Identification of Variation in the Sheep Genome Controlling Gastrointestinal Nematode Resistance

Kathryn M. McRae

B.Sc. Genetics & Zoology, M.Sc. Genetics

A thesis presented to Dublin City University for the Degree of Doctor of Philosophy

Research Supervisors:

Dr Orla M. Keane[1], Dr Barbara Good[2] and Dr Mary J. O'Connell[3]

[1]Animal & Bioscience Research Department, Teagasc Grange, Dunsany, Co. Meath, Ireland

[2]Animal Production Research Centre, Teagasc Mellows Campus, Athenry, Co. Galway, Ireland

[3]School of Biotechnology, Dublin City University Glasnevin, Dublin 9, Ireland

January 2015

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: 

I.D. No.: 10120858

Date: 14/01/2015

Acknowledgements

First and foremost thank you to my supervisors, Orla, Barbara, Mary and Seamus for your patience, guidance and support. Orla, I am incredibly thankful that you gave me the opportunity to work with you - I have learnt so much during my time in your group. Barbara, thank you for teaching me the ins and outs of parasitology, and for being the calm voice of reason when things went pear-shaped! Mary, thank you so very much for always making me feel welcome at DCU, and for always taking time to read over anything I have written (even though it was mostly about sheep and parasites). Seamus, I am incredibly grateful your your statistical advice, your patience, and of course your editing expertise. This thesis most definitely would not have been the same without your input!

I must also gratefully acknowledge the assistance of many people from Teagasc Grange, Teagasc Athenry, and the BME group at DCU, particularly Assumpta Glynn, Margaret Murray, Joe Larkin, Paul Cormican, Andrew Webb, Torres Sweeney, Simone Sebastiano, Henry Walsh, Luke O'Malley, P.J. Hastings and the farm staff at Teagasc Athenry. Without all of your help and expertise this thesis would not have been completed. A massive thank you as well to those who helped out at the slaughters, particularly the other Walsh Fellows. Funding is gratefully acknowledged from Teagasc RMIS 6001 and the Allan & Grace Kay Overseas Scholarship Trust.

I have been lucky to have been a part of two lab groups during my PhD. To Katie and Jason, it has been an absolute pleasure to share an office with you both. You've helped me stay sane during the 'long road that is a PhD', and I'm really going to miss the general hilarity, the chats, the advice, the coffee, and the Monday morning de-briefs. Andrew, Ann, Ray and Edel - thanks to you all I have always enjoyed coming in to spend time with the BME group, and I have fond memories of the banter and the laughs that went along with it.

Many people have made Ireland feel like home during my time here, and while I cannot name you all, there are a few I cannot go without mentioning. To my 'Irish Mum', Assumpta, I absolutely loved my time spent in the lab with you in Athenry. You were always on hand with stories, laughs, a listening ear, hugs, and a lot of good advice. To the girls on the Navan Ladies team, particularly Susan, Cheesey and Ciara, have been absolutely fantastic. No other rugby team will ever be the same. Last but in no way least, a special mention must go to Brian, Austen, Claire, Nicole, and Jenna. You guys are my fondest memories of my time here, and you really did make it feel like home. I'm going to miss you all.

Finally, the biggest thanks of all must go to my family for their unwavering support. Even though I have missed multiple family gatherings, engagements and births, you have been nothing but supportive of my living so far away for the past three and a half years. I cannot wait to spend some quality time with you, particularly the new additions. Tim, Matt and Joe - believe it or not, you can *finally* say your sister is no longer a student! And to Mum, this one is for you. You have always been there for me on the other side of the phone, encouraging me, and making me believe I can do it. I cannot think of a better role model for anybody to have. Thank you.

Contents

List of Figures

List of Tables

Index of Electronic Appendices

Chapter 2

Electronic Appendix 2.1: Selection differentials for all Scottish Blackface lambs

Chapter 3

Electronic Appendix 3.1: TruSeq RNA Sample Preparation v2 Guide

Chapter 4

Electronic Appendix 4.1: Scripts used during the selective pressure analysis

Electronic Appendix 4.2: Commands for the above scripts

Electronic Appendix 4.3: Multiple sequence alignments

Electronic Appendix 4.4: Maximum likelihood trees from RAxML, and the resuting alignments from the selected subtrees

Electronic Appendix 4.5: Consensus sequence for HighFEC and LowFEC Scottish Blackface animals

Electronic Appendix 4.6: Selective pressure analysis results

Chapter 5

Electronic Appendix 5.1: Buffers and solutions

Electronic Appendix 5.2: ASReml heritability code and input files

Electronic Appendix 5.3: GenABEL code and input files

Electronic Appendix 5.4: ASReml single SNP analysis code and input files

List of abbreviations

aa	Amino acid
AD	Amino-acetonitrile derivatives (Monepantel)
ALN	Abomasal lymph node
APC	Antigen presenting cell
AR	Anthelmintic resistance
bp	Base pair
BSA	Bovine serum albumin
BZ	Benzimidazole
BV	Breeding value
℃	Degree Celsius
CarLA	Carbohydrate larval surface antigen
CLR	C-type lectin receptors
cDNA	Complementary DNA
cm	Centimetre
CPM	Count per million
CT	Condensed tannins
CV	Coefficient of variation
DAMP	Damage-associated molecular pattern molecule
DE	Differentially expressed
DNA	Deoxyribonucleic acid
dpi	Day post infection
ds	Double stranded
EBV	Estimated breeding value
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
epg	Eggs per gram
EPGP	Eosinophil secondary granule protein
EtOH	Ethanol
FDR	False discovery rate
FEC	Faecal egg count
g	Gram

x g	Force of gravity
GIN	Gastrointestinal nematode
GWAS	Genome-wide association study
h	Hour
H_2O	Water
HCl	Hydrochloric acid
IFNγ	Interferon gamma
Ig	Immunoglobulin
IL	Interleukin
IPA	Ingenuity® Systems Pathway Analysis
kb	Kilobase
L	Litre
L3	3rd larval stage
L4	4th larval stage
LA	Linkage analysis
LD	Linkage disequilibrium
LMIT	Larval migration inhibition test
LPS	Lipopolysaccharides
LV	Imidazothiazole
µg	Microgram
µL	Microlitre
µm	Micrometer
µM	Micromolar
mg	Milligram
mL	Millilitre
mm	Micrometre
mM	Millimolar
ML	Macrocyclic lactone
MALDT	Micro-agar larval development test
MDS	Multi-dimensional scaling
MHC	Major histocompatibility complex
min	Minute
mRNA	Messenger RNA
NaCl	Sodium chloride
ng	Nanogram
NSC	Norwegian Sequencing Centre
OD	Optical density
OLA	Ovine lymphocyte antigen
%	Percent
PAMP	Pathogen-associated molecular pattern molecule

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PPR	Periparturient rise
PRR	Pattern recognition receptor
PS	Positively selected
qRT-PCR	Quantitative real-time PCR
QTL	Quantitative trait loci
RHM	Regional heritability mapping
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-Seq	RNA-sequencing
RO	Reverse osmosis
s	Second
SCOPS	Sustainable Control of Parasites in Sheep
s.d.	Standard deviation
s.e.	Standard error
SLA	Single locus association
SNP	Single nucleotide polymorphism
SI	Spiroindole
TCA	Trichloroacetic acid
TCR	T cell receptor
TE	Tris-EDTA
T_H	T helper cell
T_{reg}	Regulatory T cell
TLR	Toll-like receptor
TMM	Trimmed mean of M-values
TNF	Tumour necrosis factor
U/l	µM tyrosine released per 1 L plasma per minute

Abstract

The objective of this thesis was to identify Scottish Blackface lambs that differed in their resistance to gastrointestinal nematode (GIN) infection, characterise the host responses to infection in resistant and susceptible lambs, and identify genes and biological processes important for the difference in resistance status.

An animal selection model was developed that reliably identified Scottish Blackface lambs that differed in resistance to GIN. After a controlled challenge with *Teladorsagia circumcincta*, resistant (low faecal egg count; FEC) animals displayed consistently lower FEC throughout the course of infection. This was largely a result of worm fecundity differences, with resistant animals containing shorter, less fecund adult females. There was also a significant correlation between the number of adult worms and FEC at slaughter. The anti-nematode response was mediated, at least in part, by IgA, with resistant animals having significantly higher levels of serum anti-nematode IgA throughout the infection. Taken together, these results indicate lower FEC in resistant Scottish Blackface lambs is primarily a result of reduced worm fecundity, although lower adult worm burden may also play a role.

Gene expression in the abomasal lymph node was examined at 7 and 14 days post infection. A total of 194 and 144 genes were differentially expressed between resistant and susceptible lambs at 7 and 14 dpi respectively. At 7 dpi resistant animals appear to be generating a more effective immune response, whereas in susceptible animals this response is delayed until ~14 days post infection.

The genes differentially expressed between resistant and susceptible animals were examined for evidence of selective pressure. A number of genes showed evidence of *Ovis* lineage-specific positive selection. Pathway analysis revealed that these genes were involved in the inflammatory response, dermatological diseases and conditions, and connective tissue disorders. This chapter represents the first large-scale comparative genomics study of selective pressure placed on the sheep genome, in particular by gastrointestinal nematodes.

Association testing in the Scottish Blackface population was carried out using 57 markers in 13 candidate genes and 1000 markers in 7 genomic regions. A number of suggestive associations with FEC and weight traits were observed, however none were significant at the genome-wide level. The results from this study support the use of a panel of SNPs rather than individual SNPs for predicting nematode resistance, in agreement with the complexity of this polygenic trait.

Chapter 1

Gastrointestinal parasitism in sheep

1.1 Overview

Infection with gastrointestinal nematodes (GIN), resulting in clinical disease and loss of productivity, is one of the major constraints of ruminant production worldwide (Sutherland and Scott, 2009). Current control strategies in both sheep and cattle rely heavily on the use of anthelmintic treatment; in 2009 Irish farmers spent €24 million on internal parasite control (www.apha.ie/about_markets.asp), while worldwide the annual expenditure on anthelmintic treatment was over 3 billion $US (Jackson et al., 2009).

The increasing prevalence of anthelmintic resistance in sheep nematodes throughout the world (Sargison, 2012), including Ireland (McMahon et al., 2013b; Good et al., 2012; Patten et al., 2007), suggests that reliance on chemotherapy is unsustainable. Selective breeding for host resistance is an alternative, sustainable method of nematode control. The most commonly used indicator of host resistance is faecal egg count (FEC), which is moderately heritable (h^2 ~ 0.3), with a wide variability between individuals (Bishop and Morris, 2007; Safari et al., 2005). Rapid genetic progress has been demonstrated in selective breeding programmes using both experimental and commercial flocks (Morris et al., 2000, 2005; Woolaston et al., 1991; Windon, 1990; Cummins et al., 1991; Greeff and Karlsson, 2006; Morris et al., 1997). However, selection on FEC requires detailed trait measurement, which is time-consuming, expensive and unappealing. Selecting resistant animals would be simplified if animals could be selected by genotype; this could also accelerate genetic gain. A detailed understanding of the genes and mechanisms involved in protective immunity and the factors that regulate this response would aid future breeding strategies as well as the development of other effective and sustainable nematode control methods, such as immunomodulatory anthelmintics.

In Ireland, two lowland sheep breeds have been extensively studied with respect to nematode resistance and it has been demonstrated that the Texel breed is more resistant to nematode infection than the Suffolk breed (Hanrahan and Crowley, 1999; Sayers et al., 2005a; Good et al., 2006; Sayers et al., 2008; Hassan et al., 2011a). However, there have been no studies to date in Ireland on breeds that occupy hill and marginal land. Extensive studies in Scotland have shown substantial genetic variation among Scottish Blackface lambs in both FEC and in worm length (Stear et al., 1999b). This breed is commonly found on hill country in Ireland, and is also an important source of crossbred ewes for the lowland sheep sector. For these reasons the Scottish Blackface breed was the focus of this study. The aim of work reported in this thesis was to identify genes and biological process mediating the response to gastrointestinal nematodes, primarily *Teladorsagia circumcincta*, in Scottish Blackface lambs.

1.2 Gastrointestinal nematode control

1.2.1 *Teladorsagia* (*Ostertagia*) *circumcincta*

In temperate climates, such as Ireland, the economically important species of GIN in sheep are *Teladorsagia* (*Ostertagia*), *Trichostrongylus*, *Nematodirus* and increasingly *Haemonchus contortus* (Good et al., 2006; Burgess et al., 2012). In Ireland *Teladorsagia circumcincta* and *Nematodirus* are the predominant species of GIN in sheep (Good et al., 2001), with *T. circumcincta* impacting most negatively on sheep production (Good et al., 2006; Patten et al., 2011). As such, *T. circumcincta* was the focus of this thesis and therefore this review.

1.2.1.1 Life cycle

The life cycle of *T. circumcincta* is similar to other GIN, such as *Trichostrongylus* and *H. contortus*, and consists of both a free-living and a parasitic stage (Figure 1.1).

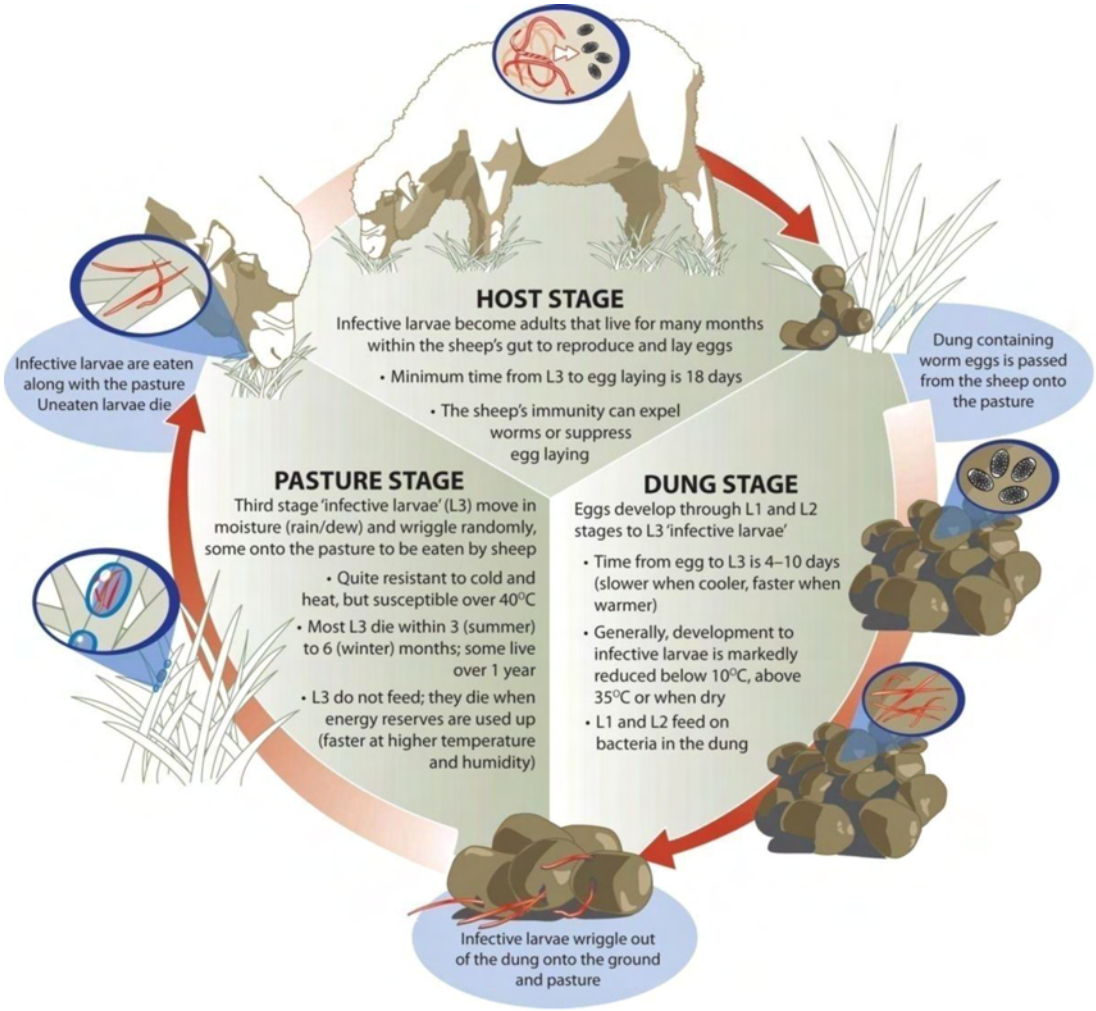


Figure 1.1: The typical life cycle of a sheep gastrointestinal nematode. Details may vary between species. Figure used with kind permission from www.WormBoss.com.au

In brief, under optimal conditions mature female worms in the gastrointestinal tract lay eggs, which are passed out in the faeces and develop into infective third stage (L3) larvae on pasture within two weeks. Free-living L3 are ingested by grazing sheep, and exsheath in the rumen before entering the gastric glands of the abomasum and undergoing two subsequent moults (Figure 1.2). Fifth stage (L5) larvae leave the abomasal gland approximately 10 days post-infection and become sexually mature on the mucosal surface. Females produce eggs, which are passed out in faeces, completing the life cycle (McNeilly et al., 2009). The entire life cycle usually takes approximately 4 to 5 weeks, although this dependant on temperature and immune status of the host. Eggs can be produced as quickly as 14 days after larval ingestion in naïve animals, but resistant hosts are able to delay larval maturation for at least a further 8 weeks (Stear et al., 1995c).

The optimal temperature range for the development of *T. circumcincta* eggs through to the infective L3 stage ranges between 16 ℃ and 30 ℃ (O'Connor et al., 2006). Development can occur at temperatures as low as 4 ℃ (Crofton, 1965), although it is greatly reduced when daily air temperature falls below 10 ℃ (Reynecke et al., 2011). Once at the L3 stage the larvae can over-winter on pasture and still remain infective (Vlassoff, 1973; Kerboeuf, 1985; O'Connor et al., 2006; Southcott et al., 1976). Once ingested, if conditions are unfavourable, early L4 can enter the abomasal mucosa (Figure 1.2) and undergo arrested development for up to 6 months in a process known as hypobiosis (Taylor et al., 2007). Hypobiosis primarily occurs when there is insufficient water in the environment or temperatures are too cold for larval development and survival (Miller and Horohov, 2006).

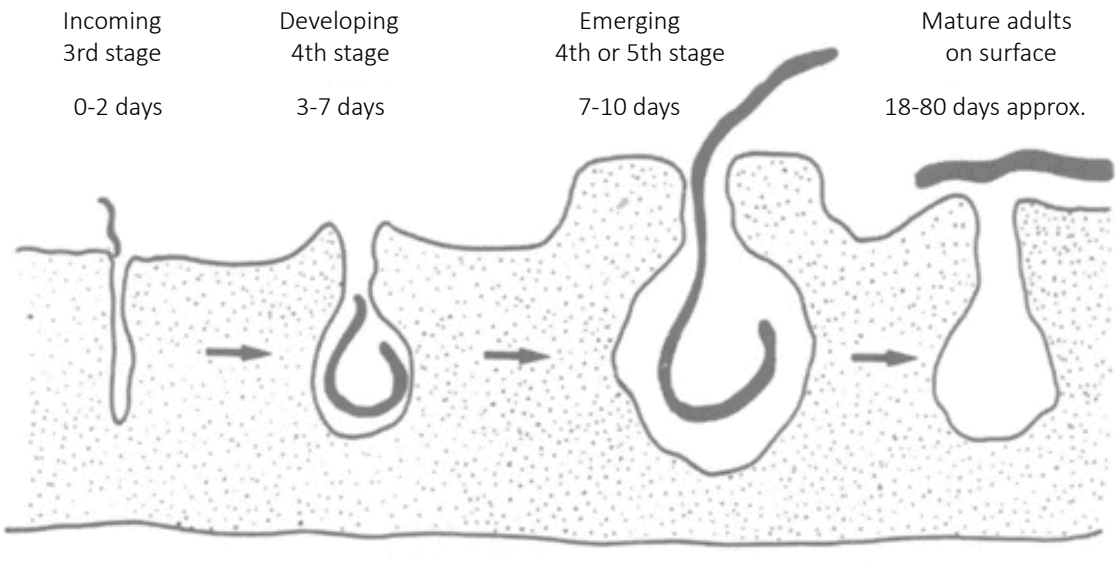


Figure 1.2: The parasitic phases of *Teladorsagia circumcincta* in the mucosa of the abomasum. Hypobiosis (arrested development) can occur at the early 4th stage sometime after the moult. Figure used with kind permission from Smith (1988).

1.2.1.2 Pathogenesis

Pathogenesis in young susceptible lambs following infection with *T. circumcincta* is reasonably well understood (Armour et al., 1966; Coop et al., 1982, 1985; Stear et al., 2003). Under normal conditions, pepsinogen is secreted into the lumen of the abomasum by chief cells, and is activated to pepsin by hydrochloric acid (HCl), which is released from parietal cells in the stomach lining. However, as mature parasites emerge from the gastric glands there is considerable damage to the abomasal mucosa. The parasites also release molecules that trigger mast cell degranulation, which results in the destruction of junctions between epithelial cells (McKellar, 1993). Therefore, infective GIN larvae damage the parietal cells, leading to a decline in HCl production and a subsequent failure of pepsinogen to convert to pepsin, resulting in increased pepsinogen passing into the circulatory system, and reduced protein digestion (Lawton et al., 1996; McKellar, 1993). Plasma pepsinogen levels can therefore be used as an indicator of the severity of infection (Lawton et al., 1996; Stear et al., 1999a).

After natural or experimental *T. circumcincta* infection there is a parasite-induced protein deficiency in the host which has multiple causes; a decrease in digestive efficiency, as outlined above, a reduction in voluntary feed intake (Coop et al., 1982; Sykes and Poppi, 1986; Coop and Kyriazakis, 1999), protein leakage through abomasal damage, and the utilisation of proteins for the immune response (Stear et al., 1997). Consequently, infection with *T. circumcincta* reduces growth rate in young lambs (Coop et al., 1982; Bishop et al., 1996; Bouix et al., 1998; Bishop and Stear, 2000a, 2001). The consequences of a high worm burden can be long-lasting, as mucosal damage can result in reduced growth rates despite clearance of infection through anthelmintic treatment (Coop et al., 1982, 1985). It has also been observed that the nutritional status of the host during infection is important, with the provision of additional protein to growing sheep during infection resulting in enhanced immunity to GIN (Brunsdon, 1964; Coop et al., 1995; Van Houtert and Sykes, 1996).

Resistance to GIN is acquired, with a significant protective immune capability developed by 10 to 12 months of age (Abbott et al., 2009; Vlassoff et al., 2001; Brunsdon, 1970). However a relaxation in host immunity to GIN is observed in ewes during the periparturient period in spring, from ~2 weeks before lambing to ~6 weeks post lambing (Taylor et al., 2007). The subsequent increase in faecal egg count (FEC) is known as the periparturient rise (PPR), and the adult nematodes that develop from hypobiotic larvae in the ewes are an important cause of pasture contamination (Abbott et al., 2009). This, combined with eggs from lambs that have ingested overwintered larvae, results in the number of L3 on pasture increasing markedly from mid-summer onwards. In temperate regions such as Europe, clinical outbreaks analogous to type I bovine ostertagiosis can occur from August to October (Taylor et al., 2007). If L3 are ingested prior to October the majority of larvae will mature; thereafter larvae may undergo hypobiosis, and resumption of development in the spring can cause clinical disease similar to type II bovine ostertagiosis in yearling sheep

(Taylor et al., 2007).

1.2.2 Anthelmintic drenches

Since the development of the benzimidazoles (1-BZ) in the 1960s (Brown et al., 1961) gastrointestinal nematode control in intensive grazing systems has relied heavily on the use of such broad spectrum anthelmintic treatments (Kettle et al., 1981, 1982; McMahon et al., 2013a; Morgan et al., 2012; Patten et al., 2011). With the recent introduction of amino-acetonitrile derivatives (4-AD) and spiroindoles (5-SI) to the market, there are now five classes of anthelmintic available for gastrointestinal nematode control in sheep (Table 1.1).

Table 1.1: Broad spectrum anthelmintic classes used in the control of gastrointestinal nematodes in sheep. [1] Year released may vary from country to country. [2] First documented case of resistance - there may be earlier published reports of suspected resistance and/or unpublished reports of resistance. Adapted from Kaplan, 2004.

Name	Reference	Release [1]	Resistance first reported [2]
Benzimidazoles (BZ)	Brown et al. (1961)	1961	Conway (1964)
Imidazothiazoles (LV)	Kates et al. (1971)	1970	Sangster et al. (1979)
Macrocyclic lactones (ML)	Chabala et al. (1980)	1981	Malan (1988)
Amino acetonitrile derivatives (AD)	Kaminsky et al. (2008)	2009	Scott et al. (2013)
Spiroindoles (SI)	Little et al. (2010)	2010	-

Since the initial detection of resistance to BZ in 1964, anthelmintic resistance (AR) has been reported worldwide for most chemical classes of anthelmintic (Table 1.1). Most recently a lack of efficacy of monepantel (4-AD) has been reported, only 4 years after the product was released to market in New Zealand, and less than 2 years after the product was first used on the farm in question (Scott et al., 2013). The only spiroindole currently on the market is Startect® (Zoetis), which combines derquantel (5-SI) and abamectin (3-ML). The combination of two classes of anthelmintic was required as derquantel alone was less than 95% effective against *T. circumcincta* adults and larvae, and *H. contortus* L4 (Little et al., 2010).

Combination anthelmintics are widely used in New Zealand and Australia (Bartram et al., 2012). Their use, however, is contentious (Coles and Roush, 1992; van, 2001), and has been restricted in the European Union until the release of Startect® (Zoetis) to the market in 2012 (Irish Medicines Board licence number: 10019/191/001). Proponents of the use of combination anthelmintics cite two main justifications: (1) to control gastrointestinal nematodes that have developed resistance to one or more classes of anthelmintics, and (2) to slow the development of resistance (Leathwick et al., 2009). Modelling studies support the use of combination anthelmintics to slow the development of resistance (Smith, 1990; Barnes et al., 1995; Dobson et al., 2011; Leathwick, 2012). This is supported by evidence from study by Leathwick et al. (2012), in which the development of anthelmintic resistance was slowed through the use of a combination anthelmintic on 90% of lambs, while leaving

10% of lambs untreated (to maintain refugia). Combination anthelmintics have also been shown to be effective against single-class AR populations of gastrointestinal nematodes (McKenna, 1990; Anderson et al., 1991b,a). Concerns over the use of combinations include the potential to select for resistance to multiple classes at the same time, and indeed resistance to dual- and triple-combination anthelmintics has been reported in New Zealand (Sargison et al., 2007; Waghorn et al., 2006; Wrigley et al., 2006; Sutherland et al., 2008). While anthelmintic combinations can play an important role in resistance management, they must be used in conjunction with other strategies, such as those set out in the Sustainable Control of Parasites in Sheep (SCOPS) manual (Abbott et al., 2009), to reduce anthelmintic use and delay the development of resistance (Bartram et al., 2012).

1.2.3 Alternative methods for GIN control

In the face of increasing anthelmintic resistance, new strategies for GIN control are called for (Vlassoff and McKenna, 1994). Alternatives to anthelmintic treatment are reviewed in Sayers and Sweeney (2005) and include bioactive forages, nutritional boost, grazing management, vaccines, micro-predacious fungi, and selection of resistant animals. These are briefly reviewed below.

Bioactive forages

Plant-based options of GIN control include the use of bioactive forages such as chicory (reviewed in Sutherland and Scott, 2009). Studies have shown that the use of chicory can result in reduced worm burden, increased weight gain, and reduced faecal egg count (FEC) (reviewed in Rattray, 2003). Bioactive forages such as chicory contain anti-parasitic condensed tannins (CT). These tannins protect dietary protein from rumen degradation, and have been shown to affect both free-living larval and adult gastrointestinal nematodes *in vitro* (Athanasiadou et al., 2001; Molan et al., 2002; Paolini et al., 2004). The observed anthelmintic activity of tannins has been hypothesised to be attributable to the capacity of tannins to bind protein (Athanasiadou et al., 2001). Aside from binding free protein and thus reducing larval nutrient availability, condensed tannins ingested by larvae may bind to the intestinal mucosa and cause autolysis. Alternatively, condensed tannins may bind the cuticle of the larvae, resulting in their death. The advantage of bioactive forages may also be in their high nutritive value and palatability, with an increase in protein ingestion resulting in increased immunity to GIN (Sutherland and Scott, 2009).

Nutritional boost

The link between the nutritional status of the host and GIN parasitism in sheep is well established (Sutherland and Scott, 2009). Protein loss due to gut damage leads to

a reduction in live weight gain, wool growth and milk production (Rattray, 2003). The Sustainable Control of Parasites in Sheep (SCOPS) manual (Abbott et al., 2009) suggests that using body condition scoring to determine the need to administer anthelmintics to mature ewes should be an integral part of a strategy to manage anthelmintic resistance. The provision of additional protein improves the development and maintenance of the immune response in both young lambs and in lactating ewes (Sykes and Coop, 2001).

Grazing management

Alternative grazing of pasture by sheep and cattle has been shown to be an effective method of parasite control, as there are very few species of parasite that infect both sheep and cattle (Bisset & Vlassoff, 1991). Once a pasture has been contaminated by sheep parasites, subsequent grazing by cattle results in both the ingestion of infective larvae and the opening up of the sward, which leads to desiccation of larvae (Rattray, 2003). Additionally, the delay in returning sheep to the paddock results in increased natural larval mortality and lower pasture burdens. This method of reducing GIN pasture contamination is only practical however in mixed farming systems.

Vaccines

Vaccination has been an effective method for control of many diseases of sheep, however there are currently no vaccines conferring protection against GIN on the market. Ideally a vaccine would be effective against multiple species of GIN, however there are a diverse range of effector mechanisms which operate against individual GIN species and their developmental stages (Shaw et al., 2013; Kemper et al., 2010; Li et al., 2012).

Recent trials at the Moredun Research Institute in Scotland have shown it is possible to successfully immunise sheep against *H. contortus* (Smith, 2014). “Barbervax” is currently being licensed in Australia, where trials indicate a reduction of FEC by 80% although repeated vaccination was required for sustained protective immunity (Smith, 2014).

Micro-predacious fungi

A number of organisms have been identified that use the free-living stages of parasites as a food source, including microarthropods, protozoa, viruses, bacteria and fungi. While there is great interest in nematode-destroying fungi, at present the biggest challenge is survival of spores through the gastrointestinal tract (Waller and Faedo, 1996). Nematophagous fungi can be divided into two groups; predacious and endoparasitic. Predacious fungi trap and kill the larvae, whereas endoparasitic fungi infect larvae by producing spore that either attach to the outside of the nematode or are ingested by the larvae (Larsen, 2000). Permanent sheep pasture is a good source of nematophagous fungi, and nematophagous

fungi including *Duddingtonia flagrans* have been detected in fresh faecal samples from Irish sheep, indicating they may have survived the gastrointestinal tract and therefore a viable option as a biological control agent (Kelly et al., 2009).

1.2.4 Selection of resistant animals

Breeding animals that are less reliant on anthelmintic drenches for control of GIN infection is a viable alternative method of disease management (Vlassoff and McKenna, 1994), particularly with the growing need to reduce drug usage in livestock (Morris, 2002). Selecting animals for increased resistance to GIN results in the host having fewer worms. Aside from the obvious health benefits of a lower worm burden, this also has the follow-on effect of reduced GIN egg output and therefore reduced pasture contamination. The combination of reduced pasture contamination and enhanced genetic resistance lowers the requirement for anthelmintic use (Bishop and Stear, 1999). Additionally, the GIN population in a flock is usually overdispersed, with a few animals harbouring the majority of the worms (Stear et al., 2009). If the few heavily infected hosts can be identified and removed from the flock, this would have a large effect on pasture contamination.

The remainder of this literature review will discuss the complexity of the host immune response, and thus the challenge of breeding sheep for resistance to GIN (primarily Trichostrongyles). The various phenotypic and genotypic markers currently available will also be discussed, along with the current efforts to identify genes associated with GIN resistance.

1.3 The host immune response

Manifestations of immunity

The development of immunity to GIN is complex and highly variable. Sheep are born naïve, and immune competence is not acquired until at least 4 months of age. Onset of immunity varies depending on the breed of sheep and the nematode species to which they are exposed (Sutherland and Scott, 2009). While lambs rapidly develop the ability to control GIN such as *Nematodirus battus* (Taylor et al., 2007), resistance to other species, such as *T. circumcincta,* is much slower to develop. Lambs start to demonstrate immunity from 4 to 5 months of age, with regular exposure to larval challenge allowing the immune response to develop until a significant protective immune capability is developed by 10 to 12 months of age (Abbott et al., 2009; Vlassoff et al., 2001; Brunsdon, 1970). Immune competence can be observed through suppressed parasite growth (and therefore fecundity), the expulsion of adult worm burden, the prevention of establishment of most incoming infective larvae, or a mixture of the three (Seaton et al., 1989; Stear et al., 1995b, 1996; Abbott et al., 2009). Adult sheep tend to remain relatively resistant to infection, harbouring only a few adult worms (Taylor et al., 2007), although regular exposure to some level of infection is required to retain immunity (Vlassoff et al., 2001). Nutritional stress, ill-health and pregnancy can all also influence an individual's immune status (Miller and Horohov, 2006).

1.3.1 The innate immune response

The immune system in vertebrates is made up of two parts, the innate (or non-specific) immune response and the adaptive (specific) response, the various cellular and chemical components of which work together to protect vertebrates from a range of threats. The first line of defence against parasites is the innate immune system, which plays a role in sensing parasites, then initiating and driving the acquired immune response (Anthony et al., 2007). Of particular relevance to GIN are innate physical barriers to the establishment and survival of parasites, and subsequently the process by which the host recognises the presence of GIN and activates a response.

1.3.1.1 Physical barriers to the establishment and survival of parasites

The surface of the gastrointestinal tract is covered with a layer of mucus, primarily produced by epithelial goblet cells (Anthony et al., 2007). This the front line of the innate defence against ingested food and pathogens in the gastrointestinal tract. The primary component of mucus is mucin, however it also contains an array of bioactive molecules. Many of these bioactive molecules have been shown to be anti-microbial, or to stimulate inflammation (Sutherland and Scott, 2009). Both increased mucus production and the presence of

inhibitory substances in the mucus have consistently been observed during the development of immunity to GIN (Balic et al., 2000b)

Smooth muscle function has also been shown to play an important role in mediating parasite resistance in mice (Vallance et al., 1997), yet its role in GIN expulsion in sheep is not clear. An up-regulation of genes related to the structure and function of the enteric smooth muscle was observed in lambs selected for resistance to GIN when compared to their susceptible counterparts (Diez-Tascon et al., 2005), however contradictory results were observed in sheep duodenal smooth muscle in response to *T. circumcincta* infection (Hassan et al., 2011a), with the susceptible lambs showing greater duodenal contractile force compared to resistant lambs.

1.3.1.2 Pattern recognition receptors (PRRs)

Among the earliest systems for the detection of pathogens are pattern recognition receptors (PRRs) such as C-type lectin receptors (CLRs) and toll-like receptors (TLRs). CLRs and TLRs are expressed by most cell types, including the cells of mucosal surfaces and immune cells such as antigen presenting cells (APCs), macrophages and dendritic cells (Geijtenbeek and Gringhuis, 2009; Glass, 2012). PRR proteins identify both pathogen-associated molecular patterns (PAMPs; pathogen molecular structures not found in the host), and damage associated molecular patterns (DAMPs; molecules released from damaged or stressed cells). Both PAMPs and DAMPs can result in the initiation and perpetuation of the inflammatory response. As well as being the first line of defence, PRRs play an important role in activating and manipulating the adaptive immune system (Hansen et al., 2011).

While viruses, bacteria and fungi are known to contain potent PAMPs, less is known about the role of PRRs in the response to nematode infection (de Veer et al., 2007). There is contrasting data on the importance of TLRs in nematode resistance, however CLRs are prime candidates for innate recognition of surface carbohydrate present on nematodes (reviewed in de Veer et al., 2007). The mannose receptor (a CLR) has been shown to bind to excretory/secretory proteins of the mouse nematode *Trichuris muris*, but was not essential for protective immunity (deSchoolmeester et al., 2009). In sheep, PRRs that detect non-host carbohydrates from the various moult stages of GIN, in addition to other signals such as tissue damage, are prime candidates for altering the host innate immune system to parasite invasion.

1.3.1.3 Innate effector cells

The mechanisms involved in innate and acquired immunity are not mutually exclusive, however innate immune defences are not pathogen-specific, and thus respond to different species in a generic way. In particular, an increase in the numbers of both eosinophils and mast cells are characteristic of infection with nematode parasites (de Veer et al., 2007).

Eosinophils

Eosinophils develop in the bone marrow from haematopoietic stem cells (Mori et al., 2009), and following infection proliferate in the blood in a process known as eosinophilia. In response to certain leukotrienes and CC family chemokines mature eosinophils migrate to the site of infection, where they are activated by T_H2 cytokines (Figure 1.5). Unlike mast cells, which appear uniformly scattered in tissue, eosinophils can show directional migration toward a parasite target (Balic et al., 2000b). Following activation, the effector functions of eosinophils include healing damaged tissue, immune regulation and resistance to parasitic invasion through degranulation and the release of eosinophil secondary granule proteins (EPGPs). This results in the damage and killing of infective larval stages of most helminth parasites (Meeusen and Balic, 2000; Balic et al., 2006; Rainbird et al., 1998). Along with mast cells, eosinophils also control mechanisms associated with allergy and asthma (Murphy et al., 2008).

Eosinophils have been shown to play a significant role in resistance to multiple species of GIN in sheep (Buddle et al., 1992; Henderson and Stear, 2006; Kanobana et al., 2002; Smith et al., 1983; Balic et al., 2006). A reduction in peripheral blood eosinophilia has been observed during primary infection with *T. circumcincta* (Sutherland et al., 1999), which was hypothesised by Sutherland and Scott (2009) to be a result of recruitment of cells into the intestinal epithelium. However the relationship between peripheral blood eosinophilia and tissue eosinophilia is reasonably weak, with only a proportion of eosinophils moving into the abomasal mucosa (Henderson and Stear, 2006). Increases in tissue eosinophils have been observed during *Haemonchus contortus* infection of both naïve (Balic et al., 2000a) and previously sensitised (Balic et al., 2002, 2006) sheep, along with resistance in Romney selection line animals with a naturally acquired mixed infection (Bisset et al., 1996).

Mast cells

Although best known for their role in the allergic response, increased numbers of mast cells are often observed during helminth infection (Anthony et al., 2007). Mast cells are connective tissue-based inflammatory cells (Murphy et al., 2008) that can both respond directly to pathogens and send signals to other tissues to modulate both the innate and adaptive immune responses (Urb and Sheppard, 2012). Two subsets of mast cells have been described based on their location: connective tissue mast cells (CTMCs) and mucosal mast cells (MMCs) (Voehringer, 2013). Activation of mast cells occurs primarily through antigen induced stimulation of the high-affinity immunoglobulin E (IgE) receptor (FcεRIs) expressed at the mast-cell surface (reviewed in Gilfillan and Tkaczyk, 2006). Mast cells can also be activated by directly interacting with pathogen-associated molecular patterns (PAMPs) through pattern recognition receptors (PRRs), including the Toll-like receptors (TLRs) (Urb and Sheppard, 2012). Mast cells store a number of inflammatory mediators (including histamine, leukotrienes and cytokines) that are released upon degranulation into

the surrounding tissues (reviewed in Abraham and St John, 2010; Dawicki and Marshall, 2007). The effects of these chemical mediators are characteristic of type 1 hypersensitivity, and include smooth muscle contraction, increased vascular permeability and local blood flow, and enhanced mucus secretion (reviewed in Balic et al., 2000b). Finally, chemotactic factors produced by mast cells can contribute to the recruitment of multiple inflammatory cells including eosinophils, natural killer (NK) cells, and neutrophils (Urb and Sheppard, 2012).

In sheep, nematode-induced activation of mast cells primarily associated with the acquired immune response (Balic et al., 2000b). The major mechanism controlling the number of adult *T. circumcincta* in previously sensitised animals appears to be IgE-dependent mast cell degranulation (Stear et al., 1995b)), with sheep mast cell proteinase systemically released during nematode infections (Huntley et al., 1987). Mast cell activation and degranulation can also be triggered by non-specific stimuli as part of the innate immune response through products of complement activation (Balic et al., 2000b).

Macrophages

Macrophages play a critical role in innate immunity, but also help initiate acquired immunity through the recruitment of other cells, such as lymphocytes (Murphy et al., 2008). M1 (classically activated) macrophages are activated through TLRs and interferon-γ, whereas M2 (alternatively activated) macrophages are stimulated by interleukin-4 (IL-4) or IL-13. Some of the molecular M2 macrophage signatures have been shown to differ between mice and humans (Martinez et al., 2013), however to date no studies have looked at the markers of macrophage subsets in sheep. M2 macrophages have three main functions during helminth infection: regulation of the immune response, healing of damaged tissue, and resistance to parasite invasion (Anthony et al., 2007). During a T_H2-type response M2 macrophages express Chitinase and FIZZ family member proteins (ChaFFs) (reviewed in Anthony et al., 2007). Chitinases degrade chitin, a molecule present in the exoskeletal elements of some animals, including helminth larvae (Fuhrman and Piessens, 1985). These proteins are therefore prime candidates for mediating host resistance to gastrointestinal nematodes (Nair et al., 2005).

1.3.2 The adaptive immune response

The major cellular components of the adaptive immune response are T and B cells (lymphocytes). T cells are involved in cell-mediated immunity, whereas B cells are involved in the humoral (antibody-mediated) immune response. The principal function of B cells is to make antibodies (immunoglobulins) against antigens (Murphy et al., 2008). Once the antigen has bound to the B-cell receptor, the lymphocyte will proliferate and differentiate into plasma cells, that secrete large amounts of antibodies.

Antigen processing and presentation

Thymus-derived T cells play a central role in the cell-mediated immune response. T cells are differentiated from other lymphocytes by the presence of a T cell receptor (TCR) on the cell surface. There are several types of T cell, including cytotoxic, helper and regulatory T cells. Cytotoxic T (T_C) cells kill cells that are infected with viruses or other intracellular pathogens. They are also known as $CD8^+$ T cells as they express the CD8 glycoprotein at their surface. Helper T (T_H) cells express the surface protein CD4, and provide essential additional signals to activate maturation of B cells, T_C cells, and macrophages. Regulatory T (T_{reg}) cells suppress the activity of other lymphocytes, and are critical for the maintenance of immunological tolerance.

Parasite antigens are trapped and processed by antigen presenting cells (APC), which present the antigen to their cognate TCR via MHC class I or II carrier molecules. APCs include dendritic cells, macrophages, and B cells, although most cells in the body can present antigen via MHC class I molecules (Murphy et al., 2008). $CD8^+$ and $CD4^+$ T cells bind MHC class I and MHC class II molecules respectively. The activation of the naïve T_H cell results in the release of cytokines, leading to T cell differentiation. The effector cells described below are regulated by the cytokine environment, generated by antigen activated T cells (Figure 1.3).

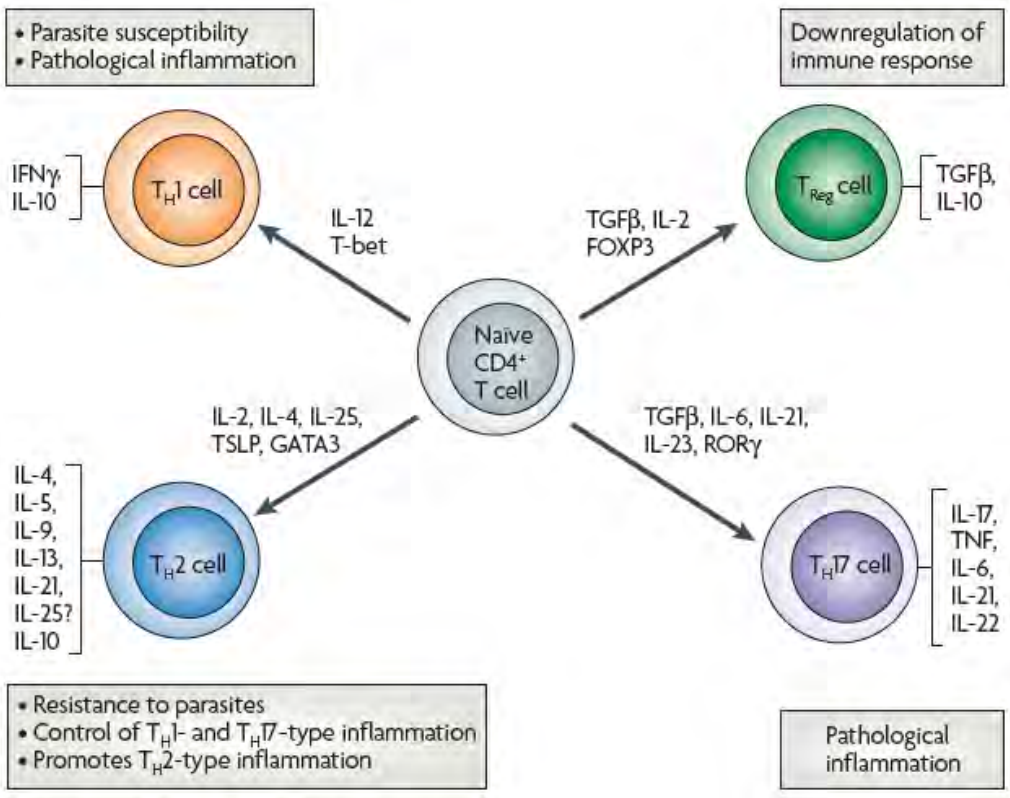


Figure 1.3: T helper and regulatory cells. Naive $CD4^+$ T cells can differentiate into several types of effector and regulatory cells during nematode infection. Specific cytokines and transcription factors contribute to differentiation, which plays a major role in determining whether the associated immune response will contribute to host protection or pathological inflammation. Figure used with kind permission from Anthony et al. (2007).

A significant number of activated antigen-specific B cells and T cells persist after an antigen has been eliminated, and these are known as memory cells. These cells form the basis of immunological memory and can be reactivated much more quickly than naïve lymphocytes, and usually provide lasting protective immunity (Murphy et al., 2008)

The T_H1 response

The T_H1 response has been traditionally associated with the immune response to intracellular bacteria and protozoa. The T_H1 cascade is triggered by the production of IL-12. It is primarily characterised by the expression of the inflammatory cytokine interferon-gamma (IFN-γ), which stimulates the production of IL-12 via a positive feedback loop. IFN-γ also inhibits the production of IL-4, while stimulating production of lymphotoxin (LT)-α (previously known as TNFß**)**. The effector molecules of the T_H1 response are specialised to stimulate proliferation of $CD8^+$ (cytotoxic/killer) T cells and activate macrophages.

The T_H2-type response

The protective T helper 2 (T_H2)-type response, elicited by helminth parasites as well as many allergic reactions, includes both innate and adaptive components (Anthony et al., 2007). Common features include expression of T_H2-type cytokines (IL-4, IL-5 and IL-13), eosinophils, basophils, mast cells (all of which can produce several types of T_H2-type cytokines) and IgE production (Figure 1.4).

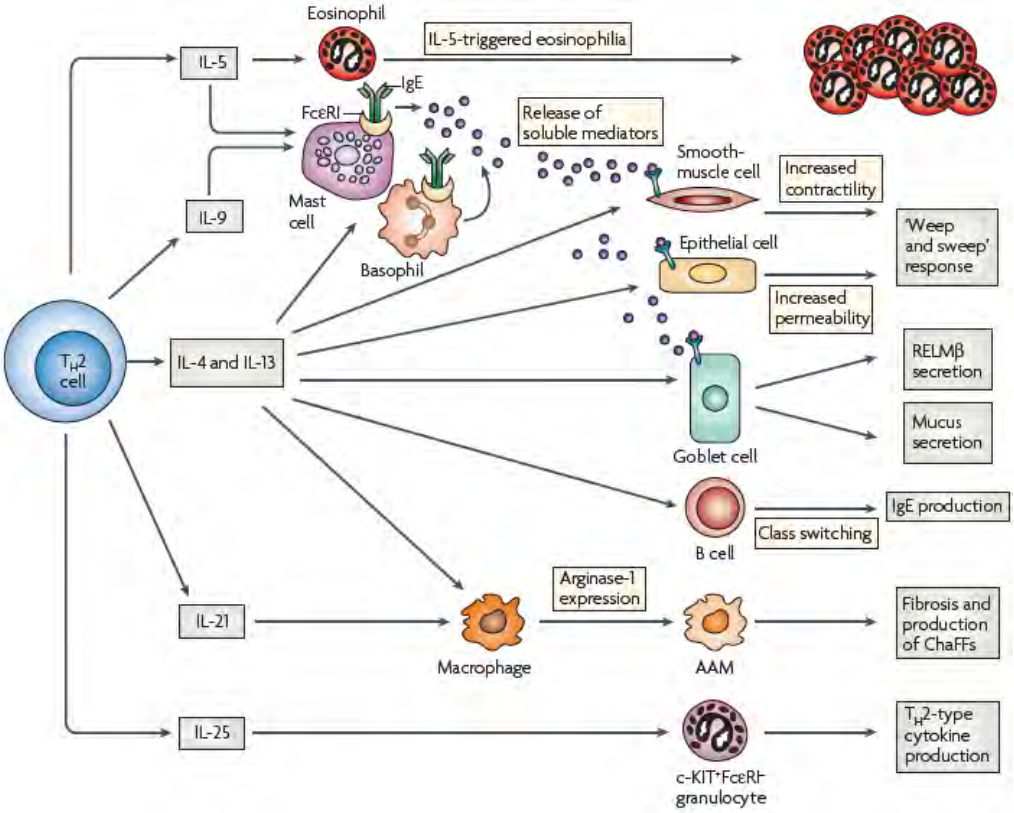


Figure 1.4: T_H2-cell functions during helminth infection. T_H2-cells orchestrate the immune response primarily through the production of cytokines. Figure used with kind permission from Anthony et al. (2007).

Interleukin-4 (IL-4) induces differentiation of naïve T_H cells to T_H2 cells, while suppressing differentiation into T_H1 cells. Upon activation, T_H2 cells produce additional IL-4 in a positive feedback loop, along with other T_H2 cytokines including IL-5, IL-9, IL-13 and IL-21 (Anthony et al., 2007). IL-4, along with the closely related IL-13, induces class switching in activated B cells, leading to production of IgE (Figure 1.4). The antibody IgE primes the IgE-mediated type 1 hypersensitivity response by binding to Fc (FcεRI and II) receptors on the surface of mast cells and basophils. When helminth antigen binds to cell bound IgE it leads to mast cell degranulation, and further production of IL-4 and IL-13 by basophils. Together the two cytokines also promote increased contractility of smooth muscle cells, increased permeability of epithelial cells, and elevated goblet-cell mucous secretion (Figure 1.4). The presence of IL-4 in extravascular tissue induces alternative activation of resident tissue macrophages, which function in wound healing and tissue repair, as mentioned above.

The T_H2 cytokine IL-13 induces epithelial cell repair and mucus production, and together with IL-9 recruits and activates mucosal mast cells (MMC). IL-5, aside from triggering eosinophilia, enhances secretion of IgA by B cells (Harriman et al., 1988). IgA, as discussed below, plays an important role in control of GIN in sheep (Lee et al., 2011b).

The T_H17 response

T_H17 cells promote acute inflammation through the recruitment of neutrophils to the site of infection. Early in infection IL-6, produced by dendritic cells, acts with TGF-β (also required for the differentiation of T_{reg} cells) to produce the T_H17 response (Murphy et al., 2008). This results in the production of IL-17 family members and IL-21, a subset of cytokines particularly important in clearing pathogens during host defence responses and in inducing tissue inflammation in autoimmune disease (Korn et al., 2009). Later, dendritic cells along with other antigen-presenting cells produce cytokines to promote either T_H1 or T_H2 development, and suppress T_H17 development.

The T_{reg} response

In the absence of pathogens, the lack of IL-6, IL-12 and IFN-γ, combined with the relative abundance of TGF-β favours the development of regulatory T cells (T_{reg}). T_{reg} are a subpopulation of $CD4^+$T cells that modulate the immune system through production of the immunosuppressive cytokines IL-10 and TGF-β. They are an important "self-check" in the immune system, and have been shown to be induced and expanded during helminth infection (Allen and Maizels, 2011).

The adaptive immune response to GIN infection in sheep

There is not a clean T_H1/T_H2 dichotomy in sheep (Sutherland and Scott, 2009). Susceptibility to GIN infection is, however, usually associated with a primarily T_H1 response, while a primarily T_H2 response is associated with resistance (Figure 1.5). Resistance is associated with a proliferation of globular leukocytes, mucosal mast cells and eosinophils, and production of parasite-specific immunoglobulin A (IgA), IgG1 and IgE. Protection against *T. circumcincta* is most strongly associated with IgE activity against L3, and IgA activity against L4. Mast cell degranulation prevents larvae from establishing (Stear et al., 1995b), while eosinophils potentially interact with IgA to regulate growth and fecundity (Henderson and Stear, 2006). Susceptibility, on the other hand, is associated with low IgA, low globular leukocytes and few mast cells and eosinophils.

