) genes were respectively found 77 kbp and 117 kbp upstream, which was the most significant SNP on OAR7 (OAR7_53177511.1) for FEC_S. The corresponding region on chromosome 7 explained approximately 16% of the genetic variance of the trait in lambs at around three months of age.

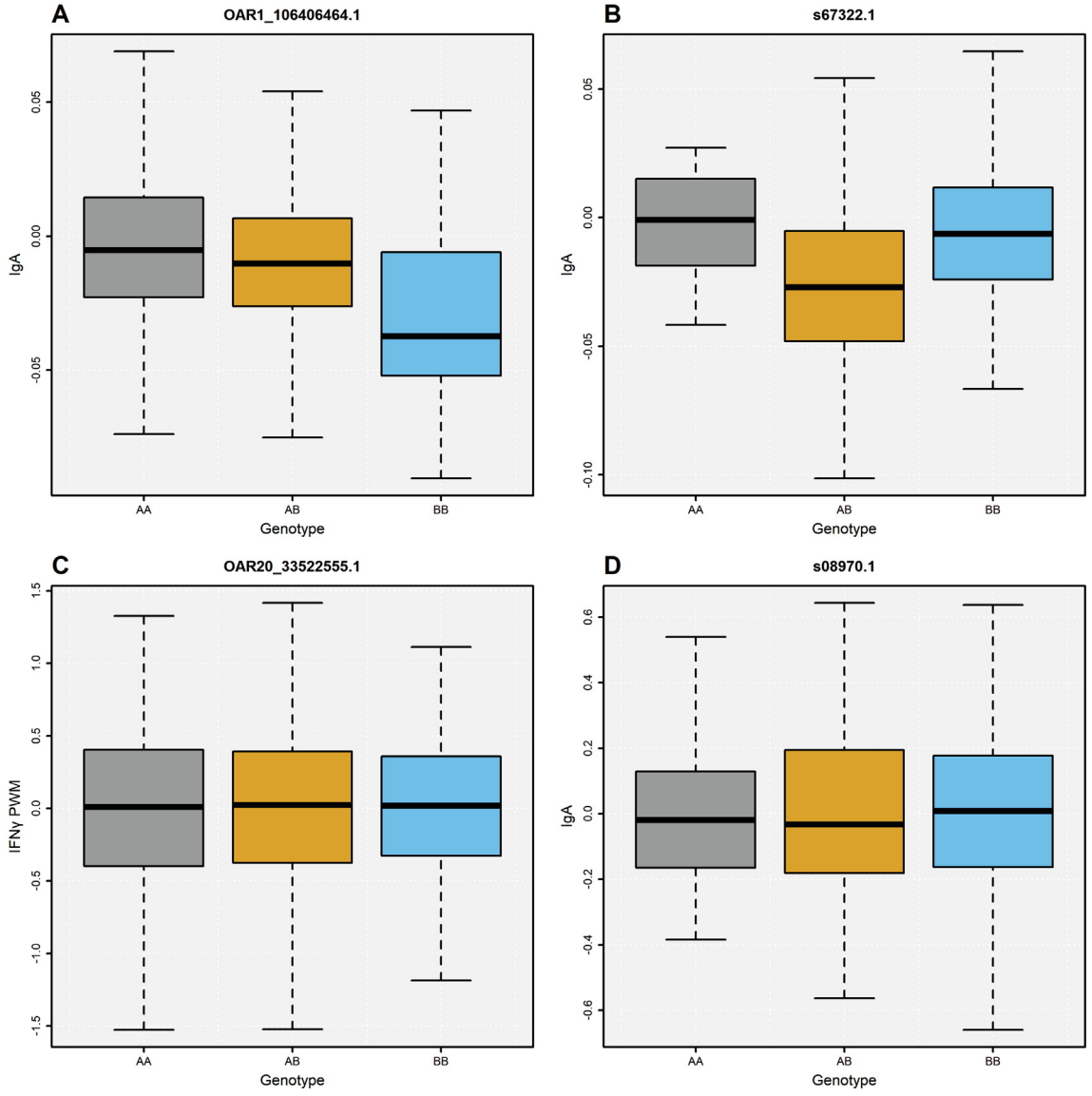

Fig. 4. Boxplots for adjusted phenotypes of (A) and (B) IgA (1st recording occasion), (C) IFN-$\gamma_{PWM)}$ and (D) IgA (2nd recording occasion) across observed genotypes for SNPs associated with these traits in Scottish Blackface sheep lambs and for which analyses showed significant dominance effects ((OAR1_106406464.1, s67322.1, OAR20_33522555.1 and s08970.1, respectively). SNP = single nucleotide polymorphism; IgA = Immunoglobulin A; IFN-γ = Interferon-gamma; PWM = Pokeweed mitogen stimulant.