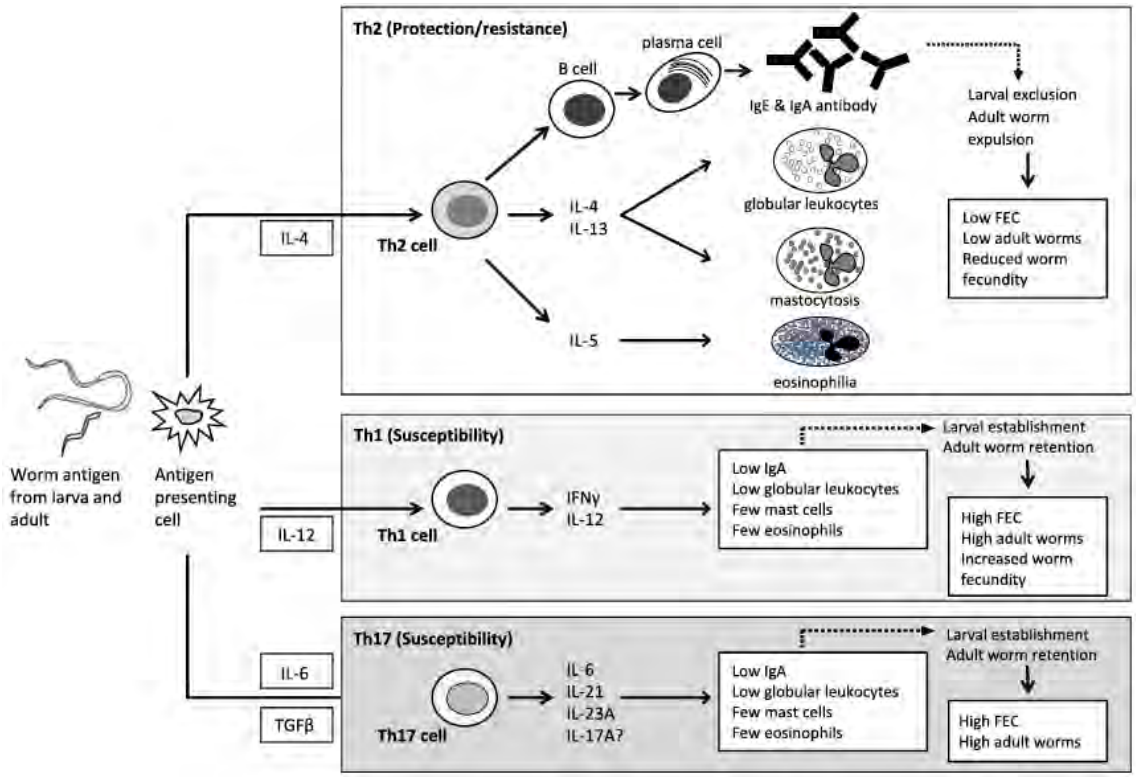


Figure 1.5: Differential activation of CD4⁺ T cell subsets associated with the adaptive immune response to *T. circumcincta* in sheep. Protective immunity (resistance) largely involves a T_H2-type response, whereas susceptibility is generally associated with a T_H1 response. The T_H17 response is a proposed pathway by which susceptibility is also favoured. Figure used with kind permission from Venturina et al. (2013).

Recent research has challenged the view that resistance to GIN is primarily a T_H2 response (Venturina et al., 2013). Gene expression studies using artificial infection of resistant and susceptible animals have indicated that the differential interplay between T_H1/T_H2 and T_{reg} genes may control the immune response to GIN rather than a straightforward T_H1 or T_H2 pathway (Hassan et al., 2011b; Ahmed, 2013).

1.4 Breeding sheep for gastrointestinal nematode resistance

1.4.1 Evidence for natural resistance to GIN infection

It is well established that there are between-breed differences in the ability of sheep to resist GIN infection (Table 1.2). There is also evidence for within-breed differences; variation in FEC has been observed within Scottish Blackface (Stear et al., 1995a), Merino (Sréter et al., 1994; Woolaston et al., 1996; Kahn, L. P. et al., 2003), and Soay (Coltman et al., 2001) animals, among others.

Table 1.2: Breed differences in resistance to GIN infection. 'Resistance' or 'susceptiblity' is relative to the other breeds in the study.

Relatively resistant breed	Relatively susceptible breed	Study
Border Leicester × Merino	Merino	Donald et al., 1982
Red Maasai	Dorper, Blackheaded Somali & Romney Marsh	Mugambi et al., 1996, 1997
Gulf Coast Native	Suffolk	Miller et al., 1998
Horro	Menz	Haile et al., 2002
Barbados Blackbelly	INRA 401	Gruner et al., 2003
Sabi	Dorper	Matika et al., 2003
Polish long-wool	Blackfaces	Nowosad et al., 2003
Santa Ines	Suffolk & Ile de France	Amarante et al., 2004
Dorper X, Katahdin & St. Croix	Hampshire	Burke and Miller, 2004
Katahdin & St. Croix × Barbados Blackbelly	Dorset X & Dorper X	Vanimisetti et al., 2004
Texel	Suffolk	Hanrahan and Crowley, 1999; Good et al., 2006

This comprehensive evidence for variation in natural resistance to GIN suggests breeding animals that are less reliant on anthelmintic drenches is a viable method of nematode control (Hunt et al., 2008).

1.4.2 Resistance versus resilience

Resistant animals are generally defined as those with an enhanced ability to acquire and mount an effective immune response to GIN, which results in reduced worm establishment (Douch et al., 1996). Animals can also be bred to have a high tolerance ('resilience') to internal parasites, where they are productive despite their worm burden (Bisset et al., 2001; Albers et al., 1987; Bisset et al., 1994; Riffkin and Dobson, 1979). Selection purely for resistance alone has been reported to result in lower live weight gain, increased breech soiling (dags) and a reduction in fleece weight (Sutherland and Scott, 2009). This can be counteracted however by using both resistance and production traits in an appropriate selection policy (Beef + Lamb New Zealand, 2008; McEwan et al., 1995). The main benefit of resistance in comparison to resilience is reduced pasture contamination. As there is no reduction in FEC when breeding for resilience, pasture contamination levels stay the same, and non-resilient animals in a flock do not benefit from reduced GIN challenge. Resistant sheep, on the other hand, have fewer adult nematodes, which leads to reduced egg output and lower pasture contamination.

1.4.3 Phenotypic markers

There are several potential manifestations of resistance to GIN (Balic et al., 2000b; Lee et al., 2011b):

- Reduced number of adult nematodes
- Reduced size of adult nematodes (resulting in reduced fecundity of females)
- Increased number of inhibited larvae

These traits do not develop at the same rate however, and all three are not observed in all resistant animals. Measurement of these traits also all require the slaughter of the animal, and therefore cannot be used in selection programmes. Correlated traits (Table 1.3) such as faecal egg count (FEC) are commonly used as secondary indicators of resistance (Sayers and Sweeney, 2005; Saddiqi et al., 2012).

Table 1.3: Parasitological, immunological and pathological phenotypic markers used to evaluate resistance to *Trichostrongyles* in sheep. Modified from Dominik (2005).

Parasitological	Immunological	Pathological
Worm burden	Antibody response (e.g. IgA, IgE, IgG1)	Plasma pepsinogen
Faecal egg count (FEC)	Blood eosinophils	Live weight
Fecundity and worm length		Dag score

1.4.3.1 Faecal egg count (FEC)

FEC is the most widely used phenotypic indicator of resistance to GIN. The trait is reliable, relatively easy to measure, and provides a direct estimate of pasture contamination. The current gold standard for measuring individual animal FEC is the modified McMaster method, with a sensitivity of 50 eggs per gram of faeces (Ministry of Agriculture Fisheries and Food, 1986). The correlation between FEC and worm burden has been reported to be good to high ($r \sim 0.7$), although this can vary depending on the species of GIN and host breed studied, as discussed below (Amarante et al., 1999, 2004; Beasley et al., 2010; Beraldi et al., 2008; Bisset et al., 1996; Good et al., 2006; Grenfell et al., 1995; Stear et al., 1995b; McKenna, 1981). Differences in FEC between animals can, however, also be due to inhibition of infective larvae, suppression of worm fecundity (discussed below) or differences in the ratio of males to females (Bricarello et al., 2004). Selection using FEC requires animals to have a relatively high GIN challenge to reliably estimate their phenotype, which can result in loss of productivity while drench is withheld (Douch et al., 1996). It is also costly to measure in a commercial farming situation and may fail to represent all of the pathways involved in internal nematode resistance due to physiological complexity (Dominik, 2005; Morris, 2009).

Despite these drawbacks, a moderate proportion of the variance in FEC has been shown to be due to host genotype, which makes it viable as a selection tool (Stear et al., 1997, 2009). The heritability of FEC as a measure of resistance varies considerably depending on both nematode species and breed surveyed. Estimates are generally moderate, ranging between 0.2 and 0.4 (Bishop and Morris, 2007; Baker et al., 1991); Safari et al. (2005) estimated from 16 published estimates a weighted mean (±S.E.) heritability of 0.27±0.02. More recently an analysis of more than 2 million pedigree-recorded animals in New Zealand by Pickering (2013) revealed that the heritability estimates for the internal parasite traits of summer FEC (FEC1 and NEM1) and autumn FEC (FEC2 and NEM2) cluster between 0.18 ± 0.01 to 0.21 ± 0.01. Selection based on low FEC has also been shown to be sustainable in the medium to long-term, with GIN shown to be slow to adapt to resistant hosts (Kemper et al., 2009).

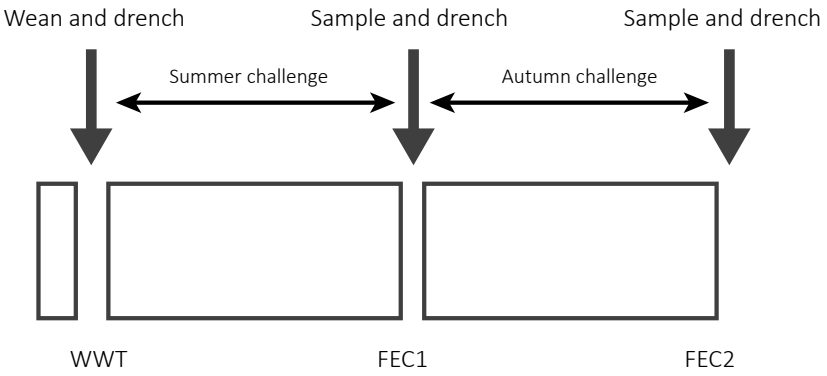


Figure 1.6: Faecal egg count (FEC) breeding value (BV) sampling method. Guidelines for the standard sampling method for the collection of faecal samples from lambs for FEC breeding value analysis. WWT = weaning weight.

Breeding programmes in both Australia (Woolaston et al., 1991; Windon, 1990; Cummins et al., 1991; Greeff and Karlsson, 2006) and New Zealand (Morris et al., 2000, 2005) have utilised FEC to select animals that are either resistant or susceptible to GIN. Both nematode species and host breed vary between studies. While these are primarily experimental flocks, this approach (Figure 1.6) has also been utilised in a commercial setting through performance recording and genetic evaluation services such a Sheep Improvement Limited (www.sil.co.nz) and Sheep Genetics (www.sheepgenetics.org.au).

1.4.3.2 Fecundity and worm length

Reduced FEC can be achieved by reducing the number of mature females and/or reducing the fecundity of the resident female population (reviewed in Balic et al., 2000b). The length of adult *T. circumcincta* females is strongly associated with both the number of eggs *in utero* (Stear et al., 1995b) and the number of eggs laid per worm per day (Stear and Bishop, 1999).

As nematode fecundity can be difficult to measure, length of adult females is therefore often used as a proxy for fecundity.

In Scottish Blackface lambs the variation in FEC following natural (primarily *T. circumcincta*) infection is reported to be primarily due to variation in worm fecundity rather than worm burden (Stear et al., 1996, 1997, 1999b). Despite this the relationship between FEC, worm burden, and worm fecundity is not clear-cut.

Differences in FEC within both Merino selection line (Kemper et al., 2010) and Scottish Blackface (Davies et al., 2005) animals showed low FEC animals regulate *T. circumcincta* burden through both reduced fecundity and reduced adult worm burden. Kemper et al. (2010) observed suppression of the development of *T. circumcincta*, with a skew towards more immature life stages. In divergent lines of Romney sheep, selected on the basis of high or low FEC, both the number and fecundity (eggs in utero) of adult *Trichostrongylus* spp. were reduced in the resistant (low FEC) animals following a natural challenge (Bisset et al., 1996). Not only were low FEC lambs limiting the number of adults establishing in the abomasum, but they were also suppressing fecundity of those worms which did manage to establish.

Reduced FEC is not always the result of reduced fecundity however. Breed differences observed in FEC of co-grazing Suffolk and Texel lambs following a natural infection were due to worm burden rather than differences in *T. circumcincta* fecundity (Good et al., 2006). In 6-month-old Scottish Blackface ewe lambs infected with *T. circumcincta* neither worm length nor egg count in utero showed a significant correlation with either worm burden or FEC (Beraldi et al., 2008). FEC did however explain approximately 60% of the variation in post-mortem adult burden. Seaton et al. (1989) demonstrated that sheep regulate worm development before they regulate worm burden in *T. circumcincta* infections, which may explain some of the differences between studies.

Fecundity can also be influenced by density-dependence, declining as worm burden increases (Bishop and Stear, 2000b). It is currently unclear whether this decline is due to competition, the immune response of the host, or a combination of both (Balic et al., 2000b; Stear et al., 2009).

While differences in both adult GIN burden and GIN fecundity can only be measured at necropsy, it is nonetheless helpful in understanding the mechanisms underpinning resistance in experimental settings, which can lead to the development of effective control strategies (Saddiqi et al., 2012).

1.4.3.3 Antibody response

A number of antibodies have been shown to be correlated with GIN resistance in sheep, including IgA, IgE and IgG1. Increased levels of the immunoglobulin IgA, the isotype closely associated with intestinal mucosal immune responses, has been positively associated

with resistance to *T. circumcincta*, regulating both worm length and fecundity (Smith et al., 1985; Stear et al., 2004; Strain et al., 2002; Stear et al., 1995b, 1999b,c; Strain and Stear, 1999; Halliday et al., 2007). This resistance is regulated through suppressed parasite growth, development and fecundity, and mediated by IgA activity against 4^{th}-stage larvae. In Scottish Blackface lambs the presence of arrested L4 larvae has been shown to be positively associated with both worm burden and the size of the local IgA immune response (Stear et al., 1995b). The association between serum anti-*T. circumcincta* IgA antibody levels and indicators of infection such as FEC and worm burden suggests that IgA levels can be used to complement selection for resistant animals based on FEC (Beraldi et al., 2008; Stear et al., 1999b; Davies et al., 2005).

A test for anti-nematode IgA antibody levels is already on the market as the CarLA Saliva Test® (www.carlasalivatest.com). The carbohydrate larval surface antigen (CarLA) is found on the epicuticle of infective-stage larvae (L3) of ruminant nematode species, and is shed during the moult to L4 (Harrison et al., 2003b). Research has shown that sheep under GIN challenge produce an IgA response against CarLA, which reduces larval establishment in the gut, and results in rapid GIN expulsion (Harrison et al., 2003a, 2008). This response has been shown to be heritable (h^2 = 0.3) and negatively correlated with FEC (r = -0.5) (Shaw et al., 2012). The CarLA Saliva Test® was subsequently developed to measure the IgA antibody response to CarLA using saliva from animals under parasite challenge. Animals with high levels of anti-CarLA IgA have lower FEC, improved growth rate post-weaning, and no tendency for increased dags (Shaw et al., 2012, 2013).

The saliva assay has several practical advantages over FEC for selection purposes: as the saliva IgA antibody response is driven by larval challenge, the test can be performed irrespective of anthelmintic drench treatment, and collection of saliva samples is rapid and relatively non-invasive. As with faecal samples, saliva samples must be kept cold or frozen, as the antibody is likely to be heat labile. Genetic gain is, however, slower than selection on FEC, providing only 17-43% of the value of two FEC measurements for increasing host resistance to gastrointestinal nematodes (McEwan and Dodds, 2009).

Increased levels of IgG1, IgE and IgM have been correlated with reduced FEC in Romney selection line sheep in New Zealand (Shaw et al., 1999; Bisset et al., 1996; Douch et al., 1994), although increased IgE was also negatively correlated with breech soiling. Elevated levels of IgG1and IgA were observed in *T. colubriformis*-challenged sheep (Cardia et al., 2011). Elevation of total and/or parasite-specific IgE serum antibodies have been reported during infection with *H. contortus* (Kooyman et al., 2000), *T. colubriformis* (Shaw et al., 1998c) and *T. circumcincta* (Pettit et al., 2005; Huntley et al., 2001). In addition, an association between a polymorphism at the 5' end of the sheep IgE gene and resistance to *T. colubriformis* has been reported (Clarke et al., 2001), although attempts to confirm this finding in other flocks failed.

1.4.3.4 Blood eosinophils

Infection with GIN can also lead to a rise in blood eosinophils. Both local and peripheral eosinophilia have been associated with resistance to infection among sheep deliberately infected with *T. circumcincta* (Stear et al., 1995b; Henderson and Stear, 2006) and *H. contortus* (Amarante et al., 1999). When measurements were taken continually during a controlled challenge, both peripheral blood eosinophilia and IgA levels peaked 8-10 days post infection, then subsequently declined (Henderson and Stear, 2006). A negative correlation has been reported between FEC and blood eosinophil values during mixed infections in some breeds (Amarante et al., 2009; Doligalska et al., 1999; Buddle et al., 1992) but not others (Woolaston et al., 1996).

Animals with high peripheral blood eosinophil levels appear to have shorter, less fecund worms (Davies et al., 2005; Terefe et al., 2007) or a reduced worm burden (Beraldi et al., 2008). It appears that eosinophilia and IgA may interact in regulating the growth of *T. circumcincta* (Henderson and Stear, 2006). Stear et al. (2002) concluded that eosinophil concentrations may be a useful indicator of resistance to GIN infection, but only in older ($\geq$ 3 months) lambs that have been continually exposed to infection.

1.4.3.5 Pepsinogen

Plasma pepsinogen level is of value in diagnosing abomasal damage, particularly during infection with GIN such as *H. contortus* and *T. circumcincta*, which damage the stomach lining, resulting in reduced protein digestion (Roeber et al., 2013). Plasma pepsinogen levels can therefore be used as an indicator of the severity of infection (Lawton et al., 1996; Stear et al., 1999a). Pepsinogen levels in sheep with heavy infections are above the normal of approximately 0.8 IU tyrosine and usually exceed 2.0 IU (Taylor et al., 2007).

Pepsinogen is routinely used as a diagnostic tool for ostertagiosis in cattle (Eysker and Ploeger, 2000; Berghen et al., 1993), with an increase in plasma pepsinogen observed after infection with *Ostertagia* (Ploeger et al., 1990a,b,c, 1994; Shaw et al., 1998a,b; Dorny et al., 1999; Charlier et al., 2011). In sheep, pepsinogen levels have been shown to be more strongly associated with variation in length and fecundity of female *T. circumcincta* than with variation in the number of nematodes present (Stear et al., 1999a; Davies et al., 2005). A negative correlation has also been observed between pepsinogen and FEC in Spanish Churra ewes (Gutierrez-Gil et al., 2009b).

1.5 Variation in the sheep genome controlling resistance to gastrointestinal nematodes

The identification of genes or linked markers that have a significant association with parasite resistance would accelerate the genetic improvement of resistance to internal nematodes; marker assisted selection would allow animals to be selected without the need for parasite challenge. Additionally a detailed understanding of the genes and mechanisms involved in protective immunity and the factors that regulate this response would also aid in informing future breeding strategies as well as the development of effective and sustainable nematode control methods, such as immunomodulatory anthelmintics.

There are four general approaches to identifying variation in the genome contributing to GIN resistance: i) association analysis, ii) positive selection analysis, iii) gene or protein expression studies and iiii) the candidate gene approach (Brown et al., 2013). These approaches are used to either compare naïve and infected animals, or resistant and susceptible individuals/breeds.

Gene/locus mapping

The first studies identifying regions of the genome associated with GIN resistance used quantitative trait locus (QTL) mapping. Identification of QTL involves whole or partial genome scans using a known pedigree and markers such as microsatellites or single nucleotide polymorphisms (SNPs) to identify regions of the genome containing loci affecting the trait of interest (Crawford et al., 2000). Once the region has been identified, it can be fine mapped using further markers to identify either the casual mutation, or a marker in close linkage disequilibrium (LD) that can be used for selection.

Studies using microsatellite-based linkage analysis (LA) identified multiple regions of the genome associated with GIN resistance (e.g. Beh et al., 2002; Crawford et al., 2006; Davies et al., 2006). The identification of candidate genes through QTL mapping has proven difficult however, as QTL often span millions of base pairs and contain hundreds of potential candidate genes. With the advent of the Illumina® OvineSNP50 BeadChip (www.sheephapmap.org) microsatellite-based linkage studies have largely been replaced with SNP-based genome-wide association studies (GWAS) (e.g. Kemper et al., 2011; Sallé et al., 2012; Riggio et al., 2013, 2014).

Table 1.4: Studies searching for genes or linked markers for GIN resistance in sheep. Studies are broadly defined as gene/locus mapping (QTL), gene or protein expression analysis, or the candidate gene approach. The total number of unique genes/proteins, QTL or SNP is given, along with host breed and gastrointestinal nematode (GIN) studied.

Study	Analysis	Genes/Proteins	QTL	SNP	DNA/RNA	Sheep breed	GIN
Sayre and Harris, 2012	Candidate gene | Literature | Systems genetics	31	-	-		N/A	N/A
Buitkamp et al., 1996	Candidate gene | Microsatellite	2	-	-	DNA	Scottish Blackface	Multi-species challenge
Charon et al., 2002	Candidate gene | Microsatellite	1	-	-	DNA	Scottish Blackface	Multi-species challenge
Coltman et al., 2001	Candidate gene | Microsatellite	1	3	-	DNA	Soay	Multi-species challenge
Janssen et al., 2002	Candidate gene | Microsatellite	1	-	-	DNA	Rhön	*H. contortus*
Outteridge et al., 1996	Candidate gene | Microsatellite	1	-	-	DNA	Merino	Multi-species challenge
Paterson et al., 1998	Candidate gene | Microsatellite	1	-	-	DNA	Soay	*Strongyle* nematodes
Paterson et al., 2001	Candidate gene | Microsatellite	1	-	-	DNA	Romney Selection Lines	Multi-species challenge
Schwaiger et al., 1995	Candidate gene | Microsatellite	1	-	-	DNA	Scottish Blackface	Multi-species challenge
Benavides et al., 2009	Candidate gene | Polymorphism	1	-	-	DNA	Corriedale & Polwarth	Primarily *H. contortus*
Clarke et al., 2001	Candidate gene | Polymorphism	1	1	-	DNA	Merino selection lines	*T. colubriformis*
Sayers et al. (2005b)	Candidate gene | Polymorphism	1	-	-	DNA	Texel & Suffolk	Multi-species challenge
Sayers et al. (2005a)	Candidate gene | Polymorphism	1	-	-	DNA	Texel & Suffolk	Multi-species challenge
Balic et al., 2006	Candidate gene | qPCR	1	-	-	RNA	Merino	*H. contortus*
Craig et al., 2007	Candidate gene | qPCR	11	-	-	RNA	Scottish Blackface-cross	*T. circumcincta*
Gossner et al., 2012, 2013	Candidate gene | qPCR	4	-	-	RNA	Scottish Blackface	*T. circumcincta*
Hassan et al., 2011b	Candidate gene | qPCR	11	-	-	RNA	Suffolk	*T. circumcincta*
Ingham et al., 2008	Candidate gene | qPCR	43	-	-	RNA	Merino selection lines	*T. colubriformis* or *H. contortus*
Ingham et al., 2011	Candidate gene | qPCR	1	-	-	RNA	Merino selection lines	*T. colubriformis* or *H. contortus*
Knight et al., 2007	Candidate gene | qPCR	1	-	-	RNA	Scottish Blackface	*T. circumcincta*

Study	Analysis	Genes/Proteins	QTL	SNP	DNA/RNA	Sheep breed	GIN
Lacroux et al., 2006	Candidate gene | qPCR	3	-	-	RNA	INRA 401	*H. contortus*
Pernthaner et al., 1997	Candidate gene | qPCR	3	-	-	RNA		
Pernthaner et al., 2005	Candidate gene | qPCR	4	-	-	RNA	Romney selection lines	*T. colubriformis*
Pernthaner et al., 2006	Candidate gene | qPCR	1	-	-	RNA	Romney selection lines	*T. colubriformis*
Shakya et al., 2009	Candidate gene | qPCR	2	-	-	RNA	Gulf Coast Native v Suffolk	*H. contortus*
Terefe et al., 2007	Candidate gene | qPCR	3	-	-	RNA	INRA 401 v Barbados Black Belly	*H. contortus*
Brown et al., 2013	Candidate gene | SNP	-	-	-	DNA	Soay	Multi-species challenge
Diez-Tascon et al., 2005	Expression | Microarray	106	-	-	RNA	Perendale selection lines	Multi-species challenge
Keane et al., 2006	Expression | Microarray	41	-	-	RNA	Perendale selection lines	Naïve
Keane et al., 2007	Expression | Microarray	297	-	-	RNA	Perendale selection lines	Multi-species challenge
Knight et al., 2010	Expression | Microarray	187	-	-	RNA	Romney	*T. colubriformis*
Knight et al., 2011	Expression | Microarray	71	-	-	RNA	Scottish Blackface-cross	*T. circumcincta*
MacKinnon et al., 2009	Expression | Microarray	37	-	-	RNA	Hair v wool	*H. contortus*
Pemberton, Beraldi, J Hopkins, *per comm*. (Brown et al., 2013)	Expression | Microarray	8	-	-	RNA	Scottish Blackface	*T. circumcincta*
Rowe et al., 2009	Expression | Microarray	58	-	-	RNA	Merino-cross	*H. contortus*
Andronicos et al., 2010	Expression | Microarray & qPCR	48	-	-	RNA	Merino selection lines	*T. colubriformis* or *H. contortus*
Gossner et al., 2013	Expression | Microarray & qPCR | RNA Seq	379	-	-	RNA	Scottish Blackface	*T. circumcincta*
Kadarmideen et al., 2011	Expression | Microarray | Systems genetics	167	-	-	RNA	Merino selection lines	*T. colubriformis* or *H. contortus*
Goldfinch et al., 2008	Expression | Protein	3	-	-	Protein	Crossbred	*T. circumcincta*
Nagaraj et al., 2012	Expression | Protein	108	-	-	Protein	Merino selection lines	*H. contortus*
Pemberton et al., 2012	Expression | Protein	11	-	-	Protein	Scottish Blackface-cross	*T. circumcincta*

Study	Analysis	Genes/Proteins	QTL	SNP	DNA/RNA	Sheep breed	GIN
Ahmed, 2013	Expression | RNA Seq	535	-	-	RNA	Suffolk & Texel	*T. circumcincta*
Pemberton et al., 2011	Expression | RNA Seq	144	-	-	RNA	Scottish Blackface	*T. circumcincta*
Brown et al, in prep. (Brown et al., 2013)	QTL | GWAS | SNP Chip	-	-	51	DNA	Resistant' v 'Susceptible'	
Kemper et al., 2011	QTL | GWAS | SNP Chip	-	-	15	DNA	Mixed breed	*T. colubriformis* & *H. contortus*
Pemberton, Beraldi, H Lee, *per comm.* (Brown et al., 2013)	QTL | GWAS | SNP Chip	-	-	-	DNA	Soay	Multi-species challenge
Pickering, 2013	QTL | GWAS | SNP Chip	-	-	15	DNA	Dual-purpose	Multi-species challenge
Riggio et al., 2013	QTL | GWAS | SNP Chip	-	11	47	DNA	Scottish Blackface	Multi-species challenge
Riggio et al., 2014	QTL | GWAS | SNP Chip	-	10	-	DNA	Multiple breeds	
Sallé et al., 2012	QTL | GWAS | SNP Chip	-	49	-	DNA	Romane*Martinik Black Belly backcross	*H. contortus*
Beh et al., 2002	QTL | Microsatellite	-	6	-	DNA	Merino selection lines	*T. colubriformis*
Beraldi et al., 2007	QTL | Microsatellite	-	2	-	DNA	Soay	Multi-species challenge
Crawford et al., 2006	QTL | Microsatellite	-	6	-	DNA	Romney*Coopworth crosses	*Trichostrongylus* spp.
Davies et al., 2006	QTL | Microsatellite	-	9	-	DNA	Scottish Blackface	*Nematodirus*
Dominik et al., 2010	QTL | Microsatellite	-	4	-	DNA	Romney*Merino backcross	*T. colubriformis* & *H. contortus*
Gutierrez-Gil et al., 2009a	QTL | Microsatellite	-	5	-	DNA	Spanish Churra	Multi-species challenge
Marshall et al., 2009	QTL | Microsatellite	-	30	-	DNA	Merino	*H. contortus*
Marshall et al., 2012	QTL | Microsatellite	-	24	-	DNA	Red Maasai, Dorper	*H. contortus*
Matika et al., 2011	QTL | Microsatellite	-	7	-	DNA	Texel	*Nematodirus*
Silva et al., 2012	QTL | Microsatellite	-	13	-	DNA	Red Maasai, Dorper	Multi-species challenge

Genome-wide association studies utilise the information provided by the large number of markers spread evenly throughout the genome to detect variants associated with a trait. These variants are unlikely to be responsible for the observed phenotype, but more likely to be in linkage disequilibrium (LD) with a causative mutation and consequently be used for selection.

While genome-wide association studies have identified hundreds of common genetic variants associated with complex disease so far, most confer relatively small increments in risk, in contrast with the initial 'common disease, common variant' hypothesis (Reich and Lander, 2001). One of the most likely explanation for this so called 'missing' heritability (Maher, 2008) is that for quantitative traits effect sizes at individual SNPs are so small that they do not reach genome-wide significance (Goddard et al., 2009; Manolio et al., 2009; Yang et al., 2010; Visscher et al., 2010).

Despite the problems with GWAS, the information provided by the Illumina® OvineSNP50 BeadChip has been used to map both QTL (Riggio et al., 2013, 2014; Sallé, 2012) and Mendelian traits (Becker et al., 2010; Johnston et al., 2011; Mömke et al., 2011; Zhao et al., 2011; Shariflou et al., 2013), and investigate patterns of LD (Kijas et al., 2009, 2012; Moradi et al., 2012). It has also been shown that sheep can be selected for low FEC using genomic selection even though there appear to be many loci of small effect controlling the trait (Kemper et al., 2011).

Detecting selection in the genome

Sheep domestication led to increased exposure to nematodes due to an increase in stocking density. Infection results in selective pressure on individuals, with the most resistant animals having increased fitness. The will result in selection for adaptive polymorphisms that increase fitness, which will potentially leave signatures in the genome associated with immunity to GIN. Evidence for positive selection can be identified by analysing allelic diversity of populations to identify genomic domains under selection. Evidence for selective sweeps in the genome can be found by comparing polymorphisms such as SNPs between populations (Kijas et al., 2012; Moradi et al., 2012; McEwan et al., 2014). Also at the population level tests such as the McDonald-Kreitman test (McDonald and Kreitman, 1991) can be used, which examine within species variation compared to between species divergence.

Alternatively, evidence for positive selection on a protein can be sought at the comparative genomic level, in the context of the evolution of the *Bovidae* or *Ovis* lineage. Genes that are involved in immunity in cattle and have undergone adaptive evolution have been identified in a number of studies (Lynn et al., 2005; Larson et al., 2006; Babiuk et al., 2007; Freeman et al., 2008; Jann et al., 2008; Takeshima et al., 2009), and as such are candidate genes for immune related traits in cattle. Studies of adaptive evolution in the sheep genome have

previously been limited by a lack of sheep coding sequence data and lack of an assembled sheep genome.

Candidate gene approach

Information from other studies can be used to focus on candidate genes chosen for their link with a trait, or their known or suspected function. Single locus association (SLA) studies search for markers such as microsatellites, SNPs or other polymorphisms in and around the gene of interest that may explain the observed phenotypic variance. Quantitative real-time PCR (qRT-PCR) has also been used to examine differential gene expression in specific genes of interest associated with trait variation. This approach has also often been used to validate differentially expressed genes discovered using microarray and RNA sequencing (RNA-Seq) approaches.

The candidate gene approach, by definition, is limited by prior knowledge of the genes involved in defence against GIN. Quantitative traits such as FEC are often under the control of many genes, indeed host resistance to GIN appears to be mediated by many genes, each with relatively small effect (Kemper et al., 2009; Crawford et al., 2006). The candidate gene approach is therefore unlikely to capture all of the variation leading to the observed phenotypic differences.

Gene expression studies

Transcriptome and proteome analysis is a powerful method for the identification and quantitation of genes and proteins expressed during GIN infection. These tools have led to a greater understanding of the molecular basis of phenotypic variation in resistance to GIN. During the past decade both bovine and ovine specific microarrays have been used in sheep to provide expression information on tens of thousands of genes at once. These studies have identified hundreds of genes differentially expressed between known resistant and susceptible animals (Table 1.4).

Recently high-throughput sequencing of cDNA (RNA-seq) has allowed researchers to look at RNA expression on an even larger scale. While microarray technology relies on prior knowledge of genomic sequence, RNA-seq allows the detection of unknown genes, alternative splice sites and novel isoforms. One advantage that is particularly relevant to sheep is the ability to re-analyse the data once more information becomes available (e.g. the release of an updated genome or transcriptome). Despite these advantages, sequencing still remains relatively costly, and computationally intensive. Microarrays on the other hand are relatively inexpensive and widely available, meaning studies of the ovine transcriptome continue to utilise both resources (Gossner et al., 2013).

Recently transcriptome analysis in sheep has been aided by the release of version 3.1 of the ovine genome (Kijas et al., 2012) and the subsequent release of the Ensembl

annotation in December 2013 (Flicek et al., 2014). To date GIN resistance using RNASeq methods has been used in only two separate studies (Pemberton et al., 2011; Gossner et al., 2013; Ahmed, 2013). Pemberton et al. (2011), and subsequently Gossner et al. (2013) examined gene expression in the abomasal lymph node of Scottish Blackface lambs resistant or susceptible to *T. circumcincta*. The number of unique, differentially expressed genes increased from 144 when reads were aligned to the bovine genome (Pemberton et al., 2011) to 379 when reads were aligned to the ovine genome (Gossner et al., 2013). Ahmed (2013) also aligned ovine reads to the bovine genome, comparing expression in the abomasal lymph node of Suffolk (relatively susceptible) and Texel (relatively resistant) lambs artificially infected with *T. circumcincta*. Other studies in sheep have examined the transcriptome in the foetal heart (Cox et al., 2012), skin (Fan et al., 2013), bone healing models (Jager et al., 2011), and high and low fecundity animals (Miao and Luo, 2013).

1.5.1 Previously identified genetic markers of GIN resistance

Studies to detect QTL for GIN resistance or detect associations with candidate genes have been based on diverse experimental approaches, sheep breeds and nematode species; it is therefore not surprising that different chromosomal regions of interest have been identified among these studies. While results are not always available in the public domain, the sheep QTL database (SheepQTLdb; www.animalgenome.org/QTLdb/sheep) contains all curated sheep QTL and associated data in the public domain (Hu et al., 2013).

Full or partial genome scans have identified QTL for FEC on almost every chromosome. QTL have also been identified for worm burden, worm length, number of eggs per worm, eosinophil counts and immunoglobulin IgA, IgG and IgE levels (SheepQTLdb). The two most consistent regions identified have been the region of the Major Histocompatibility Complex (MHC) on chromosome 20, and the region containing the interferon gamma (*IFN*γ) gene on chromosome 3 (Bishop and Morris, 2007), with a notable exception being the WormSTAR™ marker on chromosome 3p (McEwan et al., 2008). It must be noted, however, that the importance of these genes in mediating parasite resistance may be overstated, as many studies have specifically targeted these regions (Table 1.4).

The Major Histocompatibility Complex (MHC) locus

The Major Histocompatibility Complex (MHC) of sheep, also designated ovine Lymphocyte Antigen (OLA), is a multi-gene complex critical to immunity. In the mammalian genome the MHC is the most gene-dense and polymorphic region, and is associated with resistance to infectious diseases, autoimmunity, and reproductive success (The MHC Sequencing Consortium, 1999). In sheep the basic structure of the MHC region is similar to other mammals, comprising of the class I, IIa, IIb and III regions (Figure 1.7). Using shotgun sequencing, a complete ovine MHC sequence map revealed 177 genes, along with 18 predicted micro-RNA coding sequences (Gao et al., 2010).

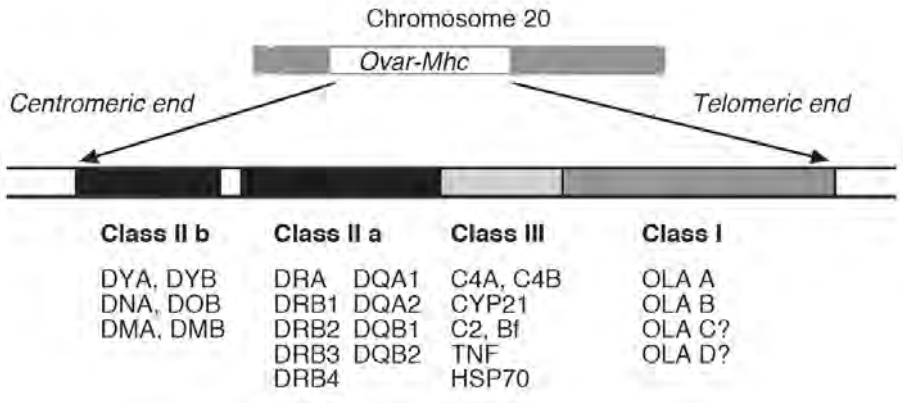


Figure 1.7: The ovine Major Histocompatibility Complex (MHC) region. Schematic presentation of the structure of the ovine MHC region on chromosome 20. Figure used with kind permission from Dukkipati et al. (2006).

The primary function of MHC genes is to code for antigen-presenting receptor glycoproteins, known as MHC molecules. The membrane-anchored MHC molecules bind to processed peptide antigens and present them to T lymphocytes, initiating the adaptive immune response (Dukkipati et al., 2006). The MHC class I molecules are expressed on the surface of all nucleated cells, and consist of a heavy α-chain and a ß-chain. The class I molecules are chiefly involved in presentation of cytosol-derived peptides to cytotoxic T lymphocytes ($CD8^+$) (Amills et al., 1998; Kumánovics et al., 2003). The MHC class II genes encode glycoproteins, which are also heterodimers consisting of two chains, α and ß. The class II molecules bind and present processed pathogen peptides to helper T cells bearing the $CD4^+$ differentiation marker (Amills et al., 1998; Kumánovics et al., 2003). Relative to the other regions of the MHC, the class III region is historically the least characterised. The genes at this locus have immunological and other functions, such as heat shock protein 70 (*HSP70*), complement cascade genes (*C4*, *C2* and *BF*) and tumour necrosis factor alpha (*TNF*) (Amills et al., 1998; Kumánovics et al., 2003).

Multiple studies have found an association between parasite resistance and regions on ovine chromosome 20, where the MHC is located (Schwaiger et al., 1995; Buitkamp et al., 1996; Paterson et al., 1998; Diez-Tascon et al., 2005; Sayers and Sweeney, 2005; Stear et al., 2005; Davies et al., 2006; Keane et al., 2007; Hassan et al., 2011a). These studies by no means provide unanimous results. Other studies using genetic marker approaches on various flocks have found no evidence for an effect of genes in the MHC on either mixed (Benavides et al., 2009), *H. contortus* (Blattman et al., 1993) or *T. colubriformis* resistance in sheep (Nicholas et al., 1993). This may be explained by the alleles themselves not causing resistance or susceptibility *per se*, but being in linkage disequilibrium (LD) with additional polymorphisms in the region (Keane et al., 2007); a combination of these polymorphisms may then contribute to resistance or susceptibility in some populations. As the extent of LD is likely to vary between breeds and populations, the MHC alleles previously implicated may not show up as being significant.

The interferon gamma (IFNγ) locus

Interferon gamma (*IFN*γ) locus, located on chromosome 3 in sheep, is a potential candidate for nematode resistance as it is associated with the host response following an immune challenge. The *IFN*γ gene codes for a cytokine secreted by T_H1 lymphocytes that plays a critical role in regulating the T_H1 versus T_H2 immune responses in vertebrates. IFNγ activates macrophages, which phagocytose intracellular pathogens (Wakelin, 1996).

A QTL for parasite resistance in Romney divergent selection lines after multi-species challenge was fine mapped to a region near the *IFN*γ gene (Paterson et al., 2001). Subsequently, a polymorphism in the region near *IFN*γ was linked to reduced FEC and increased parasite specific IgA in a wild population of Soay sheep on the island of Hirta in the St Kilda archipelago in the Outer Hebrides (Coltman et al., 2001). Sayers et al. (2005b)

also observed an association between an *IFN*γ haplotype and genetic resistance to GIN in Texel, but not Suffolk, sheep.

A partial genome scan of 139 microsatellite markers across eight chromosomes in Scottish blackface sheep identified a QTL on chromosome 3 associated with IgA activity that was very close to the interferon gamma locus (Davies et al., 2006). Likewise, using 133 markers across the genome, a QTL encompassing the *IFN*γ region was also observed in Merino divergent selection lines after challenge with *T. colubriformis* (Beh et al., 2002). Several other regions on chromosome 3 have also been found to have linkage to parasite resistance, although they are not found near *IFN*γ (Beraldi et al., 2007; Marshall et al., 2009).

Ingham et al. (2008) discovered that expression of *IFN*γ transcripts were increased in the gut mucosa of *T. colubriformis* and *H. contortus* resistant animals during the innate immune response (after primary challenge), and subsequently downregulated during the acquired immune response (after tertiary challenge).

Cytokines and Ig coding genes

The association between IgE and the response to GIN has previously been discussed. A significant association was found between an IgE allele and resistance to *T. colubriformis* in a Merino selection flock by Clarke et al. (2001), however attempts to validate the association in two other flocks assessed for resistance to either *T. colubriformis* or *H. contortus* failed to find any significant association.

Cytokines play a pivotal role in the immunity to GIN, defining the orientation and strength of the immune response (section 1.3.2). A partial association scan in Corriedale and Polwarth animals using seven markers on chromosome 5 (where *IL3*, *IL4* and *IL5* are located) showed three of the five markers were significantly associated with FEC in Corriedale animals, two of which were also associated with FEC in a Polwarth flock (Benavides et al., 2002). In a following study Benavides et al. (2009) showed polymorphisms within *IL4* was significantly associated with (primarily *H. contortus*) FEC in weaned Corriedale, but not Polwarth, ewe lambs.

Conclusions

Many other regions of the ovine genome have also been linked to parasite resistance; as the number of whole-genome scans increase, so do the number of suggestive regions, with most chromosomes being implicated in one or several studies. It can be concluded that the search for QTL or linked genes for GIN resistance in sheep is a difficult area of research. This primarily due to the physiological and phenotypic complexity of the trait, although analysis has proven to also be an issue, as most of the reported studies derive from initial low-resolution genome screens, often resulting in very wide confidence intervals. More consistency in experimental protocols, materials and analysis approaches would allow

a more precise comparison of results; the studies have differed in the breed of sheep and their immune status, nematode species used in the experiments, measurement of internal nematode resistance and the challenge regime. Comparisons between breeds of sheep, such as Soay and domestic, are also problematic for many reasons including differences in environment, age structure, treatment history, and parasitological methods.

Considering the complexity of nematode resistance, it is unsurprising that previous studies have not necessarily yielded the same results. The information gained from QTL studies can, however, be used alongside information from gene expression studies to gain a greater understanding of nematode resistance in sheep. As noted by multiple studies (Kemper et al., 2011; Riggio et al., 2014; Crawford et al., 2006), GIN resistance is a complex trait for which many genes of relatively small effect contribute.

1.6 Aims and Objectives

The objective of this study was to identify Scottish Blackface lambs that differed in their resistance to GIN infection and characterise the host responses to infection and identify genes and biological processes important for the difference in resistance. The identification of genes and alleles associated with nematode resistance would have a two-fold benefit: firstly it would aid in the implementation of a selective breeding programme in Ireland with measurable economic benefit, and secondly it will advance our fundamental understanding of the mammalian immune system and resistance to infection.

In Ireland, two sheep breeds have been extensively studied with respect to GIN resistance, and it has been demonstrated that the Texel breed is more resistant to nematode infection than the Suffolk breed. To date no studies have examined the resistance among Scottish Blackface sheep in Ireland.

The specific objectives of this study were:

i) to identify lambs that show variation in resistance to experimental nematode challenge within an Irish Scottish Blackface population,

ii) to examine gene expression in the gut lymphoid tissue of these animals using next generation sequence technology,

iii) to identify any proteins that show evidence of positive selection,

iv) to assess the role of these genes using association testing, and validate SNP associated with GIN resistance discovered in other populations,

The identification of genes or linked markers that have a significant association to host resistance to internal parasites would greatly accelerate genetic improvement. Recent advances in genome sequencing and genomic technologies provide new opportunities to understand the infection process of internal parasites in sheep at the genetic level. The aim of this project is to use these new technologies to help understand GIN resistance in sheep, and subsequently use this understanding to control parasite infection.

Chapter 2

Characterising gastrointestinal nematode resistance in Scottish Blackface lambs

2.1 Introduction

Gastrointestinal nematodes (GIN) are the most serious cause of disease in domestic sheep worldwide with symptoms ranging from cclinical disease to ill thrift. In temperate climates, such as in Ireland, the most common infective species are *Teladorsagia circumcincta*, *Trichostrongylus* spp. and *Nematodirus* spp. (Good et al., 2006; Burgess et al., 2012). There is a sizeable body of evidence for both within- and between-breed variation in the ability of sheep to resist gastrointestinal nematode infection (Table 1.2). This suggests that breeding for host resistance is a viable strategy to minimise the effects of GIN parasitism.

A number of traits can be used to identify animals that exhibit increased resistance to GIN infection, including faecal egg count (FEC), worm burden, anti-nematode antibody level and plasma pepsinogen concentration (as reviewed in Saddiqi et al., 2012). Of these the most practical and widely used indicator is FEC, which is moderately heritable and shows wide variability among individuals (Bishop and Morris, 2007; Safari et al., 2005). Resistance to GIN infection, as defined by a relatively low FEC, can manifest as a lower number of nematodes, reduced size of adult nematodes, reduced fecundity of females, increased proportion of inhibited larvae, or a combination of the foregoing elements (Balic et al., 2000b; Lee et al., 2011b). A reduction in either the number of fecund adult females or in female fecundity would have the beneficial effect of reducing worm contamination on pasture (Good et al., 2006).

In Ireland, two sheep breeds have been extensively studied with respect to nematode resistance and it has been demonstrated that the Texel breed is more resistant to nematode infection than the Suffolk breed (Hanrahan and Crowley, 1999; Sayers et al., 2005a; Good et al., 2006; Sayers et al., 2008; Hassan et al., 2011a). Previous studies have shown that there is substantial genetic variation among Scottish Blackface lambs in both FEC and worm length (Stear et al., 2009). Resistance to GIN is most likely based on the ability to develop a timely and protective immune response. To date there have been no studies of the resistance among Scottish Blackface sheep in Ireland. Therefore, the objective of the present study was to develop a robust method to identify resistant and susceptible Scottish Blackface lambs, and characterise their response to an experimental nematode challenge.

2.2 Materials and methods

2.2.1 Ethical approval

All animal procedures described in this study were conducted under experimental license from the Irish Department of Health in accordance with the Cruelty to Animals Act 1876 and the European Communities (Amendments of the Cruelty to Animals Act 1976) Regulations, 1994.

2.2.2 Animals

All lambs were sourced from the purebred Scottish Blackface flock at the Teagasc Hill Sheep Farm in Leenane, Co. Mayo over a 2 year period (2010 - 2011). The flock consisted of approximately 350 ewes and is divided between two production systems: a pure-breeding system involving about 200 ewes and a crossbreeding system in which ~150 purebred ewes are used for crossbred lamb production. Replacement ewes enter the pure-breeding system and usually remain in that system for 3 seasons after which they are moved to the crossbreeding system. The farm has both hill and lowland pasture; the unimproved hill pastures range in altitude from 15 to 275 m above sea level (Nolan et al., 2003). The pure-breeding system is based mainly on the unimproved hill pasture while the crossbreeding system is based mainly on the improved lowland grassland portion of the farm.

Two separate grazing groups of purebred lambs were run on the farm each year. Single wethers (castrated males) and all twin-born lambs grazed on improved lowland pasture (Lowland grazing group), while single-born ewe lambs grazed on the unimproved hill pasture (Hill grazing group). Three of the four grazing groups, Lowland10, Lowland11 and Hill11, were used in this study (Table 2.1). In 2011 the single-born ewe lambs were moved to lowland pasture post-weaning (14 weeks) to acquire a natural GIN infection. All lambs received an anthelmintic dose at five weeks of age for *Nematodirus*. With the closure of the Teagasc Hill Sheep Farm in September 2011 all lambs were removed to the Athenry campus in Co. Galway.

2.2.2.1 Faecal sampling

In both 2010 and 2011 flock FEC was monitored weekly from early June, when lambs were approximately 8 weeks of age, using the FECPAK system (www.fecpak.co.nz). At least ten fresh faecal samples were collected from pasture, with each sample collected from a different faecal deposit. Cold H_2O was added to the sample at a ratio of three times the weight of the samples, and homogenised to ensure an even suspension before a 30 mL aliquot was taken and added to 200 mL saturated NaCl. The solution containing the sample

and NaCl was mixed by inversion, and then passed through a 0.15 mm aperture sieve. The strained fluid was mixed, and random aliquots taken to fill both chambers of a FECPAK slide. A stereo microscope (10 x magnification) was used to count the number of eggs within the grids of both chambers. Each egg counted represented 30 eggs per gram of faeces (epg); the sensitivity of the test was therefore 30 epg.

Once flock FEC reached approximately 600 epg lambs were individually sampled twice (FEC1A and FEC1B), with samples taken one week apart. Samples were collected from the rectum of each animal, and stored in an airtight bag at 4 ℃ until processing (within 3 days). GIN burden was assessed for each animal using the modified McMaster method (Ministry of Agriculture Fisheries and Food, 1986). From each sample 3 g of faeces were homogenised with 42 mL cold H_2O, before being passed through a sieve (0.15 mm aperture). The strained fluid was agitated, and a 15 mL aliquot taken and centrifuged at 432 x g for 3 min at 4 ℃. The supernatant was removed, and the pellet resuspended in a saturated NaCl solution. Aliquots were taken from the resuspended sample to fill both chambers of a McMaster slide, and the number of eggs within the marked grids in both chambers counted using a stereo microscope (10 x magnification). Each egg counted represented 50 eggs per gram of faeces; the sensitivity of the test was therefore 50 epg.

Following FEC1B sampling, lambs were dosed with a macrocyclic lactone (ML; Oramec, Merial Animal Health Ltd) in accordance with manufacturer's recommendations. Flock FEC was monitored weekly (FECPAK), and once flock FEC again reached approximately 600 epg FEC2A and FEC2B of individual FEC counts were completed.

2.2.2.2 Blood sampling

At the last faecal sampling (FEC2B) two blood samples from each lamb were collected by jugular venipuncture into aseptic vacutainers for DNA extraction (green vacutainer; lithium heparin) and haematology analysis (purple vacutainer; EDTA). Blood samples were analysed for haematology measurements within 6 h of sampling using an ADVIA® 2120 haematology system (Siemens Healthcare Diagnostics Inc.) as per manufacturer's recommendations. Haemoglobin, red blood cells, mean corpuscular volume, platelets, haematocrit, white blood cell count, neutrophils, lymphocytes, monocytes, eosinophils, large unstained cells and basophils were measured.

2.2.3 Controlled challenge of selected HighFEC and LowFEC animals

Ten susceptible (subsequently known as "HighFEC") and 10 resistant (subsequently known as "LowFEC") lambs were selected from each of three groups: 1) 2010 born wethers, 2) 2011 born wethers and 3) 2011 born ewe lambs (from both Lowland11 and Hill11 grazing groups; Table 2.1). The selected animals from the 2011-born ewe lambs (2011E cohort; Table 2.1) were used to monitor the response to infections in HighFEC and LowFEC animals

over time (71 days after challenge). Resistant and susceptible wethers selected from the 2010 and 2011 Lowland grazing groups (2010W and 2011W cohorts; Table 2.1) were used to define the acute responses to infection; these lambs were slaughtered at either 7 or 14 days post infection (Figure 2.1).

Table 2.1: Animals selected from each grazing group for a controlled challenge with *T. circumcincta*. Lambs and their dams were managed from birth to weaning on either improved lowland pasture (Lowland group) or unimproved hill pasture (Hill group). Resistant (LowFEC) and susceptible (HighFEC) lambs were selected from within-sex (ewe or wether) groups for a controlled challenge with *T. circumcincta*.

Year	Grazing group	Sex	Selection group	Challenge cohort	Phenotype	No.
2010	Lowland10	Male (n=92)	2010-born wethers	2010W (n=20)	HighFEC	10
					LowFEC	10
2011	Lowland11	Male (n=76)	2011-born wethers	2011W (n=20)	HighFEC	10
					LowFEC	10
		Female (n=34)	2011-born ewe lambs	2011E (n=20)	HighFEC	5
					LowFEC	3
	Hill11	Female (n=56)			HighFEC	5
					LowFEC	7

2.2.3.1 Selection of resistant and susceptible animals

Individual animal values for ln(FEC+25) were used to identify the most resistant and susceptible individuals from each selection group (Table 2.1), estimated using mixed model procedures in SAS® (v9.1). Due to differences in variance estimates (among animals and residual; Table 2.2), samples from each natural infection (FEC1 and FEC2) were analysed separately.

Table 2.2: Summary of variance components for each sub analysis of ln(FEC+25). Due to differences in variance estimates, samples from each natural infection (FEC1 and FEC2) were analysed separately. [1]The Lowland11 group includes both male (n=76) and female (n=34) lambs.

Grazing group	Round	Animal	Residual	Repeatability
Lowland10	FEC1	0.3094	0.4970	0.58
(n=92)	FEC2	0.2611	0.3342	0.43
Lowland11	FEC1	1.6045	0.4917	0.77
(n=110[1])	FEC2	0.3739	0.1790	0.68
Hill11	FEC1	0.2882	0.1166	0.71
(n=56)	FEC2	0.5017	0.2696	0.65

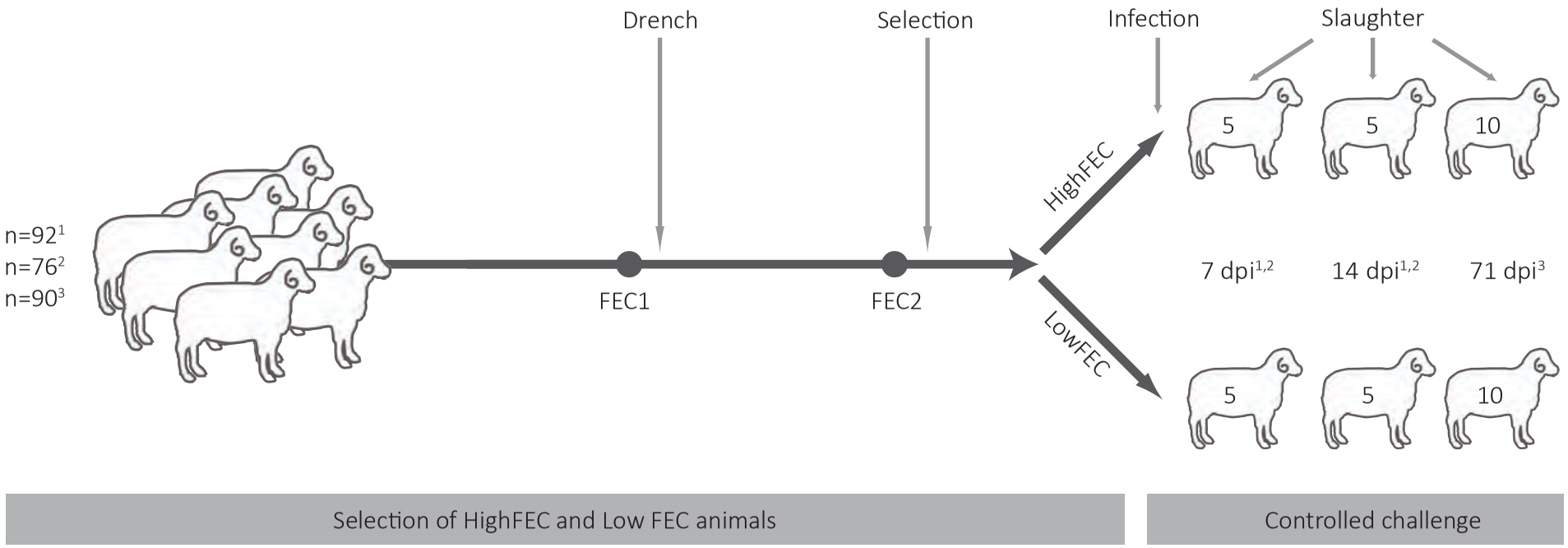


Figure 2.1: Experimental design for selection of HighFEC and LowFEC animals and subsequent controlled challenge. Lambs were selected from [1]2010 born wethers (2010W), [2]2011 born wethers (2011W) and [3]2011 born ewe lambs (2011E). Flock faecal count (FEC) was monitored weekly from when lambs were approximately 8 weeks of age. Once this reached approx. 600 epg lambs were individually sampled twice, 1 week apart, to give FEC1. Lambs were then dosed, and returned to pasture where the process was repeated for the second natural infection. FEC1 & 2 values were used to select the 10 most resistant (LowFEC) and 10 most susceptible (HighFEC) lambs, which were cleared of infection with an anthelmintic before being given a controlled challenge of 30,000 *T. circumcincta* larvae. Animals were slaughtered at either 7, 14 or 71 days post infection (dpi).

The data for each round (FEC1 and FEC2) were analysed separately by grazing group using a model that included sex (male or female), rearing type (single or twin) and sample date (A or B sample of round) as fixed effects where appropriate (Table 2.3). To get the selection differential for each animal the estimated animal effect for each round was scaled by the standard error of prediction and averaged.

Table 2.3: Fixed and random effects used for selection of resistant and susceptible animals. Fixed effects include sample date, rearing type (single or twin) and sex (male or female).

Grazing group	Fixed effects	Random effects
Lowland10	Sample date, rearing type	Animal(round)
Lowland11	Sample date, rearing type, sex	Animal(round)
Hill11	Sample date	Animal(round)

Selection differentials were used to select 10 HighFEC and 10 LowFEC animals from each group (Figure 2.1). As the 2011-born ewe lambs were raised in two separate grazing groups, selection involved combining the animal effects from two sets of analyses.

2.2.3.2 *T. circumcincta* larval culture

Naïve male lambs sourced from the Athenry campus in Co. Galway received an oral challenge of approximately 15,000 infective (L3) *T. circumcincta* larvae obtained from the Moredun Research Institute. Faeces were collected daily using bags strapped to the animal (Figure 2.2).

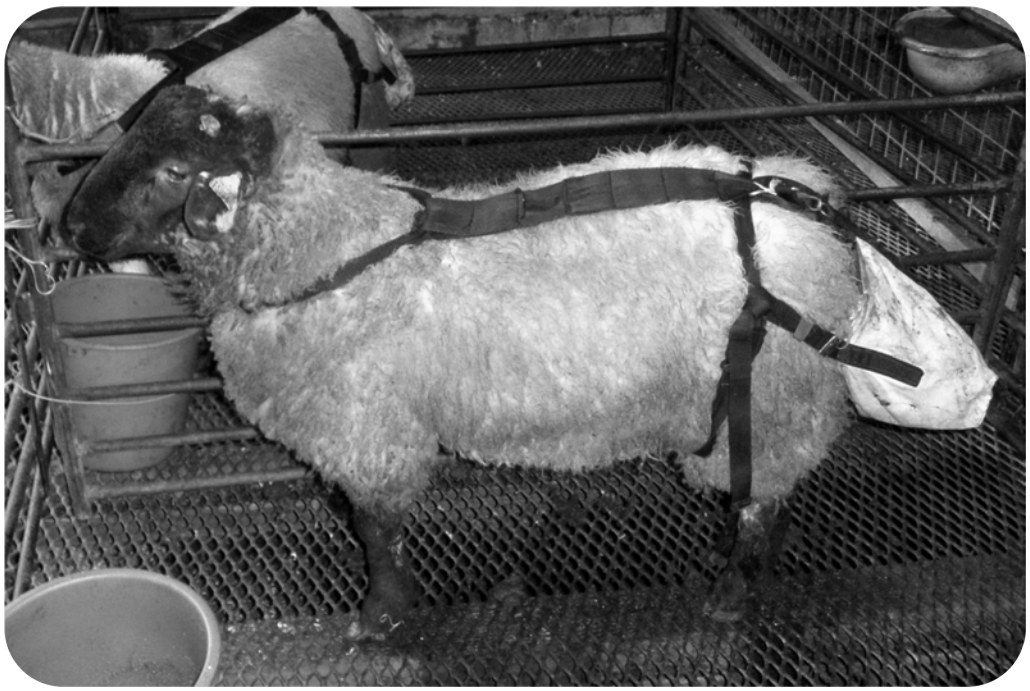


Figure 2.2: Collection of faeces for larval culture. Faeces were collected by attaching a collection bag via a harness to the rump of naïve male lambs infected with *T. circumcincta* L3.

Upon collection faeces were weighed, then mixed with vermiculite and a small quantity of water until the mixture was moist and crumbly. The mixture was placed in a container lined with plastic and the lid was replaced lightly before incubation at 27 ℃ for 14 days. To reduce mould the mixture was stirred once a day during the first 3 days. Larvae (L3) were extracted from the mixture using the Baermann technique (Hendrix, 1998) and stored in tap water in the fridge at approximately 4 ℃. Larvae were cleaned using the sucrose interface technique (Eysker and Kooyman, 1993). Using a syringe 10 mL of freshly made sucrose solution (2 g sucrose to 5 mL MilliQ H_2O) was slowly placed below up to 15 mL of a suspension of larvae in a 50 mL Falcon tube. The Falcon tube was centrifuged at 769 x g for 4 minutes, after which the larvae were removed from the interface with a Pasteur pipette and stored in tap water at 4 ℃.

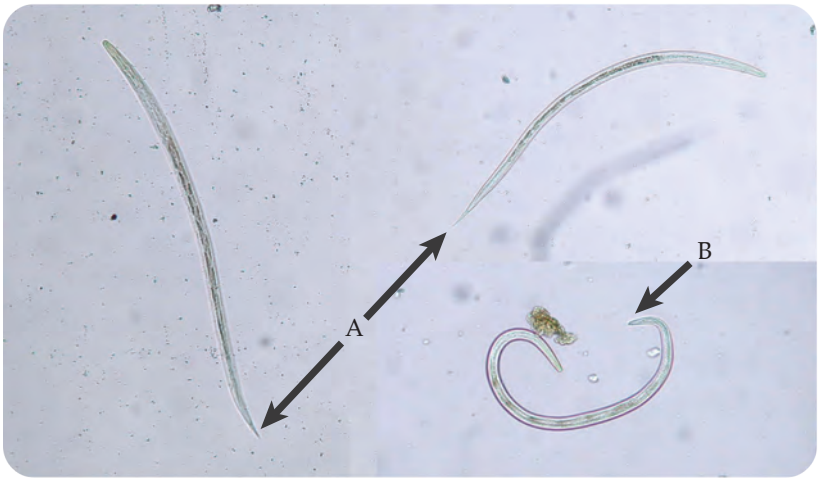


Figure 2.3: Sheathed and exsheathed *Teladorsagia circumcincta* L3 larvae. Normal larvae with tail sheath (A), and exsheathed larvae (B).

Prior to infection of the selected animals from the 2011W cohort it was observed that a substantial proportion (43%) of the cultured L3 appeared to have exsheathed, for unknown reasons. As the infectivity of these larvae could not be guaranteed selected animals from both the 2011W and 2011E cohorts were infected with L3 that had been cultured in 2010 and stored in tap water at 4 ℃.

2.2.3.3 Infection with *T. circumcincta* larvae

The selected Scottish Blackface lambs (section 2.2.3.1) were cleared of residual helminth infection with ML (Oramec, Merial Animal Health Ltd) or AD (Zolvix, Novartis), and housed on straw bedding until slaughter, with free access to water and 600 g commercial lamb ration per day. At 15 (ML) or 11 (AD) days post treatment all lambs were faecal sampled for three consecutive days to establish the absence of GIN infection. The 2011E cohort were additionally dosed with ML (Oramec, Merial Animal Health Ltd) to eliminate *S. papillosus*, a flukicide (Duotech, Norbrook Laboratories Ltd; closantel and oxfendazole combination) to clear any *Fasciola hepatica* obtained while grazing at Leenane; these lambs were faecal sampled at 19 days post treatment to establish absence of fluke infection using the fluke sedimentation test as described in Mooney et al. (2009). All chemicals were administrated in accordance with manufacturer's recommendations.

On day 0 lambs (n=60) received an oral challenge of *approximately* 30,000 *T. circumcincta* larvae (L3). Lambs were moved to Teagasc Ashtown, Co. Dublin to acclimatise for 2 days prior to slaughter. For the 2011E cohort response to infection was monitored by measuring FEC (3 times per week), along with plasma pepsinogen, serum anti-nematode IgA and IgG and haematology parameters (weekly) over the course of infection. Abomasal mucosa anti-nematode IgA and IgG, worm burden, and worm fecundity were determined from samples collected at slaughter. For the 2010W and 2011W cohorts plasma pepsinogen,

serum and mucosa anti-nematode IgA and IgG, and worm burden were determined at slaughter.

2.2.3.4 Necropsy

On the day of slaughter, 4 blood samples were collected from each lamb by jugular venipuncture into aseptic vacutainers for DNA extraction (green vacutainer; lithium heparin), pepsinogen (green vacutainer; lithium heparin), haematology analysis (purple vacutainer; EDTA) and ELISA (gold SST II vacutainer). Animals were slaughtered by electrical stunning followed immediately by exsanguination. Abomasal lymph node, abomasum, small intestinal lymph node and submandibular lymph node tissue were collected. Tissue was cut into pieces approximately 0.5 cm^3 and submerged in 10 volumes of RNAlater® (Ambion). This was stored at room temperature overnight followed by long-term storage at -80 ℃. Abomasum and small intestine muscoal scrapings were collected, snap frozen in liquid nitrogen, and stored at -80 ℃.

2.2.3.5 Nematode burden enumeration

The abomasum was removed at slaughter, and its contents recovered; the abomasum was then opened along the greater curvature, and digested in 1 L physiological saline for 4 h at 37 ℃ to recover nematodes from the tissue wall (Eysker and Kooyman, 1993). Both contents and digest were washed through a 75 μm sieve followed by a 38 μm sieve, before being preserved in 5% formalin. Adult and larval nematodes were counted from both the abomasal contents and the abomasal digest as previously described (McKenna, 2008). Total nematode burden was calculated by extrapolating from 2% (75 μm sieve samples) or 5% (38 μm sieve samples) aliquot counts.

Ten abomasa from the 2011W animals slaughtered at 7 dpi were digested further using a pepsin digest. Abomasa were digested in 500 mL of 1% pepsin (Sigma Porcine Pepsin P7125)/1% HCl solution at 38 ℃ for 7 h (Jackson et al., 2004). After digestion the contents were made up to 800 mL with H_2O and fixed with 20 mL iodine. Nematode burden was calculated by extrapolating from 10% aliquot counts.

2.2.3.6 Fecundity

All mature female worms (vulva and uterine structure present) recovered whilst enumerating worm burden were mounted on slides with lactophenol, and the number of eggs *in utero* recorded. Whenever possible, measurements on at least 30 individual worms were taken per animal. Where 30 worms had not been obtained from the worm burden aliquots, samples were taken at random from the abomasal digest (75 μm sieve) sample until 30 female worms in total had been obtained. Females were photographed with a digital camera

under the dissecting microscope, and length determined using the public domain software ImageJ 1.37 (http://rsb.info.nih.gov/ij) after calibration with a stage micrometer slide.

2.2.3.7 FEC

Individual faecal samples were taken, and coccidia*, Nematodirus* and other Trichostrongyles counts established using the modified McMaster method (Ministry of Agriculture Fisheries and Food, 1986), described previously (section 2.2.2.1). Samples with a FEC of 0 epg were subjected to an additional floatation step using the method of Kelly et al. (2009).

2.2.3.8 Pepsinogen

On the day of sampling, vacutainers containing blood samples were spun in a bench top centrifuge at 1,000 x g for 15 min at 8 ℃. Plasma was removed and stored at -20 ℃ until assayed.

Pepsinogen concentration was determined using the Ross et al. (1967) modification of the method of Hirschowitz (1955). The principle of the test is that the inactive zymogen, pepsinogen, is converted by dilute hydrochloric acid (HCl) into the enzyme pepsin, which then degrades the serum proteins into peptides possessing tyrosine end-groups which are soluble in trichloroacetic acid (TCA). Non-hydrolysed proteins are precipitated and removed. The tyrosine end-groups react colourimetrically with Folin-Ciocalteu reagent in alkaline conditions; the colour produced is proportional to the concentration of pepsinogen in the samples.

Plasma samples were removed from storage and thawed to room temperature, at which point 0.5 mL of sample was acidified with 1.5 mL of 0.1 N HCl and 1 mL H_2O, and incubated for 3 hours at 37 ℃. Post incubation 2.0 mL 10% TCA was added, and after ten minutes at room temperature the mixture was centrifuged at 1910 x g for 15 min before 2 mL of supernatant was taken. Control samples were prepared as above, however TCA was added to the sample containing plasma, HCl and H_2O immediately, with no incubation step. A standard, containing 0.2 μm/mL tyrosine, was prepared using 1 mL working standard solution (tyrosine stock standard solution (2 μm/mL) prepared by dissolving 181.2 mg tyrosine in 0.1 N HCl and diluted to 500 ml, diluted 1 in 10) added to 0.2 mL H_2O and 0.8 mL 10% TCA. Blanks were prepared using 2 mL of H_2O. To all of the above samples 4 mL 0.5 N NaOH and 1 mL Folin-Ciocalteau reagent (diluted 1 in 3) were added. Absorbance of the standard, samples and controls were measured against the blank at 560 nm. Concentration of tyrosine (U/l) was calculated by

$$\frac{A_{Test} - A_{Control}}{A_{Standard}} x\, 5.55$$

The enzyme concentration is measured in international units U/L (μmol tyrosine released per 1 L plasma per minute).

2.2.3.9 ELISA

Antibody recovery

Blood for anitbody recovery was collected by jugular venipuncture into aseptic vacutainers containing no anticoagulant. To allow for clot formation vacutainers were left in a horizontal position at room temperature for 4 h after sampling, before being refrigerated at 4 ℃ overnight. The day following collection samples were centrifuged at 1,000 x g for 15 min at 8 ℃, before the serum was removed and stored at -20 ℃ until assayed.

Mucosal samples (Table 2.2.3.4) were prepared for antibody recovery using a modified version of the method of Sinski et al. (1995). Approximately 50 mg of tissue was homogenised in 1 mL of phosphate buffered saline (PBS) and 40 uL protease inhibitor cocktail (Sigma-Aldrich). After centrifugation of the homogenate at 14,000 x g for 15 min the supernatant was removed, and protein concentration determined using a Qubit® protein assay kit (Invitrogen) before storage at -20 ℃.

CarLA-specific IgA analysis

The CarLA-specific IgA analysis was carried out by AgResearch (New Zealand). To assay specific anti–CarLA IgA antibody, EIA/RIA plates (Costar 9017, medium binding, Corning Inc, USA) were incubated overnight at 4 ℃ with 100 μL/well of purified CarLA (1.25 μg/mL) in phosphate buffered saline (PBS). CarLA was prepared from a hot water extract of the larval surface antigens of exsheathed *T. colubriformis* L3s by gel filtration on Sephadex G-100 (Harrison et al., 2008). The plates were washed twice with reverse osmosis (RO) purified water containing 0.1% (w/v) Tween 20 (RO-T20) then blocked for 30 min at room temperature with 5% skim milk powder in 10 mM phosphate buffer, 0.65 M saline, pH 7.2 containing 0.5% Tween 20. Plates were washed twice with RO-T20 then used immediately.

Serum or mucosa samples were initially diluted 1/50 in sample dilution buffer (10 mM phosphate buffer, 0.25 M NaCl, pH 7.2 containing 1.0% Tween 20 and 0.25% bovine serum albumin), then stored at -20 ℃ until assayed. Diluted serum samples were added to plate wells at two dilutions, 50 μL and 20 μL and made to 100 μL final volume with sample dilution buffer (final dilutions of 1/100 and 1/250). Mucosa samples were assayed at three dilutions (1/50, 1/100 and 1/200). Samples were then incubated for 2 h at 37 ℃. Plates were then washed 6 times with RO-T20. Rabbit anti-sheep IgA conjugated with horseradish peroxidase (Bethyl Laboratories Inc, USA), diluted 1/2000 with ELISA buffer, was added to each well (100 μL) and incubated for 2 h at 37 ℃. Plates were then washed 6 times as above, and 100 μL/well of freshly prepared substrate [0.001% (w/v) tetramethyl benzidine

(AppliChem, Germany) containing 0.003% (v/v) hydrogen peroxide in 0.1 M citrate-acetate buffer pH 5.2] added. The reaction was allowed to develop for 30 min at room temperature, and then stopped with 50 μL/well of 2 N sulphuric acid. The absorbance was measured at a wavelength of 450 nm (Versamax, Molecular Devices, USA).

A reference standard method was used to obtain concentration values for CarLA-specific IgA (Peterman and Butler, 1989). To prepare the reference standard a pool of sera with high CarLA-specific IgA levels was diluted 1/1,250 with sample dilution buffer and this was given a nominal value of 0.8 units/mL. A six-fold dilution series of the standard (1/1,250 to 1/80,000) was then prepared and loaded onto each plate in duplicate (100 μL/well). The standard curve generated was transformed by taking the natural logarithm of unit values and a logit (Peterman and Butler, 1989) of the reference standard absorbance. A linear regression was fitted to the linear region of the logit-log standard curve and the equation for this straight line remodelled to calculate sample concentration. Individual well results were multiplied by each sample dilution and the average of this calculated. To determine the minimal detectable value for the assay, the mean absorbance value plus three standard deviations was calculated for wells consisting of sample dilution buffer multiplied by 100 (serum sample dilution). This was approximately 1.6 units/mL. Samples above the standard curve were diluted further and re-assayed. Internal controls were made by spiking saliva collected from parasite-free sheep with three different amounts of standard serum. The coefficient of variation (CV) for replicates of the standards at each dilution point were $\leq$7% within assays and $\leq$10% between assays; the CV for internal controls (high, medium and low anti-CarLA IgA) were $\leq$12% within assays and $\leq$23% between assays.

Enzyme-linked immunosorbent assay (ELISA) for IgA & IgG

ELISA was used to determine *T. circumcincta*-specific IgA and IgG in serum and mucosa samples. Antigen from *T. circumcincta* L3 was freshly prepared as previously described (Sinski et al., 1995). The wells of a 96-well polystyrene ELISA plate (Thermo Fisher Scientific, USA) were coated with 100 μL of L3 antigen (2 μg/mL) in carbonate-bicarbonate buffer (Sigma-Aldrich, UK) at pH 9.6 and left overnight at 4 ℃.

Plates were run on the DSX® ELISA processing system (Dynex Technologies, USA). The plate was washed four times using PBS-T (PBS + 1% Tween 20; Sigma-Aldrich, UK). An aliquot of 100 μL of either serum sample (diluted in PBS-T + 3% BSA; Thermo Fisher Scientific, USA) or mucosal sample (diluted in PBS-T + 3% BSA) (Figure 2.4) was added to each of 3 wells and incubated at 37 ℃ for 30 min, before another 4 washes in PBS–T.

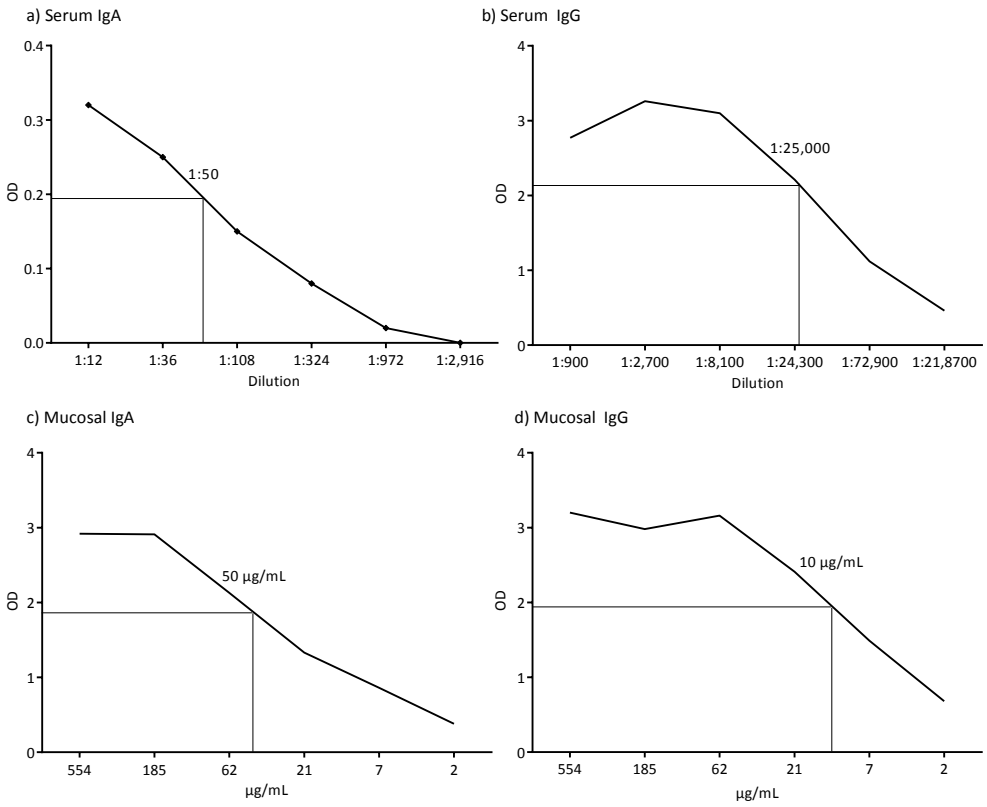


Figure 2.4: Serial dilutions for optimisation of nematode-specific ELISAs. Serial dilutions used to determine optimum serum dilution (A, B) and mucosal protein concentration (C, D) for *Teladorsagia circumcincta*-specific IgA and IgG ELISAs.

Samples were then incubated for 30 min at 37 ℃ with 100 µL of the primary antibody (diluted in PBS-T + 3% BSA), followed by 100 µL of the secondary antibody linked to horse radish peroxidase (HRP) conjugate (diluted in PBS-T + 3% BSA) (Table 2.4). Between primary and secondary antibody incubations plates were washed four times with PBS-T.

Table 2.4: Primary and secondary antibodies used for nematode-specific ELISAs. Description of antibodies used for *Teladorsagia circumcincta*-specific IgA and IgG ELISAs, along with recommended dilution factor.

	Antibody		**Dilution**
IgA	Primary	Mouse anti Bovine/Ovine IgA (AbD Serotec, UK)	1:1,000
	Secondary	Goat anti Mouse Ig:HRP (Dako, UK)	1:2,000
IgG	Primary	Rabbit anti Sheep IgG (AbD Serotec, UK)	1:20,000
	Secondary	Goat anti Rabbit IgG:HRP (AbD Serotec, UK)	1:2,000

After four final washes with PBS-T, 100 µL of chromogen tetramethylbenzidine (TMB) (Novex, UK) was added to each well and incubated for 15 min at room temperature. The reaction was stopped using 100 µL of 10% 1 N HCl and the optical density (OD) read at 450 nm. Each plate included a blank (PBS-T + 3% BSA) as a negative control as well as a pooled plasma or mucosa sample from a subset of infected animals as a positive control. All samples were assayed in triplicate.

2.2.4 Statistical Analysis

Log transformations were performed on worm burden (ln(X+1)), eggs per worm (ln(X+25)), and faecal egg counts (ln(X+25)) to stabilise the variance. Flock data were used to assess each haematology variable for normality of distribution; variables were log transformed (ln(X)) where appropriate.

All statistical analyses were performed using SAS® (v9.3) procedures. Flock haematology data from all animals (n=261) was used to determine if there was a relationship between FEC2 and the haematology variables from blood taken at the same time. A general linear model was used, with variables transformed where appropriate, and a model was fitted that had effects for rearing type, group by sex interaction, and FEC2. The interaction between group and the regression on FEC2 was examined, but was not significant. As 2011 was the only year in which data were available for both males and females within the same group, the 110 animals from the Lowland11 group (Figure 2.1) were used to estimate the effect of sex on the haematology parameters.

For the 2010W and 2011W cohorts a general linear model (GLM) was fitted with effects for phenotype (HighFEC or LowFEC), year, day post infection (7 or 14 dpi) and the interaction between year and phenotype. To account for repeated measures, the MIXED procedure was used to assess haematological and pepsinogen data in the 2011E cohort, with a model fitted with effects for dpi, phenotype, and their interaction. Only data from day 7 post infection onwards was used as this was when the animals were infected. Post-infection FEC data were classified by week, with data prior to day 26 excluded as FEC did not rise above zero. Due to extreme differences in worm burden the 2010W and 2011W cohorts were analysed separately.

2.3 Results

2.3.1 Flock-wide data

Animals were on average 19 weeks old by the time flock FEC reached 600 epg, when the first individual FEC measurements were taken (Table 2.5). Initially flock FEC was slower to rise in 2011 compared to 2010, with the Lowland11 lambs on average 41 days older than in the previous year at the time of the first individual FEC measurement. As the single-born ewe lambs (Hill11) had to be bought down from the hill, these animals were 20 days older than their Lowland11 contemporaries at the time of the initial FEC measurement.

Table 2.5: Number of days between significant events during the selection and controlled challenge of the Scottish Blackface lambs. Animals are grouped by grazing group (lowland pasture or hill pasture) and challenge cohort (wethers or ewe lambs in 2010 and 2011). Lambing is the date on which the first lamb was born, all subsequent numbers are the days from that point. Individual faecal egg counts (FEC) were taken after two natural infections (FEC1 and FEC2). These values were used to select animals from each challenge cohort (n=20), who were subsequently cleared of infection with an anthelmintic and given a controlled challenge of 30,000 *Teladorsagia circumcincta* L3. Animals were slaughtered at either 7, 14 or 71 days post infection (dpi).

Grazing group	No.	Challenge cohort	No.	Lambing	FEC1	FEC2	Cleared	Challenge	Slaughter	
Lowland10	92	2010W	20	0	105	163	178	219	226	(7 dpi)
								213	227	(14 dpi)
Lowland11	76	2011W	20				216	245	252	(7 dpi)
								239	253	(14 dpi)
	34	2011E	8	0	144	201	253	279	350	(71 dpi)
									351	(72 dpi)
Hill11	56		12	0	165	249			350	(71 dpi)
									351	(72 dpi)

There was no significant relationship between FEC2 and any of the haematology parameters on the same day. There was also no effect of sex on haematology parameters, although females tended towards a higher absolute number of circulating basophils (P = 0.053).

2.3.1.1 Selection of resistant and susceptible animals

The average selection differentials for HighFEC and LowFEC animals in the 2010W and 2011W cohorts were -2.2 and 1.8 respectively (Appendix 2.1); this corresponds to selecting the most extreme ~5% from each tail of the distribution. The difference between the average selection differentials for the HighFEC and LowFEC animals was larger for the 2011E cohort (2.7 and -3.1), with the most extreme ~1% of animals selected. As expected, the average FEC over the two natural infections (FEC1 and FEC2) of the HighFEC and LowFEC animals was markedly different (P < 0.001; Figure 2.5).

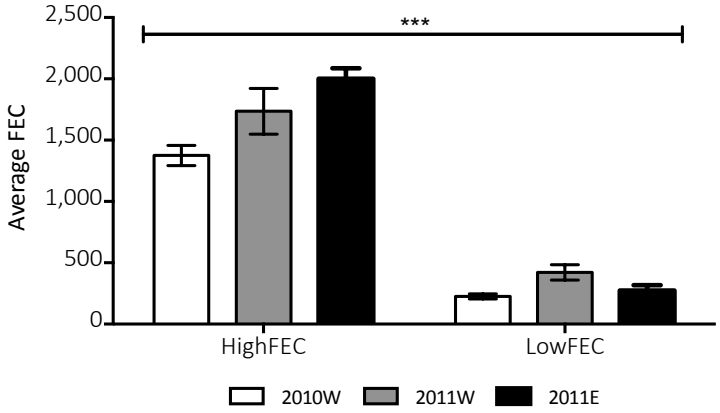


Figure 2.5: Mean faecal egg count (FEC) of selected animals over two natural nematode infections. Average FEC (± s.e.) over two natural nematode infections (FEC1 and FEC2) of HighFEC (n=10) and LowFEC (n=10) animals in the 2010W, 2011W and 2011E cohorts. (***) Indicates means differ (P < 0.001).

2.3.2 Controlled challenge

2.3.2.1 FEC over the course of infection

FEC was monitored weekly in the 2011E cohort after challenge with 30,000 *T. circumcincta* L3. Eggs were not observed in faeces until 28 days post infection (Table 2.6). Following this, egg counts rose to a maximum of 1,600 epg in the HighFEC group and 950 in the LowFEC group. Eggs were not observed at all in one of the LowFEC animals (1124975) throughout the course of infection, and in a further six animals (1 HighFEC and 5 LowFEC) egg counts did not rise above 50 epg. The difference in FEC between HighFEC and LowFEC animals was significant (P < 0.05), with a highly significant phenotype by week quadratic effect (P < 0.01; Figure 2.6) .

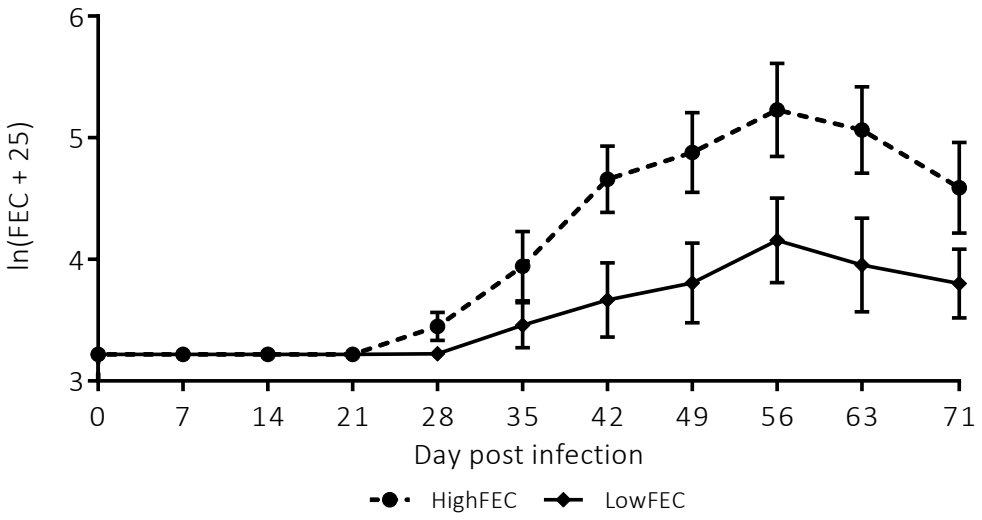


Figure 2.6: Mean faecal egg count (FEC) over the course of a controlled challenge. FEC (mean ± s.e.) in HighFEC (dashed line) and LowFEC (solid line) animals from the 2011E cohort following infection with $3x10^4$ *Teladorsagia circumcincta* L3 larvae. The mean FEC over the course of the infection was significantly different between the two groups ($P < 0.05$).

Table 2.6: Raw faecal egg counts (FEC) over the course of a controlled challenge. FEC in HighFEC and LowFEC animals from the 2011E cohort following infection with 3x10^4 *T. circumcincta* L3 larvae. Eggs were not observed in faeces until day 26. Values under 50 epg indicate eggs were observed in the floatation.

Group	Tag No	FEC1	FEC2	d26	d28	d30	d33	d35	d37	d40	d42	d44	d47	d49	d51	d54	d56	d58	d61	d63	d65	d69	d71
HighFEC	1124926	3225	1100	2	0	0	9	100	50	50	4	50	3	100	100	50	100	50	100	50	150	1	6
	1124935	1857	2275	0	0	0	2	50	100	50	100	150	100	150	150	400	600	200	200	400	250	100	100
	1124963	950	2225	2	50	5	7	50	50	100	100	50	50	150	150	100	100	1	100	25	100	50	50
	1124991	3225	650	1	1	1	0	2	50	9	100	150	300	192	200	200	200	50	150	100	4	5	13
	1125027	2500	1675	0	0	3	1	3	4	200	150	50	200	50	200	250	50	100	550	150	250	100	50
	1125048	2700	2375	1	0	0	7	50	100	5	5	50	8	50	6	50	5	100	100	50	50	0	50
	1125071	3400	675	50	50	4	100	100	10	150	100	50	26	50	250	100	21	0	50	6	1	0	6
	1125081	3675	550	2	100	200	650	600	950	600	500	1600	900	650	1350	1550	1600	900	700	375	650	50	200
	1125091	825	2575	0	0	0	1	2	0	0	2	3	1	0	1	0	1	1	0	50	1	0	0
	1125099	1825	1750	0	0	1	10	3	50	300	100	300	400	250	250	950	900	1500	1100	650	1350	150	1500
LowFEC	1124941	850	53	0	0	0	0	0	0	1	0	0	0	0	2	0	0	1	0	0	1	0	0
	1124975	850	61	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1124980	425	250	0	0	0	0	0	0	0	0	0	1	0	1	50	2	1	1	0	0	0	50
	1124987	200	75	0	0	0	0	0	0	0	0	0	2	2	6	50	100	3	6	0	50	1	6
	1125007	100	125	0	3	2	300	100	250	450	450	250	450	550	300	450	700	300	950	250	250	400	450
	1125009	474	52	0	0	0	1	50	0	150	150	100	50	350	150	450	550	400	600	150	200	0	50
	1125063	375	26	0	0	0	0	0	0	0	0	0	1	0	0	50	0	0	2	0	0	0	0
	1125068	576	175	1	0	0	1	0	0	0	0	1	0	0	50	1	0	0	0	0	0	0	0
	1125069	475	175	0	0	0	0	1	0	3	1	0	6	100	0	50	50	11	100	100	50	0	0
	1125097	100	75	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0

2.3.2.2 Nematode burden enumeration

There was no significant difference between HighFEC and LowFEC animals in worm burden, either at 7, 14 or 71 days post infection. A significantly lower number of worms was obtained from the 2011W (Figure 2.7B) and 2011E (Figure 2.7C) cohorts when compared to 2010 results (Figure 2.7A), and all data were therefore subsequently analysed separately by cohort. In the 2011E cohort the difference in adult worms as a proportion of total worms approached significance (99.7% for HighFEC and 84.6% for LowFEC; Kruskal-Wallis test; P = 0.066). No additional worms were found after re-digestion of a subset of abomasa using the pepsin digest technique, and therefore all nematode burden results presented below were those determined using the saline digestion technique.

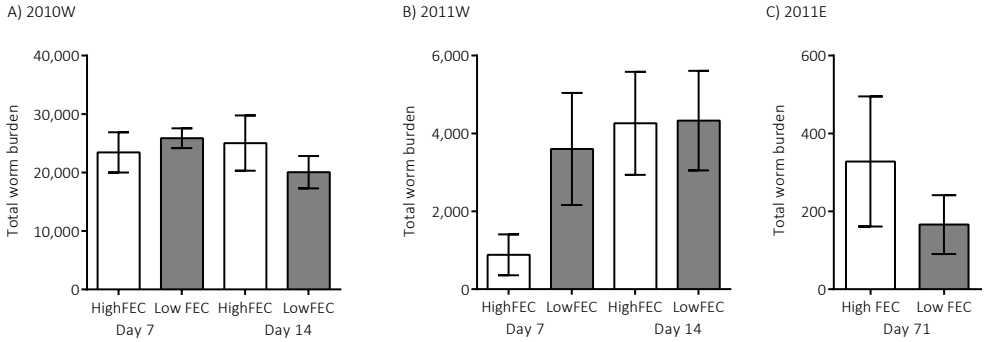


Figure 2.7: Total worm burdens. Total worm burden (mean ± s.e.) in the abomasum of HighFEC (white bars) and LowFEC (grey bars) animals following infection with $3x10^4$ *T. circumcincta* L3 larvae. Animals were slaughtered at 7 or 14 days post infection in the 2010W (A) and 2011W (B) cohorts, or at the end of a 71-day infection in the 2011E cohort (C). There were no statistically significant differences between HighFEC and LowFEC animals.

2.3.2.3 Fecundity

Adult *T. circumcincta* females were recovered from the abomasum of 13 of the 20 lambs in the 2011E cohort. The mean number of gravid female worms examined per animal was 27 (range 13–73). The within-phenotype correlation between the total number of adult *T. circumcincta* and FEC on the day of slaughter was 0.88 (P < 0.001). The mean (s.e.) length of female *T. circumcincta* in HighFEC and LowFEC animals was 11.3 (0.33) mm and 9.8 (0.49) mm, respectively (Figure 2.8A; P = 0.03). Back-transformed mean values for the number of eggs per gravid female were 30 and 15 for HighFEC and LowFEC animals, respectively (Figure 2.8B; P = 0.03). For gravid females, the correlation between length and number of eggs in utero on a within-animal basis was 0.33 (P < 0.01); the corresponding correlation for animal effects on a within-phenotype basis was 0.63 (P < 0.01).

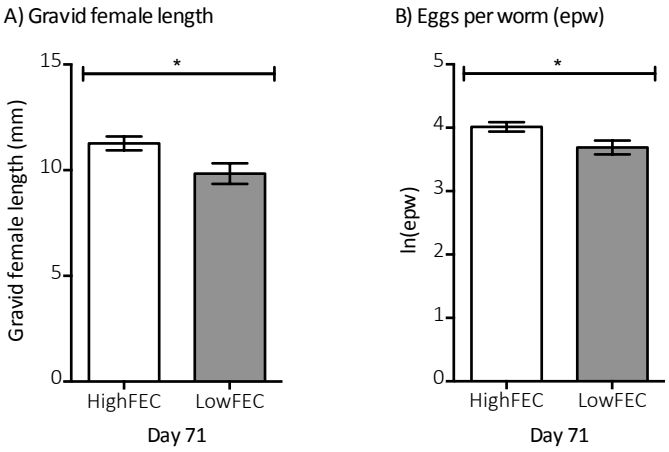


Figure 2.8: Worm fecundity traits. Mean (±s.e.) length of gravid females (A) and eggs per worm (B) from HighFEC (white bars) and LowFEC (grey bars) lambs following infection with 3x10^4 *Teladorsagia circumcincta* L3 larvae. (*) Indicates means differ (P < 0.05).

2.3.2.4 Pepsinogen

An increase in plasma pepsinogen levels was observed between day 7 and day 14 post infection in both 2010W and 2011W cohorts (Figure 2.9A), however the difference was only significant in 2010 (P < 0.01). In the 2011E cohort an increase was observed between 0 and 7 days post infection, with a subsequent decrease (Figure 2.9B).

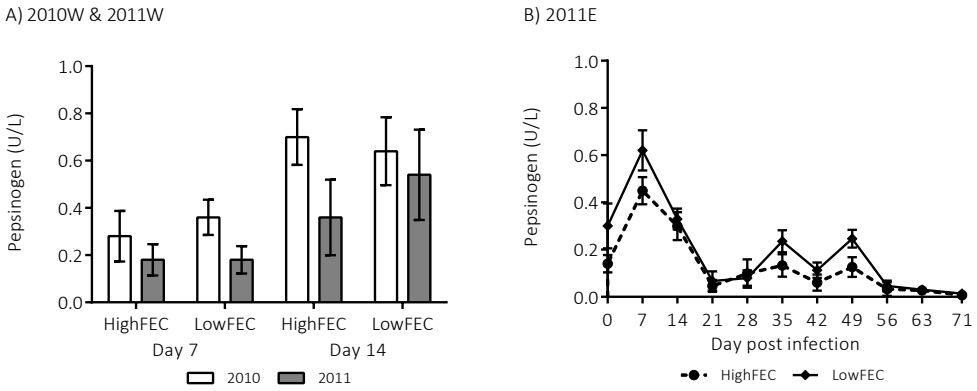


Figure 2.9: Plasma pepsinogen values. Plasma pepsinogen (mean ± s.e.) in HighFEC and LowFEC lambs following infection with 3x10^4 *Teladorsagia circumcincta* L3 larvae. Animals were slaughtered at 7 or 14 days post infection in the 2010W and 2011W (A) cohorts, or at the end of a 71-day infection in the 2011E cohort (B). There were no statistically significant differences between HighFEC and LowFEC animals.

2.3.2.5 Haematology

The mean circulating numbers of basophils, eosinophils, large unstained cells, lymphocytes, monocytes and neutrophils from both the 2010W and 2011W cohorts and the 2011E cohort

are presented in Figure 2.10 and Figure 2.11, respectively. For animals killed at 7 or 14 dpi the number of circulating basophils was significantly lower in HighFEC animals in 2010 (Figure 2.10A; P = 0.03). Day post infection was a significant source of variation for the number of circulating monocytes in both 2010 (Figure 2.10E; P < 0.01) and 2011 (P = 0.04), and large unstained cells in 2011 (Figure 2.10C; P < 0.01).

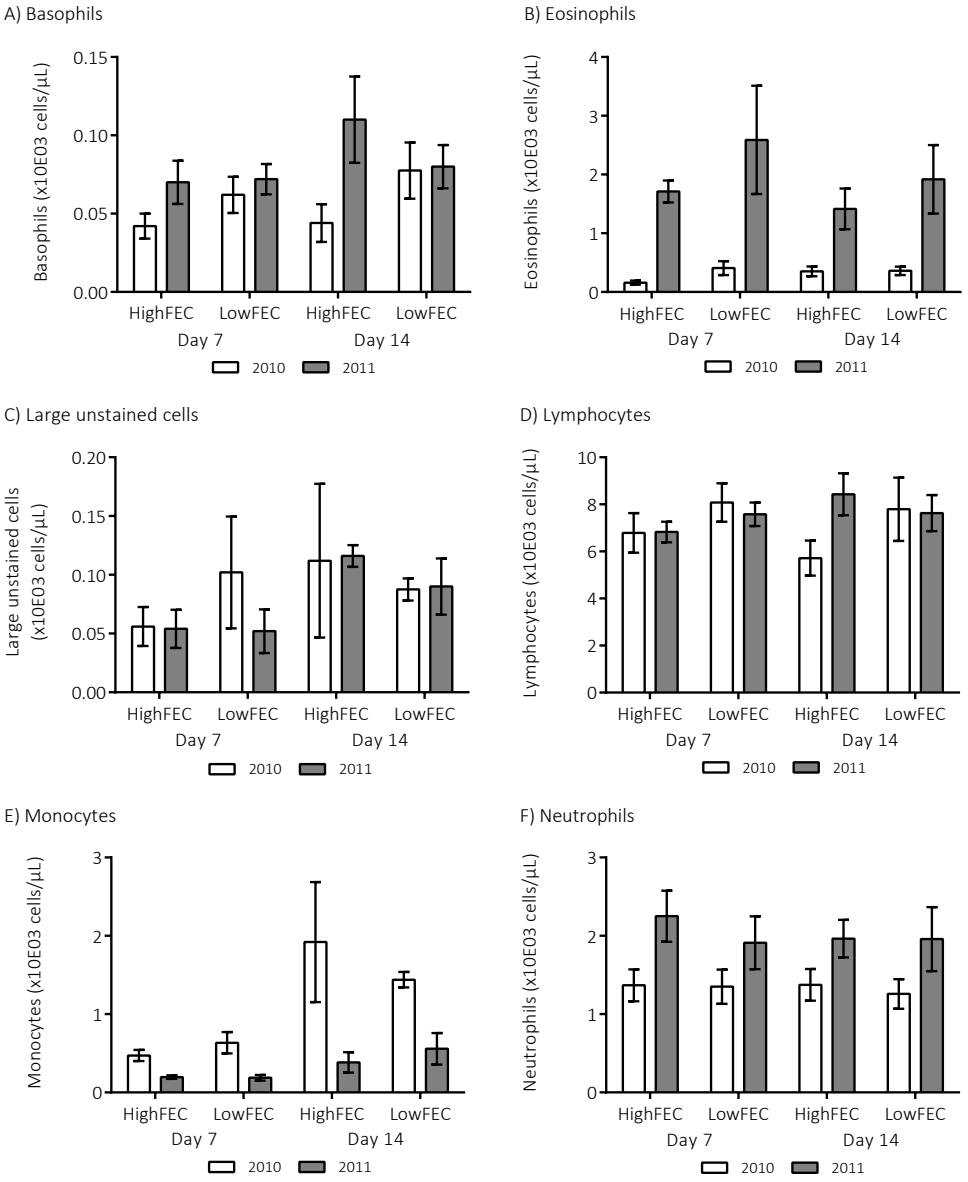


Figure 2.10: Haematological measurements at slaughter for 2010W and 2011W cohort lambs. Haematological measurements at slaughter (mean ± s.e.) for the 2010W and 2011W HighFEC and LowFEC lambs following infection with $3x10^4$ *Teladorsagia circumcincta* L3 larvae. Phenotype was a significant source of variation in the number of circulating basophils (A; P = 0.03).

For the 2011E cohort animals, slaughtered at 71 dpi, there was no effect of phenotype on any of the haematology parameters, although basophils approached significance (Figure 2.11A; P = 0.07). Day post infection was significant for all variables (P < 0.01), however there was no significant interaction with phenotype.

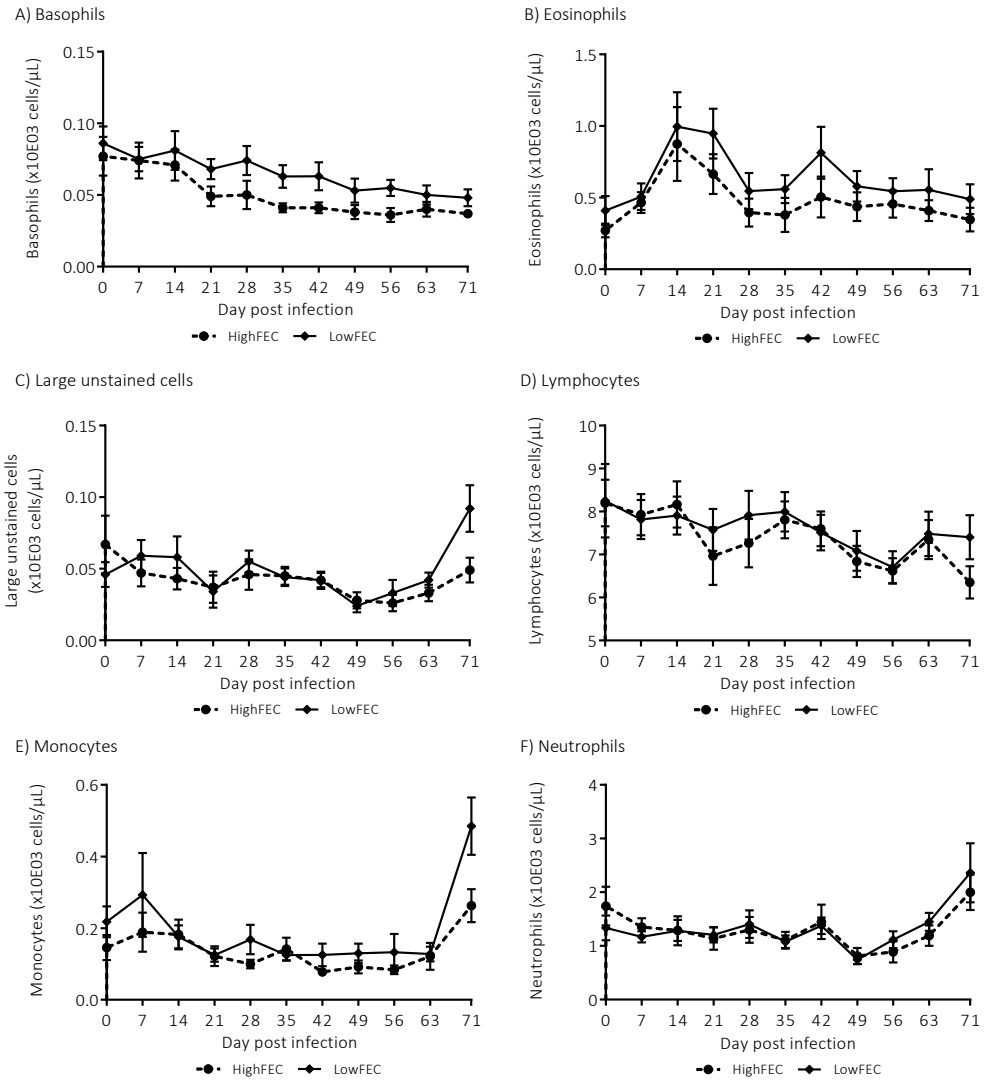


Figure 2.11: Haematological measurements at slaughter for 2010E cohort lambs.Haematological measurements (mean ± s.e.) over the course of infection in 2011E cohort HighFEC and LowFEC lambs following infection with $3x10^4$ *Teladorsagia circumcincta* L3 larvae. There were no statistically significant differences between HighFEC and LowFEC animals.