Ten genes belonging to the C-type lectin superfamily of genes were found close to SNPs associated with FEC_N. Of these, nine genes were 127–488 kbp upstream of the most significant SNP OAR3_220227474.1 (*CLEC1A*, *CLEC1B*, *CLEC2B*, *LOC101120482*, *LOC101103714*, *CLEC7A*, *CLEC9A*, *CLEC12B* and *OLR1*) and one gene (*LOC101123029*) was 363kbp downstream the same SNP. The oxidised low-density lipoprotein receptor 1 (*OLR1*) gene encodes a low-density lipoprotein receptor also belonging to the C-type lectin family. Additionally, genes belonging to the killer cell lectin-like receptor family were found near OAR3_220227474.1. Genes *LOC101116896*, *LOC101104216*, *LOC101123288*, *KLRD1*, *LOC101116641*, *LOC105614844* and *LOC105613001* were 71–295 kbp downstream of this SNP on OAR 3. Another killer cell lectin-like encoding gene (LOC101102227) was found to encompass the same SNP in chromosome 3. Furthermore, three genes were found close to SNP s15560.1 on OAR14, also associated with FEC_N. The galectin 4 (*LGALS4*) and NFKB inhibitor Beta (*NFKBIB*) genes were found 238 kbp and 156 kbp downstream of the highest SNP on OAR14, respectively, while the ZFP36 Ring Finger Protein (*ZFP36*) gene was 301 kbp upstream of the same SNP. Regions on chromosomes 3 and 14 corresponding to SNPs OAR3_220227474.1 and s15560.1 accounted for 23 and 31% of the FEC_N genetic variance, respectively. Collectively, these two SNPs, therefore account for more than half of the genetic variance. Three genes neighboured the SNP OAR1_197073884.1, associated with FEC_N. The genes immunoglobulin Superfamily Member 11 (*IGSF11*), CD80 Molecule (*CD80*) and Phospholipase A1 Member A (*PLA1A*) were found 15–463 kbp downstream of this SNP. The corresponding region of OAR1 explained 5% of the genetic variance of FEC_N in 4-month-old lambs. Additionally, a single gene, B cell linker (*BLNK*), was found 157 kbp upstream of OAR22_20459522.1 which corresponds to the most significant SNP on chromosome 22 associated with DAG and accounted for 3% of the genetic variance.

Immunological traits – Cytokines

Four genes were found around SNP OAR20_33522555.1, associated with IFN-$\gamma_{(PWM)}$. Two of these genes, belonging to the butyrophilin family of genes (Butyrophilin subfamily 1 member A1 –

Table 5
SNP information and associated genes found in proximity to significant SNPs in the population of Scottish Blackface sheep lambs analysed.

Traits	OAR	SNP	MAF	% V_A Exp.	Genes
FEC_S[1]	7	OAR7_53177511.1	0.35	16	*RNF111; CCNB2*
FEC_N[1]	3	OAR3_220227474.1	0.37	23	*CLEC1A; CLEC1B; CLEC1B; LOC101120482; LOC101103714; CLEC7A; LOC101123029; CLEC9A; CLEC12B; LOC101116896; LOC1011104216; LOC101123288; KLRD1; LOC101102227; LOC101116641; LOC105614844; LOC105613001; OLR1*
	14	s15560.1	0.09	31	*LGALS4; NFKBIB; ZFP36*
DAG[1]	22	OAR22_20459522.1	0.42	3	*BLNK*
IL-$4_{(PWM)}$[1]	14	OAR14_14894096.1	0.46	2	*MYLK3*
IgA[1]	1	OAR1_106406464.1	0.39	5	*FCGR1A; CTSS; ECM1*
	20	s67322.1	0.15	6	*IL17A; IL17F*
FEC_N[2]	1	OAR1_197073884.1	0.29	5	*IGSF11; CD80; PLA1A*
IFN-$\gamma_{(PWM)}$[2]	20	OAR20_33522555.1	0.26	8	*BTN1A1; BTN2A2; HFE; TRIM38*
IL-$4_{(PWM)}$[2]	25	OAR25_30051833.1	0.13	6	*PLA2G12B*
IL-$10_{(T\text{-}ci)}$[2]	2	OAR2_252448674.1	0.43	15	*IFI6*
IgA[2]	2	OAR2_26992982.1	0.04	4	*SYK; NFIL3*
	3	s67484.1	0.11	6	*IL1A; IL1B; IL1F10; IL1RN; IL36A; IL36B; IL36RN; IL37*
	14	s52366.1	0.04	4	*TRIM28*
	18	s08970.1	0.36	12	*PLD4; BTBD6*

[1] Recording occasion no. 1.
[2] Recording occasion no. 2. FEC_S = *Strongyles* faecal egg count; FEC_N = *Nematodirus* faecal egg count; DAG = dag scores; IFN-γ = Interferon-gamma; IL-4 = Interleukin-4; IL-10 = Interleukin-10; IgA = parasite-specific Immunoglobulin A. PWM = Pokeweed mitogen stimulant; T-ci = *T. circumcincta*-specific antigen stimulant. OAR = sheep chromosome; SNP = single nucleotide polymorphism; MAF = minor allele frequency; V_A Exp. = Proportion of genetic variance explained by individual SNP.

BTN1A1 and Butyrophilin subfamily 2 member A2 – *BTN2A2*), were located, 5kbp downstream and 8kbp upstream of the SNP on OAR20, respectively. Additionally, the Homeostatic Iron Regulator (*HFE*) and Tripartite Motif Containing 38 (*TRIM38*) genes were located 255kbp and 324kbp downstream of OAR20_33522555.1, respectively. This SNP located on OAR20 accounts for 8% of the explained genetic variance of IFN-γ when lambs are 5 months old.