2.3.2.6 ELISA

The levels of anti-CarLA IgA was measured in HighFEC and LowFEC lambs. Day post infection was a significant source of variation of CarLA-specific IgA in the serum of the 2011E (Figure 2.12C; $P < 0.0001$) and 2010W (Figure 2.12A; $P = 0.04$) cohorts. The selected 2011E LowFEC animals had significantly higher CarLA-specific IgA levels in both serum over the course of infection (Figure 2.12C; $P = 0.04$), and mucosa at slaughter (Figure 2.12D; $P = 0.04$), than their HighFEC counterparts.

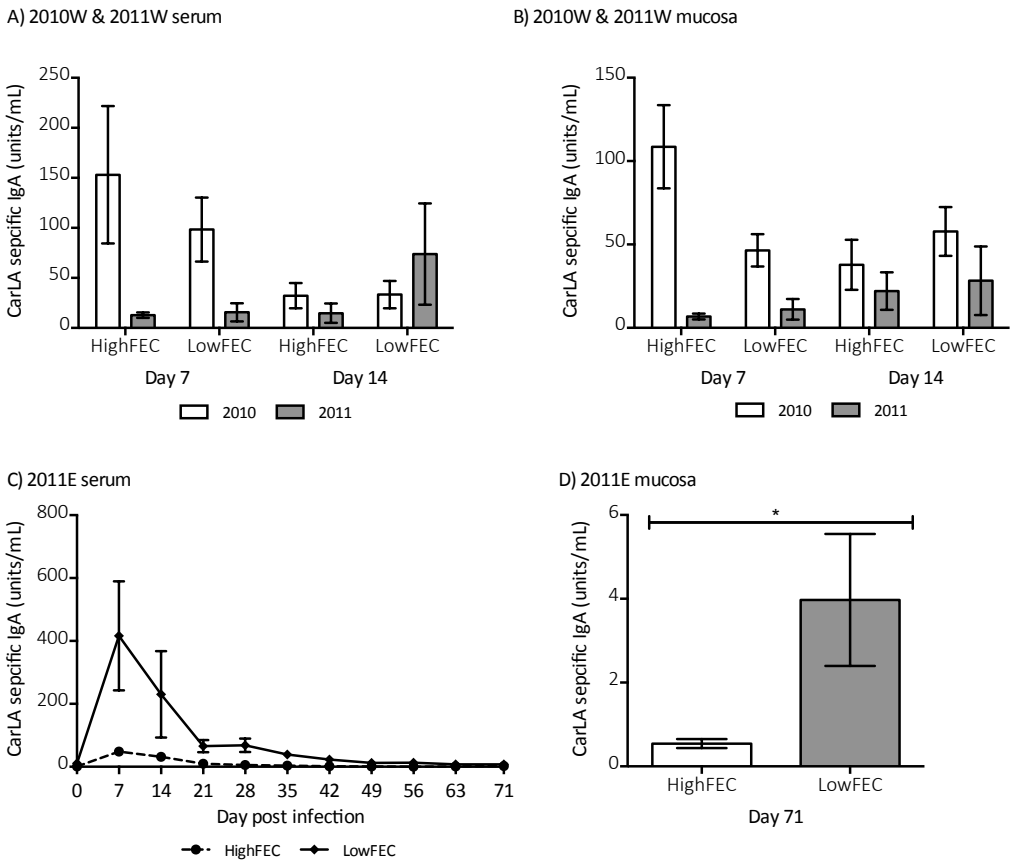


Figure 2.12: CarLA-specific IgA levels in serum and mucosa. CarLA-specific IgA levels (mean ± s.e.) in serum and abomasal mucosa of HighFEC and LowFEC lambs following infection with $3x10^4$ *Teladorsagia circumcincta* L3 larvae. Animals were slaughtered at 7 or 14 days post infection in the 2010W and 2011W (A, B) cohorts, or at the end of a 71-day infection in the 2011E cohort (C, D). Phenotype was a significant source of variation in the serum of 2011E animals over the course of infection ($P = 0.04$), and in mucosa at slaughter ($P = 0.04$).

In addition to CarLA, anti *T. circumcincta* L3 IgA levels were also determined. In the 2010W challenge cohort *T. circumcincta*-specific IgA levels were significantly higher at 7 days post infection in both serum (Figure 2.13A; $P < 0.01$) and mucosa (Figure 2.13B; $P < 0.01$). In the 2011E challenge cohort LowFEC animals had significantly higher IgA levels in serum over the course of infection ($P = 0.02$), with a spike at 7 days post infection (Figure 2.13C).

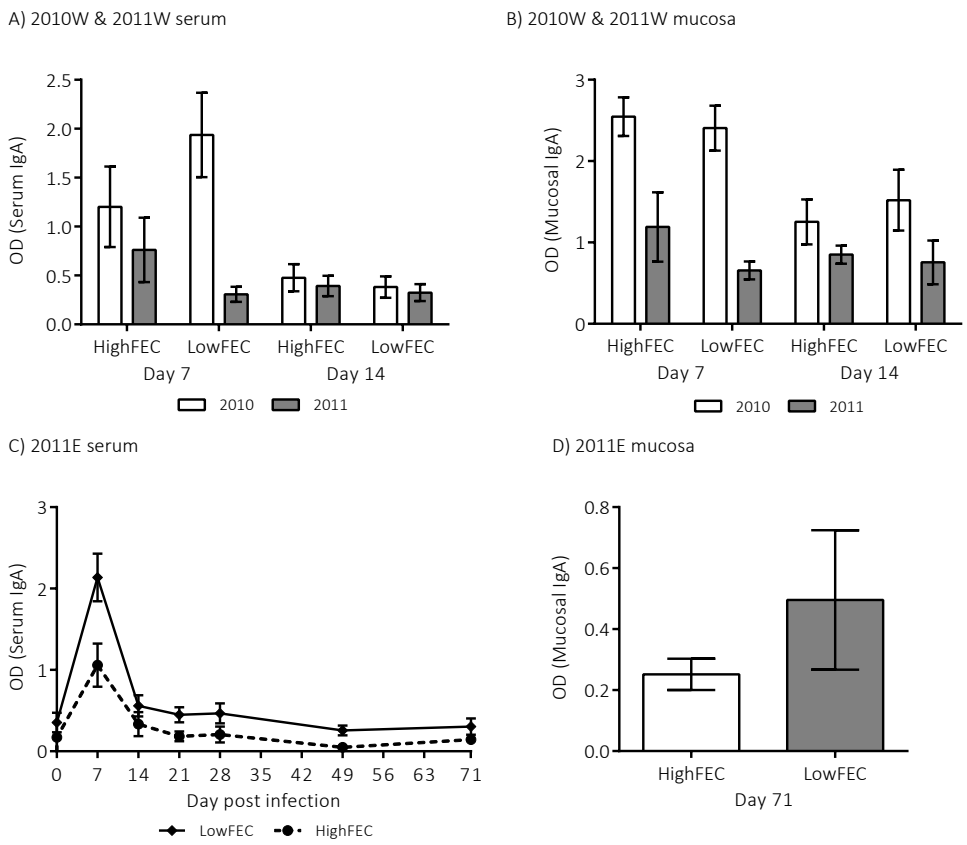


Figure 2.13: *Teladorsagia circumcincta*-specific IgA levels in serum and mucosa. *T. circumcincta*-specific IgA levels (mean ± s.e.) in serum and abomasal mucosa of HighFEC and LowFEC lambs following infection with $3x10^4$ *T. circumcincta* L3 larvae. Animals were slaughtered at 7 or 14 days post infection in the 2010W and 2011W (A, B) cohorts, or at the end of a 71-day infection in the 2011E cohort (C, D). Phenotype was a significant source of variation in the serum of 2011E cohort animals (P = 0.02).

There were no significant differences in *T. circumcincta*-specific IgG levels between HighFEC and LowFEC animals in any of the challenge cohorts (Figure 2.14). Day post infection was a significant source of variation for in the serum of the 2011E challenge cohort ($P < 0.0001$), with levels increasing from challenge until 21 days post infection, after which they declined.

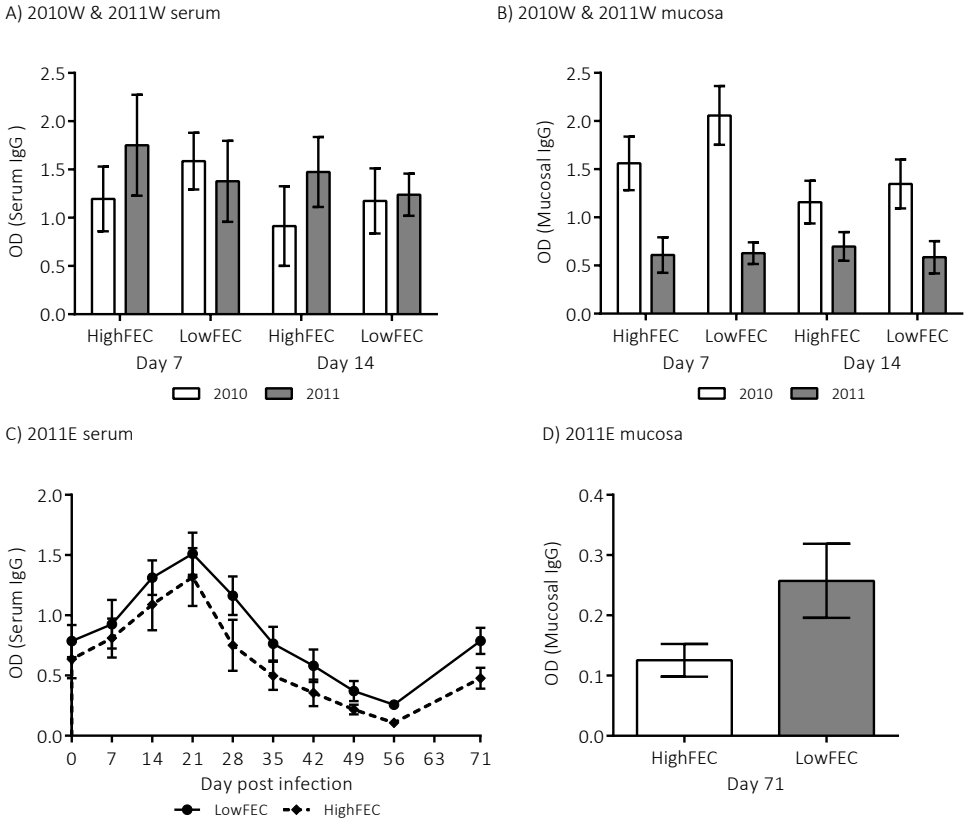


Figure 2.14: *Teladorsagia circumcincta*-specific IgG levels in serum and mucosa. *T. circumcincta*-specific IgG levels (mean ± s.e.) in serum and abomasal mucosa of HighFEC and LowFEC lambs following infection with $3x10^4$ *T. circumcincta* L3 larvae. Animals were slaughtered at 7 or 14 days post infection in the 2010W and 2011W (A, B) cohorts, or at the end of a 71-day infection in the 2011E cohort (C, D). There were no statistically significant differences between HighFEC and LowFEC animals.

2.4 Discussion

2.4.1 Validation of selection model

This is the first study in Ireland to confirm within-breed differences in the ability of Scottish Blackface lambs to resist gastrointestinal nematode infection. In the 2011E cohort, selected LowFEC animals achieved lower FEC throughout the course of infection, validating our model of selection.

Variation among hosts in nematode egg output can be a result of variation in the adult worm burden, variation in the average fecundity of each worm, or a combination of both (Stear et al., 1996). Results of Good et al. (2006) with Texel and Suffolk animals showed that differences in FEC are a result of differences in worm burden rather than in average worm fecundity. Evidence from Scottish Blackface animals however indicates that variability in FEC is largely due to differences in average worm fecundity (Stear et al., 1996), although adult worm burden is also a contributing factor (Stear et al., 1995b).

At the time of slaughter worm burden was not significantly different between the HighFEC and LowFEC animals, although it was numerically lower in the LowFEC lambs. However total worm numbers in both groups were low as animals were not reinfected during the study. From the FEC data it appears that the infection was beginning to tail off, and therefore slaughter at 70 days post infection may have been too late to capture the strongest variation in the number of adult worms.

Worm fecundity was significantly lower in the LowFEC (resistant) animals, with shorter, less fecund adult females. The correlation between worm length and eggs in utero is also indicative of the length of infection, and is in agreement with the observation from FEC data that the infection was beginning to wane. While there were no significant differences in worm burden between the HighFEC and LowFEC animals, there was a correlation between number of mature adults and FEC on day of slaughter. This indicates that adult worm burden may play a role in reduced FEC in Scottish Blackface animals, however as discussed above, the reduction in variability in FEC by the time the animals were slaughtered indicates that the peak worm burdens in these animals had passed.

While there were differences in parasitological parameters between the HighFEC and LowFEC groups, neither plasma pepsinogen nor any of the haematology variables could differentiate the two groups. Pepsinogen levels have previously been shown to be an indicator of ostertagiosis (Lawton et al., 1996; Balic et al., 2000b; Davies et al., 2005), however other studies of within-breed differences have found that it is not significant (Stear et al., 1995a). Therefore, plasma pepsinogen concentration alone may be an unreliable method of differentiating resistant or susceptible individuals (Lawton et al., 1996).

While there was no effect of phenotype on any of the haematology variables, an increase was observed in the number of large unstained cells, monocytes and neutrophils on the day

of slaughter. This could potentially be a result of stress from the animals being moved to the site of slaughter. Neutrophils have been shown to increase as a response to weaning stress in calves, however major changes in monocyte number were not observed (O'Loughlin et al., 2011).

Levels of serum anti-nematode IgA peaked at 7 days post infection. There was a significant difference between phenotypes, with resistant animals having higher levels of both anti-*T. circumcincta* and anti-CarLA serum IgA throughout the course of infection. IgA has been widely reported as mediating suppression of gastrointestinal nematode growth and fecundity during infection with *T. circumcincta* (Strain et al., 2002; Martínez-Valladares et al., 2005; Beraldi et al., 2008), with a strong IgA response often observed against 4^{th}-stage larvae in previously exposed animals (Stear et al., 1995b).

IgA is active at the site of infection, the abomasal mucosa; however plasma and mucosal IgA are correlated (Martínez-Valladares et al., 2005). In sheep, unlike in humans, plasma IgA is drawn from mucosal surfaces, and binds nematodes and the excretory-secretory molecules they release. A proportion of the unbound IgA is then transferred to the blood via the lymphatic system (Prada Jiménez de Cisneros et al., 2014). The relationship between the two parameters is strong but nonlinear, with the major determinants of plasma IgA being both worm mass (burden and size) and mucosal IgA activity (Prada Jiménez de Cisneros et al., 2014). This must therefore be taken into account when interpreting our results.

Resistant animals had significantly higher levels of serum anti-*T. circumcincta* IgA throughout the infection. While mucosal levels of anti-*T. circumcincta* IgA were numerically higher in resistant animals on the day of slaughter, this was not significant. The resistant animals therefore have larger quantities of unbound IgA entering the bloodstream, particularly at day 7. This could potentially be due to excess IgA production, or a result of more free IgA due to low worm numbers or reduced worm length.

The carbohydrate larval antigen (CarLA) is purified from *Trichostrongylus colubriformis* L3 larvae, however there is an epitope on the CarLA molecule that is common to CarLA from a wide range of gastrointestinal nematode species (Harrison et al., 2003b,a). This epitope is hidden when CarLA is present on the L3 but is available for detection by the immune system once CarLA is released as L3 larvae moult to become L4 stage larvae. The response detected in the serum and mucosa samples is most likely to this epitope. Indeed, challenge of *T. circumcincta*-immune sheep has been shown to induce a local antibody response to a molecule with very similar properties to CarLA (Balic et al., 2003). Animals identified as having 'high levels' of salivary anti-CarLA IgA have been shown to have 20–30% lower FEC during a mixed-species infection than animals with low or undetectable titres (Shaw et al., 2012). While the anti-CarLA IgA assayed by Shaw et al. was from saliva rather than serum and mucosa, these results are in agreement with our findings. The serum IgA response to CarLA was found to mirror that of the response to L3 *T. circumcincta* antigen, indicating that the CarLA test could potentially be used as a proxy if *T. circumcincta* antigen could not be sourced.

L4 larvae were observed in one animal, indicating that hypobiosis had occurred. This animal (1124975) had the highest levels of both anti-*T. circumcincta* and anti-CarLA IgA in the mucosal samples. Previous studies have shown that the number of inhibited larvae is positively associated with the size of the local IgA response to 4^{th}-stage larvae (Stear et al., 1995b). The antigen used in both ELISA's in this study was generated from L3 larvae, however it has been reported that there is a correlation ($r = 0.68$; $P < 0.001$) between the IgA response to both L3 and L4 antigen (Stear et al., 1995b). The presence of inhibited larvae is not surprising, as at the time of infection the animals were over 9 months old, by which time protective immunity is becoming established (Vlassoff et al., 2001; Abbott et al., 2009). Female lambs also have a stronger immune response than their male counterparts, who are less resistant to the establishment of infection (Smith et al., 1985; Gulland and Fox, 1992; Barger, 1993). This may also account for the animals in which there were very low FEC over the course of infection.

2.4.2 The host response to the larval stages of infection

The 2010W and 2011W cohorts were used to define the host response to the larval stages infection, with animals slaughtered at either 7 or 14 days post infection. While the plan had been to combine results from the two cohorts, a significant difference in worm burden was observed between years. The 2011W cohort were given L3 that had been stored in H_20 at 4 ℃ for 12 months, and this could have contributed to the reduced burden. There was however a large range in the total worm burden among the 2011W cohort lambs, the reason for which is unknown. To ensure that all worms were recovered from the abomasa, an additional pepsin digestion was carried out on a subset of abomasa. This digest did not result in any additional worms being detected however. Due to the differences in worm burden between the two cohorts, all data were analysed separately. Worm burden was not significantly different between the HighFEC and LowFEC groups in either cohort. This was not surprising as variation previous studies in Scottish Blackface indicate that resistance is expected to primarily manifest as reduced worm fecundity (Stear et al., 1995b, 1996), which could not be recorded in these animals. For animals that manifest resistance as a reduced worm burden it has been shown that while differences in immune response between resistant and susceptible breeds are mounted earlier in infection, differences in actual worm burden are not apparent until approximately 21 days post infection (Hassan et al., 2011a).

While there was no significant difference in plasma pepsinogen between the HighFEC and LowFEC animals, levels increased between day 7 and day 14 post infection, This is indicative of increasing gastric mucosal damage as infection progresses. A higher number of basophils were observed in the LowFEC compared to the HighFEC animals in 2010. Induction of basophils is a feature of the anti-helmintic response and drives a Th2-type immune response (Allen and Maizels, 2011). Therefore LowFEC animals may be generating

a more effective immune response to infection. This may not have been apparent in the 2011W lambs due to the reduced challenge. The number of circulating eosinophils and monocytes was significantly different between the two cohorts, which may be indicative of the difference in response to a low and high worm burden. Eosinophils were higher in the 2011W cohort. An increase of eosinophils is characteristic of helminth infection (Meeusen and Balic, 2000), however a reduction of peripheral eosinophilia has been observed during the period of primary infection following challenge of lambs with *T. circumcincta* (Sutherland et al., 1999). It was hypothesised that this may have resulted from the recruitment of cells into the intestinal epithelium (Sutherland and Scott, 2009). The higher challenge in the 2010W group may therefore have resulted in an increased migration to the site of infection. The number of circulating monocytes was higher in the 2010W cohort. Monocytes are part of the innate immune system, and have multiple roles including differentiating into macrophages and dendritic cells in response to inflammation (Murphy et al., 2008). Both macrophages and dendritic cells are foremost among the cells that recognise, process and present antigens in the gastrointestinal tract (Maizels and Yazdanbakhsh, 2003; Maizels et al., 2009). The higher levels of monocytes in the 2010W lambs may reflect the increased antigen challenge experienced by this cohort.

It has previously been reported that in 9 month old Scottish Blackface sheep infected with *T. circumcincta*, the plasma IgA response against L3 peaks at 8-10 days post infection, and subsequently declines (Henderson and Stear, 2006). This peak is despite evidence that the majority of L3 mature into L4 by 4 days post ingestion (Armour et al., 1966), therefore the observed peak may be due to cross-reactivity of anti-L3 antigen with other larval stages (Prada Jiménez de Cisneros et al., 2014). While there is a clear difference between the serum anti-nematode IgA levels in HighFEC and LowFEC animals at 7 days post infection in the 2011E cohort, as described above, this is not the case in either the 2010W or 2011W cohorts, although the LowFEC 2010W animals tend towards a higher anti-*T. circumcincta* IgA at day 7. In the 2010W cohort anti-CarLA IgA levels for both phenotypes peak at 7 dpi, whereas in the 2011W cohort animals levels were higher at 14 dpi. This may be due to these animals being sampled when IgA levels were either rising or falling, as both 7 and 14 days post infection sit just outside the range of the peak observed in the study by Henderson and Stear (2006). This could be rectified in future studies by taking a baseline blood sample at day 0, along with more regular blood samples from animals from both time points. This could also be a result of sex, or the increased power due to the larger number of animals in the 2011E cohort. When selected the 2011E cohort lambs were more divergent, representing 1% of the tails of distribution. Additionally, the 2011E cohort lambs were older at the time of slaughter, so could be mounting a more rapid immune response to infection.

2.4.3 Conclusions

In summary, this study was successful in identifying Scottish Blackface lambs with divergent phenotypes for gastrointestinal nematode resistance. Resistant lambs, as identified by a low faecal egg count (FEC) have reduced nematode fecundity, with shorter, less fecund adult females. The reduction in worm fecundity observed in these animals may be a result of anti-nematode IgA levels. There was some evidence that the HighFEC animals tended towards a higher worm burden in both the 2010W cohort animals at 14 days post infection, and the 2011E animals. The lowered FEC in these Scottish Blackface lambs may therefore be a result of not only reduced worm fecundity but also reduced worm burden, as is seen in other populations (Stear et al., 1996). Anti-CarLA IgA levels were found to mirror those of anti-*T. circumcincta* IgA, and therefore the test has the potential to be used when *T. circumcincta* antigen cannot not be sourced.

Chapter 3

Transcriptome profiling of the abomasal lymph node of Scottish Blackface lambs with divergent phenotypes for resistance to gastrointestinal nematodes

3.1 Introduction

Resistance to GIN is moderately heritable (h^2 ~ 0.3) (Safari et al., 2005; Bishop and Morris, 2007), therefore one sustainable method of nematode control is to select for genetically resistant individuals (Kemper et al., 2009). Selection using phenotypic traits such as faecal egg count (FEC) requires prior exposure to GIN, whereas selection could be simplified through the identification of molecular markers. However, any such strategy would benefit from a detailed understanding of the genes and mechanisms involved in expressing a resistant phenotype and the factors that regulate this response. The molecular mechanism of resistance may vary between breeds and populations (Brown et al., 2013; Sayre and Harris, 2012) and so investigations of the host response must be tailored to the population in question.

Transcriptome and proteome analyses are powerful methods for the identification and quantification of genes and proteins expressed during a physiological perturbation. These tools have led to a greater understanding of the molecular basis of phenotypic variation in resistance to GIN (Pemberton et al., 2011; Nagaraj et al., 2012; Pemberton et al., 2012; Ahmed, 2013; Gossner et al., 2013). During the past decade both bovine- and ovine-specific microarrays have been used in sheep to provide global gene expression information. A number of studies have used microarrays to identify genes associated with the host response to GIN in the duodenum (Diez-Tascon et al., 2005; Keane et al., 2006, 2007), abomasal mucosa (Rowe et al., 2009; Knight et al., 2011) and lymph (MacKinnon et al., 2009; Andronicos et al., 2010; Knight et al., 2010; Gossner et al., 2013) in a variety of breeds. These studies have identified hundreds of genes that are differentially expressed (DE) between known resistant and susceptible animals.

Recently high-throughput sequencing of cDNA (RNA-Seq) has allowed the examination of RNA expression on an even larger scale (Wang et al., 2009). While microarray technology relies on prior knowledge of genomic sequence, RNA-Seq allows the detection of unknown genes, alternative splice sites and novel isoforms. Another advantage that is particularly relevant to sheep is the ability to re-analyse the data once more genomic information becomes available (e.g., the release of an updated genome or transcriptome). Transcriptome analysis in sheep has been aided by the recent release of a ovine reference genome (Kijas et al., 2012; Flicek et al., 2014). To date GIN resistance using RNA-Seq has been investigated in only two separate studies (Pemberton et al., 2011; Ahmed, 2013; Gossner et al., 2013). Pemberton et al. (2011) and, subsequently Gossner et al. (2013) examined gene expression in the abomasal lymph node of resistant and susceptible Scottish Blackface lambs trickle infected with *T. circumcincta* in comparison to sham infected controls. Many of the genes with increased expression in resistant lambs were key regulators of a T_H2 immune response, although the authors concluded that resistance/susceptibility is not simply a matter of T_H1/T_H2 discrimination, as susceptible sheep also showed increased expression of T_H2-type genes (Gossner et al., 2013). A major part of the response in

resistant lambs also appeared to include control and suppression of acute inflammation (Gossner et al., 2013). Ahmed (2013) compared expression in the abomasal lymph node of Suffolk (relatively susceptible) and Texel (relatively resistant) lambs artificially infected with *T. circumcincta*. A differential polarisation of the immune response was found between the breeds, with Texel lambs have increased expression of T_H1, T_H2 and T_{reg} genes in a synchronised manner post-infection while Suffolk reduced T_H1 gene expression without significant T_H2 or T_{reg} induction (Ahmed, 2013). The aim of this project was to sample the transcriptome of Scottish Blackface lambs with divergent phenotypes for GIN resistance, in order to identify genes and biological processes associated with the host response to GIN in resistant and susceptible individuals.

3.2 Materials and Methods

3.2.1 Ethical approval

All animal procedures described in this study were conducted under experimental license from the Irish Department of Health in accordance with the Cruelty to Animals Act 1876 and the European Communities (Amendments of the Cruelty to Animals Act 1976) Regulations, 1994.

3.2.2 Experimental design

The selection and challenge of the animals used in this experiment have been described previously (section 2.2.3). Briefly, Scottish Blackface lambs, grazed together from birth, were monitored for faecal egg count (FEC; 2 samples following each of 2 independent natural infections). These observations were used to identify the most resistant (n=10, LowFEC) and susceptible (n=10, HighFEC) individuals based on individual animal values for ln(FEC+25) estimated using mixed model procedures. The selected lambs were given a controlled challenge of 30,000 *Teladorsagia circumcincta* larvae (L3) at 39 (range 29 - 50) weeks of age, and slaughtered at either 7 or 14 days post infection (dpi) to define the acute responses to infection. This procedure was replicated across 2 years (2010 and 2011).

3.2.3 Tissue samples

Immediately post slaughter abomasal lymph nodes (ALN) were removed and collected. The tissue was cut into pieces approximately 5 mm^3 and immersed in 10 volumes of RNAlater® solution (Ambion, USA). This was kept at room temperature for 24 h, before long-term storage at -80 ℃.

3.2.4 RNA extraction

Total RNA was extracted from abomasal lymph node stored in RNA later using Sigma TRI Reagent® (Sigma Aldrich, UK). Approximately 0.1 g of tissue was homogenised in 3 ml TRI Reagent®, and incubated at room temperature for 5 min before being transferred into Eppendorf tubes (1 ml/tube), to each of which 200 µl chloroform was added. Tubes were shaken vigorously, then incubated at room temperature for 2 to 3 min before centrifugation at 12,000 x g for 15 min at 4 ℃. The RNA was precipitated by transferring the resulting colourless upper aqueous phase into a clean Eppendorf tube, adding isopropanol (0.6 times the volume of the aqueous phase), mixing and centrifuging at 12,000 x g for 10 min at 4 ℃. The supernatant was removed, and the RNA pellet washed by adding 1 ml 75%

EtOH before centrifuging at 7,500 x g for 5 min at 4 ℃. Excess ethanol was aspirated before the pellet was allowed to air dry. RNA was resuspended in 50 μl nuclease free water. Total RNA was cleaned up using the RNeasy Mini Kit (Qiagen, Germany) and an in-solution DNase digestion (RNase-free DNase set; Qiagen, Germany) as per the manufacturers' recommendations. RNA quality was assessed using an Agilent® RNA 6000 Nano Assay on the 2100 Bioanalyzer, and total RNA quantified using the NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, UK). All RNA samples were stored at -80 ℃.

3.2.5 Library preparation and sequencing

Illumina TruSeq™ libraries were prepared following the TruSeq™RNA sample preparation v2 guide (Part #15026495 Rev. B; Appendix 3.1; Figure 3.1), using total RNA from cleaned abomasal lymph node samples.

Briefly, 3 μg of poly-A containing mRNA was captured and purified using poly-T oligo-attached magnetic beads. The mRNA was subsequently randomly fragmented using divalent cations under elevated temperature, which results in more uniform sequencing coverage. The cleaved RNA fragments were primed with random hexamers and reverse transcribed into cDNA. This was followed by second strand cDNA synthesis using RNase H and DNA Polymerase I, creating double stranded (ds) cDNA.

During the end repair process any remaining overhangs from the DNA Polymerase I step were fixed into blunt ends, and 5' phosphates and 3' hydroxyls added, as chemically synthesised DNA does not have 5' phosphate groups, which are needed for downstream ligation. The 3' ends of each cDNA read were then adenylated, creating an overhang, to prevent them from ligating to one another during the adapter ligation process. Unique sequencing adapters, which hybridise to the flow cell during sequencing, were ligated to the ends of the ds cDNA using the 3'-A overhang, and these fragments were enriched using PCR.

The standard protocol was followed aside from the following: in the PCR step the number of PCR cycles was reduced to 10, and to avoid bead contamination the PCR products were cleaned up using a Qiagen MinElute column rather than AMPure XP beads. Libraries were visualised using an Agilent® DNA 1000 assay on the 2100 Bioanalyzer, and quantified using the Qubit® dsDNA BR assay (Invitrogen, UK) as per the manufacturers' recommendations.

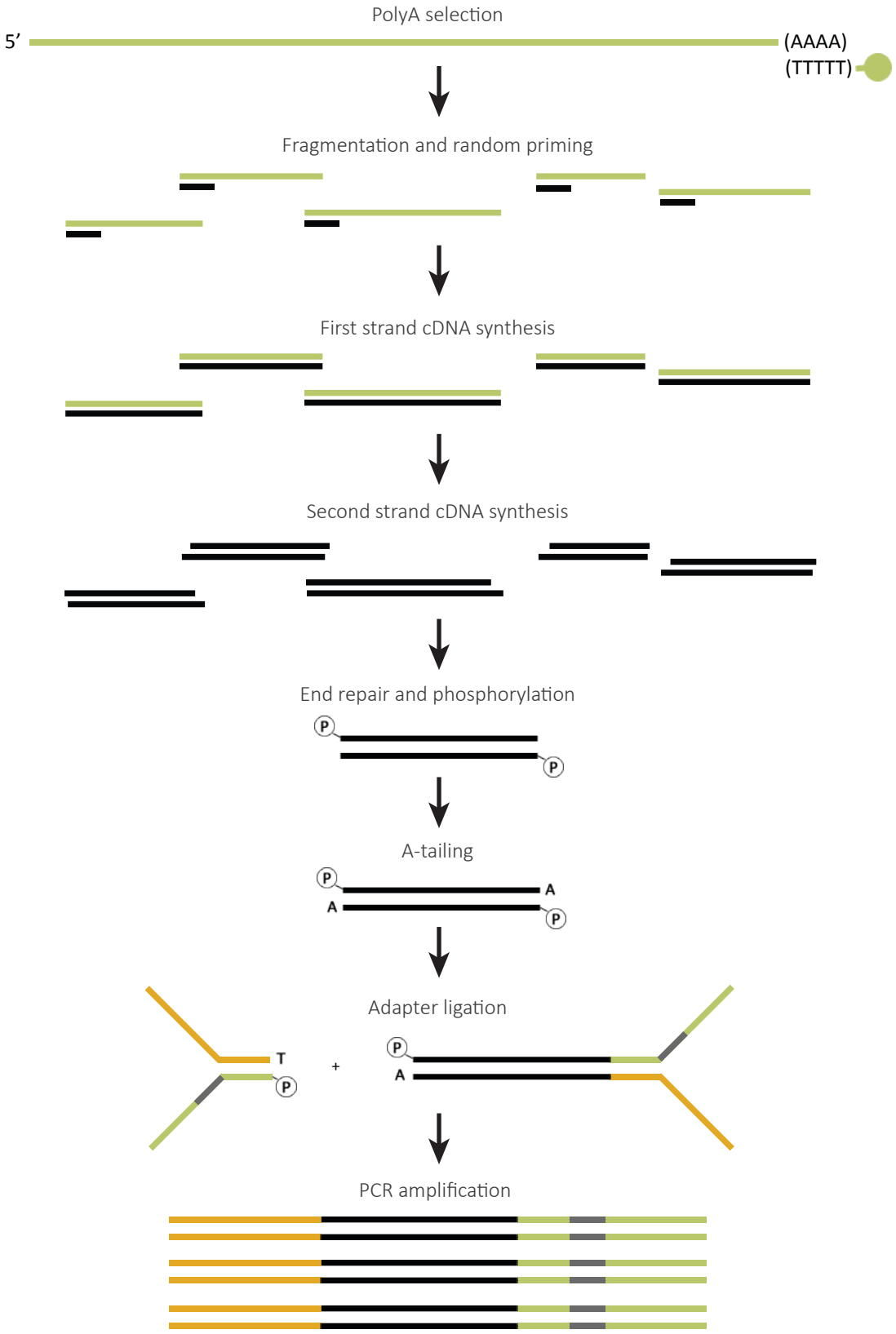


Figure 3.1: TruSeq^{TM}RNA sample preparation. Poly-A containing mRNA is captured and purified using poly-T oligo-attached magnetic beads. The mRNA is subsequently randomly fragmented, then primed with random hexamers before being reverse transcribed into cDNA. This is followed by second strand cDNA synthesis, creating double stranded (ds) cDNA. During the end repair process any remaining overhangs are fixed into blunt ends, and 5’ phosphates and 3’ hydroxyls added. The 3’ ends of each cDNA read are then adenylated, creating an overhang. Unique sequencing adapters, which hybridise to the flow cell during sequencing, are ligated to the ends of the ds cDNA using the 3’-A overhang, and these fragments are enriched using PCR.

The indexed cDNA libraries containing the specific Illumina TruSeq adapters from the 2010-born animals were sent to GATC Biotech (Kontanz, Germany), where they were pooled (Table 3.1). Each pool was sequenced on two lanes of an Illumina HiSeq2000 with 50 bp paired-end reads. The libraries from the 2011-born animals were pooled (Table 3.1) and sequenced at the Norwegian Sequencing Centre (NSC; Oslo, Norway) over four lanes of an Illumina HiSeq2000 with 100 bp paired-end reads.

Table 3.1: Illumina TruSeq™ adapter index and pooling strategy for abomasal lymph node RNA libraries. Adapters were used from both TruSeq LT Kit Set A and Set B. The libraries from the 2010-born animals were split into two pools (n = 10), with each pool sequenced on two lanes of an Illumina HiSeq2000. All (n = 20) libraries from the 2011-born animals were pooled and sequenced over four lanes of an Illumina HiSeq2000.

Animal ID	Year	Phenotype	Day post infection	Adapter	Kit	Pool
1024566	2010	High FEC	7	AR013	A	1
1024690	2010	High FEC	7	AR006	A	1
1024704	2010	High FEC	7	AR007	A	1
1024619	2010	High FEC	7	AR006	A	2
1024644	2010	High FEC	7	AR016	A	2
1024647	2010	High FEC	14	AR012	A	1
1024689	2010	High FEC	14	AR005	A	1
1024649	2010	High FEC	14	AR014	A	2
1024626	2010	High FEC	14	AR004	A	2
1024613	2010	High FEC	14	AR002	A	2
1024551	2010	Low FEC	7	AR015	A	1
1024570	2010	Low FEC	7	AR014	A	1
1024624	2010	Low FEC	7	AR015	A	2
1024630	2010	Low FEC	7	AR007	A	2
1024572	2010	Low FEC	7	AR013	A	2
1024596	2010	Low FEC	14	AR004	A	1
1024580	2010	Low FEC	14	AR002	A	1
1024659	2010	Low FEC	14	AR016	A	1
1024558	2010	Low FEC	14	AR012	A	2
1024715	2010	Low FEC	14	AR005	A	2
1124962	2011	High FEC	7	AR015	A	3
1125034	2011	High FEC	7	AR010	B	3
1125098	2011	High FEC	7	AR008	B	3
1124951	2011	High FEC	7	AR011	B	3
1124929	2011	High FEC	7	AR001	B	3
1124986	2011	High FEC	14	AR002	A	3
1125036	2011	High FEC	14	AR016	A	3
1124984	2011	High FEC	14	AR007	A	3
1124932	2011	High FEC	14	AR003	B	3
1125031	2011	High FEC	14	AR020	B	3
1124995	2011	Low FEC	7	AR014	A	3
1124924	2011	Low FEC	7	AR005	A	3
1124956	2011	Low FEC	7	AR019	A	3
1124990	2011	Low FEC	7	AR006	A	3
1124942	2011	Low FEC	7	AR004	A	3
1124967	2011	Low FEC	14	AR012	A	3
1125044	2011	Low FEC	14	AR013	A	3
1125058	2011	Low FEC	14	AR018	A	3
1125040	2011	Low FEC	14	AR009	B	3
1124946	2011	Low FEC	14	AR021	B	3

3.2.6 Data analysis

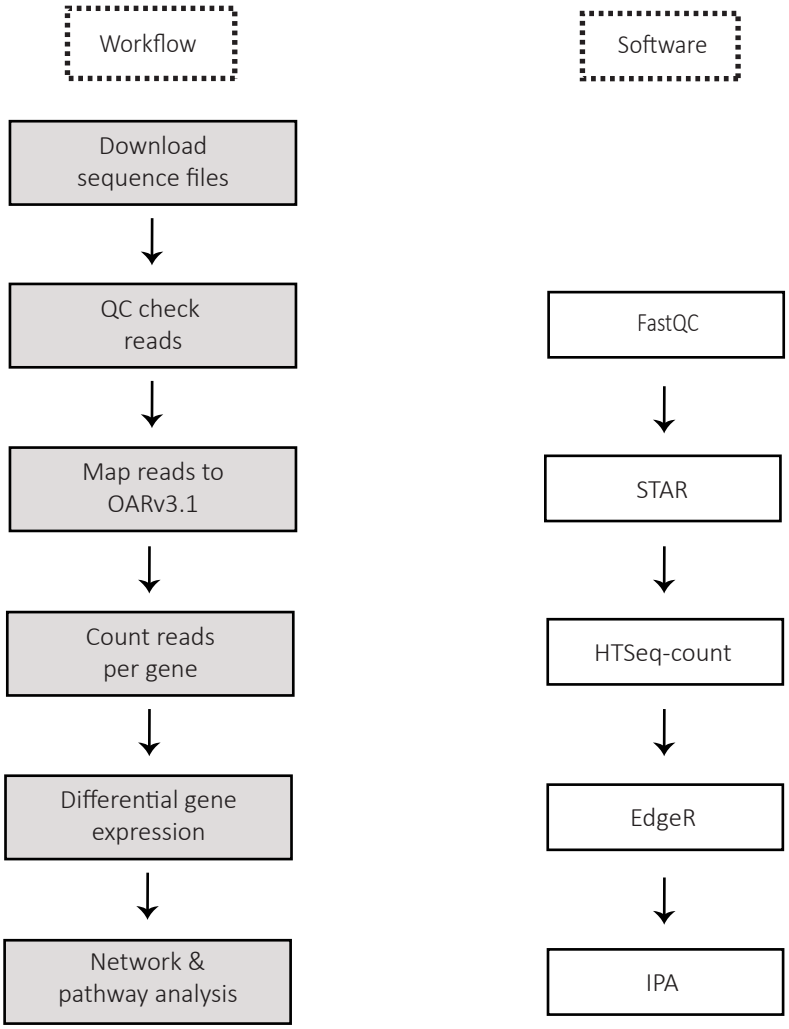


Figure 3.2: Flow diagram of the steps involved in the analysis of next generation sequencing data. The version of software used is indicated in the text.

3.2.6.1 Download and quality assessment of reads

Reads that passed filter were downloaded in .fastq format from either GATC or NSC to a Teagasc server (hcux323.teagasc.net) via file transfer protocol using either the `mget` or `wget` command. The `md5sum` command was used to check that files had downloaded correctly, after which the files were unzipped. The tools FastQC (v0.10.0) and Trim Galore (v0.3.3) (http://www.bioinformatics.babraham.ac.uk/projects/) were used to visualise the data and trim reads respectively. Trim Galore, which utilises Cutadapt (v1.2.1), was run using the default settings for paired end data, resulting in base calls with a Phred score less than 20 removed.

```
fastqc <input_file> > <output_file>
```

```
trim_galore --paired -q 20 <left_reads> <right_reads>
```

3.2.6.2 Aligning reads to the ovine genome

Trimmed reads were mapped to the ovine genome (OARv3.1) (Kijas et al., 2012) using STAR (Dobin et al., 2013), with the Ensembl (Ensembl 74) *Ovis aries* transcriptome annotation supplied. Alignments were output only if the ratio of mismatches to mapped length per read pair was less than 0.02. Only uniquely mapped reads were kept for read counts.

```
STAR --readFilesIn <trimmed_left_reads> <trimmed_right_reads> --runMode
alignReads --genomeDir <genome_dir> --runThreadN 12 00genomeLoad
NoSharedMemory --outFileNamePrefix --ouReadsUnmapped Fastx
--outFilterMismatchedNoverLmax 0.02 --outFilterMultimapNmax 1
--sjdbGTFfile <Ovis_aries_Ensembl_annotation.gtf>

samtools flagstat <accepted_hits.bam> > <flagstats.txt>
```

The resulting files were sorted by read name and converted from BAM to SAM using SAMtools (v 0.1.18; Li et al., 2009).

```
samtools sort -n <accepted_hits.bam> <accepted_hits_sorted>

samtools view -h -o <accepted_hits_sorted.sam>
<accepted_hits_sorted.bam>
```

3.2.6.3 Counting reads per gene

The mapped reads, along with the Ensembl (Ensembl 74) *Ovis aries* transcriptome annotation, were used to estimate raw counts per gene using the HTSeq (version 0.5.3p3; http://www-huber.embl.de/users/anders.HTSeq) function htseq-count, with the union overlap resolution mode. This is the recommended mode, where reads are only discarded if the read overlaps more than one feature.

```
htseq-count -m union -i gene_id -s no <tophat_accepted_hits_sorted.sam>
<Ovis_aries_Ensembl_annotation.gtf> > <htseq_count_results.txt>
```

3.2.6.4 Gene expression analysis

Comparisons were made between HighFEC and LowFEC animals at either 7 or 14 dpi or within groups over time between 7 and 14 dpi (Figure 3.3). The two years, 2010 and 2011, were analysed separately due to differences in worm burden (Figure 2.7).

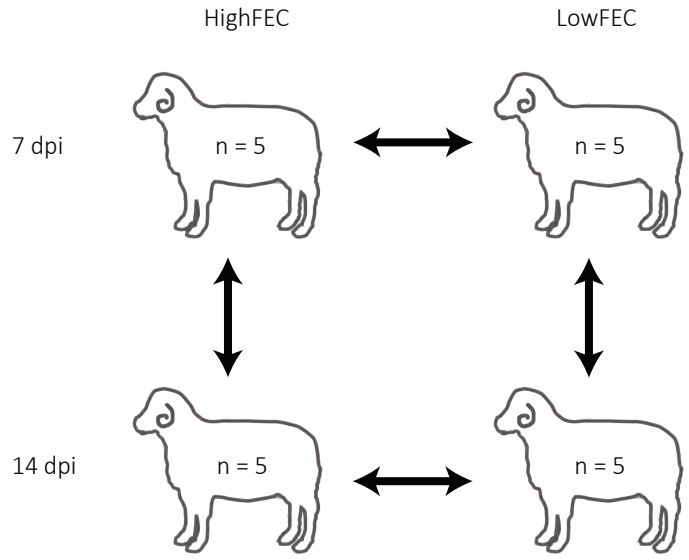


Figure 3.3: Comparisons made between groups in EdgeR analysis. Comparisons were made between phenotypes (HighFEC and LowFEC) at each time point (7 or 14 days post infection), and within phenotype (HighFEC or LowFEC) over time (7 and 14 days post infection) for both 2010- and 2011-born lambs.

The Bioconductor package edgeR (version 3.0.8) (Robinson et al., 2010) was run within R software (version 3.0.2) to analyse differential expression of read counts.

```
require(EdgeR)
D <- as.matrix(read.table("<comparison>.txt", header=TRUE, row.names = 1))
head(D)
```

Firstly, low expression tags were filtered, keeping only genes that achieved at least one count per million (CPM) in at least five samples.

```
keep <- rowSums(D)>1)>=5
D <- D[keep,]
table(keep)
```

Trimmed mean of M-values (TMM) (Robinson and Oshlack, 2010) normalisation was used to account for differences in RNA composition between samples. TMM normalisation methods are robust to the presence of different library sizes and widely different library compositions, both of which are typical of real RNA-seq data (Dillies et al., 2013).

```
f <- calcNormFactors(D, method=c("TMM"))
g <- gsub("[0-9]", "", colnames(D))
g2 <- gsub("_", "", g)
d <- DGEList(counts = D, group = g2, lib.size = colSums(D) * f)
```

Data were analysed using both common and moderated tagwise dispersions. Tagwise dispersion ranks genes more highly when counts are consistent between replicates, as opposed to those with highly variable read numbers.

```
dc <- estimateCommonDisp(d, verbose=T)
dtw <- estimateTagwiseDisp(dc)
```

MDS plots were created to show distances, in terms of leading log-fold-changes between samples; the leading log-fold-change is the average of the largest absolute log-fold-change between each pair of samples, and can be viewed as a type of unsupervised clustering.

```
plotMDS(dc/dtw)
```

Exact genewise tests for differential expression between groups were computed; to account for multiple testing genes were filtered using a Benjamini & Hochberg false discovery rate (FDR; Benjamini and Hochberg, 1995) of ≤ 0.1.

```
de.com <- exactTest(dc/dtw)
edgeR.results <- (topTags(de.com, n=row(de.com$table), adjust.method="BH",
sort.by="p.value")$table)
de.gene.nos <- summary(de<-decideTestsDGE(de.com, adjust.method="BH",
p.value=0.1))
```

Log-fold-changes, highlighting the DE genes, were visualised using a smear plot.

```
edgeR.results.sig <- subset(edgeR.results, FDR<0.1)
plotSmear(de.com,de.tags=edgeR.results.sig)
```

Heat maps of individual RNA-seq samples were created using moderated log-counts-per-million of the DE genes in each group.

```
logCPM <- cpm(dc/dtw, prior.count=2, log=TRUE)
edgeR.results.sig$GeneID=rownames(edgeR.results.sig)
sig.genes=edgeR.results.sig$GeneID
myPalette <- colorRampPalette(c("dodgerblue4","white","darkgreen"))(n=299)
groups <- d$samples$group
colour.map <- function(groups) { if(groups=="HF") "lightgrey" else "darkgrey"
}
colour.bar <- unlist(lapply(groups,colour.map))
heatmap.2(logCPM[c(sig.genes), ],scale="none",trace="none", density.info="none",
offsetRow=0.5, offsetCol=0.5, col=myPalette, cexRow=0.8, cexCol=0.8, srtCol=0,
adjCol=0.5,margins=(c(4,9)), keysize=1)
heatmap.2(logCPM[c(sig.genes), ], Rowv=T, Colv=F, scale="none", trace="none",
density.info="none", offsetRow=0.5, offsetCol=0.5, col=myPalette, cexRow=0.8,
cexCol=0.8, srtCol=0, adjCol=0.5,margins=(c(5,9)), ColSideColors=colour.bar,
keysize=1)
```

All genes differentially expressed (FDR $\leq$ 0.05) using common dispersion estimates were included in downstream pathway analysis. For “novel protein coding genes” *Bos taurus* orthologs were examined, and 1-to-1 orthologs were used to predict gene function.

3.2.7 Pathway analysis

Human Ensembl (version 74) 1-to-1 orthologs were obtained using Ensembl’s Biomart tool (www.ensembl.org/biomart/martview/). Ingenuity® Systems Pathway Analysis (IPA; Ingenuity Systems, Redwood City, CA, USA; www.ingenuity.com; v18030641) was used to identify networks of interacting genes and other functional groups from DE genes. The comparison analysis was used to identify the unique and common molecules across the two years.

3.3 Results

3.3.1 RNA isolation & library preparation

High quality RNA (Table 3.2) was isolated from ALN for all animals; the average concentration was of RNA 390 ng/μL. Bioanalyzer analysis of RNA showed an average RNA integrity number (RIN) of 9.2 (range 8.5-9.5).

A secondary peak around 1500 bp was observed in the Bioanalyzer trace of some libraries (Figure 3.4a). Possible explanations for this peak were 1) bead contamination or 2) over-cycling of PCR leading to concatemers. To avoid bead contamination the PCR product was cleaned up using a Qiagen MinElute column rather than AMPure XP beads, and a titration of PCR cycles was performed. When the PCR was run with 10 cycles no additional peaks were observed (Figure 3.4b), and subsequently all samples were prepared using 10 cycles of PCR.

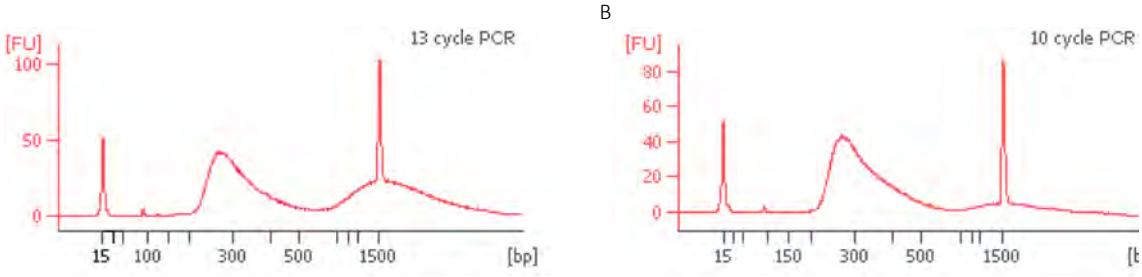


Figure 3.4: Titration of PCR cycles for TruSeq™RNA library sample preparation. Bioanalyzer trace from sample amplified using 10 (A) or 13 (B) PCR cycles.

Table 3.2: Extraction and quantification of RNA from abomasal lymph nodes. Animalas are defined by year (2010 or 2011), Phenotype (HighFEC or LowFEC) and day post infection (dpi; 7 or 14). Weight of abomasal lymph node (ALN) sample used for RNA extraction is given, along with concentration of extracted RNA, RIN value, and 28:18S ratio.

Animal ID	Year	Phenotype	dpi	Weight of ALN (g)	RNA conc. (ng/uL)	RIN	28S:18S
1024566	2010	HighFEC	7	0.1253	261	9.1	1.9
1024690	2010	HighFEC	7	0.1498	253	8.9	1.7
1024704	2010	HighFEC	7	0.0932	232	9.1	1.8
1024619	2010	HighFEC	7	0.1534	315	9.2	2.1
1024644	2010	HighFEC	7	0.1423	254	9.0	1.8
1024647	2010	HighFEC	14	0.1256	257	9.3	2.0
1024689	2010	HighFEC	14	0.0986	213	9.1	1.8
1024649	2010	HighFEC	14	0.1258	191	9.4	2.0
1024626	2010	HighFEC	14	0.0903	203	9.3	1.9
1024613	2010	HighFEC	14	0.1140	220	9.0	1.8
1024551	2010	LowFEC	7	0.1394	407	8.5	1.5
1024570	2010	LowFEC	7	0.1318	416	8.9	2.0
1024624	2010	LowFEC	7	0.1489	338	9.0	1.7
1024630	2010	LowFEC	7	0.1386	233	9.1	1.8
1024572	2010	LowFEC	7	0.1212	423	9.1	1.8
1024596	2010	LowFEC	14	0.1068	184	9.1	1.8
1024580	2010	LowFEC	14	0.1170	204	9.1	1.7
1024659	2010	LowFEC	14	0.1056	191	9.4	1.8
1024558	2010	LowFEC	14	0.1349	269	9.2	2.0
1024715	2010	LowFEC	14	0.1106	241	9.2	1.8
1124962	2011	HighFEC	7	0.1250	658	9.2	1.7
1125034	2011	HighFEC	7	0.1540	493	9.2	1.9
1125098	2011	HighFEC	7	0.1370	660	9.1	1.9
1124951	2011	HighFEC	7	0.1180	524	9.2	2.0
1124929	2011	HighFEC	7	0.1030	423	9.5	2.0
1124986	2011	HighFEC	14	0.1060	636	9.5	2.1
1125036	2011	HighFEC	14	0.1210	704	9.4	1.8
1124984	2011	HighFEC	14	0.1200	545	9.1	1.7
1124932	2011	HighFEC	14	0.1290	479	9.5	2.1
1125031	2011	HighFEC	14	0.1010	425	9.3	2.0
1124995	2011	LowFEC	7	0.1130	420	9.4	2.1
1124924	2011	LowFEC	7	0.1070	248	9.1	1.8
1124956	2011	LowFEC	7	0.1370	623	9.1	1.7
1124990	2011	LowFEC	7	0.1080	393	9.0	1.7
1124942	2011	LowFEC	7	0.1020	464	9.2	1.7
1124967	2011	LowFEC	14	0.1280	688	9.2	1.8
1125044	2011	LowFEC	14	0.1040	418	9.4	1.8
1125058	2011	LowFEC	14	0.1240	461	9.3	2.1
1125040	2011	LowFEC	14	0.1350	500	9.3	1.9
1124946	2011	LowFEC	14	0.1200	518	9.2	1.8

3.3.2 Data analysis

High throughput sequencing of the ovine lymph node RNA resulted in 790,415,623 paired-end reads (50bp) in 2010 and 645,282,528 paired-end reads (100bp) in 2011. Approximately 1% of reads were removed due to low-quality (Phred score <20; Table 3.3 & Table 3.4). There was greater variability in the number of reads per animal for the libraries sequenced at GATC Biotech (19,613,960 to 73,744,229) compared with those sequenced at the NSC (26,090,324 to 36,640,772), indicating a more even pooling of libraries at the NSC.

Using the alignment software program STAR, an average of 30,050,657 (84%) reads per sample mapped to a unique region of the ovine genome. STAR has been shown to compare favourably to other software for short read alignment (Engstrom et al., 2013), and was chosen primarily because of the speed of alignment combined with its accuracy. Of the aligned reads per sample, an average of 13,090,566 (36% of sequenced reads) did not align to a known gene, with 16,670,465 (46% of sequenced reads) aligning to a known feature (Table 3.3 & Table 3.4). This is consistent with a study in cattle, where 46% of reads aligned to a known feature (Foley, 2014).

The top 100 most expressed genes in each comparison (Figure 3.3) were examined using $\log_2$CPM values from EdgeR, with genes ranked from most highly expressed to least highly expressed. The top two most highly expressed gene in all comparisons were ENSOARG00000009143, a novel protein coding gene orthologous to the human immunoglobulin heavy constant gamma family of genes, and ENSOARG00000012585, a novel protein coding gene orthologous to immunoglobulin lambda human genes. The presence of these genes, alongside genes coding for known immunoglobulins (*IGHM*), ribosomal proteins (*RPLP0*, *RPS11* and *RPL3*), and MHC components (*B2M* and *CD74*) in the top 20 most highly expressed genes in the abomasal lymph node over both years acts as a quality control step, indicating that the RNA-Seq worked. The similarity between the 2 years is of particular importance, due to the observed differences in worm burden (Figure 2.7) between the years.

Table 3.3: Summary of all reads from 2010-born animals prior to and post aligning to the ovine genome. Reads were trimmed using TrimGalore before being aligned to the ovine genome (OARv3.1). The number of properly paired reads (both reads in pair map to the same chromosome and point towards one another) and singletons (only one read mapped) are given. Reads were further defined as those that did not map to a known gene (No feature), those that could have been assigned to more than one gene (Ambiguous), and those that aligned uniquely to a gene (Aligned to feature). Percentages are of total reads sequenced.

			FastQC	*TrimGalore*	*Flagstats*			*HTSeq-Count*			
Animal ID	**Phenotype**	**dpi**	**Reads sequenced**	**Reads post trimming**	**Properly paired**	**%**	**Singletons**	**No feature**	**Ambiguous**	**Aligned to feature**	**%**
1024566	HighFEC	7	63,639,385	62,980,542	52,375,518	82.3%	24,403	23,850,311	357,279	28,029,436	44.0%
1024619	HighFEC	7	25,486,216	25,239,263	21,287,051	83.5%	9,680	9,479,472	148,553	11,612,814	45.6%
1024690	HighFEC	7	28,411,597	28,114,067	23,489,695	82.7%	11,230	10,457,307	164,925	12,814,035	45.1%
1024644	HighFEC	7	63,495,636	62,824,805	51,857,863	81.7%	25,050	21,551,894	415,155	29,782,668	46.9%
1024704	HighFEC	7	25,463,738	25,198,770	21,082,961	82.8%	10,008	9,791,851	146,045	11,088,162	43.5%
1024647	HighFEC	14	20,639,918	20,431,564	17,152,234	83.1%	8,129	7,464,572	130,797	9,515,174	46.1%
1024649	HighFEC	14	73,744,229	72,934,510	60,886,446	82.6%	29,497	25,550,332	487,050	34,710,314	47.1%
1024626	HighFEC	14	19,613,950	19,399,981	16,138,006	82.3%	7,887	6,948,450	124,846	9,024,996	46.0%
1024689	HighFEC	14	31,727,789	31,423,401	26,148,309	82.4%	11,593	11,526,176	197,218	14,351,893	45.2%
1024613	HighFEC	14	34,779,451	34,425,516	28,972,686	83.3%	12,893	13,142,718	197,729	15,552,250	44.7%
1024624	LowFEC	7	52,592,394	52,085,087	43,242,692	82.2%	19,306	18,851,940	338,394	23,949,689	45.5%
1024630	LowFEC	7	28,216,606	27,928,412	23,172,127	82.1%	10,849	10,335,226	169,770	12,605,994	44.7%
1024551	LowFEC	7	65,802,221	65,131,958	53,766,653	81.7%	24,914	22,873,525	436,623	30,319,175	46.1%
1024570	LowFEC	7	62,994,686	62,390,171	51,933,414	82.4%	23,303	22,483,555	415,214	28,912,510	45.9%
1024572	LowFEC	7	39,150,816	38,784,021	32,392,528	82.7%	14,547	13,267,805	253,167	18,798,796	48.0%
1024596	LowFEC	14	22,687,497	22,448,917	18,463,078	81.4%	8,758	7,825,044	145,272	10,442,510	46.0%
1024580	LowFEC	14	25,074,182	24,826,911	20,815,320	83.0%	9,363	9,418,330	143,786	11,199,748	44.7%
1024558	LowFEC	14	24,188,956	23,952,802	20,111,310	83.1%	9,088	9,079,103	143,614	10,834,896	44.8%
1024715	LowFEC	14	34,847,920	34,510,171	28,889,990	82.9%	13,369	12,759,363	219,303	15,832,707	45.4%
1024659	LowFEC	14	47,858,436	47,349,049	39,285,931	82.1%	19,109	17,451,340	290,481	21,436,653	44.8%

Table 3.4: Summary of all reads from 2011-born animals prior to and post aligning to the ovine genome. Reads were trimmed using TrimGalore before being aligned to the ovine genome (OARv3.1). The number of properly paired reads (both reads in pair map to the same chromosome and point towards one another) and singletons (only one read mapped) are given. Reads were further defined as those that did not map to a known gene (No feature), those that could have been assigned to more than one gene (Ambiguous), and those that aligned uniquely to a gene (Aligned to feature). Percentages are of total reads sequenced.

			FastQC	*TrimGalore*	*Flagstats*			*HTSeq-Count*			
Animal ID	**Phenotype**	**dpi**	**Reads sequenced**	**Reads post trimming**	**Properly paired**	**%**	**Singletons**	**No feature**	**Ambiguous**	**Aligned to feature**	**%**
1125034	HighFEC	7	33,921,462	33,675,615	28,872,518	85.1%	38,402	12,524,016	229,472	16,083,455	47.4%
1124962	HighFEC	7	30,229,494	30,018,588	25,910,395	85.7%	27,516	11,692,907	186,045	13,985,755	46.3%
1125098	HighFEC	7	30,053,307	29,854,263	25,671,184	85.4%	31,086	11,559,342	195,958	13,882,886	46.2%
1124951	HighFEC	7	32,656,518	32,438,245	27,989,479	85.7%	35,376	12,715,926	217,189	15,018,935	46.0%
1124929	HighFEC	7	36,336,563	36,089,415	31,077,183	85.5%	35,057	13,496,039	234,810	17,302,061	47.6%
1124932	HighFEC	14	36,000,106	35,718,948	30,772,980	85.5%	38,121	13,073,765	255,713	17,425,112	48.4%
1124986	HighFEC	14	36,640,772	36,418,498	31,265,705	85.3%	32,265	13,782,271	226,219	17,212,741	47.0%
1125031	HighFEC	14	35,547,874	35,335,670	30,421,350	85.6%	31,877	13,364,013	229,960	16,766,501	47.2%
1125036	HighFEC	14	31,318,516	31,103,271	26,547,979	84.8%	28,741	11,043,012	206,417	15,262,758	48.7%
1124984	HighFEC	14	27,900,647	27,736,708	23,920,417	85.7%	26,429	10,791,890	173,133	12,912,069	46.3%
1124995	LowFEC	7	32,314,478	32,120,919	27,384,573	84.7%	28,601	10,726,774	252,370	16,374,234	50.7%
1124924	LowFEC	7	32,934,872	32,741,152	28,032,894	85.1%	32,723	12,291,705	203,384	15,490,078	47.0%
1124956	LowFEC	7	29,108,521	28,932,074	24,860,058	85.4%	25,240	10,882,418	203,565	13,734,396	47.2%
1124990	LowFEC	7	33,917,831	33,705,413	28,923,887	85.3%	32,107	12,996,104	210,067	15,679,646	46.2%
1124942	LowFEC	7	32,353,660	32,139,168	27,797,222	85.9%	34,953	12,463,648	203,591	15,086,224	46.6%
1124967	LowFEC	14	26,090,324	25,934,084	22,300,334	85.5%	24,086	9,597,952	169,719	12,502,132	47.9%
1125040	LowFEC	14	31,607,418	31,378,429	26,786,160	84.7%	34,149	11,641,207	203,370	14,918,070	47.2%
1125044	LowFEC	14	27,858,382	27,710,553	23,981,990	86.1%	21,624	10,615,277	178,296	13,136,645	47.2%
1125058	LowFEC	14	36,316,002	36,091,390	30,761,849	84.7%	32,208	12,712,962	242,754	17,760,575	48.9%
1124946	LowFEC	14	32,175,781	31,960,566	27,284,309	84.8%	33,814	11,543,086	241,375	15,469,897	48.1%

3.3.3 Differential gene expression

Data were analysed using both common and moderated tagwise dispersions in EdgeR. Genes detected as DE using tagwise dispersion and a FDR of <0.1 are reported. The list of genes differentially expressed (FDR <0.05) using common dispersion were used for pathway analysis, as this method is more robust to false positives.

3.3.3.1 HighFEC vs. LowFEC animals

Differences in the transcriptional profiles of HighFEC and LowFEC animals were examined at both 7 and 14 days post infection (dpi). As animals born in 2010 and 2011 had different worm burdens at slaughter (Figure 2.7) and separated on the MDS plot by year (Figure 3.19), data from the two years could not be combined. In fact, when the data from both years were combined no DE genes were detected (Table 3.5), most likely due to the substantial gene expression variation between animals from 2010 and 2011.

When the data were analysed within year a total of 41 genes were identified as DE (tagwise dispersion; FDR < 0.1) between HighFEC and LowFEC animals (Table 3.5). The highest number of genes was at 7 dpi in 2011, while the lowest number of DE genes was at 7 dpi in 2010.

Table 3.5: Number of differentially expressed genes between HighFEC and LowFEC animals using common and tagwise dispersion estimates in EdgeR. Differential expression was examined at 7 and 14 days post infection (dpi). The false discovery rate (FDR) was set to 0.1 or 0.05. Genes with a FDR < 0.05 and a log fold change (logFC) of >±2 are reported in the third column (logFC >±2).

		Common dispersion			*Tagwise dispersion*		
		FDR <0.1	**FDR <0.05**	**logFC >±2**	**FDR <0.1**	**FDR <0.05**	**logFC >±2**
2010	7 dpi	217	194	4	7	5	2
	14 dpi	179	144	8	6	4	2
2011	7 dpi	245	196	2	16	10	1
	14 dpi	245	190	5	12	8	2
Both	7 dpi	0	0	0	0	0	0
	14 dpi	0	0	0	0	0	0

Multi-dimensional scaling (MDS) plots were produced, in which distances correspond to leading log-fold-changes between each pair of RNA samples (Figure 3.5). This is the average of the largest absolute log-fold-changes between each pair of samples, and can be viewed as a type of unsupervised clustering showing the relative similarities between samples. The low number of DE genes is reflected in the MDS plots, which fail to decisively separate animals by phenotype (Figure 3.5) at any point, although in 2011 only one animal (LF_3) did not group by phenotype (Figure 3.5c).

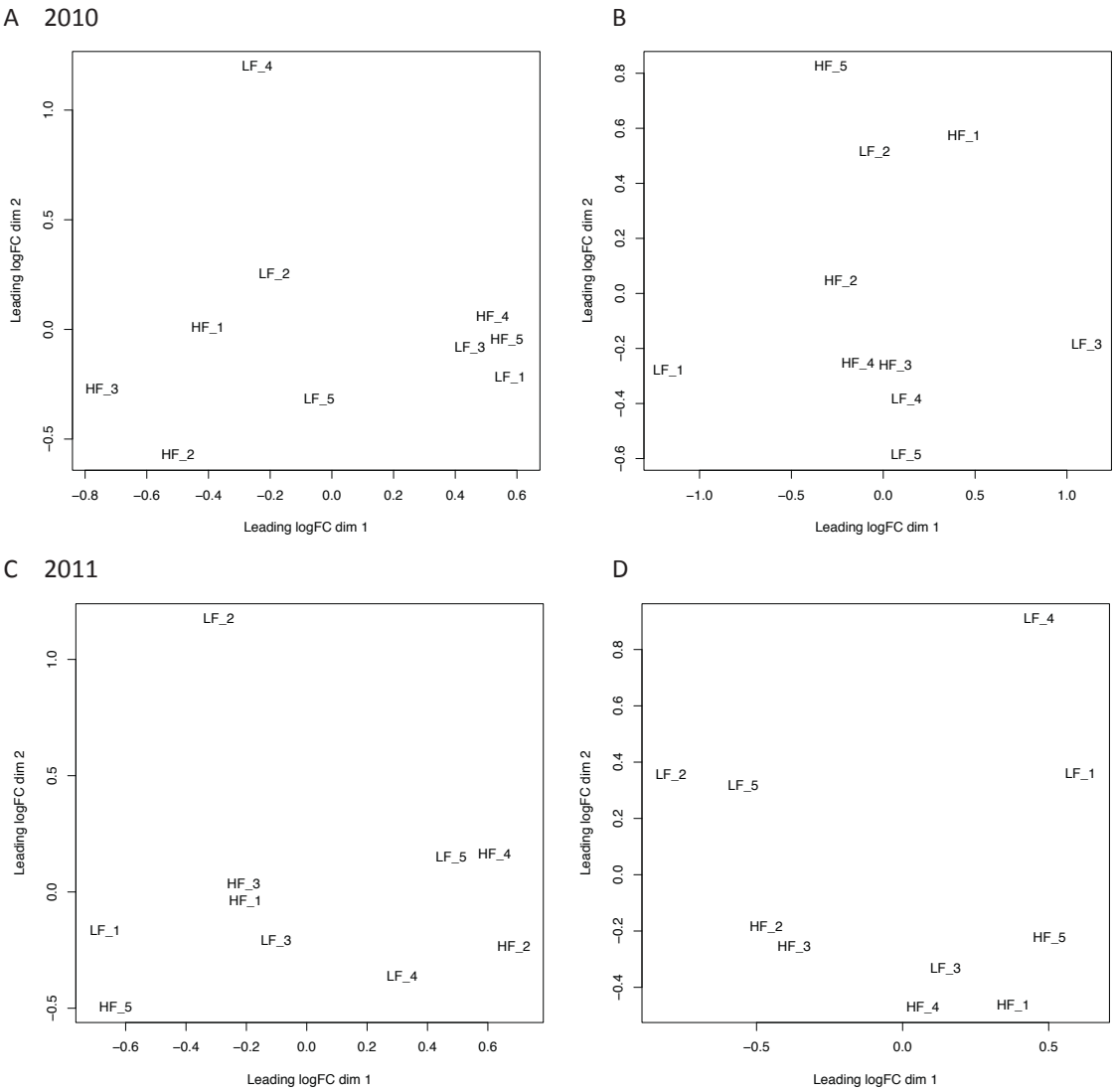


Figure 3.5: Multi-dimensional scaling (MDS) plots comparing RNA-Seq results from HighFEC and LowFEC animals. Animals were slaughtered at 7 (A & C) and 14 (B & D) days post infection in 2010 and 2011 respectively. Distances correspond to leading log-fold-changes between each pair of samples.

Of the 41 genes DE (tagwise dispersion; FDR <0.1) between HighFEC and LowFEC animals, 15 (37%) were novel protein coding genes, and 11 (27%) had been identified in previous studies (Table 3.6 & Table 3.7). These numbers are in line with those found in the common dispersion analysis (FDR <0.05); of the 724 total DE genes, 265 (37%) were novel protein coding genes and 203 (28%) had been identified in previous studies.

Table 3.6: Genes differentially expressed between HighFEC and LowFEC animals at 7 days post infection. Differentially expressed genes were determined in EdgeR using tagwise dispersion estimates, and a false discovery rate (FDR) of 0.1. Log fold change (logFC), log counts-per-million (logCPM) and P-value are given. Studies in which the gene has perviously been implicated in resistance or susceptibility to gastrointestinal nematodes are also reported.

Year	Up in	Ensembl gene ID	logFC	logCPM	P-Value	FDR	Gene	Description	Studies
2010	HighFEC	ENSOARG00000013782	-2.2006	0.9257	0.0000	0.0002	*ALB*	Albumin	Keane et al., 2006; Knight et al., 2010; Nagaraj et al., 2012; Pemberton et al., 2012
		ENSOARG00000020224	-1.1814	1.7317	0.0000	0.0061	*COL9A2*	Collagen, type IX, alpha 2	
		ENSOARG00000011275	-2.8653	1.9095	0.0000	0.0061	-		
		ENSOARG00000003000	-1.3552	2.4414	0.0000	0.0078	-		
		ENSOARG00000005312	-0.7054	3.6815	0.0000	0.0078	*ZNF461*	Zinc finger protein 461	
		ENSOARG00000001778	-1.2806	2.3723	0.0000	0.0207	-		
		ENSOARG00000005490	-1.2052	1.0490	0.0000	0.0270	*SLC30A2*	Solute carrier family 30, member 2	
2011	HighFEC	ENSOARG00000024912	-0.6073	4.9465	0.0000	0.0255	-	Novel miRNA	
		ENSOARG00000013341	-1.3033	1.1194	0.0000	0.0255	-		
		ENSOARG00000006960	-0.4592	7.5102	0.0000	0.0363	*FCRL1*	Fc receptor-like 1	Gossner et al., 2013
		ENSOARG00000019308	-0.6187	6.4914	0.0000	0.0372	*BTLA*	B and T lymphocyte associated	
		ENSOARG00000002342	-1.0555	1.1125	0.0000	0.0550	*SH3RF2*	SH3 domain containing ring finger 2	
		ENSOARG00000002175	-0.5099	4.4734	0.0001	0.0577	*DNASE1*	Deoxyribonuclease I	
	LowFEC	ENSOARG00000000857	0.7262	11.1760	0.0000	0.0255	-		
		ENSOARG00000009194	1.7458	1.3182	0.0000	0.0255	*ATP10A*	ATPase, class V, type 10A	Ahmed, 2013
		ENSOARG00000001393	3.3696	5.1678	0.0000	0.0255	*GIMAP8*	GTPase, IMAP Family Member 8	Keane et al., 2006; Ahmed, 2013
		ENSOARG00000007987	1.2141	3.0913	0.0000	0.0326	*GZMK*	Granzyme K (granzyme 3; tryptase II)	
		ENSOARG00000016543	0.9063	5.6238	0.0000	0.0363	*CXCL9*	Chemokine (C-X-C motif) ligand 9	Ahmed, 2013
		ENSOARG00000009143	0.6979	12.4807	0.0000	0.0419	-		
		ENSOARG00000004253	1.1847	2.8893	0.0000	0.0550	-	Putative MIP1-beta protein (CCL4)	
		ENSOARG00000020386	1.1194	3.8968	0.0001	0.0577	*APOD*	Apolipoprotein D	Gossner et al., 2013
		ENSOARG00000018254	1.1247	2.5238	0.0001	0.0577	*FRAS1*	Fraser syndrome 1	
		ENSOARG00000002035	0.3887	8.4324	0.0001	0.0777	*WARS*	Tryptophanyl-tRNA synthetase	Andronicos et al., 2010; Nagaraj et al., 2012; Ahmed, 2013

Table 3.7: Genes differentially expressed between HighFEC and LowFEC animals at 14 days post infection. Differentially expressed genes were determined in EdgeR using tagwise dispersion estimates, and a false discovery rate (FDR) of 0.1. Log fold change (logFC), log counts-per-million (logCPM) and P-value are given. Studies in which the gene has perviously been implicated in resistance or susceptibility to gastrointestinal nematodes are also reported.

Year	Up in	Ensembl gene ID	logFC	logCPM	P-Value	FDR	Gene	Description	Studies
2010	HighFEC	ENSOARG00000020373	-1.9749	0.9808	0.0000	0.0020	*MFI2*	Melanoma associated antigen p97	
		ENSOARG00000011529	-1.5539	2.1248	0.0000	0.0465	*ZFR2*	Zinc finger RNA binding protein 2	Ahmed, 2013
	LowFEC	ENSOARG00000005549	2.2422	5.9907	0.0000	0.0020	-		
		ENSOARG00000001701	0.8207	8.4599	0.0000	0.0020	-	MHC class I antigen	
		ENSOARG00000004858	1.8298	0.6533	0.0000	0.0427	-		
		ENSOARG00000000928	2.3652	1.8282	0.0000	0.0427	*ECT2L*	Epithelial cell transforming sequence 2 oncogene-like	Ahmed, 2013
2011	HighFEC	ENSOARG00000001898	-0.8183	3.3457	0.0000	0.0149	*FOLR4*	Folate receptor 4, delta (putative)	
		ENSOARG00000021044	-0.6843	4.6090	0.0000	0.0149	*CLEC2B*	C-type lectin domain family 2, member B	
		ENSOARG00000013498	-1.6378	0.8169	0.0000	0.0151	*DNAH5*	Dynein, axonemal, heavy chain 5	Ahmed, 2013
		ENSOARG00000010719	-0.7466	4.2008	0.0000	0.0166	*LPL*	Lipoprotein lipase (LPL)	Ahmed, 2013
		ENSOARG00000018806	-1.4248	0.7170	0.0001	0.0869	*WBP2NL*	WBP2 N-terminal like	
		ENSOARG00000012882	-0.6801	3.4899	0.0001	0.0869	*CACNB2*	Calcium channel, voltage-dependent, beta 2 subunit	
		ENSOARG00000001718	-0.8650	5.8732	0.0001	0.0954	*MPPE1*	Metallophosphoesterase 1	Ahmed, 2013
	LowFEC	ENSOARG00000000895	2.2066	3.6449	0.0000	0.0013	-		
		ENSOARG00000007871	1.1754	2.5970	0.0000	0.0149	*SCRN2*	Secernin 2	Ahmed, 2013
		ENSOARG00000002371	2.7435	1.1785	0.0000	0.0149	-		
		ENSOARG00000018173	0.9573	8.3535	0.0000	0.0430	-		
		ENSOARG00000001286	1.2979	1.3900	0.0000	0.0755	-		

Hierarchical clustering was performed on DE genes (tagwise dispersion; FDR <0.1) using moderated $\log_2$ counts-per-million (logCPM) to observe the similarities between samples (Figure 3.6). It is expected that normalisation removes variation that is not due to biological differences (Rapaport et al., 2013). Dendrograms illustrate the similarities between both samples and genes. In 2011, animals clustered on the basis of their phenotype. In 2010, the majority of animals clustered according to their phenotype, however at each dpi one animal did not group within phenotype in 2010; HF_4 at day 7 and LF_2 at day 14 post infection.

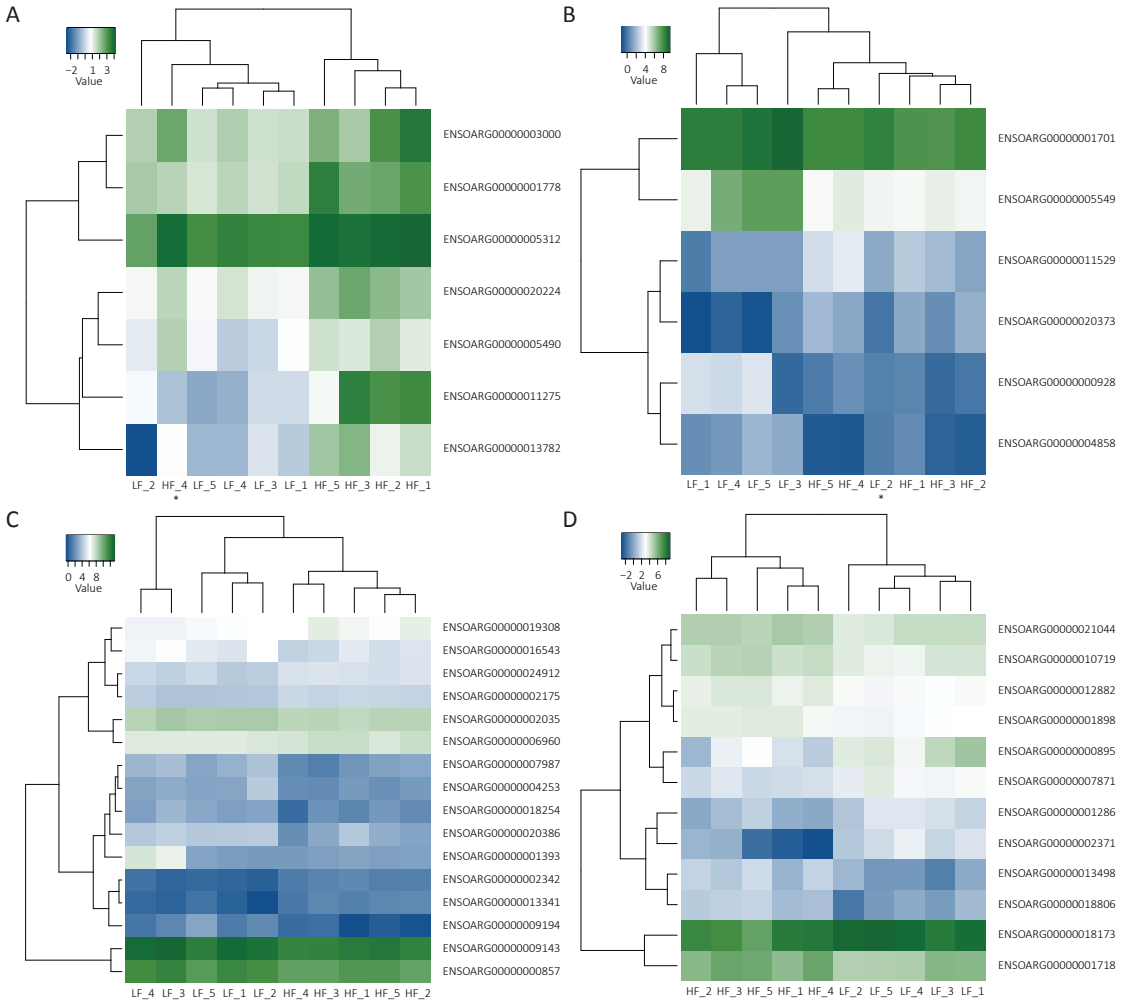


Figure 3.6: Heat map of hierarchical clustering of genes differentially expressed between HighFEC and LowFEC animals. Differentially expressed genes were determined in EdgeR using tagwise dispersion estimates, and a false discovery rate (FDR) of 0.1. Animals were slaughtered at 7 (A & C) and 14 (B & D) days post infection in 2010 and 2011 respectively. Animals that do not group within phenotype are indicated with an asterisk (*).

Pathway analysis

Human 1-to-1 orthologs were found for approximately 60% of the DE genes (common dispersion; FDR < 0.05) between HighFEC and LowFEC animals (Table 3.8).

Table 3.8: Number of genes differentially expressed between HighFEC and LowFEC animals for which 1-to-1 human orthologs could be found. Differentially expressed genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05. Orthologs were obtained from Ensembl release 74.

Comparison		No. genes	1-to-1 ortholog		No 1-to-1 ortholog	
2010	7 dpi	194	118	61%	76	39%
	14 dpi	144	84	58%	60	42%
2011	7 dpi	196	129	66%	67	34%
	14 dpi	190	107	56%	83	44%

IPA analysis was carried out using the 1-to-1 orthologs. For each comparison the top networks (Table 3.9), diseases, and biological functions (Table 3.10) were determined. Functions associated with the top networks in each comparison included cell-to-cell signalling and interaction, molecular transport, cellular movement and the hypersensitivity response (Table 3.9). Diseases and disorders associated with the DE genes included the hypersensitivity response, organismal injury and abnormalities, and metabolic disease (Table 3.10). Significant physiological system development and function categories include immune cell trafficking, tissue development, connective tissue development and function and haematological system development and function (Table 3.10).

Table 3.9: Top IPA networks at 7 or 14 days post infection (dpi) when comparing HighFEC and LowFEC animals. Score represents the number of differentially expressed genes in the network.

Year	Time point	Associated network functions	Score
2010	7 dpi	Cell-To-Cell Signalling and Interaction, Cellular Movement, Immune Cell Trafficking	30
		Cell Signalling, Molecular Transport, Vitamin and Mineral Metabolism	30
		Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry	26
		Cell-To-Cell Signalling and Interaction, Cellular Function and Maintenance, Inflammatory Response	21
		Cell Cycle, Hair and Skin Development and Function, Cancer	21
	14 dpi	Cellular Assembly and Organisation, Lipid Metabolism, Small Molecule Biochemistry	38
		Gene Expression, Cell Death and Survival, Embryonic Development	30
		Molecular Transport, Small Molecule Biochemistry, Amino Acid Metabolism	28
		Cell-To-Cell Signalling and Interaction, Nervous System Development and Function, Gene Expression	16
		Cardiovascular System Development and Function, Organ Morphology, Cellular Movement	15
2011	7 dpi	Lipid Metabolism, Small Molecule Biochemistry, Metabolic Disease	39
		Cellular Movement, Haematological System Development and Function, Hypersensitivity Response	37
		Cellular Movement, Haematological System Development and Function, Hypersensitivity Response	28
		Cell-To-Cell Signalling and Interaction, Haematological System Development and Function, Tissue Development	25
		Lipid Metabolism, Nucleic Acid Metabolism, Small Molecule Biochemistry	24
	14 dpi	Cellular Function and Maintenance, Hypersensitivity Response, Haematological Disease	39
		Connective Tissue Disorders, Hereditary Disorder, Immunological Disease	30
		Cell-To-Cell Signalling and Interaction, Cellular Movement, Haematological System Development and Function	26
		Developmental Disorder, Neurological Disease, Skeletal and Muscular Disorders	24
		Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry	24

Table 3.10: IPA analysis of differentially expressed (DE) genes between HighFEC and LowFEC animals. Listed are the top five diseases, biological functions and canonical pathways that the DE genes affect. DE genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05.

Year	Time point	Diseases and disorders	p-value	Molecules
2010	7 dpi	Hypersensitivity Response	1.29E-06	9
		Inflammatory Response	1.29E-06	20
		Haematological Disease	3.18E-05	13
		Metabolic Disease	3.18E-05	19
		Organismal Injury and Abnormalities	3.74E-05	31
	14 dpi	Cancer	2.98E-07	55
		Organismal Injury and Abnormalities	7.53E-07	28
		Reproductive System Disease	7.53E-07	15
		Neurological Disease	2.62E-06	19
		Gastrointestinal Disease	3.22E-06	40
2011	7 dpi	Dermatological Diseases and Conditions	8.77E-10	48
		Hypersensitivity Response	2.93E-09	12
		Endocrine System Disorders	1.56E-08	24
		Metabolic Disease	1.56E-08	35
		Cancer	4.82E-08	86
	14 dpi	Hypersensitivity Response	2.31E-05	10
		Connective Tissue Disorders	2.82E-05	4
		Hereditary Disorder	2.82E-05	25
		Immunological Disease	2.82E-05	13
		Inflammatory Disease	2.82E-05	18

Year	Comparison	Physiological system development and function	p-value	Molecules
2010	7 dpi	Haematological System Development and Function	1.77E-07	19
		Immune Cell Trafficking	1.77E-07	16
		Tissue Development	1.29E-06	18
		Organismal Survival	3.49E-06	13
		Connective Tissue Development and Function	6.08E-06	6
	14 dpi	Nervous System Development and Function	2.75E-05	15
		Organismal Development	1.84E-04	22
		Organ Development	2.11E-04	10
		Organismal Survival	2.25E-04	27
		Connective Tissue Development and Function	2.94E-04	9
2011	7 dpi	Haematological System Development and Function	6.29E-10	32
		Tissue Development	6.36E-10	27
		Immune Cell Trafficking	2.93E-09	22
		Cell-mediated Immune Response	2.67E-07	8
		Haematopoiesis	2.67E-07	3
	14 dpi	Haematological System Development and Function	2.09E-06	25
		Tissue Morphology	2.09E-06	29
		Haematopoiesis	7.84E-05	6
		Tissue Development	8.42E-05	21
		Embryonic Development	1.13E-04	12

Year	Time point	Canonical pathways	p-value	Ratio
2010	7 dpi	Pathogenesis of Multiple Sclerosis	1.2E-07	4/10 (0.4)
		Hepatic Fibrosis / Hepatic Stellate Cell Activation	1.17E-03	5/155 (0.032)
		Granulocyte Adhesion and Diapedesis	3.12E-03	5/182 (0.027)
		Agranulocyte Adhesion and Diapedesis	4.05E-03	5/192 (0.026)
		LXR/RXR Activation	4.97E-03	4/139 (0.029)
	14 dpi	GABA Receptor Signalling	1.37E-02	2/56 (0.036)
		AMPK Signalling	1.47E-02	3/181 (0.017)
		Protein Citrullination	1.88E-02	1/6 (0.167)
		Agrin Interactions at Neuromuscular Junction	2.75E-02	2/70 (0.029)
		PPARa/RXRa Activation	2.98E-02	3/200 (0.015)
2011	7 dpi	Pathogenesis of Multiple Sclerosis	1.31E-09	5/10 (0.5)
		LXR/RXR Activation	1.39E-05	7/139 (0.05)
		Interferon Signalling	6.73E-05	4/36 (0.111)
		Granulocyte Adhesion and Diapedesis	1.41E-04	7/182 (0.038)
		Agranulocyte Adhesion and Diapedesis	2.06E-04	7/192 (0.036)
	14 dpi	Histamine Biosynthesis	5.34E-03	1/3 (0.333)
		Triacylglycerol Degradation	8.42E-03	2/33 (0.061)
		Interferon Signalling	1.41E-02	2/36 (0.056)
		tRNA Splicing	1.5E-02	2/46 (0.043)
		Uracil Degradation II (Reductive)	2.12E-02	1/11 (0.091)

A comparison analysis was undertaken using IPA to identify the unique and common molecules across both years. Heat maps were generated of the 20 most significant diseases and biological functions (Figure 3.7), canonical pathways (Figure 3.8) and upstream regulators (Figure 3.9).

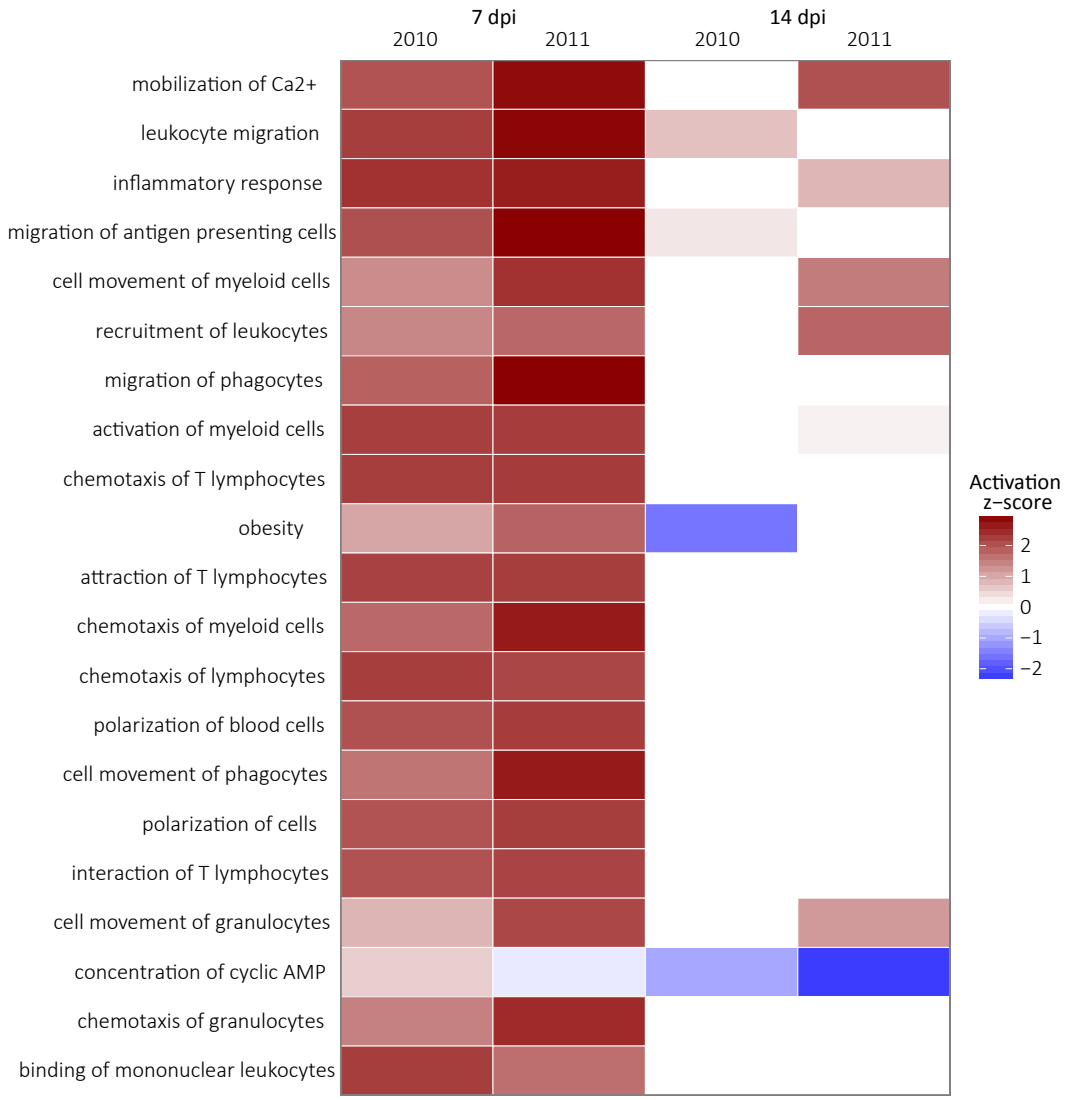


Figure 3.7: Heat map of the top 20 IPA-derived diseases and biological functions affected by genes differentially expressed between HighFEC and LowFEC animals. Heat map is sorted by activation z-score. Red: predicted activation in LowFEC animals; Blue: predicted activation in HighFEC animals.

Figure 3.8: Heat map of the top 20 IPA-derived canonical pathways affected by genes differentially expressed between HighFEC and LowFEC animals. Heat map is sorted by sorted by -log(p-value). Expression of genes in each pathway may be increased in either HighFEC or LowFEC animals.

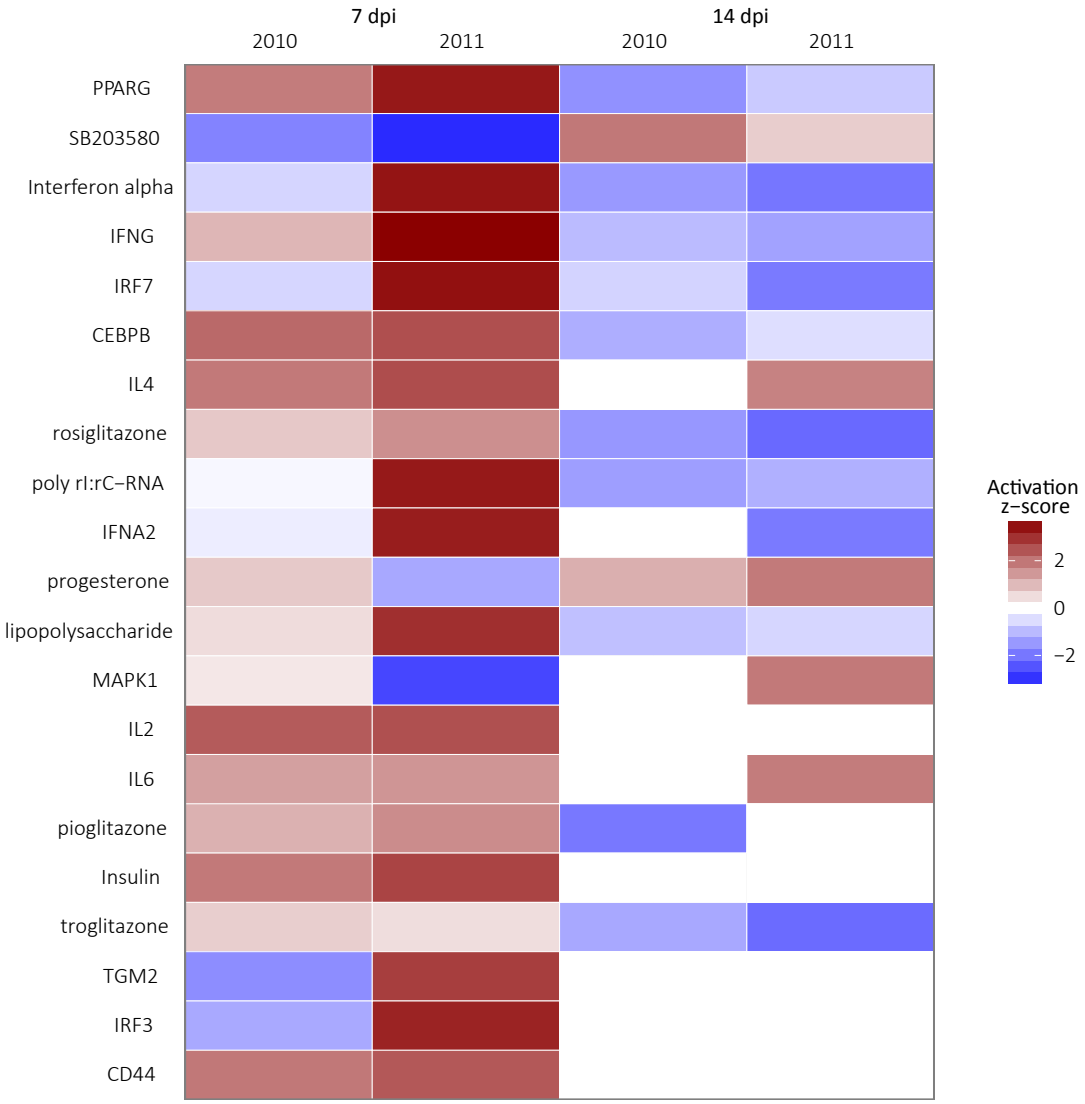
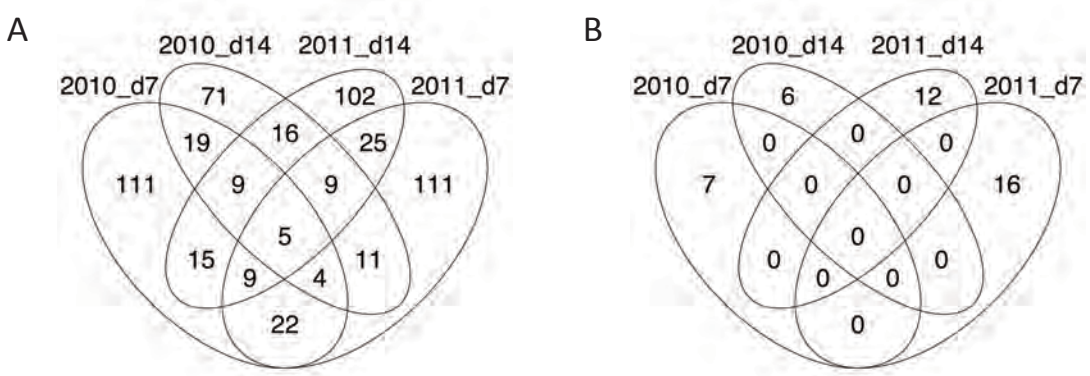


Figure 3.9: Heat map of the top 20 IPA-derived upstream regulators affected by genes differentially expressed between HighFEC and LowFEC animals. Heat map is sorted by sorted by activation z-score. Red: expression of downstream genes increased in LowFEC animals; Blue: expression of downstream genes increased in HighFEC animals.

Genes in common between 2010 and 2011 analyses

Using common dispersion estimates, 5 genes (Table 3.5) were differentially expressed (FDR <0.05) between HighFEC and LowFEC animals at both time points over both years (Figure 3.10).

Figure 3.10: Venn diagram comparing differentially expressed genes in HighFEC vs. LowFEC analyses in 2010 and 2011. Venn diagram of genes in common between common (A; FDR <0.05) and tagwise (B; FDR<0.1) dispersion analyses comparing gene expression in the abomasal lymph node of HighFEC and LowFEC animals at 7 or 14 days post infection in 2010 and 2011 born animals.



Both adiponectin (*ADIPOQ*) and a novel protein coding gene, which was a 1-to-1 ortholog of bovine carboxylesterase 1 (*CES1*), were upregulated in LowFEC animals at 7 dpi, and HighFEC animals at 14 dpi. Interferon stimulated gene 17 (*ISG17*) and g-coupled protein receptor associated sorting protein 1 (*GPRASP1*) did not follow a particular pattern - both were upregulated in HighFEC animals at all time points apart from 7 dpi and 14 dpi in 2011 respectively. A novel gene (ENSOARG00000019179) orthologous to human chemokine motifs *CCL23* and *CCL15* was upregulated in all LowFEC animals compared to their HighFEC counterparts, albeit with an average logFC of 0.68.

Table 3.11: Genes differentially expressed between HighFEC and LowFEC animals at 7 and 14 days post infection (dpi) in both 2010 and 2011. Log fold changes (logFC) in genes significantly (common dispersion; FDR <0.05) differentially expressed between HighFEC and LowFEC animals at 7 and 14 days post infection in both 2010 and 2011. Positive logFC (**bold**) indicates gene is upregulated in LowFEC animals, whereas negative logFC (regular) indicates gene is upregulated in HighFEC group.

Gene ID	Gene name	2010		2011	
		7 dpi	14 dpi	7 dpi	14 dpi
ENSOARG00000018232	-	**1.07**	-0.82	**1.67**	-1.17
ENSOARG00000020509	*ADIPOQ*	**0.79**	-0.57	**1.43**	-0.93
ENSOARG00000007233	*ISG17*	-0.68	-0.93	**1.37**	-0.83
ENSOARG00000002182	*GPRASP1*	-0.64	-0.85	-0.77	**1.09**
ENSOARG00000004774	-	**0.58**	**0.59**	**0.70**	**0.86**

Of the DE genes (common dispersion; FDR < 0.05) for which human 1-to-1 orthologs were found, 3 remained in common between all comparisons (*ADIPOQ*, *ISG17* and *GPRASP1*).

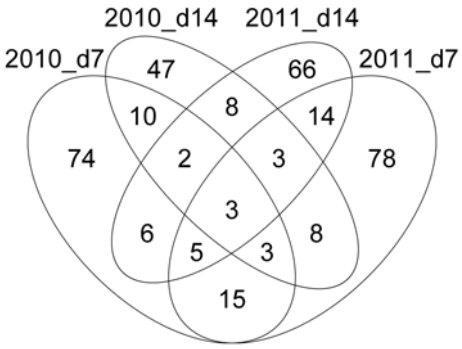


Figure 3.11: Venn diagram comparing differentially expressed genes in HighFEC vs. LowFEC analyses in 2010 and 2011 after conversion to 1-to-1 Human Ensembl orthologs. Differentially expressed genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05.

3.3.3.2 Day 7 vs. day 14 post infection

The immune response to gastrointestinal nematodes over time can be examined by looking at the transcriptional profiles of animals slaughtered at 7 days post infection (dpi) compared to those slaughtered at 14 dpi. As expected, combining the data from 2010 and 2011 resulted in a very low number of DE genes (Table 3.12). When the data was analysed within year a total of 68 genes were identified as DE (tagwise dispersion; FDR <0.1) between 7 and 14 days post infection (Table 3.12).

Table 3.12: Number of differentially expressed genes between 7 and 14 days post infection using common and tagwise dispersion estimates in EdgeR. Differential expression was examined in HighFEC (HF) and LowFEC (LF) animals. The false discovery rate (FDR) was set to 0.1 or 0.05. Genes with a FDR < 0.05 and a log fold change (logFC) of >±2 are reported in the third column (logFC >±2).

		Common dispersion			*Tagwise dispersion*		
		FDR <0.1	**FDR <0.05**	**logFC >±2**	**FDR <0.1**	**FDR <0.05**	**logFC >±2**
2010	HF	272	224	2	7	3	2
	LF	213	163	6	14	12	2
2011	HF	326	247	5	26	10	2
	LF	285	204	12	21	13	6
Both	HF	1	0	0	1	0	0
	LF	2	1	0	2	1	0

Multi-dimensional scaling (MDS) plots were produced, in which distances correspond to leading log-fold-changes between each pair of RNA samples (Figure 3.12). This is the average of the largest absolute log-fold-changes between each pair of samples, and can be viewed as a type of unsupervised clustering showing the relative similarities between samples. The low number of DE genes is reflected in the MDS plots, which fail to decisively separate animals by day post infection (Figure 3.12), however in the 2010 HighFEC animals there is only one animal (SEVEN_1) that does not cluster within day post infection.

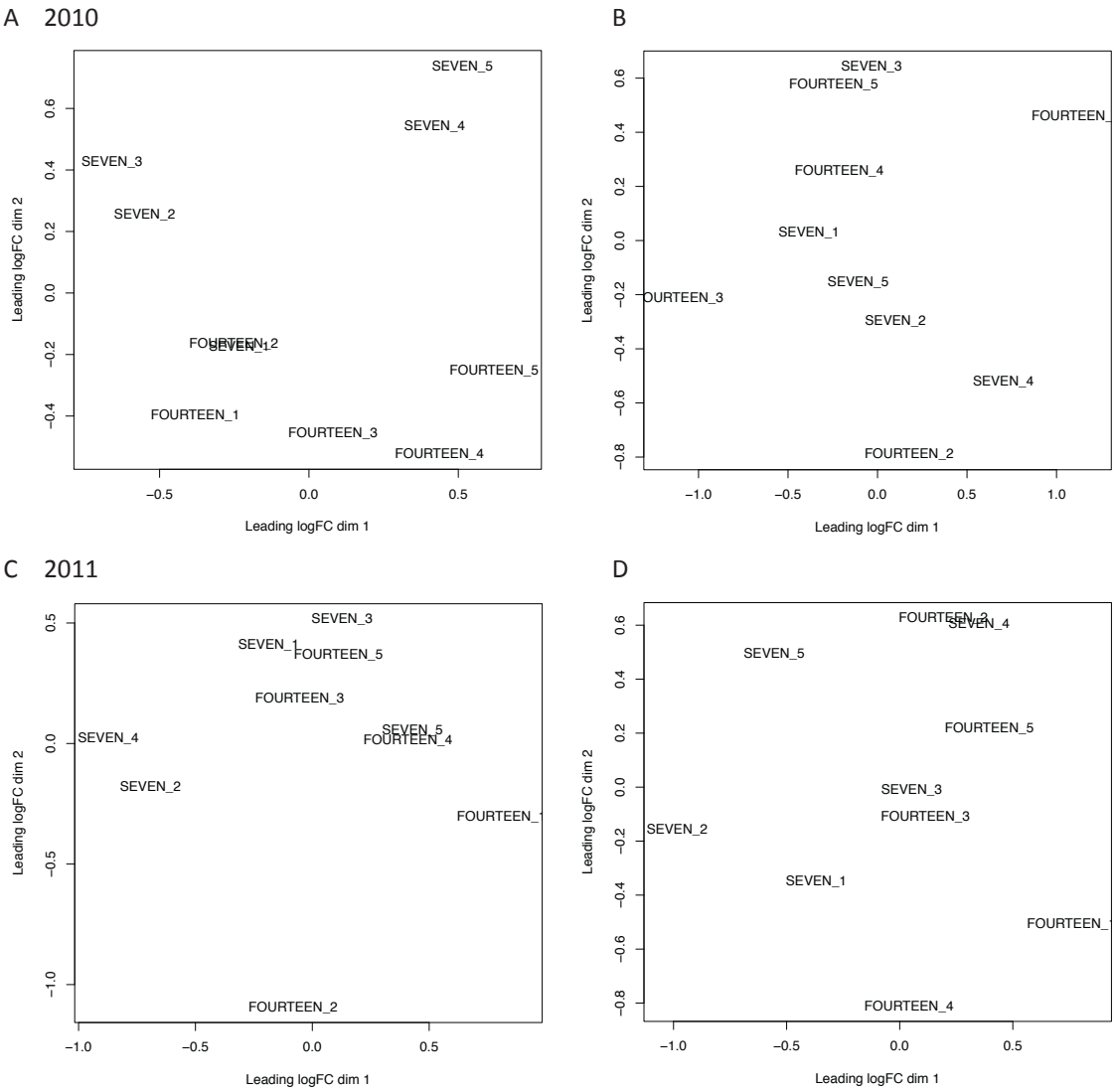


Figure 3.12: Multi-dimensional scaling (MDS) plots comparing RNA-Seq results from animals slaughtered at 7 and 14 days post infection. Animals had a HighFEC (A & C) or LowFEC (B & D) phenotype, and were born in 2010 or 2011 respectively. Distances correspond to leading log-fold-changes between each pair of samples.