The Myosin Light Chain Kinase 3 (*MYLK3*) gene was found to be associated with IL-$4_{(PMW)}$ and located 96 kbp downstream of the OAR14_14894096.1 on chromosome 14, which accounts for 2% of genetic variance for this trait. The Phospholipase A2 Group XIIB (*PLA2G12B*) gene is located 98kbp upstream of the OAR_253001833.1 SNP associated with IL-$4_{(PWM)}$. This SNP explained around 6% of the genetic variance of this trait. Finally, a significant SNP found on OAR2 (OAR2_252448674.1) accounted for approximately 15% of the genetic variance of the trait IL-$10_{(T\text{-}ci)}$. The Interferon Alpha Inducible Protein 6 (*IFI6*) was found 491 kbp downstream of this SNP.

Immunological traits – Immunoglobulin A

Eighteen immune-related genes were found in proximity to six SNPs associated with IgA, namely OAR1_106406464.1, s67322.1, OAR2_26992982.1, s67484.1, s52366.1 and s08970.1. Three of these genes were located on OAR1 around the OAR1_106406464.1 SNP: the Fc Gamma receptor Ia (*FCGR1A*) gene is located 302kbp downstream of this SNP, while the Extracellular Matrix protein 1 (*ECM1*) and Cathepsin S (*CTSS*) genes are located 311 kbp and 493 kbp upstream of the same SNP, respectively. Two further genes belonging to the IL-17 receptor family (Interleukin 17A - *IL17A* and Interleukin 17F - *IL17F*) on OAR20 were found in proximity to the SNP s67322.1. This SNP is found within the *IL17F* gene, whereas *IL17A* is located 25 kbp downstream. The OAR1_106406464.1 and s67322.1 SNPs together explain 11% of the genetic variance of IgA in 2-month-old lambs (5 and 6%, respectively). The Spleen Associated Tyrosine Kinase (*SYK*) and nuclear factor Interleukin 3 (*NFIL3*) genes were respectively located 243kbp upstream and 298kbp downstream of SNP OAR2_26992982.1 on OAR2, which accounted for 4% of genetic variance explained for this trait. Eight potentially interesting genes were found on chromosome 3, around SNP s67484.1 accounting for 6% of the trait variance. All eight genes belong to the interleukin (IL)-1 family of cytokines and were 74–443 kbp upstream of this SNP. These are the Interleukin 1 Alpha and Beta (*IL1A* and *IL1B*, respectively), the Interleukin 1 Receptor Antagonist (*IL1RN*), the Interleukin 1 Family Member 10 (*IL1F10*), the Interleukin 36 Alpha and Beta (*IL36A* and *IL36B*, respectively), the Interleukin 36 Receptor Antagonist (**IL36RN**), and the Interleukin 37 gene (*IL37*). The Tripartite motif 28 (*TRIM28*) gene was found 119kbp upstream of the s52366.1 SNP on OAR14, which explained 4% of the genetic variance. Lastly, two further genes were found in proximity to the s08970.1 SNP on OAR18, representing the only significant SNP at genome-wide level accounting for 12% of the trait genetic variance. The Phospholipase D Family Member 4 (*PLD4*) and the BTB Domain Containing 6 (*BTBD6*) genes are located 372kbp and 182kbp downstream of the SNP, respectively.

Gene ontology results

Gene ontology (GO) terms were extracted for each of the genes discussed above, located within 500kbp of the corresponding SNPs. This analysis is used as a gene functional classification system that allows us to describe the properties of genes. GO terms are functionally separated into three domains: biological processes, cellular components and molecular function. Most of the GO terms found in the present study were associated with biological processes. GO terms and corresponding genes relating to the various terms are summarised in Supplementary Table S3. All terms described relate to immune function.

Parasitic infection traits

There were several GO terms associated with disease traits. The C-type lectin encoding gene *CLEC7A* is associated with the signalling of pattern recognition receptor activity (GO:0008329). *LOC101116896*, encoding a killer cell lectin-like receptor, has important roles in the natural killer cell activation, positive regulation of killer cell-mediated cytotoxicity and the MHC class Ib receptor activity (GO:0030101, GO:0045954 and GO:0032394). The gene *ZFP36* is an important negative regulator of inflammatory responses (GO:0050728), while CD80 is involved in T cell co-stimulation and positive alpha–beta T cell proliferation (GO:0031295 and GO:0046641, respectively).

Immunological traits – Cytokines

MYLK3, associated with IL-4, has a role in the regulation of vascular permeability involved in acute inflammatory responses (GO:0002528). *BTN2A2* is a positive regulator of T cell differentiation (GO:0045591) but also has roles as a negative regulator of cytokine secretion and negative regulator of the T cell signalling pathway (GO:0050710 and GO:0050860), respectively. *HFE* is a negative regulator of T cell antigen processing and presentation, T cell cytokine production and antigen processing and presentation of endogenous peptides antigens via MHC class I (GO:0002626, GO:0002725 and GO:1904283, respectively).

Immunological traits – Immunoglobulin A

FCGR1A is implicated in the antigen processing and presentation of exogenous peptide antigen via MHC class I and defence responses to bacterium (GO:0042590 and GO:0042742, respectively). *ECM1* is involved in the regulation of Th2 immune responses (GO:0002828), while *CTSS* has roles in adaptive immune responses and antigen processing and presentation of peptide antigen (GO:0002250 and GO:0030574, respectively). Two genes coding for IL-17 cytokines, *IL17A* and *IL17F*, are both involved in inflammatory responses (GO:0006954) and in the positive regulation of cytokine production involved in inflammatory response (GO:1900017). *SYK* is involved in the macrophage activation involved in immune responses, as well as in innate immune responses, and is a positive regulator of B cell differentiation (GO:0002281, GO:0045087 and GO:0045579, respectively). *IL1A* has a role in connective tissue replacement involved in inflammatory response to wound healing and immune responses (GO:0002248 and GO:0006955). *IL1B*, *IL36A*, *IL36B* and *IL37* genes are all associated with GO terms related to inflammatory responses and immune responses (GO:0006954 and GO:0006955, respectively). *IL1RN* serves as a negative regulator of the cytokine-mediated signalling pathway and a negative regulator of the IL-1-mediated signalling pathway (GO:0001960 and GO:2000660), while *IL36RN* serves as a negative regulator of the cytokine-mediated signalling pathway and a negative regulator of IL-17 production and IFN-γ secretion (GO:0001960, GO:0032700 and GO:1902714, respectively).

Pathways of the Kyoto encyclopedia of genes and genomes

In addition to GO terms, relevant KEGG pathways were also identified. KEGG pathway analysis results can be consulted in Supplementary Table S4. Here, we summarise the most relevant results.

Parasitic infection traits

Four members of the killer cell lectin-like receptor family, which are *LOC101104216*, *LOC101123288*, *LOC101116896* and *LOC101102227*, are involved in the antigen process and presentation pathway (oas04612), with the latter two genes also being involved in the natural killer cell-mediated cytotoxicity pathway (oas04650). *NFKBIB* gene is involved in both T and B cell receptor signalling pathways (oas04660 and oas04662, respectively). The *CD80* gene is involved in the intestinal immune network for the IgA production pathway (oas04672).

Immunological traits – Immunoglobulin A

The *CTSS* gene is also involved in the antigen processing and presentation pathway (oas04612), while *SYK* is involved in the B cell receptor pathway (oas04662). *IL17A*, *IL1A* and *IL1B* are all involved in the cytokine-cytokine receptor interaction pathway (oas4060). Additionally, *IL1A* and *IL1B* are involved in the mitogen-activated protein kinase (MAPK) pathway (oas4010) and the hematopoietic cell lineage pathway (oas04640). *IL1B* is also involved in the Toll-like receptor signalling pathway NF-kappa B signalling pathway (oas04620 and oas04064, respectively).

Discussion

The present study set out to investigate the genomic architecture of lamb traits related to parasitic disease resistance, immune profile, and production. Multiple significant associations were identified for the first two trait categories, but not for production (LWT), and are discussed below.

Parasitic infection traits

Of the five SNPs associated with parasitic infection traits, four SNPs found to be associated with faecal counts account for a range of 5 and 31% of genetic variance. It is important to note that, together, SNPs OAR3_220227474.1 and s15560.1 accounted for over 50% of the genetic variance of FEC_N during the 1st recording occasion. This result is consistent with Davies et al (2006), who reported several QTLs associated with FEC traits at varying time points with respective estimates of proportions of genetic variance varying from 26 to 79%. In contrast, in a study by Keane et al. (2018), the proportion of genetic variance explained by significant SNPs associated with *Trichostrongyle* FEC and *Nematodirus* FEC varied between 1.6 and 12.7%. Estrada-Reyes et al., (2019) found that individual SNPs explained a proportion of genetic variance of FEC ranging from 10 to 15% in a Florida native sheep breed. While it is possible that the size of SNP effects in genomic studies is overestimated (Mucha et al., 2015), the results reported here reveal the presence of notable signals in the respective genomic regions and may harbour genes that play a substantial role in animal resistance to parasitic infection.

Indeed, multiple genes were identified in our post-GWAS analyses that are linked with the above-mentioned SNPs. Among them are several genes located on OAR3 and in proximity to SNP OAR3_220227474.1 (associated with FEC_N), encoding C-type lectins (**CTLs**), which are involved in the recognition of pathogen glycans as innate immune receptors and are involved in the induction of inflammation, and therefore responsible for pro-inflammatory responses. CTLs are involved in immune responses, from initial recognition and uptake to the modulation of adaptive immunity (Drummond and Brown, 2013). The influence of other CTL-associated alleles on immune responses against parasitic infection has been described in semi-feral Soay sheep, with the *CLEC16A* gene being strongly correlated with specific IgA levels against *T. circumcincta* in both lambs and mature sheep (Sparks et al., 2019). CTLs were shown to be up-regulated in abomasal mucosa of immune cattle following infection with *Ostertagia ostertagi* and *Cooperia oncophora*, including *CLEC12A* (Li et al., 2011), supporting their role in invoking host immune responses and the development of resistance against pathogens. Also, in proximity to SNP OAR3_220227474.1, we report several receptors for natural killer cell genes, represented by members of the killer cell lectin-like receptor family, capable of transmitting activating or inhibitory signals. Associations between polymorphisms within the CD94/NKG2 cluster and various diseases have been reported (Iwaszko and Bogunia-Kubik, 2011). Innately activated NK cells produce Interleukin (IL)-13 (De Veer et al., 2007), a cytokine involved in Th2 immune responses.