Of the 71 genes DE (tagwise dispersion; FDR <0.1) between 7 and 14 days post infection, 25 (35%) were novel protein coding genes, and 24 (34%) had been identified in previous studies (Table 3.13 & Table 3.14). This is similar to the common dispersion analysis; of the 838 DE genes found, 278 (33%) were predicted genes of unknown function and 239 (29%) had been identified in previous studies.

Table 3.13: Genes differentially expressed between 7 and 14 days post infection (dpi) in HighFEC animals. Differentially expressed genes were determined in EdgeR using tagwise dispersion estimates, and a false discovery rate (FDR) of 0.1. Log fold change (logFC), log counts-per-million (logCPM) and P-value are given. Studies in which the gene has perviously been implicated in resistance or susceptibility to gastrointestinal nematodes are also reported.

Year	Up at	Ensembl gene ID	logFC	logCPM	P-Value	FDR	Gene name	Description	Studies
2010	7 dpi	ENSOARG00000013782	7.3780	0.7569	0.0000	0.0000	*ALB*	Albumin (ALB)	Keane et al., 2006; Knight et al., 2010; Nagaraj et al., 2012; Pemberton et al., 2012
		ENSOARG00000010231	2.6685	4.0990	0.0000	0.0369	*MX2*	Myxovirus (influenza virus) resistance 2 (mouse)	Nagaraj et al., 2012; Ahmed, 2013
		ENSOARG00000004858	2.0583	0.7903	0.0000	0.0761	-		
	14 dpi	ENSOARG00000010111	-0.7418	3.6434	0.0000	0.0299	-		
		ENSOARG00000011072	-0.5857	5.1376	0.0000	0.0675	*LYVE1*	Lymphatic vessel endothelial hyaluronan receptor 1	
		ENSOARG00000019641	-1.1580	1.0998	0.0000	0.0774	*LSAMP*	Limbic system-associated membrane protein	
		ENSOARG00000019517	-0.9168	5.5266	0.0000	0.0774	*CHI3L2*	Chitinase 3-like 2	Gossner et al., 2013
2011	7 dpi	ENSOARG00000000070	1.0007	5.3637	0.0000	0.0035	-		
		ENSOARG00000005011	0.6499	5.2459	0.0000	0.0246	*TTC21B*	Tetratricopeptide repeat domain 21B	
		ENSOARG00000001140	0.9718	8.6963	0.0000	0.0250	-		
		ENSOARG00000018398	1.4112	3.8246	0.0000	0.0256	*MMRN1*	Multimerin 1	Ahmed, 2013
		ENSOARG00000010473	0.7628	5.3399	0.0000	0.0427	*EPB41L3*	Erythrocyte membrane protein band 4.1-like 3	Ahmed, 2013
		ENSOARG00000002851	0.7699	5.4066	0.0001	0.0775	*CYP4B1*	Cytochrome P450, family 4, subfamily B, polypeptide 1	Ahmed, 2013
		ENSOARG00000012377	2.8183	2.4880	0.0001	0.0775	*GSDMA*	Gasdermin A	Ahmed, 2013
		ENSOARG00000012211	0.7576	3.3896	0.0001	0.0775	*RPS25*	Ribosomal protein S25 (RPS25)	
		ENSOARG00000007534	0.7424	3.9636	0.0001	0.0775	*LACC1*	Laccase (multicopper oxidoreductase) domain containing 1	
		ENSOARG00000004611	2.1113	0.5350	0.0001	0.0826	*MCP-3*	Mast cell proteinase-3 (MCP-3)	
		ENSOARG00000015996	0.7429	7.4676	0.0002	0.0912	*STAB2*	Stabilin 2	
	14 dpi	ENSOARG00000001393	-4.4887	6.2205	0.0000	0.0000	*GIMAP8*	GTPase, IMAP Family Member 8	Keane et al., 2006; Ahmed, 2013
		ENSOARG00000013846	-0.4379	6.2367	0.0000	0.0250	*NAPSA*	Napsin A aspartic peptidase	Gossner et al., 2013
		ENSOARG00000016543	-0.7346	5.5200	0.0000	0.0272	*CXCL9*	Chemokine (C-X-C motif) ligand 9	Ahmed, 2013
		ENSOARG00000000101	-2.0574	4.8237	0.0000	0.0371	-		
		ENSOARG00000003661	-0.5648	4.2727	0.0000	0.0454	*TOX2*	TOX high mobility group box family member 2	
		ENSOARG00000004253	-1.0222	2.7854	0.0000	0.0501	-	Putative MIP1-beta protein (CCL4)	
		ENSOARG00000001131	-0.7488	6.3769	0.0001	0.0775	-		
		ENSOARG00000003547	-0.5483	7.6500	0.0001	0.0775	*MYBL2*	V-myb avian myeloblastosis viral oncogene homolog-like 2	
		ENSOARG00000004198	-0.6130	4.7987	0.0001	0.0775	*SCIMP*	SLP adaptor and CSK interacting membrane protein	
		ENSOARG00000009389	-0.5034	5.5362	0.0001	0.0775	*KIF18B*	Kinesin family member 18B	
		ENSOARG00000013068	-1.3952	2.0931	0.0001	0.0775	*SLC9A4*	Solute carrier family 9, subfamily A (NHE4, cation proton antiporter 4), member 4	
		ENSOARG00000017587	-0.6208	3.8368	0.0001	0.0775	*PDCD1*	Programmed cell death 1	
		ENSOARG00000012936	-0.4881	5.6610	0.0001	0.0826	*CD83*	CD83 molecule	Kadarmideen et al., 2011
		ENSOARG00000010272	-0.7069	3.8748	0.0002	0.0931	*DUSP4*	Dual specificity phosphatase 4	
		ENSOARG00000015815	-1.3691	1.2008	0.0002	0.0956	*CNKSR2*	Connector enhancer of kinase suppressor of Ras 2	

Table 3.14: Genes differentially expressed between 7 and 14 days post infection (dpi) in LowFEC animals. Differentially expressed genes were determined in EdgeR using tagwise dispersion estimates, and a false discovery rate (FDR) of 0.1. Log fold change (logFC), log counts-per-million (logCPM) and P-value are given. Studies in which the gene has perviously been implicated1 in resistance or susceptibility to gastrointestinal nematodes are also reported.

Year	Up at	Gene id	logFC	logCPM	P-Value	FDR	Gene name	Description	Studies
2010	7 dpi	ENSOARG00000010231	1.8939	4.2686	0.0000	0.0000	*MX2*	Myxovirus (influenza virus) resistance 2 (mouse)	Nagaraj et al., 2012; Ahmed, 2013
		ENSOARG00000020801	1.3258	6.6374	0.0000	0.0174	-		
		ENSOARG00000020811	1.5164	3.2723	0.0000	0.0196	-		
		ENSOARG00000020789	1.0816	10.1503	0.0000	0.0340	-		
		ENSOARG00000020194	1.8702	2.7320	0.0000	0.0361	*MAB21L3*	Mab-21-like 3 (*C. elegans*)	
	14 dpi	ENSOARG00000016098	-2.2864	7.3832	0.0000	0.0045	-		
		ENSOARG00000002985	-2.6871	7.6617	0.0000	0.0174	-	DQA	
		ENSOARG00000010572	-1.2532	5.8947	0.0000	0.0260	-		
		ENSOARG00000020792	-1.5821	6.2244	0.0000	0.0260	-		
		ENSOARG00000001720	-1.3436	0.7634	0.0000	0.0344	-		
		ENSOARG00000000058	-1.7761	4.8124	0.0000	0.0344	-		
		ENSOARG00000015866	-0.5651	7.2474	0.0000	0.0344	-	-	
		ENSOARG00000001279	-1.1375	5.2625	0.0001	0.0894	-	Galectin-14 (LOC443162)	
		ENSOARG00000005126	-0.9353	3.0375	0.0001	0.0991	*EMR3*	Egf-like module containing, mucin-like, hormone receptor-like 3	Ahmed, 2013; Gossner et al., 2013
2011	7 dpi	ENSOARG00000007090	2.0560	1.9721	0.0000	0.0114	*CYP2F1*	Cytochrome P450, family 2, subfamily F, polypeptide 1	
		ENSOARG00000018232	2.8764	2.6187	0.0000	0.0202	-		
		ENSOARG00000010539	1.2049	2.8536	0.0000	0.0249	*HIF3A*	Hypoxia inducible factor 3, alpha subunit	
		ENSOARG00000010719	1.9267	5.0369	0.0000	0.0261	*LPL*	Lipoprotein lipase (LPL)	Ahmed, 2013
		ENSOARG00000005878	2.4984	1.9040	0.0000	0.0261	*CIDEC*	Cell death-inducing DFFA-like effector c	Ahmed, 2013
		ENSOARG00000013340	1.6160	2.8130	0.0000	0.0261	*CCL26*	Chemokine (C-C motif) ligand 26	Gossner et al., 2013
		ENSOARG00000020509	2.2821	5.7219	0.0000	0.0299	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	Ahmed, 2013
		ENSOARG00000020386	1.0129	3.9219	0.0000	0.0380	*APOD*	Apolipoprotein D	Gossner et al., 2013
		ENSOARG00000004875	1.4507	0.6305	0.0000	0.0380	-		
		ENSOARG00000010344	0.5034	6.3457	0.0000	0.0381	*LTBP1*	Latent transforming growth factor beta binding protein 1	
		ENSOARG00000007233	2.3750	4.1483	0.0001	0.0734	*ISG17*	Interferon stimulated gene 17	Nagaraj et al., 2012
		ENSOARG00000005151	1.4471	2.9896	0.0001	0.0734	*HMCN1*	Hemicentin 1	Ahmed, 2013
		ENSOARG00000011709	2.2558	2.3920	0.0001	0.0861	*PLIN1*	Perilipin 1	
		ENSOARG00000013498	2.3739	1.3710	0.0001	0.0878	*DNAH5*	Dynein, axonemal, heavy chain 5	Ahmed, 2013
		ENSOARG00000014842	1.6731	4.1759	0.0001	0.0994	*COL6A5*	Collagen, type VI, alpha 5	Gossner et al., 2013
	14 dpi	ENSOARG00000012273	-2.6203	1.0941	0.0000	0.0066	*PLK5*	Polo-like kinase 5	Ahmed, 2013
		ENSOARG00000002418	-2.5601	0.4516	0.0000	0.0249	-		
		ENSOARG00000002371	-1.7440	1.3515	0.0000	0.0261	-		
		ENSOARG00000014218	-1.4438	0.2888	0.0001	0.0806	-		
		ENSOARG00000009963	-0.7684	5.0979	0.0001	0.0861	-		
		ENSOARG00000008738	-0.7452	4.2223	0.0001	0.0861	*SCRN1*	Secernin 1	Pemberton et al., 2011; Ahmed, 2013

Hierarchical clustering was performed on DE genes (tagwise dispersion; FDR <0.1) using moderated $\log_2$ counts-per-million (logCPM) to observe the similarities between samples (Figure 3.13). It is expected that normalisation removes variation that is not due to biological differences (Rapaport et al., 2013). Dendrograms illustrate the similarities between both samples and genes, with all animals grouping within day post infection.

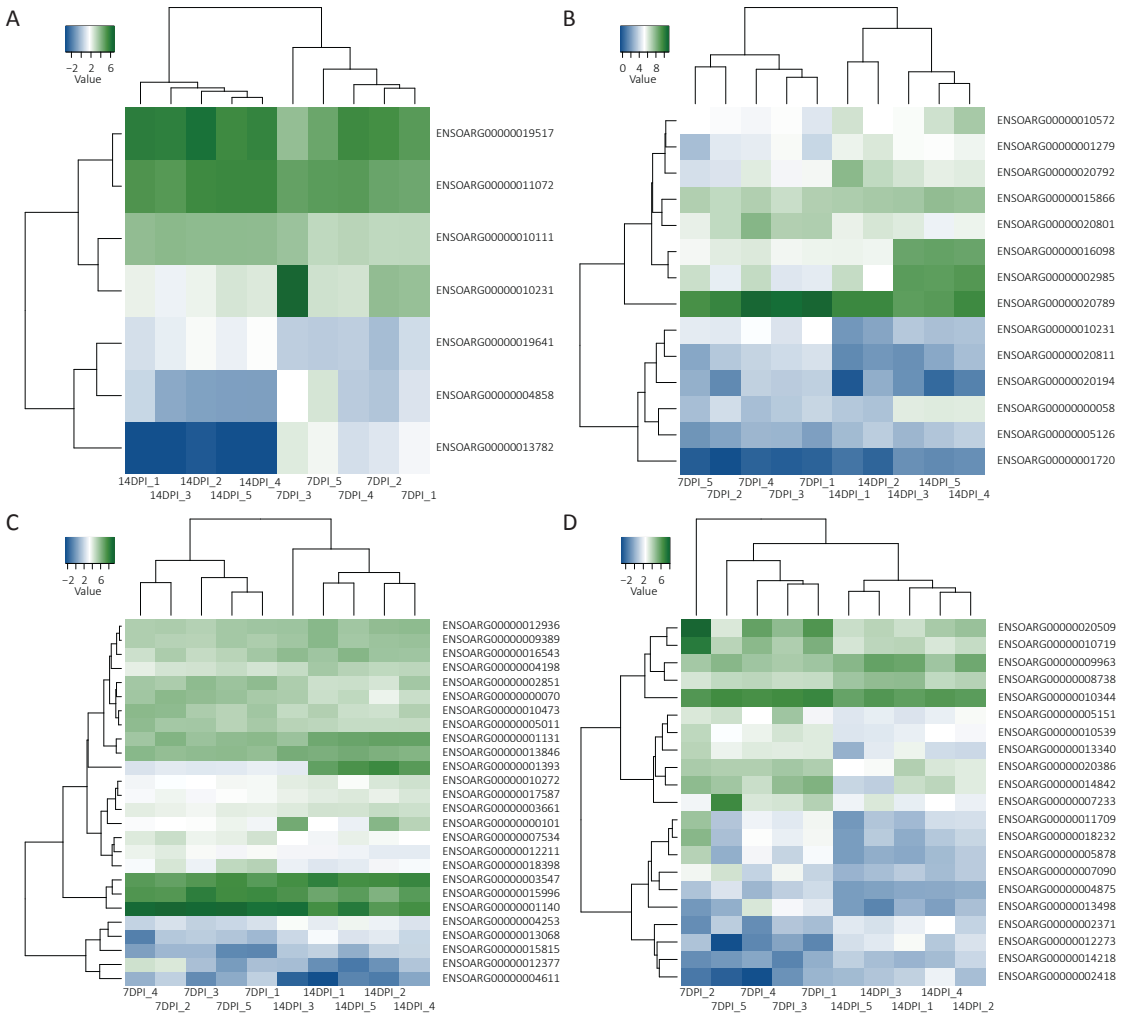


Figure 3.13: Heat map of hierarchical clustering of genes differentially expressed between animals slaughtered at 7 and 14 days post infection (dpi). Differentially expressed genes were determined in EdgeR using tagwise dispersion estimates, and a false discovery rate (FDR) of 0.1. Animals had a HighFEC (A & C) or LowFEC (B & D) phenotype, and were born in 2010 or 2011 respectively. Animals that do not group within phenotype are indicated with an asterisk (*).

Pathway analysis

Human 1-to-1 orthologs were found for approximately 64% of the DE genes (common dispersion; FDR < 0.05) between 7 and 14 days post infection (Table 3.15).

Table 3.15: Number of genes differentially expressed between 7 and 14 days post infection for which 1-to-1 human orthologs could be found. Differentially expressed genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05. Orthologs were obtained from Ensembl release 74.

Comparison		No. genes	1-to-1 ortholog		No 1-to-1 ortholog	
2010	HF	224	139	62%	85	38%
	LF	163	98	60%	65	40%
2011	HF	247	158	64%	89	36%
	LF	204	139	68%	65	32%

IPA analysis was carried out using the 1-to-1 orthologs. For each comparison the top networks (Table 3.16), diseases, and biological functions (Table 3.17) were analysed. None of the associated network functions were in common between any of the comparisons (Table 3.16). Associated diseases and disorders included infection, inflammatory and immunological disease (Table 3.17). Significant physiological system development and function categories included tissue development and haematological system development and function (Table 3.17).

Table 3.16: Top IPA networks in HighFEC or LowFEC animals when comparing 7 and 14 days post infection. Score represents the number of differentially expressed genes in the network.

Year	Phenotype	Associated network functions	Score
2010	HighFEC	Connective Tissue Development and Function, Embryonic Development, Organ Development	43
		Carbohydrate Metabolism, Small Molecule Biochemistry, Respiratory System Development and Function	35
		Infectious Disease, Dermatological Diseases and Conditions, Inflammatory Disease	31
		Developmental Disorder, Drug Metabolism, Energy Production	29
		Dermatological Diseases and Conditions, Developmental Disorder, Hereditary Disorder	24
	LowFEC	Cardiovascular System Development and Function, Organismal Development, Cell-To-Cell Signalling and Interaction	32
		Cell-To-Cell Signalling and Interaction, Cellular Movement, Immune Cell Trafficking	29
		Connective Tissue Disorders, Hereditary Disorder, Metabolic Disease	27
		Cell Cycle, Cellular Growth and Proliferation, Embryonic Development	24
		Cellular Movement, Drug Metabolism, Lipid Metabolism	24
2011	HighFEC	Cell Death and Survival, Cancer, Organismal Injury and Abnormalities	43
		Humoral Immune Response, Protein Synthesis, Cell-To-Cell Signalling and Interaction	40
		Connective Tissue Disorders, Hereditary Disorder, Immunological Disease	26
		Cell Morphology, Cellular Assembly and Organisation, Cellular Function and Maintenance	26
		Embryonic Development, Organismal Development, Carbohydrate Metabolism	25
	LowFEC	Antimicrobial Response, Inflammatory Response, Dermatological Diseases and Conditions	43
		Lipid Metabolism, Small Molecule Biochemistry, Endocrine System Disorders	41
		Cell Morphology, Cell-To-Cell Signalling and Interaction, Nervous System Development and Function	27
		Organismal Development, Cellular Movement, Carbohydrate Metabolism	26
		Embryonic Development, Nervous System Development and Function, Organ Development	25

Table 3.17: IPA analysis of differentially expressed (DE) genes between 7 and 14 days post infection. Listed are the top five diseases, biological functions and canonical pathways that the DE genes affect. DE genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05.

Year	Phenotype	Diseases and disorders	p-value	Molecules
2010	HighFEC	Infectious Disease	1.33E-04	8
		Inflammatory Disease	1.50E-04	7
		Neurological Disease	1.50E-04	22
		Skeletal and Muscular Disorders	1.50E-04	19
		Gastrointestinal Disease	2.83E-04	52
	LowFEC	Inflammatory Response	9.86E-09	29
		Immunological Disease	6.07E-07	28
		Infectious Disease	1.56E-06	15
		Antimicrobial Response	1.76E-06	8
		Dermatological Diseases and Conditions	1.17E-05	25
2011	HighFEC	Cancer	1.16E-05	99
		Immunological Disease	2.83E-05	32
		Hypersensitivity Response	4.26E-05	10
		Connective Tissue Disorders	5.56E-05	25
		Hereditary Disorder	5.56E-05	25
	LowFEC	Dermatological Diseases and Conditions	3.01E-08	32
		Inflammatory Disease	3.20E-07	9
		Neurological Disease	3.20E-07	32
		Skeletal and Muscular Disorders	3.20E-07	19
		Endocrine System Disorders	4.77E-07	29

Year	Comparison	Physiological system development and function	p-value	Molecules
2010	HighFEC	Nervous System Development and Function	1.70E-04	26
		Tissue Development	2.83E-04	26
		Tumour Morphology	2.83E-04	2
		Organismal Functions	3.96E-04	9
		Digestive System Development and Function	4.70E-04	9
	LowFEC	Immune Cell Trafficking	3.89E-11	20
		Haematological System Development and Function	9.86E-09	26
		Cardiovascular System Development and Function	9.64E-07	21
		Organismal Survival	1.88E-06	13
		Cell-mediated Immune Response	3.29E-06	7
2011	HighFEC	Humoral Immune Response	1.05E-04	11
		Haematological System Development and Function	1.37E-04	30
		Tissue Development	1.66E-04	37
		Embryonic Development	3.08E-04	22
		Organ Development	3.08E-04	16
	LowFEC	Organismal Development	5.30E-07	34
		Endocrine System Development and Function	3.41E-06	17
		Connective Tissue Development and Function	5.79E-06	19
		Skeletal and Muscular System Development and Function	5.79E-06	12
		Respiratory System Development and Function	2.56E-05	13

Year	Phenotype	Canonical pathways	p-value	Ratio
2010	HighFEC	Agranulocyte Adhesion and Diapedesis	9.38E-03	5/192 (0.026)
		IL-17A Signalling in Gastric Cells	1.29E-02	2/28 (0.071)
		Glutamine Degradation I	1.38E-02	1/5 (0.2)
		STAT3 Pathway	1.42E-02	3/80 (0.038)
		Glutathione-mediated Detoxification	1.6E-02	2/44 (0.045)
	LowFEC	Pathogenesis of Multiple Sclerosis	5.25E-08	4/10 (0.4)
		IL-17A Signalling in Gastric Cells	2.03E-04	3/28 (0.107)
		Tryptophan Degradation to 2-amino-3-carboxymuconate Semialdehyde	5.8E-04	2/18 (0.111)
		Granulocyte Adhesion and Diapedesis	1.27E-03	5/182 (0.027)
		Agranulocyte Adhesion and Diapedesis	1.67E-03	5/192 (0.026)
2011	HighFEC	Histamine Biosynthesis	7.48E-03	1/3 (0.333)
		Agranulocyte Adhesion and Diapedesis	1.28E-02	5/192 (0.026)
		Glutamine Degradation I	1.49E-02	1/5 (0.2)
		Bupropion Degradation	1.61E-02	2/33 (0.061)
		Cellular Effects of Sildenafil (Viagra)	1.61E-02	4/155 (0.026)
	LowFEC	Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	6.52E-05	6/109 (0.055)
		Interferon Signalling	9.8E-05	4/36 (0.111)
		LXR/RXR Activation	1.77E-03	5/139 (0.036)
		AMPK Signalling	2.86E-03	5/181 (0.028)
		Nicotine Degradation II	9.51E-03	3/85 (0.035)

A comparison analysis was undertaken using IPA to identify the unique and common molecules across the two years. Heat maps were generated of the 20 most significant diseases and biological functions (Figure 3.14) , canonical pathways (Figure 3.15) and upstream regulators (Figure 3.16).

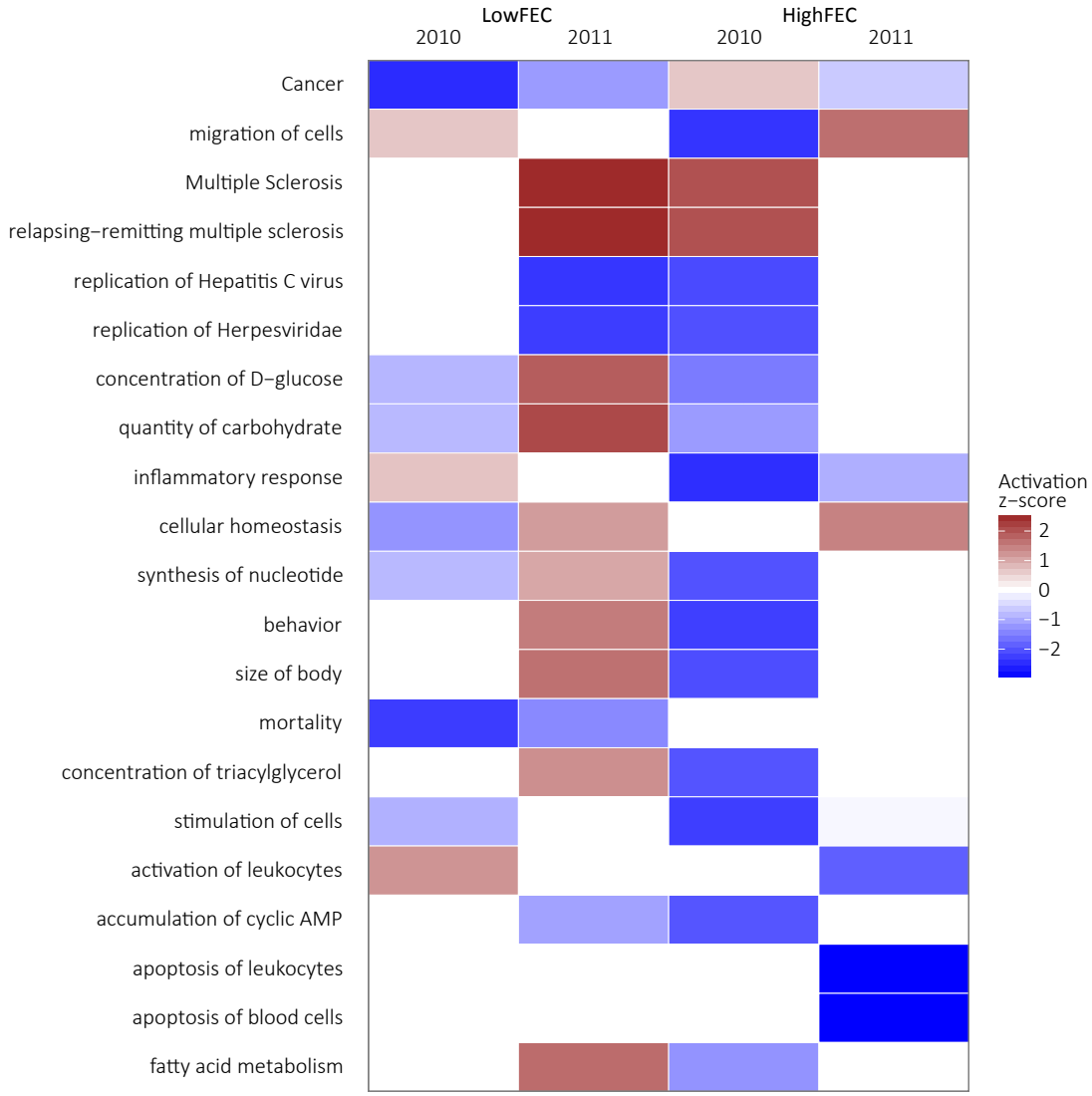


Figure 3.14: Heat map of the top 20 IPA-derived diseases and biological functions affected by genes differentially expressed between animals killed at 7 or 14 days post infection (dpi). Heat map is sorted by activation z-score. Red: predicted activation at 7 dpi; Blue: predicted activation 14 dpi.

Figure 3.15: Heat map of the top 20 IPA-derived canonical pathways affected by genes differentially expressed between animals killed at 7 or 14 days post infection (dpi). Heat map is sorted by -log(p-value). Expression of genes in each pathway may be increased at either 7 or 14 dpi.

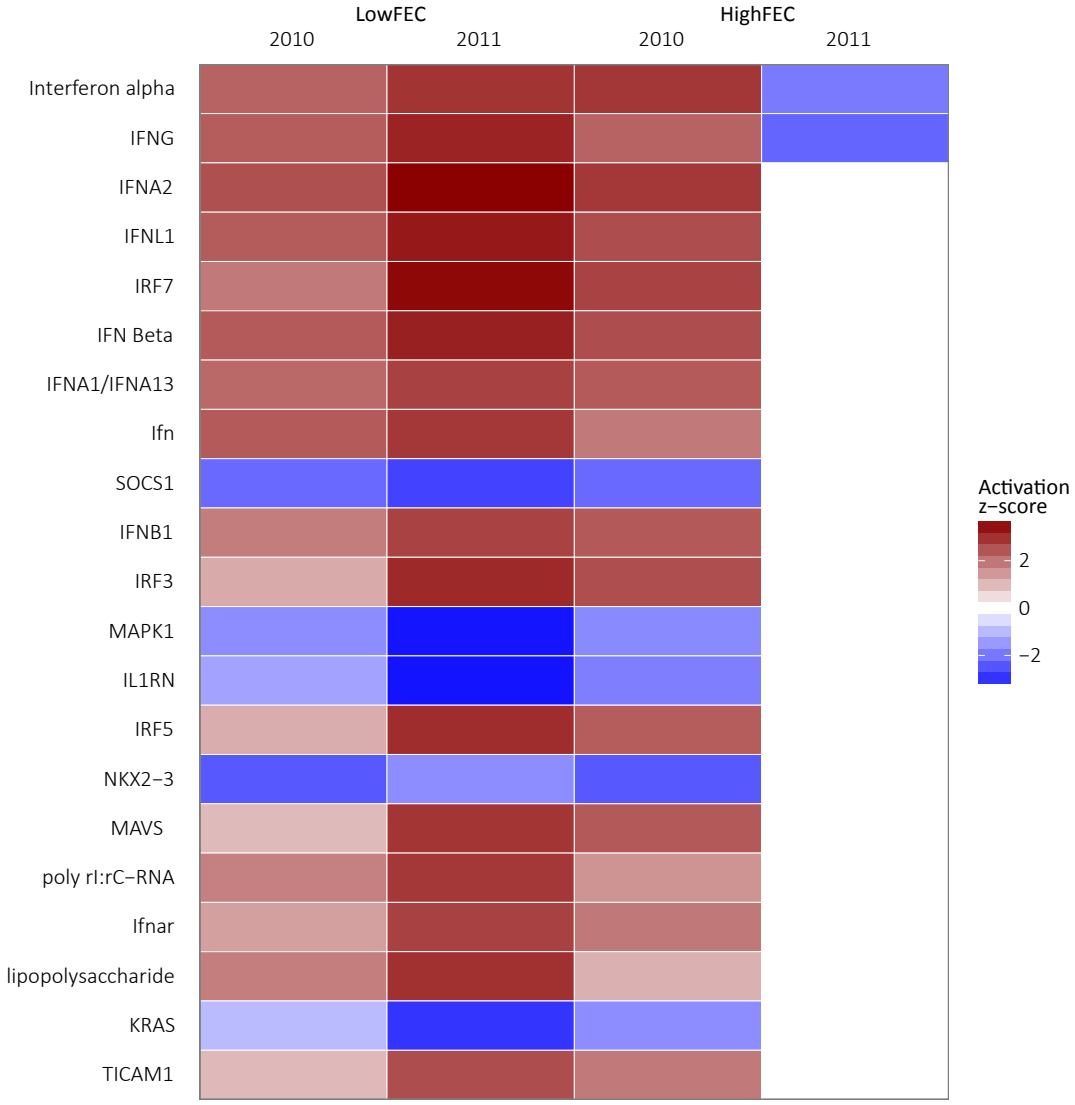
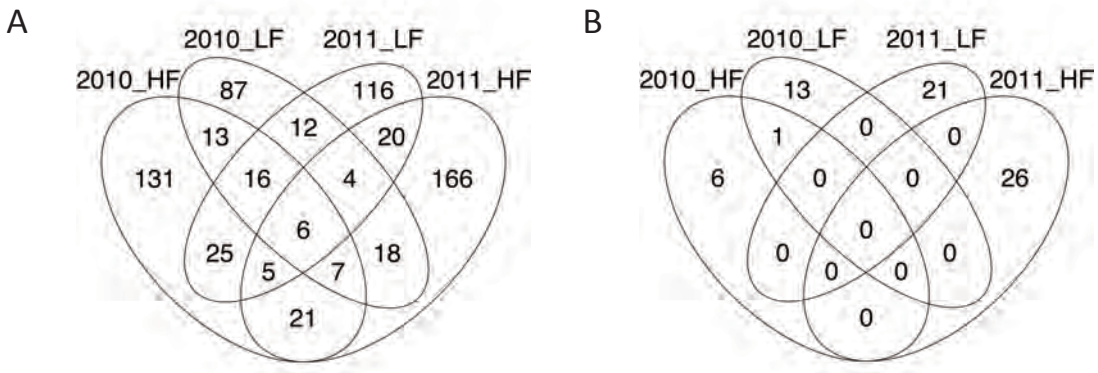


Figure 3.16: Heat map of the top 20 IPA-derived upstream regulators affected by genes differentially expressed between animals killed at 7 or 14 days post infection (dpi). Heat map is sorted by activation z-score. Red: expression of downstream genes increased at 7 dpi; Blue: expression of downstream genes increased at 14 dpi.

Genes in common between 2010 and 2011 analyses

Using common dispersion estimates, 6 genes (Table 3.12) were differentially expressed (FDR <0.05) in all 4 comparisons (Table 3.18).

Figure 3.17: Venn diagram comparing differentially expressed genes in 7 vs. 14 days post infection (dpi) analyses in 2010 and 2011. Genes differentially expressed between 7 and 14 dpi for each phenotype and year using common (A; FDR <0.05) and tagwise (B; FDR<0.1) dispersion.



Of these genes only one, *TRYPTASE-1* was consistent across group, being upregulated at 7 dpi in all comparisons. Four genes, *GZMA*, *MMRN1*, *COL6A5* and a novel protein pseudogene similar to GTPase IMAP family member 7, were up/downregulated according to year.

Table 3.18: Genes differentially expressed between 7 and 14 days post infection (dpi) in HighFEC and LowFEC animals in both 2010 and 2011. Log fold changes (logFC) in genes significantly (common dispersion; FDR <0.05) differentially expressed between day 7 and 14 dpi in HighFEC and LowFEC animals in both 2010 and 2011. Positive logFC indicates gene is upregulated at 7 dpi, whereas negative logFC indicates gene is upregulated at 14 dpiß.

Gene ID	Gene name	2010		2011	
		HighFEC	LowFEC	HighFEC	LowFEC
ENSOARG00000007970	*GZMA*	**1.05**	**0.85**	-0.86	-0.90
ENSOARG00000018398	*MMRN1*	-0.92	-0.73	**1.42**	**1.18**
ENSOARG00000014842	*COL6A5*	-0.87	-0.59	**1.83**	**1.67**
ENSOARG00000013119	*SNTG1*	**1.20**	-1.49	-1.06	-1.19
ENSOARG00000014689	*TRYPTASE-1*	**0.79**	**0.95**	**1.04**	**0.76**
ENSOARG00000001424	-	-0.58	-1.10	**1.36**	**0.87**

Of the DE genes (common dispersion; FDR < 0.05) for which human 1-to-1 orthologs were found, 4 remained in common between all comparisons (*GZMA*, *MMRN1*, *COL6A5* & *SNTG1*) (Figure 3.18).

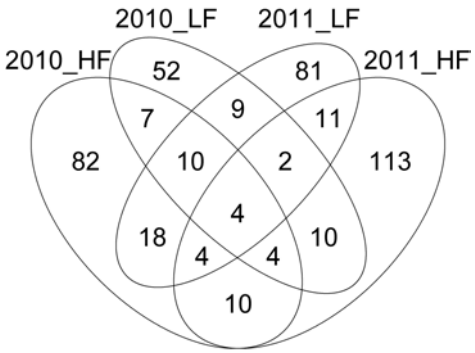


Figure 3.18: Venn diagram comparing differentially expressed genes in 7 vs. 14 days post infection analyses in 2010 and 2011. Differentially expressed genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05.

3.3.3.3 2010- vs. 2011-born animals

Overall variation in gene expression among all animals was examined using a multi-dimensional scaling plot to determine if the data from both years could be combined. An MDS plot (Figure 3.19) showed that the samples separated clearly by year. This most likely represents the extreme differences in worm burden (Figure 2.7) observed between animals challenged in 2010 and those challenged in 2011. Gene expression differences between animals with a heavy infection (2010-born) and a light infection (2011-born) with *T. circumcincta* were therefore examined.

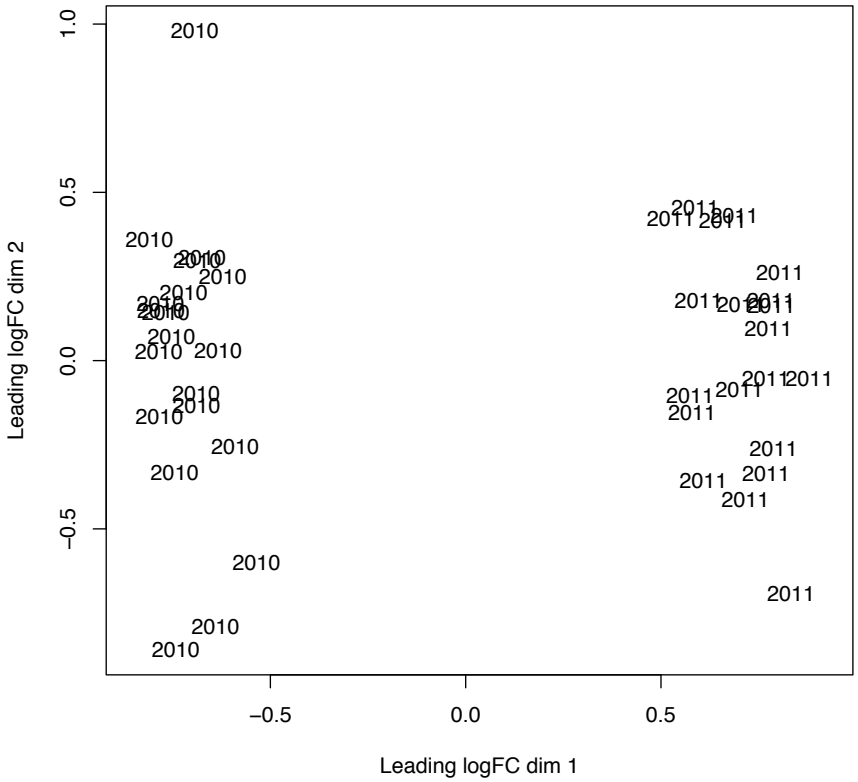


Figure 3.19: Multi-dimensional scaling (MDS) plots comparing RNA-Seq results from both 2010- and 2011-born animals. Distances correspond to leading log-fold-changes between each pair of samples.

Using tagwise dispersion estimates 4,411 genes were identified as differentially expressed (FDR < 0.1) between 2010- and 2011-born animals (Table 3.19). The high number of DE genes is reflected in the MDS plots, which decisively separates animals by year on the first dimension (Figure 3.19).

Table 3.19: Number of differentially expressed genes between 2010- and 2011-born animals using common and tagwise dispersion estimates in EdgeR. The false discovery rate (FDR) was set to 0.1 or 0.05. Genes with a FDR < 0.05 and a log fold change (logFC) of >±2 are reported in the third column (logFC >±2).

	Common dispersion			*Tagwise dispersion*		
Comparison	**FDR <0.1**	**FDR <0.05**	**logFC >±2**	**FDR <0.1**	**FDR <0.05**	**logFC >±2**
2010v2011	3,386	2,776	68	4,411	3,328	68

Of the DE genes (tagwise dispersion; FDR <0.1), 869 (26%) were novel protein coding genes. None of the remaining genes had been implicated in previous studies. Human 1-to-1 orthologs were found for 2,409 (72%) genes (Table 3.20), which were then used as input for IPA analysis.

Table 3.20: Number of genes differentially expressed between 2010- and 2011-born animals for which 1-to-1 human orthologs could be found. Differentially expressed genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05. Orthologs were obtained from Ensembl release 74.

Comparison	**No. genes**	**1-to-1 ortholog**		**No 1-to-1 ortholog**	
2010v2011	3,328	2,409	72%	919	28%

For each comparison the top networks (Table 3.21), diseases and biological functions (Table 3.22) and canonical pathways (Table 3.23) were analysed. The top network, “Hereditary Disorder, Ophthalmic Disease, Molecular Transport”, is centred around the expression of ubiquitin C (*UBC*). Expression of *UBC* was increased in the 2010-born animals in comparison to those born in 2011.

Table 3.21: Top IPA networks when comparing 2010- and 2011-born animals. Score represents the number of differentially expressed genes in the network.

Associated Network Functions	**Score**
Hereditary Disorder, Ophthalmic Disease, Molecular Transport	42
DNA Replication, Recombination, and Repair, Gene Expression, Connective Tissue Development and Function	39
RNA Post-Transcriptional Modification, Cardiovascular System Development and Function, Embryonic Development	37
Amino Acid Metabolism, Small Molecule Biochemistry, Cell Cycle	34
RNA Post-Transcriptional Modification, Metabolic Disease, Neurological Disease	34

The top diseases and disorders from the IPA analysis include cancer (a subset of which was gastrointestinal tract cancer), infectious disease and gastrointestinal disease (Table 3.22). Under the *Haematological System Development and Function* category the most significant

over-represented annotations include quantity of blood and haematopoietic cells, and quantity and homing of haematopoietic progenitor cells.

Table 3.22: IPA analysis of differentially expressed (DE) genes between 2010- and 2011-born animals. Listed are the top five diseases, molecular and cellular functions and physiological system functions that the DE genes affect. DE genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05.

Comparison	Top diseases & bio functions	p-value	Molecules
Diseases & disorders	Cancer	1.96E-12	1170
	Developmental Disorder	4.67E-10	333
	Infectious Disease	6.13E-10	306
	Gastrointestinal Disease	7.18E-10	626
	Organismal Injury and Abnormalities	7.39E-09	596
Molecular & cellular functions	Cellular Growth and Proliferation	8.73E-27	689
	Cell Death and Survival	6.45E-18	650
	Cell Cycle	2.64E-12	270
	Cell Morphology	1.34E-11	440
	Gene Expression	6.67E-11	401
Physiological system development & function	Organismal Survival	1.02E-17	475
	Cardiovascular System Development and Function	6.02E-12	275
	Tissue Morphology	5.99E-11	417
	Organismal Development	1.93E-10	582
	Haematological System Development and Function	2.19E-08	349

The most significant canonical pathway was the protein ubiquitination pathway (Table 3.23).

Table 3.23: Top IPA canonical pathways when comparing 2010- and 2011-born animals. Listed are the top five canonical pathways that the DE genes affect. DE genes were determined in EdgeR using common dispersion estimates, and a false discovery rate (FDR) of 0.05.

Canonical pathways	p-value	Ratio
Protein Ubiquitination Pathway	7.04E-08	55/270 (0.204)
Mitochondrial Dysfunction	1.08E-05	36/215 (0.167)
Epithelial Adherens Junction Signalling	6.55E-05	31/154 (0.201)
Oxidative Phosphorylation	7.54E-05	25/120 (0.208)
EIF2 Signalling	1.07E-04	36/201 (0.179)

A number of the genes involved in the protein ubiquitination pathway are part of the proteasome and immunoproteasome (Figure 3.20).

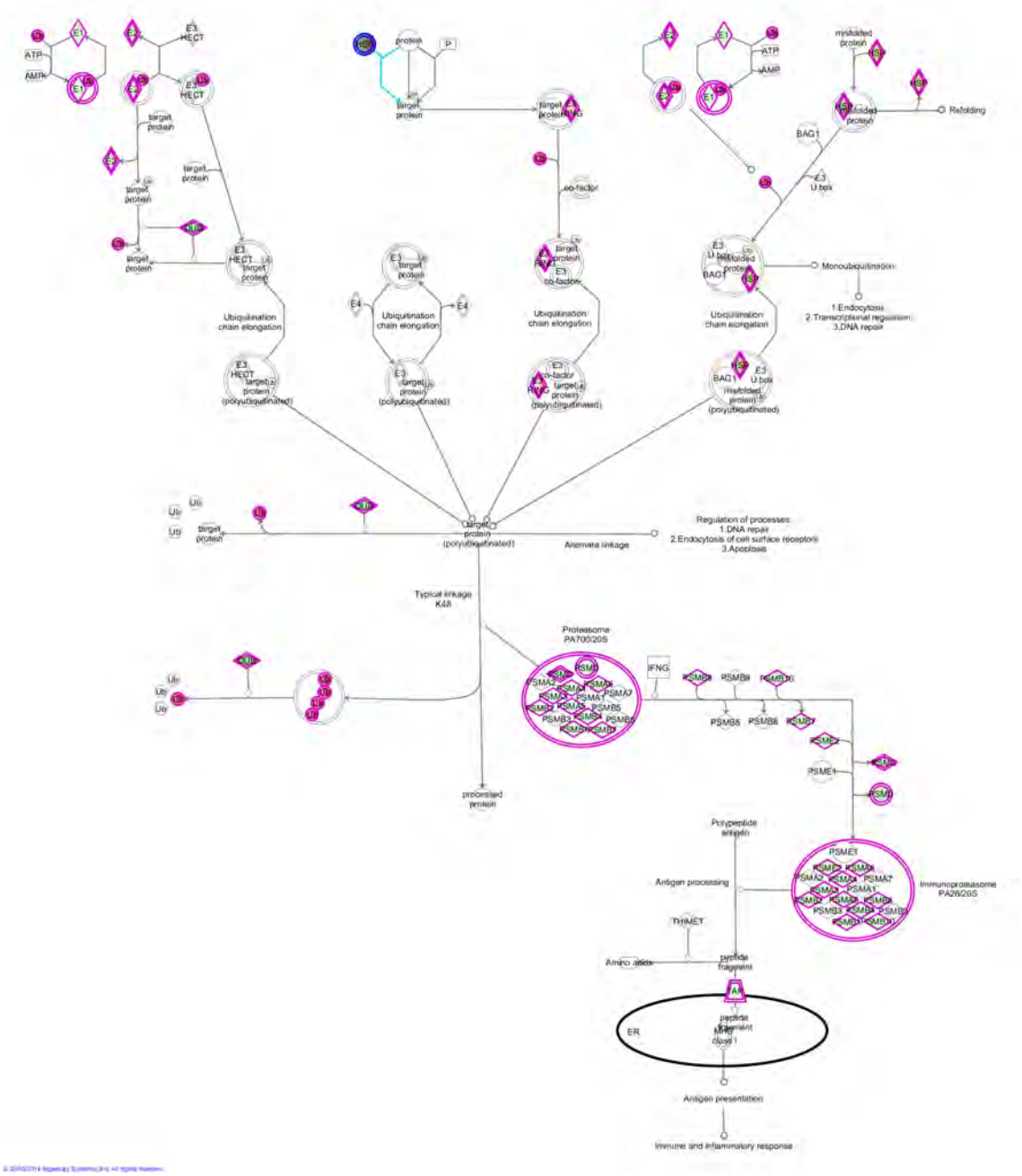


Figure 3.20: The protein ubiquitination pathway. Genes differentially expressed between 2010- and 2011-born animals are highlighted in pink (figure from IPA).

3.4 Discussion

3.4.1 Response to gastrointestinal nematode infection in Scottish Blackface lambs with divergent phenotypes for resistance

3.4.1.1 Day 7 post infection

Significantly differentially expressed genes

In 2010 the only genes differentially expressed (tagwise dispersion; FDR < 0.1) at 7 dpi were more highly expressed in HighFEC animals. Of particular interest is albumin (*ALB*), which has been found to be associated with GIN infection in four separate studies. Serum albumin, the main protein of plasma, is a carrier protein for steroids, fatty acids, and thyroid hormones, and functions as a regulator of the colloidal osmotic pressure of blood. *ALB* has been reported to be more highly expressed in the duodenum of helminth-naïve genetically resistant animals compared to susceptible individuals (Keane et al., 2006), and downregulated in the abomasum of sheep during the course of repeated truncated immunising infections with *T. colubriformis* larvae (Knight et al., 2010). Protein albumin levels were also increased in the abomasal mucosa of resistant selection-line animals compared to their susceptible counterparts three days after experimental challenge with *Haemonchus contortus* (Nagaraj et al., 2012). In a separate study on changes in abomasal protein expression following trickle infection with *T. circumcincta*, *ALB* was relatively highly expressed in the mucosa of helminth-naïve animals in comparison to both their immune and immune-waning counterparts (Pemberton et al., 2012). Immunohistochemistry suggests the site of expression may be mucus-producing epithelial cells. In the same study serum albumin was also significantly decreased in GIN immune sheep compared to the naïve controls. Pemberton et al. (2012) hypothesised that albumin may be constitutively released into the gastric mucus, and may therefore play an innate protective role. In 2010 the response to GIN infection in the HighFEC animals appears to be similar to that previously observed in naïve animals, suggesting that the HighFEC animals may not be generating an effective immune response to GIN infection.

In 2011 HighFEC animals had elevated expression of *BTLA* and *FCRL1* at 7 dpi in comparison to their LowFEC counterparts. B-and T-lymphocyte attenuator (*BTLA*) encodes for a protein that relays inhibitory signals to suppress the immune response. *BTLA* is not expressed by naive T cells, but it is induced during activation, and remains expressed on T_H1 but not T_H2 cells (Watanabe et al., 2003), indicating the HighFEC animals had a more T_H1-type response to GIN infection. BTLA is a high affinity co-receptor for herpes virus entry mediator (HVEM) (Sedy et al., 2005; Gonzalez et al., 2005), and the BTLA-HVEM complex negatively regulates T-cell immune responses by inhibiting T-cell proliferation (Gonzalez et al., 2005). The immunoglobulin receptor superfamily member Fc receptor-

like 1 (*FCRL1*) is expressed primarily, although not exclusively, by B cells (Davis et al., 2001). *FCRL1* expression begins in pre-B cells, reaches peak levels on naive B cells, and is down-regulated after B cells are activated (Leu et al., 2005). This could be a result of the HighFEC animals having fewer activated B cells than their LowFEC counterparts. FCRL1 is an intrinsic activation molecule, that acts as an activation co-receptor to augment B-cell antigen receptor (BCR)-induced activation of B-cells. The ligand for this receptor is currently unknown; unlike Fc receptors FCRL1 does not bind IgM, IgA, or IgGs (Leu et al., 2005). *FCRL1* expression has been shown to be significantly higher in peripheral blood B lymphocytes of patients with certain autoimmune diseases including multiple sclerosis (Baranov et al., 2012). In agreement with our study *FCRL1* has been previously reported to be more highly expressed in the abomasal lymph node of susceptible Scottish Blackface lambs in comparison to their resistant counterparts (Gossner et al., 2013).

In 2011 LowFEC animals had increased expression of *GIMAP8*, *GZMK*, *WARS*, *CCL4* and *CXCL9* at 7 dpi. GTPase, IMAP family member 8 (*GIMAP8*) belongs to the immune-associated nucleotide (IAN) subfamily of nucleotide-binding proteins, and is expressed in T cells and thymocytes (Dion et al., 2005). In agreement with our study Keane et al. (2006) found *GIMAP8* expression to be increased in genetically resistant helminth naïve animals compared to their susceptible counterparts. However, Ahmed (2013) found expression to be higher in Suffolk (relatively susceptible) animals compared to Texel (relatively resistant) during a controlled challenge with *T. circumcincta.* Granzyme K (*GZMK*) is a trypsin-like serine protease that highly expressed in peripheral blood leukocytes, spleen, thymus, and lung tissues (Przetak et al., 1995). Extracellular GZMK is capable of activating Protease-Activated Receptor 1 (PAR-1) and inducing fibroblast cytokine secretion and proliferation (Cooper et al., 2011). The expression of tryptophan-tRNA synthetase (*WARS*, previously known as *TrpRS*) has been shown to be stimulated by interferon gamma (*IFN-*γ) in humans and mice (Bange et al., 1992; Rubin et al., 1991; Miyanokoshi et al., 2013). Despite being associated with resistance in this study, *WARS* has previously been associated primarily with susceptibility of lambs to GIN (Andronicos et al., 2010; Nagaraj et al., 2012; Ahmed, 2013). Thus, WARS protein levels were increased in the abomasal mucosa of susceptible selection line animals compared to their resistant counterparts 3 days after experimental challenge with *Haemonchus contortus* (Nagaraj et al., 2012). The expression of *WARS* was increased in previously-challenged uninfected Suffolk lambs when compared to uninfected Texel lambs (Ahmed, 2013). Expression of *WARS* has also been shown to be increased in the jejunum of resistant compared to susceptible animals, from lines of a flock selected on *Haemonchus* (HSF), after three infections with *T. colubriformis*, in agreement with our study. However, it was also found to be upregulated in the jejunum of animals from a flock selected for resistance to *Trichostrongylus* (TSF) animals after a single infection (Andronicos et al., 2010). The chemotactic cytokines (chemokines) *CCL4* and *CXCL9* are both part of the chemokine control of T cell migration in the lymph node during the immune response (Bromley et al., 2008). The pro-inflammatory chemokine *CCL4*, previously

known as macrophage inflammatory protein-1β (*MIP-1β*), can be induced in most mature haematopoietic cells, particularly macrophages, dendritic cells, and lymphocytes (Menten et al., 2002; Maurer and von Stebut, 2004). *CCL4* binds to the T cell-chemoattractant *CCR5*, and promotes beneficial leukocyte recruitment to infected tissues (Broxmeyer et al., 1990). However, chronic expression of this chemokine contributes to inflammatory disease (Menten et al., 2002). While expression of *CCL4* was initially thought to be only associated with a T_H1 immune response (Schrum et al., 1996), it has subsequently been associated both with T_{reg} cells (Bystry et al., 2001) and a T_H2 response (Meagher et al., 2007; Ashenafi et al., 2014). The T_H1 inflammatory chemokine *CXCL9* binds to the chemokine receptor *CXCR3* and preferentially recruits *IFN-y* producing CD4+ T cells (Debes et al., 2006). There is an overlap of the chemokine receptor expression profiles in T_{reg} and T helper subsets; the receptors for *CCL4* and *CXCL9*, *CCR5* and *CXCR3*, are expressed together on both T_{reg} and T_H1 cells (Bromley et al., 2008). The source of the differentially expressed chemokines is potentially dendritic cells (DCs); mature T_H1-promoting DCs have been shown to constitutively express both inflammatory (*CCL4*) and T_H1-associated (*CXCL9*) chemokines (Lebre et al., 2005).