The gene *LGALS4*, an inhibitor of inflammation found in proximity to SNP s15560.1 (associated with FEC_N) codes Galectin-4, a protein selectively expressed within the GI tract, which has been shown to ameliorate signs of inflammation (Paclik et al., 2008). *LGALS4* has been identified as one of three genes belonging to the galectin family to be GIN-activated and up-regulated in resistant

sheep infected with *T. circumcincta* (Chitneedi et al., 2018). Similarly, Galectin-4 was shown to be significantly up-regulated in naïve yearling lambs after being infected with *T. circumcincta* L3 (Chitneedi et al., 2018). The gene *CD80* (close to SNP OAR1_197073884.1) is responsible for encoding CD80 molecules, which are involved in T-cell activation by antigen-presenting cells (Tatari-Calderone et al., 2002). Down-regulation of *CD80* on the surface of antigen-presenting cells is capable of interfering with the generation and maintenance of T cell responses (McNeilly et al., 2013).

The *BLNK* gene in proximity to an SNP associated with DAG is a regulator of B cell receptor signalling. This gene encodes a central linker protein, with a function in regulating the biological outcome of B cell function and development (Fu et al., 1998). *BLNK* was identified as being one of the genes associated with super-shedding of E. coli in cattle (Wang et al., 2017).

Immunological traits – Cytokines

The proportion of genetic variance explained by the four SNPs in association with cytokines ranges from 2 to 15%. These results are within the range of the results reported by Ahola-Olli et al., (2017), where proportions of genetic variance of multiple cytokines explained by SNPs varied from 1 to 34%. Three noteworthy genes were identified on OAR20 near the SNP OAR20_33522555.1, which was associated with IFN-$\gamma_{(PWM)}$ in the present study. *BTN1A1* and *BTN2A2*, known inhibitors of CD4 and CD8 T cell proliferation, belong to the Butyrophilin group of MHC-associated protein-encoding genes. These proteins are capable of mediating complex interactions between antigen-presenting cells and T cells (Rhodes et al., 2016). Evidence suggests that butyrophilin-like proteins may play a role as regulators of intestinal inflammation (Yamazaki et al., 2010). *BTN1A1* and *BTN2A2* have been shown to inhibit the expression of cytokines involved in T cell metabolism and activation, in particular the secretion of IFN-γ (Arnett and Viney, 2014). Additionally, the *HFE* gene, with a role in the negative regulation of T cell antigen processing and presentation and T cell cytokine production, and also a negative regulator of antigen processing and presentation of endogenous peptide antigen has been primarily described for its role in iron metabolism and possesses a remarkable structural homology with MHC molecules (Reuben et al., 2017). Restriction of iron absorption by pathogens is one of the host's defences against parasitic and bacterial growth, with evidence suggesting microbes are capable of regulating their host's iron metabolism to escape immune surveillance (Liu et al., 2021). In mice infected with *Salmonella*, IFN-γ expression limits iron availability to the microbes and strengthens macrophage immune function (Nairz et al., 2008). IFN-γ is one of the main cytokines deployed in immune responses to pathogens, being responsible for the up-regulation of MHC I expression to enhance cytotoxic T lymphocytes and may also induce components of the antigen processing pathway (Reuben et al., 2017).

Immunological traits – Immunoglobulin A

The proportion of IgA genetic variation explained by six SNPs with significant effects ranged from 4% to 12%, results comparable to the reported proportions of genetic variance of IgA explained by three SNPs (between 10 and 20%) in a population of Soay sheep (Sparks et al., 2019). In an earlier study, Davies et al., (2006) reported a higher proportion of genetic variance explained by two significant QTLs for IgA that varied between 41 and 51%.

Multiple genes were found to be associated with IgA in the present study, including *IL17A* and *IL17F*, which encode Th17-related cytokines that are up-regulated as a consequence of Th17 differentiation. In the GI tract, *IL17A* is necessary for protective immunity against several pathogens, being an important regulator of mucosal immune defence associated with inflammation (Dann et al., 2015) and regulates the secretion of parasite-specific intestinal IgA (Paerewijck et al., 2019). *IL17F* shares the highest sequence homology with *IL17A* and is a weaker inducer of pro-inflammatory cytokine expression (Samiei et al., 2018). IL-17 has been shown to be a contributing cytokine involved in the increased secretion of IgA in the intestine of chickens (Karaffová et al., 2015).

Finally, several genes encoding members for the Interleukin-1 family of cytokines were also associated with IgA. This cytokine is implicated in acute and chronic inflammatory diseases, playing a significant role in immune regulation and inflammatory responses (Hahn et al., 2009). *IL1A* and *IL1B*, involved in pro-inflammatory responses, respectively encode IL-1α, which represents a known agonist of the IL-1 receptor and has a pro-inflammatory action (Hahn et al., 2009), and IL-1β, one of the typical cytokines mediating Th1 immune responses, alongside with cytokines such as IFN-γ (Rad et al., 2004). The *IL1RN* gene encodes a receptor antagonist which inhibits IL-1α and IL-1β function, neutralising the activity of these cytokines in immune responses (Cruz-Robles et al., 2009). The *IL1F10* gene regulates adaptive and innate immune responses by inhibiting the production of T cell cytokines (Fonseca-Camarillo et al., 2018) driving the proliferation of Tregs, preventing them from transforming into Th17 (Xie et al., 2019). Moreover, the *IL36A*, *IL36B* and *IL36RN* genes encode two IL-36 cytokines with agonistic activity (IL-36α and IL-36β) and IL-36Ra, a receptor antagonist, which are responsible for the regulation of IL-36 signalling (Saito et al., 2020). The *IL37* gene, also a member of the IL-1 cytokine family**,** encodes IL-37 cytokine with a role in limiting innate inflammation and in suppressing acquired immunity (Dinarello et al., 2016). The IL-1 family of cytokines is considered to play crucial roles in the progression of IgA nephropathy (Hahn et al., 2009).

Conclusions

The present study identified multiple potential candidate genes with known biological functions related to the studied animal traits. We note the importance of several C-type lectins associated with animal resistance to *Nematodirus* infection and several genes encoding a number of cytokines belonging to the IL-1 family and IL-17 were associated with IgA levels. The present study revealed a largely complex and polygenic genetic control on resistance to parasitic infection and immunological traits in the studied population. Nevertheless, certain genomic regions identified may hold larger genomic loads than others, thus requiring further attention. Insights into these regions may increase the accuracy of genomic evaluation and selection towards enhancing disease resistance. Future work should examine the significance of our results on independent data and further investigate the reported genomic regions associated with each trait.

Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.animal.2023.101069.

Ethics approval

All experiments were approved by Scotland's Rural College Animal Experiments Committee and were performed according to Home Office Guidelines under Project Licence numbers 60/4358 and P90111799.

Data and model availability statement

The genetic model used for the analyses was not deposited in an official repository. Data on animal performance, pedigree and health traits are maintained in a secure SQL database at Scotland's Rural College and are available upon request from the corresponding author.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors did not use any artificial intelligence-assisted technologies in the writing process.

Author ORCIDs

António Pacheco: https://orcid.org/0000-0002-7920-9362.
Georgios Banos: https://orcid.org/0000-0002-7674-858X.
Nicola Lambe: https://orcid.org/0000-0002-4251-4742.
Ann McLaren: https://orcid.org/0000-0001-8515-9974.
Tom N. McNeilly: https://orcid.org/0000-0001-6469-0512.
Joanne Conington: https://orcid.org/0000-0002-2387-3555.

CRediT authorship contribution statement

A. Pacheco: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **G. Banos:** Writing – review & editing, Supervision, Software, Methodology, Funding acquisition, Conceptualization. **N. Lambe:** Writing – review & editing. **A. McLaren:** Writing – review & editing. **T.N. McNeilly:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **J. Conington:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Data curation, Conceptualization.

Declarations of interest

None.

Acknowledgements

The technical expertise of Maureen Steel is gratefully acknowledged along with the support from other Scotland's Rural College farm and technical staff.

Financial support statement

AP was supported by a PhD studentship jointly funded by SRUC and the Moredun Foundation. The resource flock, data collection and analyses were funded by the Rural & Environment Science & Analytical Services (RESAS) Division of the Scottish Government. GB & JC were also supported by the European Commission SMARTER project (grant number 772787).

References

Acevedo-Whitehouse, K., Cunningham, A.A., 2006. Is MHC enough for understanding wildlife immunogenetics? Trends in Ecology and Evolution 21, 433–438.

Ahola-Olli, A.V., Würtz, P., Havulinna, A.S., Aalto, K., Pitkänen, N., Lehtimäki, T., Kähönen, M., Lyytikäinen, L.-P., Raitoharju, E., Seppälä, I., Sarin, A.-P., Ripatti, S., Palotie, A., Perola, M., Viikari, J.S., Jalkanen, S., Maksimow, M., Salomaa, V., Salmi, M., Kettunen, J., Raitakari, O.T., 2017. Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. The American Journal of Human Genetics 100, 40–50.

Amin, N., van Duijn, C.M., Aulchenko, Y.S., 2007. A genomic background based method for association analysis in related individuals. PLoS ONE 2, e1274.

Arnett, H.A., Viney, J.L., 2014. Immune modulation by butyrophilins. Nature Reviews Immunology 14, 559–569.