Together these results indicate than in 2011 the HighFEC animals potentially have lower T cell proliferation, along with an increased T_H1-type response and an increased number of naive B cells. The LowFEC animals however appear to be expressing cytokines and chemokines that generate an inflammatory immune response to GIN infection, and this may result in an influx of immune cells to the site of infection.

Pathway analysis

The list of genes differentially expressed using common dispersion (FDR <0.05) were used for pathway analysis, as this method is more robust to false positives. Using this criterion a number of chemokine (C-C and C-X-C motif) ligands (CCL and CXCL) and receptors (CCR and CXCR) were significantly increased at 7 dpi in LowFEC animals in both 2010 and 2011. The main function of chemokines is to control leukocyte migration. It appears that, in LowFEC animals, increased cytokine expression leads to an increase in the inflammatory response, including migration of leukocytes and antigen presenting cells. These chemokine ligands and receptors are also present in the pathogenesis of multiple sclerosis pathway, which was the top canonical pathway in the comparison analysis.

Primarily as a result of the increase in the above chemokines, in both 2010 and 2011 IPA analysis showed an increase in molecules associated with *haematological system development and function*, *immune cell trafficking* and *tissue development*. The top physiological system development and function category in both years was *haematological system development and function*; the majority of genes in this category in both 2010 (16/19 genes) and 2011 (28/32 genes) were upregulated in LowFEC animals. This is also true of the disease and disorder category *hypersensitivity response*; 7/9 genes in 2010 and 11/12

genes in 2011 were upregulated in LowFEC animals.

3.4.1.2 Day 14 post infection

Significantly differentially expressed genes

At 14 dpi, 18 genes were differentially expressed between HighFEC and LowFEC animals, 6 in 2010 and 12 in 2011. Of note is the epithelial cell transforming sequence 2 oncogene-like (*ECT2L*), which was increased in LowFEC animals in 2010. Recurrent somatic mutations in *ECT2L* have been associated with early T-cell precursor acute lymphoblastic leukaemia (Zhang et al., 2012). However none of the other differentially expressed genes, in either 2010 or 2011, had a known immune system function. Despite this, 6 genes were in common with the study by Ahmed (2013). This study compared Suffolk (relatively susceptible) and Texel (relatively resistant) animals over the course of a controlled challenge with *T. circumcincta*. Of the genes in common, 2 were consistent in their association with resistance between the two studies (*SCRN2*, more highly expressed in resistant animals, and *LPL*, more highly expressed in susceptible animals), while four were not (*ECT2L*, *ZFR2*, *DNAH5* and *MPPE1*).

Pathway analysis

IPA analysis did not reveal any diseases or biological functions common to both years. However, the network *Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry*, which was significant at 7 dpi in 2010, was significant at 14 dpi in 2011.

3.4.1.3 Comparison analysis of all HighFEC and LowFEC animals

While the differentially expressed genes may be different between the two years, examining upstream regulators allows us to see what pathways and biological mechanisms may be regulating the response to GIN in Scottish Blackface sheep. At 7 dpi genes downstream from the regulators *PPARG*, *IFNG* and *IL4* were activated in LowFEC animals, whereas by day 14 genes downstream of these regulators were upregulated in HighFEC animals.

Peroxisome proliferator-activated receptor gamma (*PPARG*) is an important modulator of the inflammatory response in various tissues, including the intestinal tract and colonic mucosa (Mansén et al., 1996; Dubuquoy et al., 2002). PPARγ is strongly induced by the T_H2 cytokine IL-4 (Huang et al., 1999). It therefore appears that a major transcriptional role of PPARγ is negative regulation of specific subsets of genes that are activated by LPS (lipopolysaccharides) and IFNγ, supporting a physiologic role of PPARγ in regulating both innate and acquired immune responses (Welch et al., 2003). Previous gene expression profiling of naïve sheep representing groups genetically resistant and susceptible to

gastrointestinal nematodes revealed motifs for PPARG binding in the promoter regions of the genes more highly expressed in susceptible animals (Keane et al., 2006). Interferon gamma (*IFN*γ) is secreted by T_H1 lymphocytes, and plays a critical role in regulating the type 1 versus type 2 immune responses in vertebrates. IFNγ activates macrophages, which can kill intracellular pathogens, and display increased ability to present antigens (Wakelin, 1996). It also helps to determine whether a humoral or cell mediated response predominates. Interleukin 4 (*IL4*) is typically associated with a T_H2-type response. IL-4 induces differentiation of naïve T_H cells to T_H2 cells, while suppressing differentiation into T_H1 cells (Anthony et al., 2007).

While the development of resistance to GIN is associated with the development of a $CD4^+$ T_H2-type immune response, and susceptibility with a $CD4^+$ T_H2-type response, recent research has challenged the view that it is simply a matter of T_H2/T_H1 dichotomy (Venturina et al., 2013). Several studies have indicated that it appears to be the differential interplay between $T_H1/T_H2/T_{reg}$ genes that control the response to gastrointestinal nematodes in resistant compared to susceptible breeds (Hassan et al., 2011b; Ahmed, 2013). This is supported by the recognition of new subsets of T cells, such as T_H17 and Treg cells (Zhu et al., 2010), and increased recognition of their plasticity (Nakayamada et al., 2012). Through analysis of upstream regulators in IPA, it appears that in both 2010 and 2011 the LowFEC animals are generating an immune response to *T. circumcincta* that is a combination of T_H1 and T_H2-related genes. In the HighFEC animals this response appears to be delayed until 14 days post infection, so therefore the LowFEC (resistant) animals may be generating an effective immune response more quickly.

3.4.2 Changes in gene expression over time in the abomasal lymph node of Scottish Blackface lambs challenged with *T. circumcincta*

3.4.2.1 HighFEC animals

Significantly differentially expressed genes

Using tagwise dispersion estimates, expression of *ALB* and *MX2* were significantly higher at 7 dpi in the 2010 HighFEC animals. The increased expression of *ALB* was by far the most significant change, with a log fold change of 7.3. Expression of myxovirus resistance 2 (*MX2*) is strongly induced by interferon-α (IFN-α) (Melén et al., 1996; Kane et al., 2013), and shows antiviral activity (Sasaki et al., 2014). Despite the association between this gene and viral infections, it has been shown to be upregulated at both the protein and mRNA level in GIN susceptible animals (Nagaraj et al., 2012; Ahmed, 2013).

At 14 dpi, *LYVE1* and *CHI3L2* were upregulated. Lymphatic vessel endothelial hyaluronan receptor 1 (*LYVE1*) is a major receptor for hyaluronan on the lymph vessel wall (Banerji et al., 1999). Hyaluronan is an abundant component of skin and mesenchymal tissues, where it

facilitates cell migration during wound healing, inflammation, and embryonic morphogenesis. Chitinase-like proteins such as *CHI3L2* have a role in inflammation, tissue remodelling, and injury (Lee et al., 2011a). Upregulated expression of *CHI3L2* has been observed in the abomasum of 18 and 21 week old steers exposed to *Ostertagia ostertagi*, and the abomasal lymph node of resistant and susceptible Scottish Blackface lambs infected with *T. circumcincta* in comparison to sham-infected controls (Gossner et al., 2013). In human macrophages *CHI3L2* has been found to be upregulated by *IL-4* and *TGF-*β (Gratchev et al., 2008).

In 2011 upregulated genes at 7 dpi include *GSDMA*, *MMRN1*, *LACC1*, *MCP-3* and *STAB2*. Human gasdermin A (*GDSMA*) is preferentially expressed in the gastric epithelium (Saeki et al., 2000). *GDSMA* is a component of TGF-β signalling, that induces of apoptosis in the pit cells of human gastric epithelium (Saeki et al., 2007) and has been associated with gastric cancer (Saeki et al., 2000). Asthma-associated polymorphisms have also been shown to lead to increased *GDSM* expression (Lluis et al., 2011). Ahmed (2013) observed increased expression of *GSDMA* in Texel animals compared with their Suffolk counterparts throughout the course of infection with *T. circumcincta*. Multimerin 1 (*MMRN1*) is expressed in megakaryocytes, platelets and endothelial cells (cells that lines the interior surface of blood vessels and lymphatic vessels), and plays a role in the storage and stabilisation of factor V in platelets (Jeimy et al., 2008). This gene has been shown to be more highly expressed in Suffolk compared to Texel animals at the early stages of infection with *T. circumcincta* (Ahmed, 2013), and downregulated in the mesenteric lymph nodes during the preclinical stages of scrapie in sheep (Filali et al., 2014). Variants in laccase (multicopper oxidoreductase) domain containing 1 (*LACC1*, formerly known as *C13ORF31*) have been related to susceptibility to leprosy (Zhang et al., 2009), ulcerative colitis, and Crohn's disease (Barrett et al., 2008). However, the function of this gene in the innate immune response remains unknown. Sheep mast cell proteinase-3 (*MCP-3*) is mast-cell-derived and may be co-expressed with *MCP-1* in abomasal mast cells of sheep infected with *T. circumcincta* (McAleese et al., 1998). The scavenger receptor Stabilin-2 (STAB2) is primarily expressed in organs with specialised endothelia such as liver, spleen and lymph node (Politz et al., 2002). SATB2 plays an important role in lymphocyte recruitment in the hepatic vasculature (Jung et al., 2007), and binds to and mediates endocytosis of hyaluronic acid (Politz et al., 2002).

Expression of *CCL4*, *CXCL9*, and *GIMAP8* was upregulated at 14 dpi in the 2011 HighFEC animals. These three genes were also upregulated in LowFEC compared to HighFEC animals at 7 dpi in the same year, indicating the HighFEC animals may be having a similar response to their LowFEC counterparts, but delayed by seven days. Other genes upregulated at 14 dpi include *SCIMP*, *SLC9A4*, *PDCD1*, *CD83*, *DUSP4* and *CNKSR2*. *SCIMP* is a lipid tetraspanin-associated transmembrane adapter/mediator involved in major histocompatibility complex (MHC) class II signalling transduction (Draber et al., 2011). *SCIMP* is expressed in B cells and other professional APCs, and serves as a regulator of

antigen presentation and other APC functions. The Na+/H+ exchanger (NHE) isoform 4, encoded by *SLC9A4*, is most abundant in the stomach of rats, followed by intermediate levels in small intestine and colon (Orlowski et al., 1992). *SLC9A4* knockout mice show sharply reduced numbers of parietal cells, a loss of mature chief cells, increased numbers of mucous and undifferentiated cells, and an increase in the number of necrotic and apoptotic cells. These changes indicate that *SLC9A4* is important for normal levels of gastric acid secretion and gastric epithelial cell differentiation (Gawenis et al., 2005). Programmed cell death-1 (*PDCD1*) is a surface receptor critical for the regulation of T cell function during immunity and tolerance (Fife and Pauken, 2011). Upon ligand binding, PDCD1 inhibits T-cell effector functions in an antigen-specific manner (Fife and Pauken, 2011). The type I Ig superfamily glycoprotein *CD83* is a well-known marker for mature dendritic cells (Zhou and Tedder, 1995). The primary function of *CD83* lies in the regulation of T- and B-lymphocyte maturation and in the regulation of their peripheral responses (reviewed in Breloer and Fleischer, 2008). Scholler et al. (2001) showed that CD83 bound to monocytes, but not lymphocytes, and that the binding was enhanced by stress. The mitogen- and stress-inducible *DUSP4* belongs to a class of dual-specificity phosphatases (DUSPs), designated MKPs, that negatively regulate members of the mitogen-activated protein kinase (MAPK) superfamily (MAPK/ERK, SAPK/JNK, p38) (Caunt and Keyse, 2013). The p38MAPK and ERK protein kinase pathways are contributors to intestinal hyper-contractility under T_H2 mediated inflammatory events such as colitis and inflammatory bowel disease (IBD) (Ihara et al., 2009). *Lanigan et al. (2003)* showed that *CNKSR2* (connector enhancer of kinase suppressor of Ras 2) mediates MAPK pathways downstream from Ras. The ability to interact with both Ras effector proteins Raf and Rlf, suggesting that *CNKSR2* may integrate signals between MAPK and Ral pathways through a complex interplay of components (Lanigan et al., 2003).

These results indicate that in both 2010 and 2011, HighFEC animals were developing a more effective immune response against GIN infection at 14 days post infection, with an increase in genes associated with immune cell function.

Pathway analysis

In both 2010 and 2011 the canonical pathways *Agranulocyte Adhesion and Diapedesis* and *Glutamine Degradation I* were both significant in HighFEC animals. In both years genes encoding for myosin and chemokines are differentially expressed between 7 and 14 dpi. These genes form part of the *Agranulocyte Adhesion and Diapedesis* pathway. The migration of agranulocytes (lymphocytes and monocytes) from the vascular system to sites of pathogenic exposure is a key event in the process of inflammation. The process of *Glutamine Degradation*, known as glutaminolysis, is mediated by glutaminase. The phosphate-activated mitochondrial protein glutaminase (*GLS2*) was upregulated at 7 dpi in both 2010 and 2011. It has been suggested that glutamate has a role as a key

immunomodulator in the initiation and development of T-cell-mediated immunity in peripheral tissues (Pacheco et al., 2007).

3.4.2.2 LowFEC animals

Significantly differentially expressed genes

Like their HighFEC counterparts, 2010-born LowFEC animals expressed higher levels of *MX2* at 7 days post infection than at 14 dpi. At 14 dpi LowFEC animals had increased expression of the epidermal growth factor-7-transmembrane (EGF-TM7) family member *EMR3*, which displays a predominantly leukocyte-restricted expression pattern, with highest levels in neutrophils, monocytes, and macrophages (Stacey et al., 2001). The EMR3 ligand is present at the surface of both monocyte-derived macrophages and activated human neutrophils, suggesting a potential role for EMR3 in myeloid-myeloid interactions during immune and inflammatory responses (Stacey et al., 2001). *EMR3* has been shown to be differentially expressed in the abomasal lymph node in two separate studies, one comparing Scottish Blackface lambs infected with *T. circumcincta* to sham-infected controls (Gossner et al., 2013), and Texel animals in comparison to Suffolk (Ahmed, 2013). While the rest of the DE genes 14 dpi are "novel protein coding" genes in Ensembl (Ensembl Release 74), several are orthologous to *Bos taurus* major histocompatibility complex genes. ENSOARG00000016098 and ENSOARG00000015866 are 1-to-1 orthologs of *BOLA-DRB3* and *BOLA-DQB* respectively. ENSOARG00000002985 is a one-to-many ortholog of *BOLA-DQA1* and ENSOARG00000010572 is a many-to-many ortholog of *BOLA*.

Many of the genes upregulated at 7 dpi in 2011 are involved in lipid metabolism. Lipoprotein lipase (*LPL*) is a key enzyme of lipid metabolism that acts to hydrolyse triglycerides, providing free fatty acids for cells and affecting the maturation of circulating lipoproteins (Wion et al., 1987). Apolipoprotein D (*APOD*) is a protein component of high density lipoprotein (HDL) (Fielding and Fielding, 1980). Perilipin 1 (*PLIN1*) coats lipid droplets in adipocytes, the fat-storing cells in adipose tissue, and is thus an important regulator of lipid storage (Greenberg et al., 1991). The adipose-tissue specific protein adiponectin (*ADIPOQ*) modulates a number of metabolic processes, including glucose regulation and fatty acid oxidation, and has been shown to have anti-inflammatory properties (Diez and Iglesias, 2003). The protein encoded by cell death-inducing DFFA-like effector C (*CIDEC*) promotes lipid droplet formation in adipocytes (Puri et al., 2007; Keller et al., 2008). A reduction in *ADIPOQ* expression is associated with obesity and insulin resistance in some animal models (reviewed in Diez and Iglesias, 2003), whereas expression of *CIDEC* is positively correlated with insulin sensitivity (Ito et al., 2010). Expression of both of these genes indicates there is an increase in the sensitivity of cells in the abomasal lymph node of LowFEC animals to insulin at 7 dpi compared to 14 dpi. Additional genes upregulated at 7 dpi include *CCL26*, *LTBP1* and *ISG17*. The chemokine (C-C motif) ligand 26 (*CCL26*) is expressed by several cell types, including endothelial cells that have been stimulated with

IL-4 (Shinkai et al., 1999). CCL26 binds to the cell surface chemokine receptor CCR3, and is chemotactic for both eosinophils and basophils (Kitaura et al., 1999), and has previously been reported to be upregulated in the abomasal lymph node of resistant Scottish Blackface lambs compared to control animals (Gossner et al., 2013). Latent transforming growth factor beta binding protein 1 (*LTBP1*) belongs to the family of latent TGF-beta binding proteins (LTBPs). LTBP1 forms a complex with latent TGF-β and targets it to the extracellular matrix, where the latent cytokine is subsequently activated (Miyazono et al., 1988); TGF-β plays an important role in controlling the immune system (Letterio and Roberts, 1998). The interferon-stimulated gene 17 (*ISG17*) shares an 81% identity to human ISG15 (Nighswonger et al., 2000), which is secreted from monocytes in response to type I interferons, leading to NK proliferation and augmentation of non-MHC-restricted cytotoxicity (Meraro et al., 2002). When the data from both 2010 and 2011 were analysed together, *ISG17* was still more highly expressed in LowFEC animals at 7 dpi compared to 14 dpi.

Secerin 1 (*SCRN1*) was upregulated at 14 dpi relative to 7 dpi. Secerin 1 regulates exocytosis in mast cells (Way et al., 2002), and is frequently expressed in gastric cancer and colon cancer tissues (Suda et al., 2006). *SCRN1* has previously been reported to be upregulated both in resistant compared to control animals (Pemberton et al., 2011) and Suffolk compared with Texel animals at 7 and 14 days post infection (Ahmed, 2013).

Pathway analysis

There were no pathways in common between LowFEC animals in 2010 and 2011. In 2010 the top diseases and biological functions related to the inflammatory response, cell-to-cell signalling and interaction, and immune cell trafficking. In 2011 this was replaced by dermatological diseases and conditions, lipid metabolism and organismal development.

3.4.2.3 Comparison analysis from 7 to 14 days post infection in all animals

Of the DE genes (common dispersion) in common over the four comparisons, fold changes were consistent within year rather than phenotype. When comparing the 7 vs 14 day post infection IPA analyses of both phenotypes across the two years, it appears that the 2010 HighFEC and 2011 LowFEC animals are the most similar, with 25 genes uniquely in common between the two groups. It is important to remember that despite the genes being common between the two groups the fold changes can be in opposite directions; of the 52 total genes that are DE in both groups, only 22 are in the same direction.

When comparing all HighFEC and LowFEC animals, we hypothesised that the HighFEC animals had a delayed immune response in comparison to their LowFEC counterparts. The similarities between the 2010 HighFEC and 2011 LowFEC animals could be a result of the reduced larval challenge given to the 2011 animals; the LowFEC animals in 2011 may have mounted a delayed immune response compared to the 2010 animals as they received a

much lower challenge. This is backed up by upstream analysis, where the 2010 (LowFEC and HighFEC) and the 2011 LowFEC animals appear similar. In comparison, the 2011 HighFEC animals have a different set of upstream regulators.

3.4.3 Response to high and low *T. circumcincta* burden in Scottish Blackface lambs

The number of genes differentially expressed between lambs with a high worm burden (2010) and those with a low worm burden (2011) was large. Significant diseases and biological functions in the comparison between 2010 and 2011 animals included *Infectious Disease, Gastrointestinal Disease*, *Organismal Injury and Abnormalities* and *Haematological System Development*. The top canonical pathway was *Protein Ubiquitination Pathway*. The genes involved in this pathway included for heat shock proteins (expressed in both years), proteosome genes (expressed in 2010) and ubiquitin genes (expressed in both years).

The primary function of the proteasome (PA700/20S) is the degradation of abnormal or foreign proteins. This involves firstly the conjugation of multiple ubiquitin moieties (Ub) to the target protein, then secondly degradation of the polyubiquitinated protein by the 26S proteasome complex (Lecker et al., 2006). Induction of the immunoproteasome (PA28/20S) by IFN-γ has been associated classically with the processing of MHC class I peptides. However, recent research has uncovered unanticipated functions for IP in innate immunity and non-immune processes, including the regulation of protein homeostasis, cell proliferation, and cytokine gene expression (Ebstein et al., 2012). Accelerated protein turnover helps to prevent the accumulation of harmful protein aggregates during inflammation (van Deventer and Neefjes, 2010). The increased expression of proteasome genes in 2010 compared to 2011 therefore fits with the lower larval challenge in 2011.

3.4.4 Conclusions

The number of genes differentially expressed for each comparison was reasonably low when compared to other transcriptome studies of the same tissue (Pemberton et al., 2011; Ahmed, 2013; Gossner et al., 2013). This could be a result of the within-breed basis of the comparison. However, results from a study by Pemberton et al. (2011) (and subsequently Gossner et al., 2013) on the variation in the abomasal lymph node transcriptome in related Scottish Blackface animals, following a trickle infection with *T. circumcincta*, showed that 146 genes were found to be differentially expressed between resistant and susceptible animals. An alternative suggestion is that there may be variation between animals in how they manifest resistance. This between-animal variation within group is shown by the increase in the number of significant genes found using the less stringent common dispersion estimate in EdgeR. Additionally this study is hampered by lack of statistical

power. While the plan had been to combine the analysis of 2010 and 2011, the difference in worm burdens at slaughter and clear separation of the two groups on the MDS plot by year meant that data from the two years could not be combined.

Despite this, some significantly differentially expressed genes were found. Many of these genes were found to be in agreement with other studies of resistance to gastrointestinal nematodes in sheep. The results were most similar to those of Ahmed (2013), Gossner et al. (2013) and Nagaraj et al. (2012). Ahmed (2013) and Gossner et al. (2013) used transcriptome analysis on the abomasal lymph node to quantify the response of Scottish Blackface, and Suffolk and Texel, respectively, to *T. circumcincta*. As the same species of gastrointestinal nematode was investigated, the similarities in results between the three studies are not surprising. Nagaraj et al. (2012) looked at the proteome of the abomasal mucosa following infection with *H. contortus*. Despite looking at a different GIN species, there were still multiple genes in common with the present. This suggests that there may be some pathways in common to GIN resistance over multiple breeds of sheep and species of gastrointestinal nematode, which is in agreement with the systems genetics study reported by Sayre and Harris (2012).

Additionally, gene expression was examined in the abomasal lymph node, the site from which the immune response is generated, rather than the abomasal mucosa, the site of interaction. It was shown in a study of Johne's disease in calves that cytokine levels were generally higher in the mesenteric lymph nodes than in the intestine, and that the cytokines expressed had different profiles depending on the type of tissue examined and the time of sampling following infection (Wu et al., 2007).

The differentially expressed genes, along with the results from the pathway analysis, indicate that it appears to be the differential interplay between T_H1/T_H2 genes that controls the response to gastrointestinal nematodes in resistant compared to susceptible animals. This is in agreement with previous studies using Texel and Suffolk animals (Hassan et al., 2011b; Ahmed, 2013). The LowFEC (resistant) Scottish Blackface lambs are generating an immune response to *T. circumcincta* at 7 days post infection, whereas in their HighFEC counterparts this response is delayed until ~14 days post infection. The immune response generated by the LowFEC animals may therefore be influencing the larval stages of *T. circumcincta*, which could result in shorter, less fecund adults, as reported in Chapter 2.

Chapter 4

Detecting selective pressure variation in the sheep genome

4.1 Introduction

Domestication of livestock occurred during the Neolithic revolution, a defining step in human history where the transition was made from the hunter-gatherer lifestyle to settled farming communities. Archaeozoological evidence suggests that the domestication of sheep (*Ovis aries*) occurred in the Near East approximately 8,000– 9,000 years ago (Legge, 1996), second in history only to dogs (Vilà et al., 1997). The presence of multiple mitochondrial lineages in the extant sheep species suggest that domestication occurred several times, in the same way as for other livestock species such as cattle, goat and pig (Meadows et al., 2007). Since domestication, sheep have been established in a wide geographic range due to their adaptability to withstand nutrient poor diets and tolerate extreme climatic conditions (Kijas et al., 2009). In addition their modest size has aided intensive husbandry, which has resulted in diverse populations (Figure 4.1), with over 1,400 recorded breeds (Scherf, 2000).

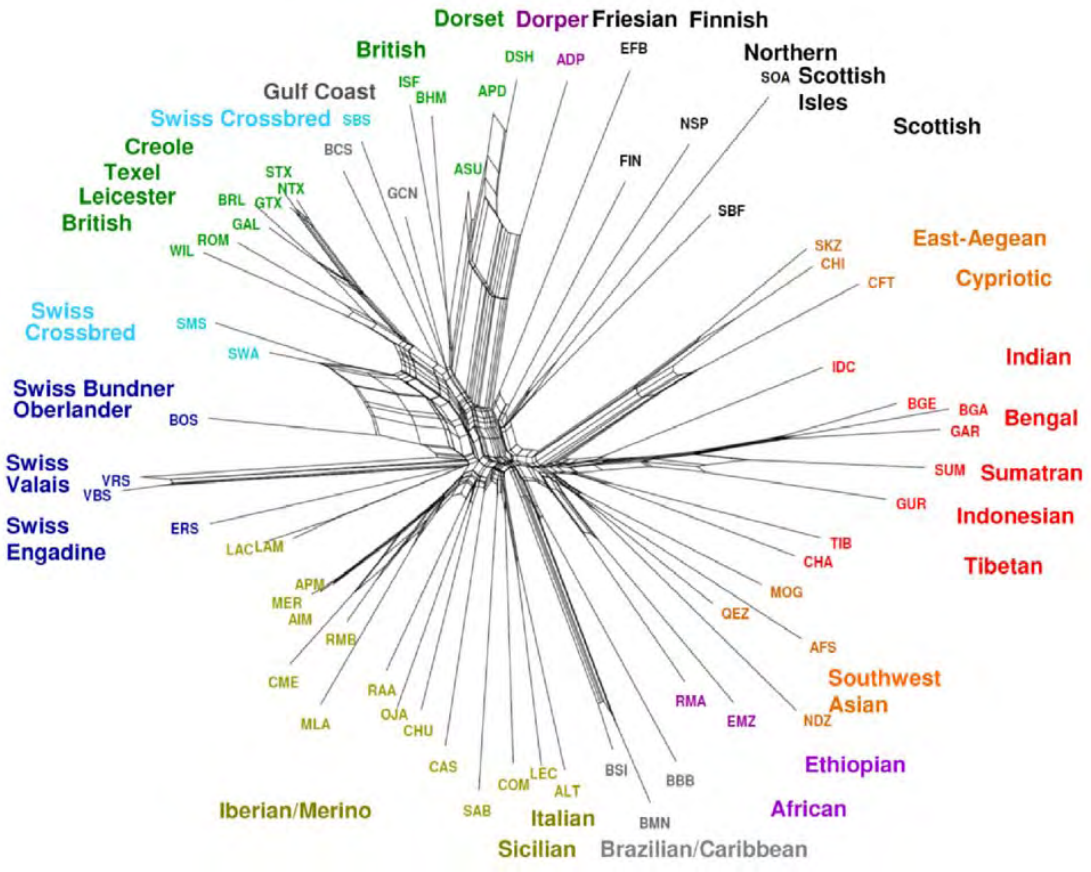


Figure 4.1: Relationship between sheep breeds based on divergence time. The divergence time between breeds (in generations) estimated using LD was used to draw a NeighborNet graph. Figure used with kind permission from Kijas et al. (2012).

Since domestication, sheep have been subject to increased exposure to infectious diseases such as nematodes, due to increased population density and proximity to other domesticated ruminants such as goats and cattle. Indeed, of all the infections of sheep, gastrointestinal

nematode parasitism is considered the most ubiquitous and the most direct threat to their health (Sutherland and Scott, 2009). The selective pressure placed on natural populations by pathogenic infection can have measurable and detectable impact on the genome of a species (Nielsen et al., 2005; Akey, 2009).

As a result of regular challenge by pathogens, positive selection can occur in key immune genes; over time this can result in the appearance of alleles conferring a selective advantage (Vallender and Lahn, 2004). The interplay, sometimes referred to as a 'molecular arms race' between host and pathogen, is caused by the selective pressure placed on the host by the pathogen for continued recognition and clearance in the face of continued pressure on the pathogen to evade detection (Van Valen, 1973).

Positive selection can either be directional or non-directional. The MHC multigene family for example is known to be under non-directional positive selection (Hughes and Nei, 1988), where the fitness advantage may be gained by constantly changing. Directional positive selection on the other hand occurs where a change in an amino acid is beneficial and is subsequently fixed in the population or retained at high frequency. This has been illustrated in various systems previously (Obbard et al., 2009; Levi-Acobas et al., 2009; Little et al., 2007; Flajnik and Kasahara, 2001). Searching the sheep genome for such molecular signatures of positive selection can help to identify key genes, amino acid positions and pathways important in resistance to gastrointestinal nematodes (Nielsen et al., 2005).

Evolutionary theory

Once a new mutation arises in a gene, it is subject to a number of evolutionary forces. The neutral theory of molecular evolution argues that the majority of mutations are neutral or nearly neutral and the rate of fixation of these mutations is based on the rate of random genetic drift and the effective population size (Ohta, 1973; Kimura, 1984). Therefore, it is assumed that only a small minority of mutations face selection, and these will have prevailed over genetic drift by natural or artificial selection (Lynch and Conery, 2000).

While the neutral/nearly-neutral theory applies to molecular evolution, natural or artificial selection will act on the molecular changes governing the phenotype of an organism. If a mutant arises that causes a corresponding protein change, this can be considered to have changed the phenotype of that individual. If the phenotypic effect of the mutation is too small to affect the individual's fitness, it will not be acted upon by natural selection, and is selectively neutral. The neutral and nearly neutral theories state that the mutant may however still be fixed at random (Ohta, 1973; Kimura, 1984). If the protein change sufficiently affects the fitness of the individual, natural selection will then act to select for or against the phenotype (Loughran et al., 2008, 2012). The primary difference between the neutral and the nearly neutral theories is that the nearly neutral theory predicts a relationship between evolutionary rate and species population size: in larger populations genetic drift

is a weaker force, and therefore evolution occurs more slowly than in smaller populations (Ohta, 1973; Lynch and Conery, 2000).

Mutations that impact the function of a gene, or the phenotype of an individual are significantly affected by natural selection (Hurst, 2009). Natural selection can be subdivided into three different types: (i) balancing selection, which favours diversity (such as the non-directional positive selection seen in the MHC region (Hughes and Nei, 1988)), (ii) purifying selection, which eliminates deleterious mutations (e.g. widespread constraint acting on synonymous sites in Drosophila (Lawrie et al., 2013)), and, less commonly, (iii) positive selection or adaptive evolution, where an advantageous allele spreads to fixation in the population (for example the positive selection observed in mammalian reproductive proteins (Morgan et al., 2010)).

Gene duplication

Homologous genes are defined as genes inherited from a common ancestor, and can be further classified as either the product of a speciation event (ortholog), or the result of a gene duplication event (paralog) (Figure 4.2).

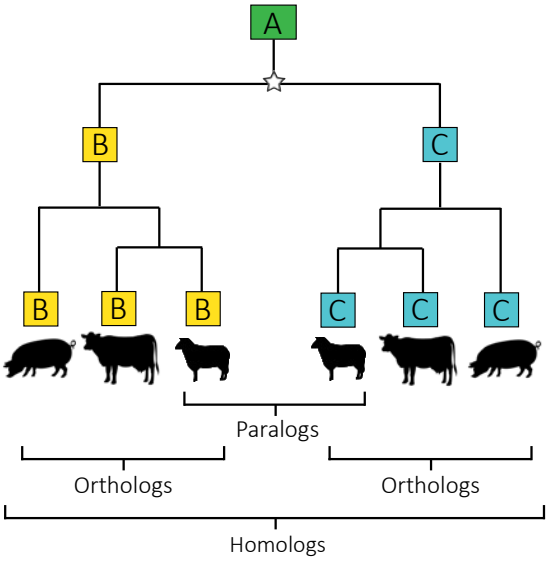


Figure 4.2: Gene phylogeny depicting the relationships among homologs. Genes resulting from a speciation event are orthologs (e.g. 'B' or 'C' genes in all species), while genes resulting from a duplication event (star symbol) are paralogs (e.g. sheep 'B' and 'C' genes).

Gene duplication, the mechanism by which paralogs arise, is believed to be the main source of new genes (Ohno, 1970), although new genes may also arise through recombination (Long et al., 2003). Gene duplication (Figure 4.3) can result in either (i) functional redundancy (pseudogene), (ii) neofunctionalisation, where one duplicate evolves a new function, while the other copy retains the original function, or (iii) subfunctionalisation, the ancestral function is partitioned between the two duplicates (summarised in Long et al.,

2003). Duplication events can lead to the evolution of families of multiple genes with similar functions, such as the vertebrate MHC family (The MHC Sequencing Consortium, 1999).

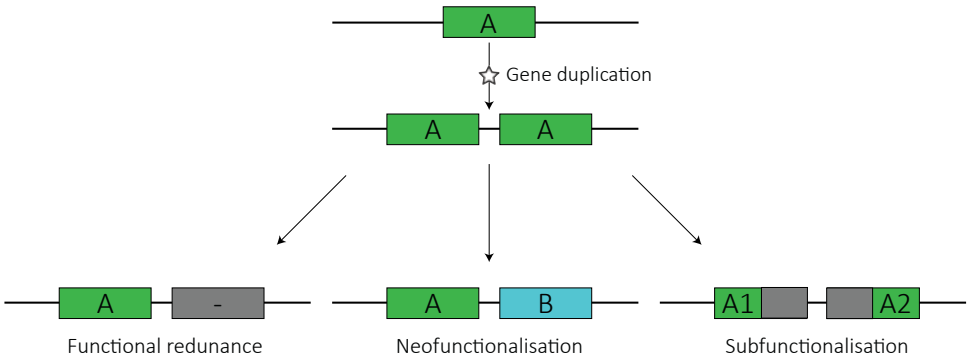


Figure 4.3: Potential outcomes of a gene duplication event (star symbol). Gene duplication can result in either (i) functional redundancy (pseudogene), (ii) neofunctionalisation, where one duplicate evolves a new function, while the other copy retains the original function, or (iii) subfunctionalisation, the ancestral function is partitioned between the two duplicates

Positive selection and protein functional shift

While positive selection is generally considered to be synonymous with a change in protein function (functional shift), empirical evidence was slow to emerge, highlighting the need for experimental validation of predicted positively selected sites. Yokoyama et al. (2008) identified sites under positive selective pressure in vertebrate rhodopsin genes, however repeated directed site mutagenesis of these sites revealed no significant influence on the adaptation of rhodopsin sensitivity. More recent studies, employing maximum likelihood methods of phylogenetic reconstruction and codon based models of evolution (such as those employed in this chapter), have shown definitive links between positive selection predicted *in silico* and protein functional shift, validated *in vitro* (Loughran et al., 2012; Levasseur et al., 2006).

Methods for detecting selective pressure variation

By comparing orthologous genes between species, it is possible to determine the mode of selection acting upon a particular species or gene. For a particular protein coding sequence, the selective pressure can be estimated by comparing synonymous (unchanged amino acid sequence) substitutions per synonymous site (Ds) to non-synonymous (resulting in an amino acid change) substitutions per non-synonymous site (Dn; Figure 4.4). Most synonymous mutations occur at the third position of a codon, whereas mutations in the first two codons are typically non-synonymous (Erives, 2011).

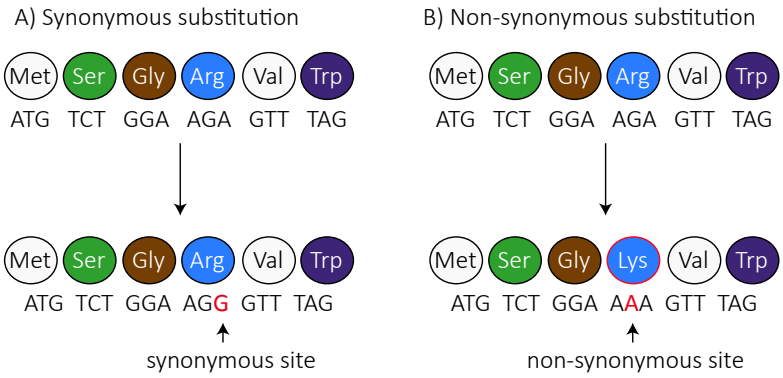


Figure 4.4: Consequences of mutations in the exon of a gene on the resulting amino acid sequence. Diagrammatic representation of a synonymous substitution at a synonymous site (A) and a non-synonymous substitution at a non-synonymous site, resulting in a protein change (B; Arg > Lys).

The ratio between the two numbers (Dn/Ds = ω) can be used to identify the mode of selection: (i) if $\omega = 1$, this is indicative of neutral evolution, (ii) if $\omega < 1$, this is indicative of negative or purifying selection, (ii) if $\omega > 1$, this is indicative of positive selection (Yang and Bielawski, 2000).

The first methods for detecting positive selection were distance-based, assessing Dn/Ds across a gene in a pairwise comparison (Li et al., 1985; Nei and Gojobori, 1986). These methods were incapable of measuring lineage-specific selective pressure variation, whereas subsequent methods incorporating phylogenetic trees can place selection in a phylogenetic context (Yang, 1997, 2007). The McDonald-Kreitman test (McDonald and Kreitman, 1991) uses population level data together with a closely related outgroup species to take into account polymorphisms and divergence within and between species at synonymous and non-synonymous sites. Signatures of adaptive evolution occur where there are more fixed mutations than polymorphisms observed between species than within the study population (McDonald and Kreitman, 1991).

A more sophisticated approach to detect adaptive evolution uses a maximum-likelihood (ML) and Bayesian-based approach, which incorporates the data, the phylogenetic relationship between species, and the complexity of evolutionary (codon) models (Goldman and Yang, 1994; Yang and Nielsen, 2002). CodeML, within the Phylogenetic Analysis by Maximum Likelihood (PAML) package (Yang, 1997, 2007), incorporates a number of models to test for selective pressure variation across branches of a phylogeny (lineage-specific models), across specific codon positions (site-specific models), or both simultaneously (lineage-site models).

For this chapter the aim was to detect evidence of sites that were only positively selected in our lineages of interest (Bovidae, *Ovis*, and HighFEC or LowFEC animals). Two lineage-site models have been developed in CodeML to detect positive selection at individual sites along specified lineages (Yang and Nielsen, 2002; Yang et al., 2005). Model A and model B each work on a phylogeny that has been separated into “foreground” and “background”

lineages, and they permit ω to vary among sites and lineages. Using these nested codon based models of evolution we can assess whether the foreground lineages have signatures of positive selection that are unique in comparison to the background lineages.

The parameters estimated under Model A are summarised in Table 4.1. For the background lineages, there are two classes of sites: conserved sites under purifying selection with $0 < \omega_0 < 1$ and sites undergoing neutral or weak purifying selection with $\omega_1 = 1$. Along the foreground lineages, a proportion $(1 - p_0 - p_1)$ of sites are permitted to evolve under positive selection with $\omega_2 \geq 1$ (Yang, 2006). Model A fixes ω_1, and estimates ω_0 from the data (Yang et al., 2005), whereas in model B the ω values are estimated from the data, and are free to vary (Yang and Nielsen, 2002), thus making Model B the most parameter-rich model.

Table 4.1: The ω ratios permitted in lineage-site model A. Model A involves four parameters: p_0, p_1, ω_0, ω_2, with $\omega_1$fixed at 1. The 3 ω value estimates are across 4 site classes: are 0, 1, 2a and 2b. The ω estimates for site classes 0 and 1 are respectively ω_0 and ω_1, in both foreground and background lineages. For site class 2a, the background lineage is constrained to ω_0, while the foreground lineage is constrained to ω_2. For site class 2b, the background lineage is constrained to ω_1, while the foreground lineage is again constrained to ω_2.

Site class	Proportion	Background ω	Foreground ω	Foreground
0	p_0	$0 < \omega_0 < 1$	$0 < \omega_0 < 1$	Purifying selection
1	p_1	$\omega_1 = 1$	$\omega_1 = 1$	Neutral/weak purifying selection
2a	$(1 - p_0 - p_1)p_0/(p_0 + p_1)$	$0 < \omega_0 < 1$	$\omega_2 > 1$	Positive selection
2b	$(1 - p_0 - p_1)p_1/(p_0 + p_1)$	$\omega_1 = 1$	$\omega_2 > 1$	Positive selection

To prevent the CodeML models reporting results from a local minimum on the likelihood plane, a variety of starting omega values (i.e., 0, 1, 2, 10) are employed, as in previous publications (Loughran, 2010; Morgan et al., 2010). The models are analysed using a likelihood ratio test (LRT; Table 4.2), which assess the significance of the more parameter-rich models compared to their less parameter-rich counterparts (Nielsen and Yang, 1998; Yang and Bielawski, 2000).

Table 4.2: Likelihood ratio tests used in CodeML analysis. The χ 2 critical value must be exceeded by the LRT test statistic D in order for the null model to be rejected at a 5% significance level.

Null model	Alternative model	d.f.	χ 2 critical value
Model 1a	Model A	2	5.99
Model A Null	Model A	2	3.84

When CodeML infers that positive selection has occurred, as confirmed by LRT, Empirical Bayesian (EB) methods are used to estimate the posterior probability (PP) that a given site belongs to the positively selected site category. Two EB methods are employed, Naïve Empirical Bayes (NEB; Yang and Nielsen, 1998) and Bayes Empirical Bayes (BEB; Yang et al., 2005). NEB is particularly sensitive to errors in small datasets where ML estimates

may have large sampling errors (Anisimova et al., 2002), and therefore BEB approaches have been developed to account for these uncertainties, reducing the rate of false positive detection (Yang et al., 2005). The BEB method has not however been implemented for all models of codon evolution.

Detecting selective pressure variation in sheep

The selective pressure caused by gastrointestinal nematodes in ruminants may leave a genetic footprint for directional positive selection, quantifiable by measures of sequence change (as in Morgan et al., 2010). The aim of this study was to take a subset of genes, identified as being differentially expressed in the abomasal lymph node of resistant and susceptible Scottish Blackface lambs after infection with the gastrointestinal nematode *Teladorsagia circumcincta*, and look for selective pressure variation that may be associated with resistance or susceptibility to gastrointestinal nematodes.

4.2 Materials and Methods

4.2.1 Genes of interest

Candidate genes were chosen if they were differentially expressed (tagwise dispersion; FDR<0.1) in Scottish Blackface animals during a controlled challenge with the gastrointestinal nematode *Teladorsagia circumcincta* (Chapter 3). In total 109 genes were differentially expressed (DE); 41 genes between HighFEC and LowFEC animals, and 68 genes between 7 and 14 days post infection. Removing genes that were significant in multiple comparisons resulted in 97 unique genes. Six MHC genes were removed from the analysis, as these genes are highly polymorphic (Stear et al., 2005) and thought to be maintained by balancing selection (Hughes and Yeager, 1998). One additional gene was removed as it was a miRNA (Table 4.3), leaving 90 genes available for positive selection analysis.

Table 4.3: Genes differentially expressed between HighFEC and LowFEC animals removed from analysis.

Ensembl Gene ID	Gene Name	Gene Set	Reason for removal
ENSOARG00000002985	-	RNA-Seq	MHC family
ENSOARG00000010572	-	RNA-Seq	MHC family
ENSOARG00000015866	*HLA-DQB2*	RNA-Seq	MHC family
ENSOARG00000016098	-	RNA-Seq	MHC family
ENSOARG00000000058	-	RNA-Seq	MHC family
ENSOARG00000001701	-	RNA-Seq	MHC family
ENSOARG00000024912	-	RNA-Seq	miRNA

4.2.2 Species of interest

The genomes of 20 species with over 6x genome coverage (Table 4.4) were selected, from which gene family members were obtained through Ensembl Biomart (Smedley et al., 2009). It has been shown that genes with less than 3x sequencing coverage result in higher inferred rates of positive selection than those with more than 3x coverage (Schneider et al., 2009), and selection of high quality genomes reduces the rate of potential false positives.

These species were chosen to represent taxa from a diverse range of clades in order to ensure any selected sites were not a reversal to an ancestral state. Mammalian Orders represented in this study are; the Cetartiodactyla (sheep, cow and pig), the Carnivora (dog and cat), the Perissodactyla (horse), the Chiroptera (microbat), the Primates (gorilla and human), the Proboscidea (elephant), the Didelphimorphia (opossum) and the Monotremata (platypus). The outgroup species were from the Classes Aves, containing the Orders Galliformes (turkey and chicken) and Passeriformes (zebra finch); Reptilia, containing the Order Squamata (anole lizard), and Actinopterygii, containing the orders Cypriniformes

(zebrafish) and Tetraodontiformes (fugu). The species tree used in this study (Figure 4.8) was based on previous work by Morgan et al. (2013) with the addition of sheep, which was placed next to cow (Jiang et al., 2014). Throughout this thesis, Ensembl data has been obtained from Ensembl release 74 unless otherwise specified.

Table 4.4: Full list of species and Ensembl genome assemblies (release 74) used in this analysis. Coverage is approximate.

Common name	Scientific name	Species code	Genome build	Coverage (x)
Anole lizard	*Anolis carolinensis*	ACA	AnoCar2.0	6
Cat	*Felis catus*	FCA	Felis_catus_6.2	6
Chicken	*Gallus gallus*	GAL	Galgal4	6.6
Cow	*Bos taurus*	BTA	UMD3.1	9.5
Dog	*Canis lupus familiaris*	CAF	CanFam3.1	7.6
Elephant	*Loxodonta africana*	LAF	loxAfr3	7
Fugu	*Takifugu rubripes*	TRU	FUGU4	8.7
Gorilla	*Gorilla gorilla gorilla*	GGO	gorGor3.1	35
Horse	*Equus caballus*	ECA	EquCab2	6
Human	*Homo sapiens*	-	GRCh37	High
Microbat	*Myotis lucifugus*	MLU	Myoluc2.0	7
Mouse	*Mus musculus*	MUS	GRCm38	High
Opossum	*Monodelphis domestica*	MOD	BROADO5	7.3
Pig	*Sus scrofa*	SSC	Sscrofa10.2	10
Platypus	*Ornithorhynchus anatinus*	OAN	OANA5	6
Rat	*Rattus norvegicus*	RNA	Rnor_5.0	7
Sheep	*Ovis aries*	OAR	Oar_v3.1	75
Turkey	*Meleagris gallopavo*	MGA	UMD2	30
Zebra Finch	*Taeniopygia guttata*	TGU	taeGut3.2.4	7
Zebrafish	*Danio rerio*	DAR	Zv9	Not available

4.2.3 Analysis of heterogeneous selective pressures

4.2.3.1 Overview of analysis

An overview of the entire analysis is given in Figure 4.5. All Python scripts used in the analysis were coded by Andrew Webb (BME group, DCU), and have been provided (Appendix 4.1) along with instructions for their use (Appendix 4.2).

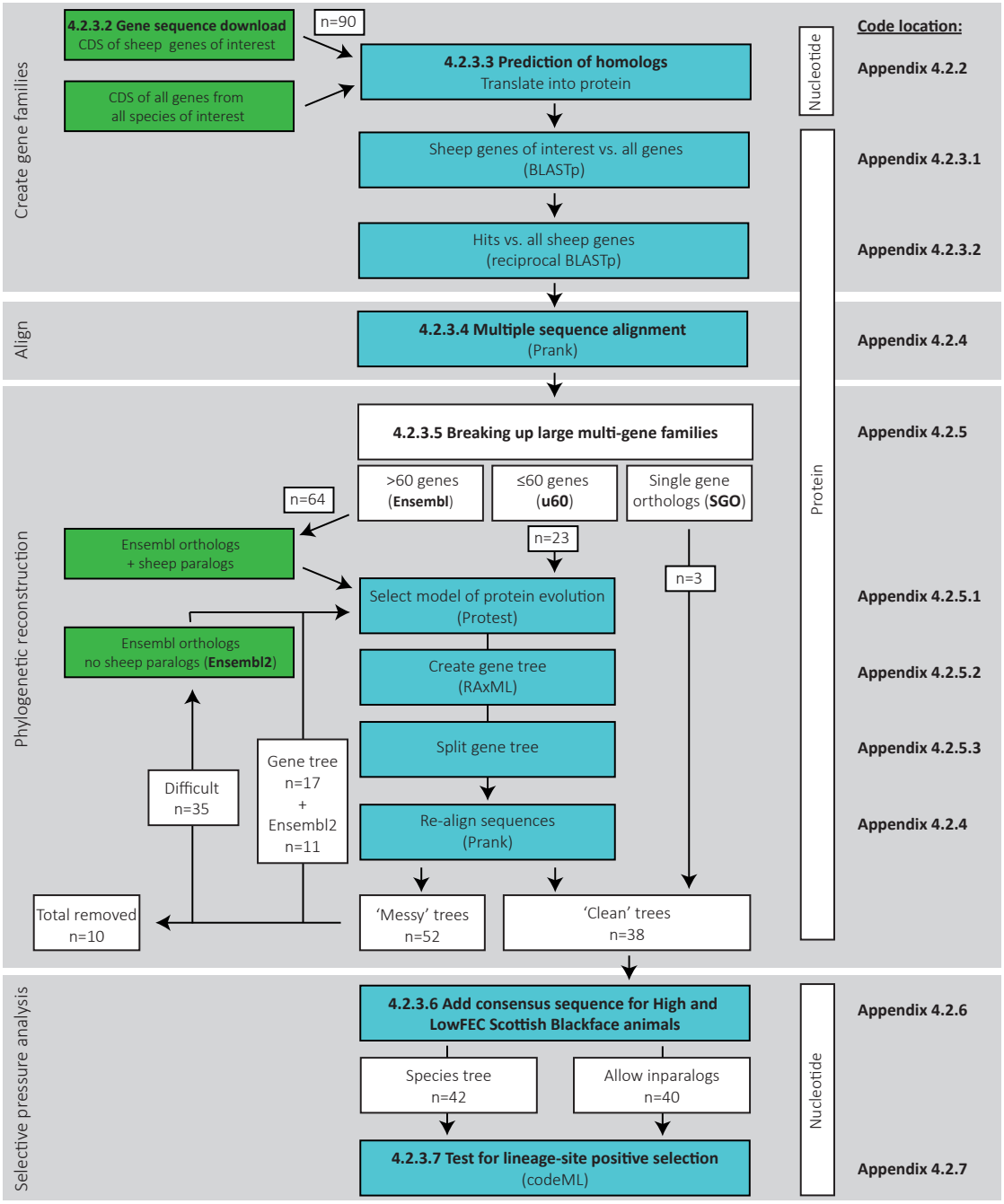


Figure 4.5: Overview of selective pressure heterogeneity analysis. Numbers refer to methodology sections below.

4.2.3.2 Gene sequence download

Protein coding DNA sequences (CDS) for all genes (Ensembl Genes 74; Table 4.4) were downloaded from Ensembl through BioMart (Smedley et al., 2009). Where multiple transcripts were available for each gene, the longest transcript was selected. Any sequences that were not divisible by three were discarded, and then all remaining sequences were translated into their amino acid counterparts (Figure 4.5).

4.2.3.3 Prediction of homologs

A reciprocal BLASTp approach (Altschul et al., 1997) was employed to predict homologs (Figure 4.5). Sheep protein sequences were searched against protein sequences from all other species (Table 4.4) using a cut-off threshold of E^{-11}. Three other thresholds (E^{-5}, $E^{-7}$and E^{-9}) were also tested, but E^{-11} was chosen as it is stricter, and therefore reduced the number of genes per species present in multigene families. A reciprocal BLASTp approach was then used with the results from the initial BLASTp against the *Ovis aries* CDS. The results from both BLASTp runs were concatenated, and used to form multigene families containing putative orthologs.