Atlija, M., Arranz, J.J., Martinez-Valladares, M., Gutiérrez-Gil, B., 2014. Search of genomic regions influencing faecal egg count, as an indicator of resistance to gastrointestinal nematode infections, based on the analysis of the OvineSNP50 BeadChip. In: Proceedings of 10th World Congress of Genetics Applied to Livestock Production, 17–22 August 2014, Vancouver, Canada, 523.

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. Molecular Ecology 22, 757–773.

Chitneedi, P.K., Suárez-Vega, A., Martínez-Valladares, M., Arranz, J.J., Gutiérrez-Gil, B., 2018. Exploring the mechanisms of resistance to Teladorsagia circumcincta infection in sheep through transcriptome analysis of abomasal mucosa and abomasal lymph nodes. Veterinary Research 49, 39.

Conington, J., Bishop, S.C., Grundy, B., Waterhouse, A., Simm, G., 2001. Multi-trait selection indexes for sustainable UK hill sheep production. Animal Science 73, 413–423.

Conington, J., Bishop, S.C., Lambe, N.R., Bünger, L., Simm, G., 2006. Testing selection indices for sustainable hill sheep production – lamb growth and carcass traits. Animal Science 82, 445–453.

Cruz-Robles, D., Chávez-Gonzlez, J.P., Cavazos-Quero, M.M., Prez-Mndez, O., Reyes, P.A., Vargas-Alarcn, G., 2009. Association between IL-1B and IL-1RN gene polymorphisms and chagas' disease development susceptibility. Immunological Investigations 38, 231–239.

Dann, S.M., Manthey, C.F., Le, C., Miyamoto, Y., Gima, L., Abrahim, A., Cao, A.T., Hanson, E.M., Kolls, J.K., Raz, E., Cong, Y., Eckmann, L., 2015. IL-17A promotes protective IgA responses and expression of other potential effectors against the lumen-dwelling enteric parasite Giardia. Experimental Parasitology 156, 68–78.

Davies, G., Stear, M.J., Benothman, M., Abuagob, O., Kerr, A., Mitchell, S., Bishop, S.C., 2006. Quantitative trait loci associated with parasitic infection in Scottish blackface sheep. Heredity 96, 252–258.

De Veer, M.J., Kemp, J.M., Meeusen, E.N.T., 2007. The innate host defence against nematode parasites. Parasite Immunology 29, 1–9.

Dinarello, C.A., Nold-Petry, C., Nold, M., Fujita, M., Li, S., Kim, S., Bufler, P., 2016. Suppression of innate inflammation and immunity by interleukin-37. European Journal of Immunology 46, 1067–1081.

Drummond, R.A., Brown, G.D., 2013. Signalling C-type lectins in antimicrobial immunity. PLoS Pathogens 9, e1003417.

Estrada-Reyes, Z.M., Rae, O., Postley, C., Jiménez Medrano, M.B., Leal Gutiérrez, J.D., Mateescu, R.G., 2019. Association study reveals Th17, Treg, and Th2 loci related to resistance to Haemonchus contortus in Florida Native sheep. Journal of Animal Science 97, 4428–4444.

Fonseca-Camarillo, G., Furuzawa-Carballeda, J., Iturriaga-Goyon, E., Yamamoto-Furusho, J.K., 2018. Differential expression of IL-36 family members and IL-38 by immune and nonimmune cells in patients with active inflammatory bowel disease. BioMed Research International 2018, 1–12.

Fu, C., Turck, C.W., Kurosaki, T., Chan, A.C., 1998. BLNK: a central linker protein in B cell activation. Immunity 9, 93–103.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., 2009. ASReml user guide release 3.0. In: VSN International Ltd, Hemel Hempstead, United Kingdom. www.vsni.co.uk.

Gutiérrez-Gil, B., Pérez, J., Álvarez, L., Martínez-Valladares, M., de la Fuente, L.-F., Bayón, Y., Meana, A., Primitivo, F.S., Rojo-Vázquez, F.-A., Arranz, J.-J., 2009. Quantitative trait loci for resistance to trichostrongylid infection in Spanish Churra sheep. Genetics Selection Evolution 41, 46.

Hahn, W.H., Cho, B.S., Kim, S.D., Kim, S.K., Kang, S., 2009. Interleukin-1 cluster gene polymorphisms in childhood IgA nephropathy. Pediatric Nephrology 24, 1329–1336.

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols 4, 44–57.

Iwaszko, M., Bogunia-Kubik, K., 2011. Clinical Significance of the HLA-E and CD94/NKG2 Interaction. Archivum Immunologiae et Therapiae Experimentalis 59, 353–367.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Research 42, D199–D205.

Karaffová, V., Bobíková, K., Husáková, E., Levkut, M., Herich, R., Revajová, V., Levkutová, M., Levkut, M., 2015. Interaction of TGF-β4 and IL-17 with IgA secretion in the intestine of chickens fed with *E. faecium* AL41 and challenged with *S. Enteritidis*. Research in Veterinary Science 100, 75–79.

Keane, O.M., Hanrahan, J.P., McRae, K.M., Good, B., 2018. An independent validation study of loci associated with nematode resistance in sheep. Animal Genetics 49, 265–268.

Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., Kersey, P., Flicek, P., 2011. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database : The Journal of Biological Databases and Curation 2011, bar030.

Li, R.W., Rinaldi, M., Capuco, A.V., 2011. Characterization of the abomasal transcriptome for mechanisms of resistance to gastrointestinal nematodes in cattle. Veterinary Research 42, 114.

Liu, J., Wu, X., Wang, H., Wei, J., Wu, Q., Wang, X., Yan, Y., Cui, J., Min, J., Wang, F., Zhou, J., 2021. HFE inhibits type I IFNs signaling by targeting the SQSTM1-mediated MAVS autophagic degradation. Autophagy 17, 1962–1977.

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., Matthews, J.B., 2013. Suppression of ovine lymphocyte activation by Teladorsagia circumcincta larval excretory-secretory products. Veterinary Research 44, 1–18.

Meuwissen, T., Hayes, B., Goddard, M., 2016. Genomic selection: A paradigm shift in animal breeding. Animal Frontiers 6, 6.

Mucha, S., Bunger, L., Conington, J., 2015. Genome-wide association study of footrot in Texel sheep. Genetics Selection Evolution 47, 35.

Nairz, M., Fritsche, G., Brunner, P., Talasz, H., Hantke, K., Weiss, G., 2008. Interferon-γ limits the availability of iron for intramacrophage Salmonella typhimurium. European Journal of Immunology 38, 1923–1936.

Paclik, D., Danese, S., Berndt, U., Wiedenmann, B., Dignass, A., Sturm, A., 2008. Galectin-4 controls intestinal inflammation by selective regulation of peripheral and mucosal T cell apoptosis and cell cycle. PLoS ONE 3, e2629.

Paerewijck, O., Maertens, B., Gagnaire, A., De Bosscher, K., Geldhof, P., 2019. Delayed development of the protective IL-17A response following a Giardia muris infection in neonatal mice. Scientific Reports 9, 8959.

Purcell, S., Neale, B., Todd-brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., 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. The American Journal of Human Genetics 81, 559–575.

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Rad, R., Dossumbekova, A., Neu, B., Lang, R., Bauer, S., Saur, D., Gerhard, M., Prinz, C., 2004. Cytokine gene polymorphisms influence mucosal cytokine expression, gastric inflammation, and host specific colonisation during Helicobacter pylori infection. Gut 53, 1082–1089.

Reuben, A., Chung, J.W., Lapointe, R., Santos, M.M., 2017. The hemochromatosis protein HFE 20 years later: An emerging role in antigen presentation and in the immune system. Immunity, Inflammation and Disease 5, 218–232.

Rhodes, D.A., Reith, W., Trowsdale, J., 2016. Regulation of Immunity by Butyrophilins. Annual Review of Immunology 34, 151–172.

Rupp, R., Mucha, S., Larroque, H., McEwan, J., Conington, J., 2016. Genomic application in sheep and goat breeding. Animal Frontiers 6, 39.

Saito, K., Iwata, Y., Fukushima, H., Watanabe, S., Tanaka, Y., Hasegawa, Y., Akiyama, M., Sugiura, K., 2020. IL-36 receptor antagonist deficiency resulted in delayed wound healing due to excessive recruitment of immune cells. Scientific Reports 10, 14772.

Samiei, G., Yip, W.K., Leong, P.P., Jabar, M.F., Dusa, N.M., Mohtarrudin, N., Seow, H.F., 2018. Association between polymorphisms of interleukin-17A G197A and interleukin-17F A7488G and risk of colorectal cancer. Journal of Cancer Research and Therapeutics 14, S299–S305.

Sparks, A.M., Watt, K., Sinclair, R., Pilkington, J.G., Pemberton, J.M., McNeilly, T.N., Nussey, D.H., Johnston, S.E., 2019. The genetic architecture of helminth-specific immune responses in a wild population of Soay sheep (Ovis aries). PLOS Genetics 15, e1008461.

Tatari-Calderone, Z., Semnani, R.T., Nutman, T.B., Schlom, J., Sabzevari, H., 2002. Acquisition of CD80 by human T cells at early stages of activation: Functional involvement of CD80 acquisition in T cell to T cell interaction. The Journal of Immunology 169, 6162–6169.

The Gene Ontology Consortium, 2021. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Research 49, D325–D334.

VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. Journal of Dairy Science 91, 4414–4423.

VanRaden, P.M., O'Connell, J.R., Wiggans, G.R., Weigel, K.A., 2011. Genomic evaluations with many more genotypes. Genetics Selection Evolution 43, 10.

Wang, O., McAllister, T.A., Plastow, G., Stanford, K., Selinger, B., Guan, L.L., 2017. Host mechanisms involved in cattle Escherichia coli O157 shedding: A fundamental understanding for reducing foodborne pathogen in food animal production. Scientific Reports 7, 1–15.

Xie, L., Huang, Z., Li, H., Liu, X., Zheng, S., Su, W., 2019. IL-38: A new player in inflammatory autoimmune disorders. Biomolecules 9, 345.

Yamazaki, T., Goya, I., Graf, D., Craig, S., Martin-Orozco, N., Dong, C., 2010. A butyrophilin family member critically inhibits T cell activation. The Journal of Immunology 185, 5907–5914.

Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics 44, 821–824.