Multigene families were split into three categories, according to size (Figure 4.5): 'single gene orthologs' (*SGO* group), groups with $\leq$60 genes (*u60* group), and groups with >60 genes. For the larger multigene families (>60 genes) the gene(s) of interest were determined, and known orthologs downloaded using the Ensembl BioMart tool rather than using the reciprocal BLAST results (*Ensembl* group).

4.2.3.4 Multiple sequence alignment

Alignment errors may lead to high false positive numbers in the lineage-site test (Fletcher and Yang, 2010), and therefore it is important to use reliable alignment methods. PRANK v.130820 (Loytynoja and Goldman, 2008) has been demonstrated to outperform other alignment programs in the alignment of mammalian and vertebrate genes (Anisimova et al., 2008; Fletcher and Yang, 2010), and was therefore used for multiple sequence alignment (MSA) of amino acids (Figure 4.5). Multiple sequence alignments have been provided (Appendix 4.3).

4.2.3.5 Reduction of large multigene families

The large multigene families (*u60* and *Ensembl* groups) could not be analysed using selective pressure analysis tools such as CodeML (Yang, 1997; Yang and Nielsen, 1998), and therefore were reduced using the Maximum-likelihood based phylogenetic inference tool RAxML (Stamatakis, 2014) (Figure 4.5).

Selecting model of protein evolution Each alignment file was converted to PHYLIP format using readAl v1.2 (Capella-Gutierrez et al., 2009). ProtTest3 (Darriba et al., 2011) was used to select the model of protein evolution that best fitted each given set of sequences, which was then used to form gene trees using RAxML (Stamatakis, 2014).

Creation of subtrees A rapid Bootstrap analysis in RAxML (Stamatakis, 2014) was used to search for the best-scoring maximum-likelihood (ML) gene tree, which could

be subsequently pruned (Figure 4.5). RAxML was chosen for the highly optimised and efficient likelihood functions, in addition to the low memory consumption. Each tree was pruned manually using the *create_subtrees* function in the bmeTools software package (in preparation for publication), and the genes in the resulting subtree were realigned using PRANK. The subtrees were split into three groups (Figure 4.5), those that could be run through positive selection software with a species tree (*species_tree*), those that could be run with the *-allow_in-paralogs* option during the CodeML setup in bmeTools (*allow_in-paralogs*), and those that needed to be re-run in RAxML (*rerun_RAxML*). A subset of trees (n=36) were also re-run through the RAxML pipeline excluding sheep paralogs (*Ensembl2*). The annotation of the sheep genome is very recent, and therefore some genes have a large number (>20) of paralogs, which makes the gene tree difficult to interpret.

4.2.3.6 Obtaining consensus sequence for HighFEC and LowFEC Scottish Blackface animals

Consensus sequence from the HighFEC (n=20) and LowFEC (n=20) Scottish Blackface animals was obtained from the RNA-Seq reads (Chapter 2). RNA reads for each sample were aligned to the CDS sequence for each gene in the sheep (OAR3.1) genome (Appendix 4.5). Following removal of duplicate reads, the alignments for all within group animals (HighFEC and LowFEC) were merged and a read pileup at each base in the CDS was generated using Samtools (v 0.1.18; Li et al., 2009). An in-house PERL script was used to generate a consensus FASTA sequence for each gene in both the HighFEC and LowFEC animals. Regions of genes for which no coverage was detected were substituted with “N”’s to preserve reading frame integrity of the CDS. The consensus sequence for each group was then added to each gene group MSA using the sheep gene as a reference (Figure 4.5).

4.2.3.7 Testing for lineage-site selective pressure

Selective pressure analyses were performed using CodeML (Yang, 1997; Yang and Nielsen, 1998), from the PAML package (version 4.4c Yang, 2007). Sensitivity to taxa number is a known limitation when detecting positive selection, and therefore more than 6 taxa are required to get statistically robust estimations of selective pressure (Anisimova et al., 2002). Additionally, simulations have also shown that if sequence length is less that 50 codons, prediction of positively selected sites is unreliable (Anisimova et al., 2001). Selective pressure variation was tested in a lineage-site-specific manner across the dataset using model A (Figure 4.5), and a site specific manner using model 8, although these results were not analysed further for this thesis. Model 8 was used as this is the most reliable site specific model (Yang et al., 2000). The multiple sequence alignments of inferred lineage-specific positively selected sites were reviewed, and the Swiss-Prot (www.uniprot.org/) annotated feature examined for possible functional shifts.

It has been shown in humans that polygenic adaptation to pathogens has left signatures in the genome that can be detected using gene set enrichment methods (Daub et al., 2013). Gene set enrichment tests wether the distribution of statistics computed across all genes of a given biological pathway statistically differs from genome-wide expectations (Daub et al., 2013). This approach can be used with SNP data from GWAS studies (Holden et al., 2008; Wang et al., 2007), as well as genes detected as being positively selected using CodeML (Roux et al., 2014). To look at evidence of enrichment at the biological pathway level genes with lineage-specific positive selection were examined using Ingenuity® Systems Pathway Analysis (IPA; Ingenuity Systems, Redwood City, CA, USA; www.ingenuity.com; v18030641). Human Ensembl (version 74) 1-to-1 orthologs for all genes with lineage-specific positive selection were obtained using Ensembl's Biomart tool (www.ensembl.org/biomart/martview/). IPA was used to identify networks of interacting genes and other functional groups.

4.3 Results and discussion

4.3.1 Alignment

In total there were 90 alignments, with an average length of 1,392 aa (113 aa - 10,272 aa; Table 4.5). To determine the quality of the sheep transcripts, the average length of each gene of interest was compared to the background average for each alignment. Of the 90 multiple sequence alignments, 67 were within one standard deviation of the background average, with the remaining 23 outside of this range. The full list of alignments are available in Appendix 4.3.

4.3.2 Gene families

Of the 91 gene families, only 3 were single gene orthologs (*SGOs*). The remaining 88 families were split between i) multigene families with under 60 genes (*u60*, n=23), ii) large groups for which Ensembl-defined orthologs and sheep paralogs were used (*Ensembl*, n=29), and iii) large groups for which Ensembl-defined orthologs only were selected (*Ensembl2*, n=35). During the first round of analysis, one group was removed as there were under seven species represented in the multiple sequence alignments. A further two were removed as there was an ancestral duplication present in the gene tree, which made the MGF too complex for the scope of this analysis (Figure 4.6).

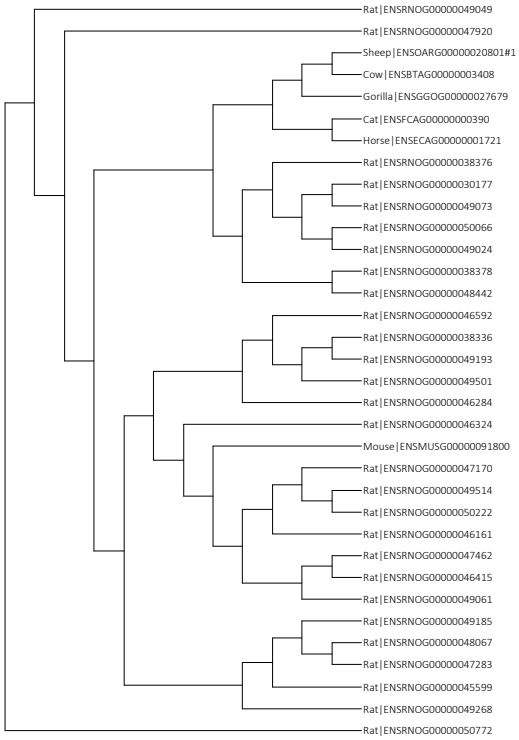


Figure 4.6: An example of a tree removed from the analysis due to ancestral duplication pattern in rat.

In total 28 MSA's were pruned (as demonstrated in Figure 4.7), with the reduced datasets realigned and analysed in ProtTest3 and RAxML (Table 4.5).

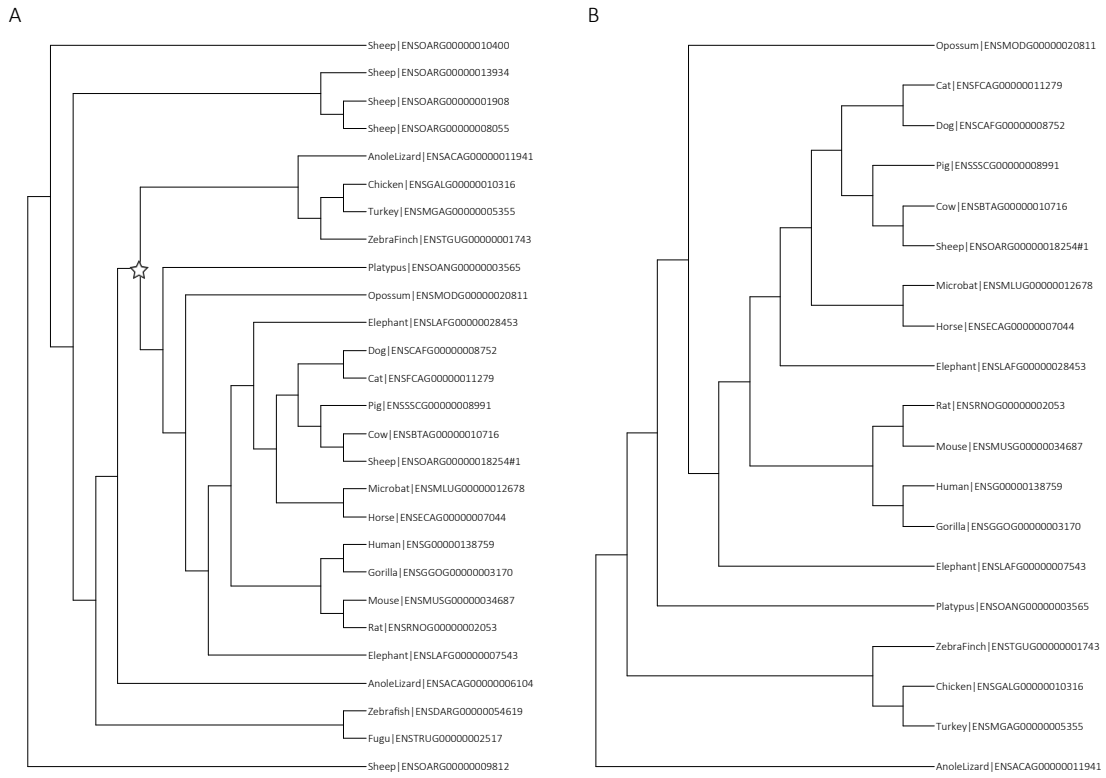


Figure 4.7: Pruning gene trees.Example of gene tree from RAxML (ENSOARG00000018254) before (A) and after (B) pruning (at star symbol).

Table 4.5: Summary of gene families analysed for selective pressure variation. Gene families were split into four groups: 1. SGO: single gene orthologs; 2. u60: multigene families with under 60 genes; 3. Ensembl: large groups for which Ensembl-defined orthologs and sheep paralogs were used; 4. Ensembl2: large groups for which Ensembl-defined orthologs only were selected. Amino acid substitution models ([1]) were obtained from ProtTest3 analysis. These models were used to create gene trees in RAxML. CodeML options allow the use of in-paralogs, or a straight species tree. Alternatively, those marked "Gene tree" were reanalysed in ProtTest3 and RAxML.

Group	Ensembl Gene ID	Name	MSA length (aa)	Species	Model[1]	CodeML option	Model[1]	CodeML option2
1. SGO	ENSOARG00000004198	SCIMP	288	12	-	Species tree	-	-
1. SGO	ENSOARG00000012936	CD83	339	15	-	Species tree	-	-
1. SGO	ENSOARG00000019308	BTLA	346	12	-	Species tree	-	-
2. u60	ENSOARG00000013782	ALB	997	16	JTT+G	Allow in-paralogs	-	-
2. u60	ENSOARG00000001286	-	526	15	JTT+G	Allow in-paralogs	-	-
2. u60	ENSOARG00000017587	PDCD1	668	14	JTT+G	Allow in-paralogs	-	-
2. u60	ENSOARG00000001720	-	1,630	16	JTT+G	Allow in-paralogs	-	-
2. u60	ENSOARG00000004253	-	351	14	JTT+G	Allow in-paralogs	-	-
2. u60	ENSOARG00000002035	WARS	531	20	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
2. u60	ENSOARG00000011072	LYVE1	684	16	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
2. u60	ENSOARG00000007534	LACC1	503	20	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
2. u60	ENSOARG00000001718	MPPE1	432	19	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
2. u60	ENSOARG00000020386	APOD	255	18	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
2. u60	ENSOARG00000003547	MYBL2	2,937	16	JTT+G	Gene tree: rerun RAxML	JTT+G	Species tree
2. u60	ENSOARG00000005011	TTC21B	1,663	16	JTT+G	Gene tree: rerun RAxML	JTT+G	Species tree
2. u60	ENSOARG00000018398	MMRN1	1,922	15	JTT+G	Species tree	-	-
2. u60	ENSOARG00000020373	MFI2	1,827	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000014218	-	368	15	JTT+G	Species tree	-	-
2. u60	ENSOARG00000008738	SCRN1	629	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000007871	SCRN2	629	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000005878	CIDEC	532	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000003661	TOX2	1,388	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000001898	FOLR4	584	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000020194	MAB21L3	762	16	JTT+G	Species tree	-	-
2. u60	ENSOARG00000016543	CXCL9	150	14	JTT+G	Species tree	-	-
2. u60	ENSOARG00000001778	OSBPL5	1,283	16	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000010272	DUSP4	2,320	20	JTT+G	Allow in-paralogs	-	-
3. Ensembl	ENSOARG00000013068	SLC9A4	2,034	18	JTT+G	Allow in-paralogs	-	-
3. Ensembl	ENSOARG00000013498	DNAH5	7,273	19	LG+G	Allow in-paralogs	-	-

Group	Ensembl Gene ID	Name	MSA length (aa)	Species	Model[1]	CodeML option	Model[1]	CodeML option2
3. Ensembl	ENSOARG00000020509	ADIPOQ	1,601	20	WAG+G	Allow in-paralogs	-	-
3. Ensembl	ENSOARG00000010231	MX2	1,510	16	JTT+G	Allow in-paralogs	-	-
3. Ensembl	ENSOARG00000011709	PLIN1	1,497	20	JTT+G	Allow in-paralogs	-	-
3. Ensembl	ENSOARG00000009194	ATP10A	2,447	20	JTT+G	Allow in-paralogs	-	-
3. Ensembl	ENSOARG00000002851	CYP4B1	835	20	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
3. Ensembl	ENSOARG00000015996	STAB2	3,124	20	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
3. Ensembl	ENSOARG00000002342	SH3RF2	2,279	20	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
3. Ensembl	ENSOARG00000012377	GSDMA	664	17	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
3. Ensembl	ENSOARG00000019517	CHI3L2	797	12	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
3. Ensembl	ENSOARG00000010344	LTBP1	10,272	19	JTT+G	Gene tree: rerun RAxML	JTT+G	Removed: too messy
3. Ensembl	ENSOARG00000014842	COL6A5	5,569	15	JTT+G	Gene tree: rerun RAxML	JTT+G	Species tree
3. Ensembl	ENSOARG00000018254	FRAS1	5,215	20	JTT+G	Gene tree: rerun RAxML	JTT+G	Species tree
3. Ensembl	ENSOARG00000010719	LPL	1,039	20	WAG+G	Gene tree: rerun RAxML	JTT+G	Species tree
3. Ensembl	ENSOARG00000020224	COL9A2	3,145	19	JTT+G	Gene tree: rerun RAxML	JTT+G	Species tree
3. Ensembl	ENSOARG00000005151	HMCN1	9,746	19	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000006960	FCRL1	1,434	16	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000000928	ECT2L	1,061	16	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000015815	CNKSR2	1,325	19	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000007090	CYP2F1	750	13	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000012273	PLK5	1,407	15	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000009389	KIF18B	2,786	17	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000019641	LSAMP	631	19	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000005490	SLC30A2	840	20	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000010539	HIF3A	2,394	15	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000007233	ISG17	446	19	JTT+G	Species tree	-	-
3. Ensembl	ENSOARG00000010473	EPB41L3	2,930	20	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000009143	-	1,065	13	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000018232	-	2,010	19	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000018173	-	212	8	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000010111	-	502	19	JTT+G	Allow in-paralogs	-	-

Group	Ensembl Gene ID	Name	MSA length (aa)	Species	Model[1]	CodeML option	Model[1]	CodeML option2
4. Ensembl2	ENSOARG00000004611	MCP-3	285	11	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000005549	-	267	16	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000003000	-	470	5	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000005126	EMR3	925	14	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000013341	-	320	9	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000001140	-	2,542	16	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000004875	-	1,966	12	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000021044	CLEC2B	263	8	JTT+G	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000004858	-	182	17	JTT	Allow in-paralogs	-	-
4. Ensembl2	ENSOARG00000009963	-	233	15	VT+G	Gene tree: rerun RAxML	VT+G	Allow in-paralogs
4. Ensembl2	ENSOARG00000007987	GZMK	277	14	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
4. Ensembl2	ENSOARG00000011275	-	189	17	JTT	Gene tree: rerun RAxML	JTT	Allow in-paralogs
4. Ensembl2	ENSOARG00000002371	-	1,108	13	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
4. Ensembl2	ENSOARG00000001279	-	511	17	JTT+G	Gene tree: rerun RAxML	JTT+G	Allow in-paralogs
4. Ensembl2	ENSOARG00000002418	-	1,220	13	JTT+G	Gene tree: rerun RAxML	JTT	Allow in-paralogs
4. Ensembl2	ENSOARG00000020792	-	257	10	JTT+G	Gene tree: rerun RAxML	JTT+G	Removed: too messy
4. Ensembl2	ENSOARG00000000895	-	903	12	JTT+G	Gene tree: rerun RAxML	JTT+G	Removed: too messy
4. Ensembl2	ENSOARG00000020789	-	270	7	JTT+G	Gene tree: rerun RAxML	JTT+G	Removed: too messy
4. Ensembl2	ENSOARG00000000101	-	295	7	JTT+G	Gene tree: rerun RAxML	JTT+G	Removed: too messy
4. Ensembl2	ENSOARG00000011529	ZFR2	1,502	16	JTT+G	Gene tree: rerun RAxML	JTT+G	Species tree
4. Ensembl2	ENSOARG00000000070	-	2,582	16	JTT+G	Removed: not enough species	-	-
4. Ensembl2	ENSOARG00000020801	-	274	7	JTT+G	Removed: not enough species	-	-
4. Ensembl2	ENSOARG00000000857	-	818	12	WAG+G	Species tree	-	-
4. Ensembl2	ENSOARG00000005312	ZNF461	942	10	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000012211	RPS25	184	18	JTT	Species tree	-	-
4. Ensembl2	ENSOARG00000020811	-	253	7	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000018806	WBP2NL	622	19	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000013846	NAPSA	592	20	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000013340	CCL26	113	11	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000001393	GIMAP8	818	14	JTT+G	Species tree	-	-
4. Ensembl2	ENSOARG00000001131	GIMAP8	702	5	WAG+G	Species tree	-	-

Of these groups, five were removed as the multigene family was too large, with too many independent duplications and losses to infer reliable evolutionary history ('complexity'; Table 4.6). The full list of maximum likelihood trees and the resulting alignment from the selected subtree are provided in Appendix 4.4.

Table 4.6: Genes of interest from multigene families removed during the analysis.

Ensembl gene ID	name	Gene length	Reason for removal
ENSOARG00000001131	GIMAP8	6,459	Not enough species
ENSOARG00000000070	-	878	Ancestral duplication pattern
ENSOARG00000020801	-	2,304	Ancestral duplication pattern
ENSOARG00000010344	LTBP1	458,787	Complexity
ENSOARG00000020792	-	512	Complexity
ENSOARG00000000895	-	7,670	Complexity
ENSOARG00000020789	-	10,646	Complexity
ENSOARG00000000101	-	559	Complexity

4.3.3 Detecting selective pressure variation

Following the various filtering steps, 85 gene families remained for positive selection analysis. The *Ovis* lineage showed the largest number (22) of gene families with lineage-specific positive selection (Figure 4.8 & Table 4.7). Signals of positive selection were detected in the LowFEC lineage in only one of these gene families (OSBPL5). The Bovidae lineage, which included all sheep plus cattle, showed evidence of positive selection in 12 gene families. The full set of selective pressure analysis results are available in Appendix 4.6.

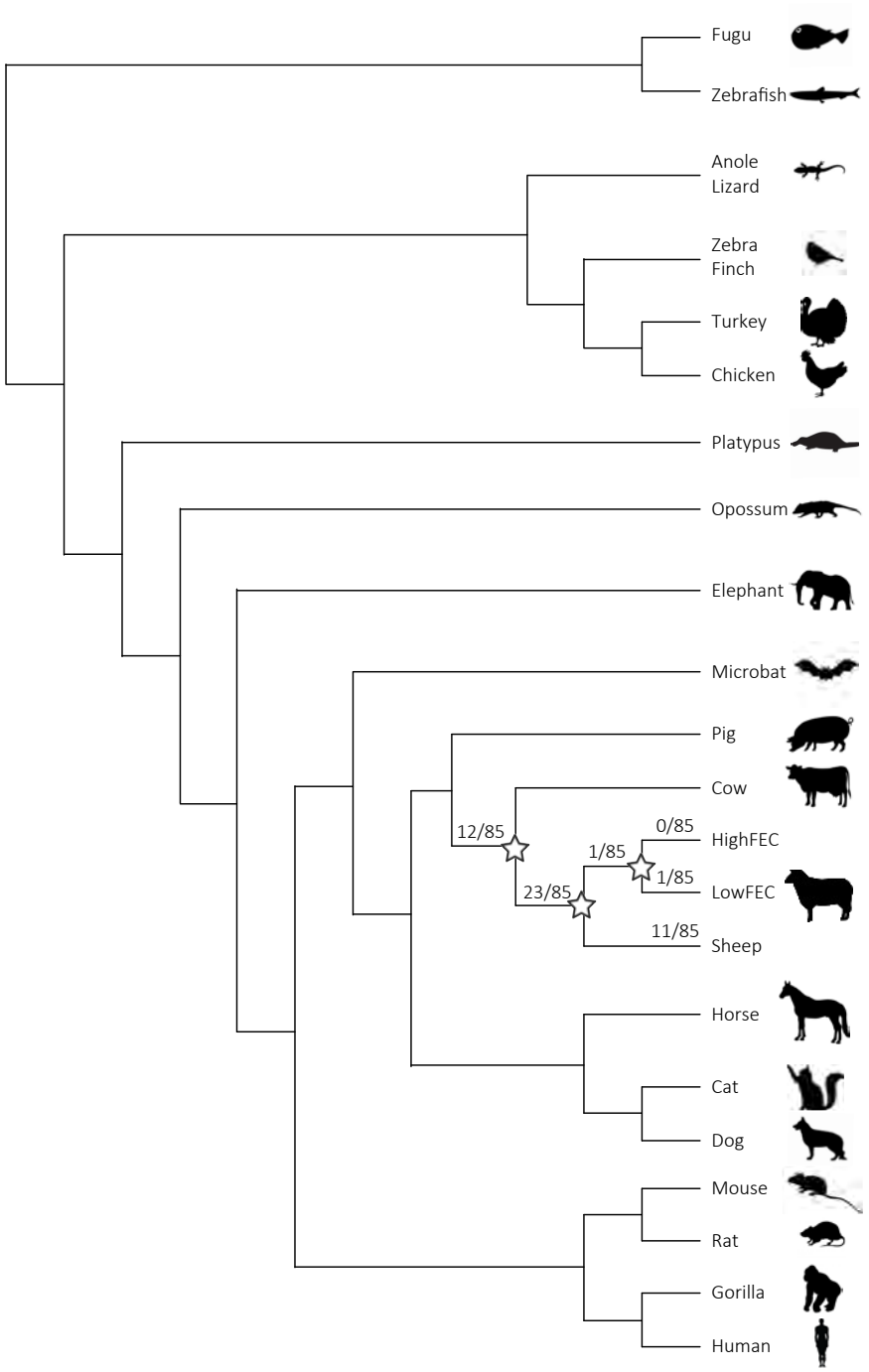


Figure 4.8: Summary of lineage specific positive selection results imposed onto the species phylogeny used in this chapter. Lineage specific positive selection was examined in the Bovidae (all sheep and cattle), *Ovis* (all sheep), sheep reference, Scottish Blackface (HighFEC and LowFEC), HighFEC and LowFEC lineages. In total 85 gene families were examined for selective pressure.

Table 4.7: Results of selective pressure analysis. Results for site-specific selective pressure are shown along with lineage-specific results for HighFEC (H), LowFEC (L), Scottish Blackface (SB), reference sheep (S), *Ovis* (O) and Bovidae (B) lineages.

Ensembl Gene ID	Gene Name	Gene length	Site	H	L	SB	S	O	B
ENSOARG00000004198	*SCIMP*	6,608	-	-	-	-	-	-	-
ENSOARG00000012936	*CD83*	18,537	-	-	-	-	-	+	-
ENSOARG00000019308	*BTLA*	33,061	-	-	-	-	-	-	-
ENSOARG00000013782	*ALB*	22,576	+	-	-	-	-	-	+
ENSOARG00000001286	-	1,505	+	-	-	-	-	+	-
ENSOARG00000017587	*PDCD1*	7,763	-	-	-	-	-	+	-
ENSOARG00000001720	-	2,762	-	-	-	+	-	+	-
ENSOARG00000004253	-	1,421	-	-	-	-	-	-	-
ENSOARG00000002035	*WARS*	20,533	-	-	-	-	-	-	-
ENSOARG00000011072	*LYVE1*	13,901	+	-	-	-	-	-	-
ENSOARG00000007534	*LACC1*	7,803	-	-	-	-	-	-	-
ENSOARG00000001718	*MPPE1*	13,251	-	-	-	-	-	+	-
ENSOARG00000020386	*APOD*	12,496	+	-	-	-	-	-	-
ENSOARG00000003547	*MYBL2*	31,799	+	-	-	-	-	-	-
ENSOARG00000005011	*TTC21B*	93,463	-	-	-	-	-	-	-
ENSOARG00000018398	*MMRN1*	84,917	-	-	-	-	-	-	-
ENSOARG00000020373	*MFI2*	25,778	-	-	-	-	-	-	-
ENSOARG00000014218	-	42,784	-	-	-	-	+	+	-
ENSOARG00000008738	*SCRN1*	42,085	-	-	-	-	-	-	-
ENSOARG00000007871	*SCRN2*	3,179	-	-	-	-	-	+	-
ENSOARG00000005878	*CIDEC*	7,777	-	-	-	-	-	-	-
ENSOARG00000003661	*TOX2*	122,422	-	-	-	-	-	-	-
ENSOARG00000001898	*FOLR4*	4,279	+	-	-	-	-	-	-
ENSOARG00000020194	*MAB21L3*	30,418	+	-	-	-	-	-	-
ENSOARG00000016543	*CXCL9*	4,700	+	-	-	-	-	-	-
ENSOARG00000001778	*OSBPL5*	56,065	-	-	+	-	+	+	+
ENSOARG00000010272	*DUSP4*	13,283	-	-	-	-	-	-	-
ENSOARG00000013068	*SLC9A4*	50,331	-	-	-	-	+	+	-
ENSOARG00000013498	*DNAH5*	263,825	-	-	-	-	+	+	-
ENSOARG00000020509	*ADIPOQ*	12,544	-	-	-	-	-	-	-
ENSOARG00000010231	*MX2*	28,730	+	-	-	-	-	-	-
ENSOARG00000011709	*PLIN1*	11,487	-	-	-	-	-	+	-
ENSOARG00000009194	*ATP10A*	183,350	-	-	-	-	-	-	-
ENSOARG00000002851	*CYP4B1*	23,436	-	-	-	-	-	-	+
ENSOARG00000015996	*STAB2*	169,758	-	-	-	-	+	+	-
ENSOARG00000002342	*SH3RF2*	131,416	-	-	-	-	-	-	-
ENSOARG00000012377	*GSDMA*	13,104	-	-	-	-	-	-	-
ENSOARG00000019517	*CHI3L2*	18,328	-	-	-	-	-	-	-
ENSOARG00000014842	*COL6A5*	164,242	+	-	-	-	-	-	-
ENSOARG00000018254	*FRAS1*	514,386	+	-	-	-	-	+	-
ENSOARG00000010719	*LPL*	23,784	-	-	-	-	-	-	-
ENSOARG00000020224	*COL9A2*	18,385	+	-	-	-	+	-	-

Ensembl Gene ID	Gene Name	Gene length	Site	H	L	SB	S	O	B
ENSOARG00000018254	*FRAS1*	514,386	+	-	-	-	-	+	-
ENSOARG00000010719	*LPL*	23,784	-	-	-	-	-	-	-
ENSOARG00000020224	*COL9A2*	18,385	+	-	-	-	+	-	-
ENSOARG00000005151	*HMCN1*	538,202	-	-	-	-	-	-	-
ENSOARG00000006960	*FCRL1*	15,697	+	-	-	-	-	-	-
ENSOARG00000000928	*ECT2L*	64,564	-	-	-	-	-	-	-
ENSOARG00000015815	*CNKSR2*	222,620	-	-	-	-	-	-	-
ENSOARG00000007090	*CYP2F1*	9,762	-	-	-	-	-	-	-
ENSOARG00000012273	*PLK5*	7,182	+	-	-	-	-	+	+
ENSOARG00000009389	*KIF18B*	15,934	+	-	-	-	-	-	-
ENSOARG00000019641	*LSAMP*	316,557	-	-	-	-	-	-	-
ENSOARG00000005490	*SLC30A2*	10,307	+	-	-	-	+	+	-
ENSOARG00000010539	*HIF3A*	33,107	-	-	-	-	-	+	-
ENSOARG00000007233	*ISG17*	946	-	-	-	-	-	-	+
ENSOARG00000010473	*EPB41L3*	89,355	-	-	-	-	+	+	-
ENSOARG00000009143	-	3,509	+	-	-	-	-	-	-
ENSOARG00000018232	-	26,795	+	-	-	-	-	-	-
ENSOARG00000018173	-	654	+	-	-	-	-	-	-
ENSOARG00000010111	-	905	-	-	-	-	-	-	-
ENSOARG00000004611	*MCP-3*	2,920	+	-	-	-	-	-	-
ENSOARG00000005549	-	4,756	-	-	-	-	-	-	-
ENSOARG00000003000	-	4,250	+	-	-	-	-	-	-
ENSOARG00000005126	*EMR3*	62,795	+	-	-	-	-	-	+
ENSOARG00000013341	-	288,484	+	-	-	-	-	+	+
ENSOARG00000001140	-	175,629	+	-	-	-	-	-	-
ENSOARG00000004875	-	27,125	+	-	-	-	-	-	-
ENSOARG00000021044	*CLEC2B*	15,904	+	-	-	-	-	+	-
ENSOARG00000004858	-	305	-	-	-	-	-	-	-
ENSOARG00000009963	-	6,471	+	-	-	-	-	-	+
ENSOARG00000007987	*GZMK*	10,023	+	-	-	-	-	-	-
ENSOARG00000011275	-	305	-	-	-	-	-	-	-
ENSOARG00000002371	-	10,875	+	-	-	-	+	-	-
ENSOARG00000001279	-	9,277	+	-	-	-	-	-	-
ENSOARG00000002418	-	7,055	+	-	-	-	-	-	-
ENSOARG00000011529	*ZFR2*	22,497	-	-	-	-	-	+	+
ENSOARG00000000857	-	505	-	-	-	-	-	-	-
ENSOARG00000005312	*ZNF461*	20,568	+	-	-	-	-	-	-
ENSOARG00000012211	*RPS25*	1,748	-	-	-	-	-	-	-
ENSOARG00000020811	-	754	-	-	-	-	-	-	+
ENSOARG00000018806	*WBP2NL*	19,300	+	-	-	-	+	+	-
ENSOARG00000013846	*NAPSA*	11,091	-	-	-	-	-	-	+
ENSOARG00000013340	*CCL26*	4,033	-	-	-	-	-	-	-
ENSOARG00000001393	*GIMAP8*	2,960	+	-	-	-	-	+	+

A total of 13 genes showed both site and lineage-site signals of positive selection (Figure 4.9).

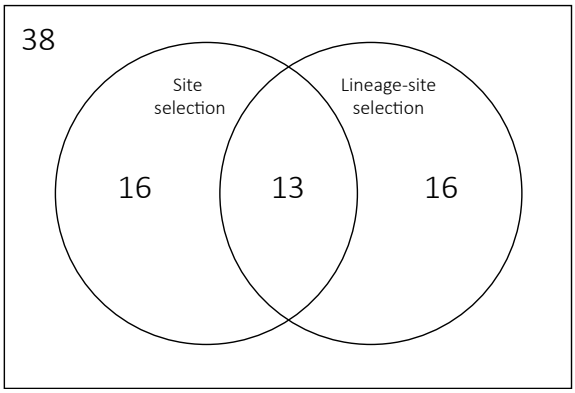


Figure 4.9: Summarisation of results from site and lineage-specific selective pressure analysis. Numbers shown are genes that show site (16) or lineage-site (16) signals of positive selection, both (13), or none (38).

Looking for selective pressure variation in the sheep reference proteins separately to the Scottish Blackface allows analysis of regions where there is no coverage from the HighFEC and LowFEC RNA-Seq reads. One such example is DNAH5 (ENSOARG00000013498), for which 17 sites were positively selected with respect to the *Ovis* lineage. The sheep reference lineage also showed positive selection (Figure 4.10), however upon further examination of the site under positive selection, this was a region for which there was no coverage in the Scottish Blackface.

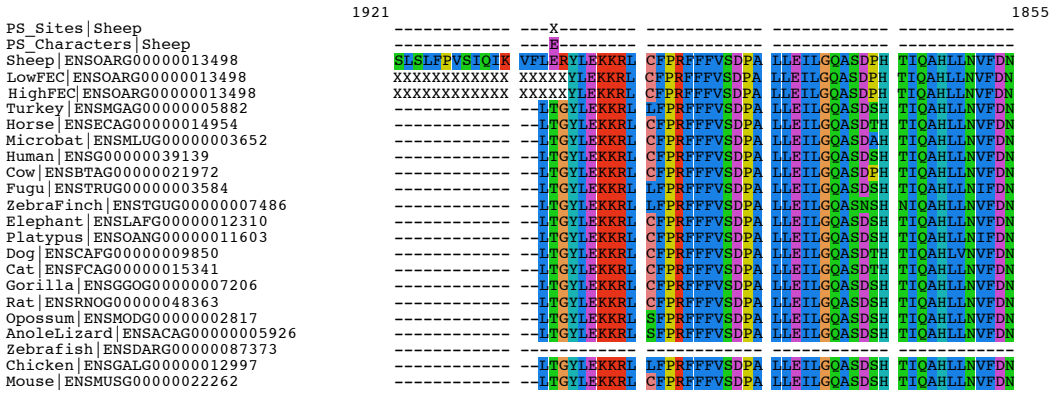


Figure 4.10: Example of positive selection inferred in the reference sheep gene at a site where there is no coverage in either Scottish Blackface group (HighFEC or LowFEC). Positive selection is indicated by an X.

4.3.3.1 Positive selection of genes in the Bovidae lineage

As an example of the 12 gene families inferred to be under positive selection in the Bovidae (cow, sheep and Scottish Blackface) lineage, the CodeML estimates and LRT results for albumin (ALB) are presented below. Five sites were inferred to be under positive selection using BEB (Table 4.8), although examination of the alignments revealed one site (position 482 in the alignment; Figure 4.11) where the amino acid under positive selection was also

present in dog. This site was subsequently discarded as the observed changes may be a result of reversal rather than a unique result. The remaining 4 sites were compared to homologous positions in the human (P02768) and bovine (P02769) Swiss-Prot entries.

The two positively selected (PS) sites at the beginning of the alignment (alignment position 94 and 96; Figure 4.11) are located within the albumin 1 domain, at amino acid 83 and 85 respectively. This region contains a modified residue at position 82, which results in serine being phosphorylated to become phosphoserine (Han et al., 2008). The positively selected sites at amino acid 528 and 530 (in the human form; alignment position 546 and 548; Figure 4.11) are located within the albumin 3 domain. The sheep reference sequence contains a phenylalanine at amino acid 530 (alignment position 548; Figure 4.11), whereas the Scottish Blackface and cattle reference contain a leucine. In the background species the predominant residue is threonine. This may suggest that there are a limited number of alternative amino acids at this location that are acceptable for the protein to remain functional.

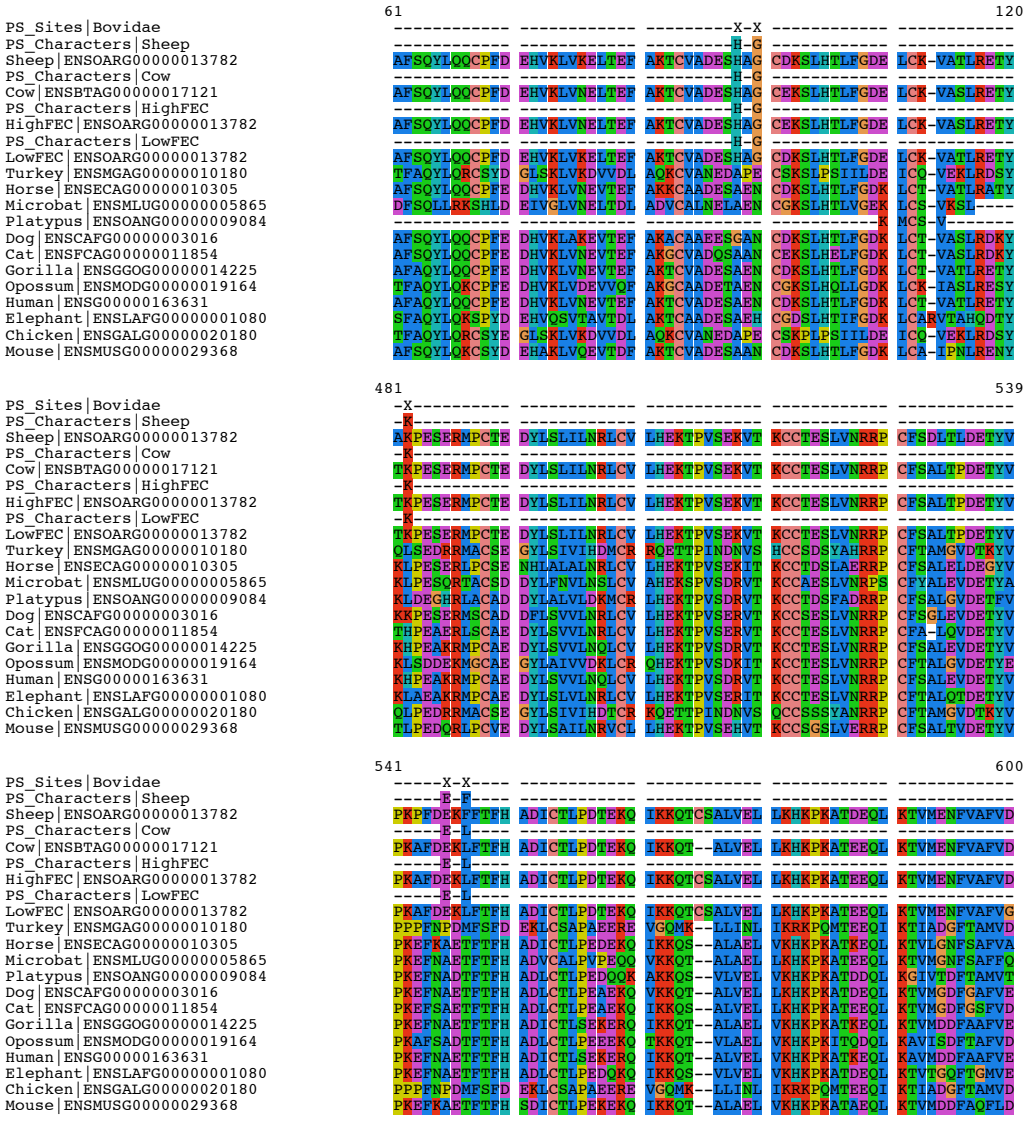


Figure 4.11: Selected regions of ALB multiple sequence alignment. Amino acid sites predicted to be under positive selection in Bovidae (sheep and cattle) by model A are marked with an X.

Table 4.8: Results of lineage-site selective pressure analysis on ALB. The P column shows the number of free parameters in the ω distribution that are estimated under the given model. The ω (t=0) column shows the initial ω value used in the CodeML run from which results were taken. The lnL column shows the log-likelihood of the given model. The LRT Result column shows the result of LRTs (if any) for the given model. The Parameter Estimates column shows the parameter estimates of each given model for the current dataset. The Positive Selection column indicates whether positive selection was predicted under the given model. Finally, the Positively Selected Sites column shows the number of positively selected sites (if any).

Tree	Model	p	ω (t=0)	lnL	LRT Result	Parameter Estimates	Positive Selection	Positively Selected Sites (P(ω>1) > 0.5)
Bovidae	modelA	3	2	-15650.65276	modelA	p0=0.67316 p1=0.31312 p2=0.00936 p3=0.00435 w0=0.17872 w1=1.00000 w2=26.67145	Yes	5 BEB sites
	modelAnull	3	1	-15654.28751	N/A	p0=0.66009 p1=0.30978 p2=0.02051 p3=0.00962 w0=0.17831 w1=1.00000 w2=1.00000	Not Allowed	
Ovis	modelA	3	10	-15653.92144	m1Neutral, modelAnull	p0=0.67603 p1=0.31929 p2=0.00318 p3=0.00150 w0=0.17914 w1=1.00000 w2=99.92513	No	
	modelAnull	3	1	-15654.53717	N/A	p0=0.67816 p1=0.32184 p2=0.00000 p3=0.00000 w0=0.17898 w1=1.00000 w2=1.00000	Not Allowed	
Scottish Blackface	modelA	3	2	-15654.53717	m1Neutral, modelAnull	p0=0.67290 p1=0.31934 p2=0.00527 p3=0.00250 w0=0.17898 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-15654.53718	N/A	p0=0.55530 p1=0.26353 p2=0.12287 p3=0.05831 w0=0.17898 w1=1.00000 w2=1.00000	Not Allowed	
Sheep reference	modelA	3	2	-15654.53717	m1Neutral, modelAnull	p0=0.67816 p1=0.32184 p2=0.00000 p3=0.00000 w0=0.17898 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-15654.53717	N/A	p0=0.67816 p1=0.32184 p2=0.00000 p3=0.00000 w0=0.17898 w1=1.00000 w2=1.00000	Not Allowed	
HighFEC	modelA	3	2	-15654.53717	m1Neutral, modelAnull	p0=0.67704 p1=0.32130 p2=0.00112 p3=0.00053 w0=0.17898 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-15654.53717	N/A	p0=0.67755 p1=0.32154 p2=0.00062 p3=0.00029 w0=0.17898 w1=1.00000 w2=1.00000	Not Allowed	
LowFEC	modelA	3	2	-15654.53717	m1Neutral, modelAnull	p0=0.67816 p1=0.32184 p2=0.00000 p3=0.00000 w0=0.17898 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-15654.53717	N/A	p0=0.67816 p1=0.32184 p2=0.00000 p3=0.00000 w0=0.17898 w1=1.00000 w2=1.00000	Not Allowed	

4.3.3.2 Positive selection of genes in the *Ovis* lineage

The *Ovis* lineage showed the largest number of genes with evidence of positive selection, with 22 gene families showing evidence of positively selected sites (Table 4.7). The two proteins with the most positively selected sites in the *Ovis* lineage were FRAS1 and DNAH5.

There were 20 positively selected sites identified in the extracellular matrix protein FRAS1 using BEB (Table 4.9). All 20 sites were compared to homologous positions in the human Swiss-Prot entry (Q86XX4), and were found to be contained within an extracellular topological domain predicted by UniProt (27-3,901 aa). The positively selected sites were distributed throughout this domain, but no further functional annotation could be gleaned from available data. The Fraser syndrome 1 (*FRAS1*) gene encodes an extracellular matrix protein that appears to function in the regulation of epidermal-basement membrane adhesion and organogenesis during development (McGregor et al., 2003; Short et al., 2007). The FRAS1 protein contains a series of N-terminal cysteine-rich repeat motifs previously implicated in BMP metabolism, and it has been suggested that it may play a role in both structure and signal propagation in the extracellular matrix (McGregor et al., 2003). *In vitro* studies have suggested a role for extracellular matrix proteins in regulating the response to chemokines (Young, 1999).

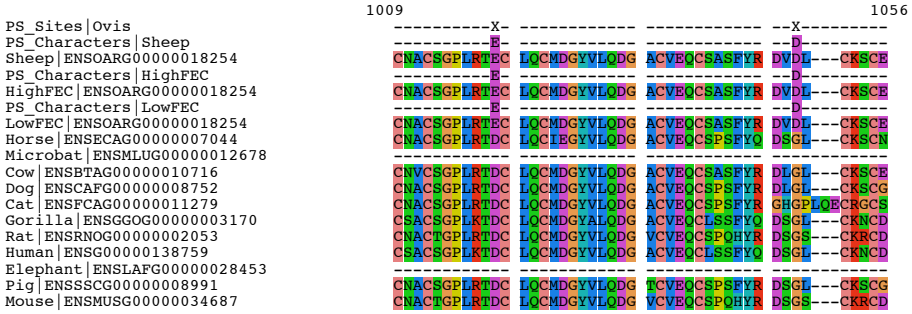


Figure 4.12: Selected region of FRAS1 multiple sequence alignment. Amino acid sites predicted to be under positive selection in *Ovis* (sheep reference and Scottish Blackface) by model A are marked with an X.

Positive selective pressure was identified in 32 BEB sites in DNAH5 (dynein heavy chain 5, axonema). Of these, 18 were confirmed as unique amino acids through visual inspection of protein alignments. These 18 sites were compared to homologous positions in the human Swiss-Prot entry (Q8TE73), and were all found to be contained within the dynein heavy chain 5, axonemal (1-4,624 aa). Dyneins are microtubule-associated motor protein complexes composed of several heavy, light, and intermediate chains, and are required for structural and functional integrity of cilia and flagella (Ibañez-Tallon et al., 2003). Cilia have been shown to play a role in the gastrointestinal tract of mice, where they modulate gastrin secretion and gastric acidity (Saqui-Salces et al., 2012). DNAH5 has been shown to be decreased in the intestine of pigs on a protein restricted diet compared to controls, indicating that DNAH5 may play a role in nutrition absorption (Ren et al., 2014). Expression

Table 4.9: Results of lineage-site selective pressure analysis on FRAS1. The P column shows the number of free parameters in the ω distribution that are estimated under the given model. The ω (t=0) column shows the initial ω value used in the CodeML run from which results were taken. The lnL column shows the log-likelihood of the given model. The LRT Result column shows the result of LRTs (if any) for the given model. The Parameter Estimates column shows the parameter estimates of each given model for the current dataset. The Positive Selection column indicates whether positive selection was predicted under the given model. Finally, the Positively Selected Sites column shows the number of positively selected sites (if any).

Tree	Model	p	ω (t=0)	lnL	LRT Result	Parameter Estimates	Positive Selection	Positively Selected Sites (P(ω>1) > 0.5)
Bovidae	modelA	3	0	-48737.2336	m1Neutral, modelAnull	p0=0.83953 p1=0.16047 p2=0.00000 p3=0.00000 w0=0.07994 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-48737.2336	N/A	p0=0.83953 p1=0.16047 p2=0.00000 p3=0.00000 w0=0.07994 w1=1.00000 w2=1.00000	Not Allowed	
Ovis	modelA	3	2	-48708.44652	modelA	p0=0.83954 p1=0.15725 p2=0.00271 p3=0.00051 w0=0.08035 w1=1.00000 w2=999.00000	Yes	20 BEB sites
	modelAnull	3	1	-48735.97353	N/A	p0=0.81015 p1=0.15433 p2=0.02984 p3=0.00568 w0=0.07927 w1=1.00000 w2=1.00000	Not Allowed	
Scottish Blackface	modelA	3	2	-48737.16423	m1Neutral, modelAnull	p0=0.81926 p1=0.15667 p2=0.02021 p3=0.00386 w0=0.07984 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-48737.16423	N/A	p0=0.81925 p1=0.15667 p2=0.02021 p3=0.00387 w0=0.07984 w1=1.00000 w2=1.00000	Not Allowed	
Sheep reference	modelA	3	2	-48737.2336	m1Neutral, modelAnull	p0=0.83953 p1=0.16047 p2=0.00000 p3=0.00000 w0=0.07994 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-48737.2336	N/A	p0=0.83954 p1=0.16046 p2=0.00000 p3=0.00000 w0=0.07994 w1=1.00000 w2=1.00000	Not Allowed	
HighFEC	modelA	3	2	-48734.86298	m1Neutral, modelAnull	p0=0.74056 p1=0.14119 p2=0.09932 p3=0.01894 w0=0.07969 w1=1.00000 w2=2.81033	No	
	modelAnull	3	1	-48734.89708	N/A	p0=0.55240 p1=0.10529 p2=0.28750 p3=0.05480 w0=0.07970 w1=1.00000 w2=1.00000	Not Allowed	
LowFEC	modelA	3	2	-48734.39968	m1Neutral, modelAnull	p0=0.50520 p1=0.09665 p2=0.33422 p3=0.06394 w0=0.07948 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-48734.39968	N/A	p0=0.50520 p1=0.09665 p2=0.33422 p3=0.06394 w0=0.07948 w1=1.00000 w2=1.00000	Not Allowed	

of DNAH5 was increased in the abomasal lymph node of HighFEC Scottish Blackface lambs in comparison to their LowFEC counterparts (Chapter 3), despite the lack of ciliated epithelial cells in the gastrointestinal tract of ruminants. The DNAH5 paralog DNAH9 had low F_{ST} values in a survey of the genetic diversity of 19 diverse cattle breeds (The Bovine HapMap Consortium et al., 2009), indicating that genes in this family may play an important role in ruminants, despite the lack of evidence for their role in the gastrointestinal system.

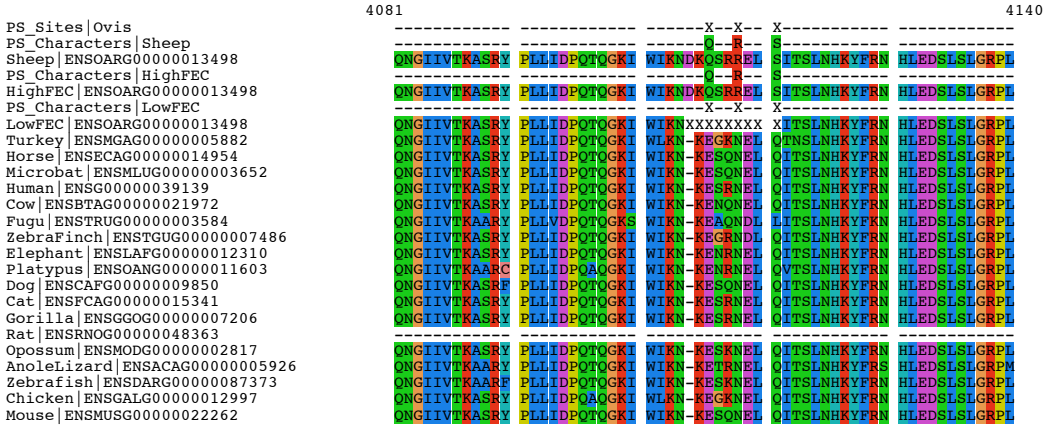


Figure 4.13: Selected region of DNAH5 multiple sequence alignment. Amino acid sites predicted to be under positive selection in *Ovis* (sheep reference and Scottish Blackface) by model A are marked with an X.

4.3.3.3 Putative positive selection in the LowFEC Scottish Blackface lineage

The *Ovis* lineage was predicted to be undergoing strong positive selection at specific sites in OSBPL5, with the Bovidae, sheep reference, and LowFEC Scottish Blackface lineages all showing evidence of positive selection (Table 4.11). Upon further examination of the the two positively selected sites in the LowFEC lineage identified using BEB (Figure 4.14) however, it appears that this may not actually be the case. In the region containing these two sites the sheep alignment does not accurately match those of the other species, including cattle (Figure 4.14). This may be the result of an error in either the assembly or the annotation of the sheep genome, and is one of the consequences of working on a 'new' genome. This result emphasises that the rate of false positives in selective pressure analyses is high, and it is particularly prone to errors in genome/annotation quality and alignment (Fletcher and Yang, 2010; Redelings, 2014).

Table 4.10: Results of lineage-site selective pressure analysis on DNAH5. The P column shows the number of free parameters in the ω distribution that are estimated under the given model. The ω (t=0) column shows the initial ω value used in the CodeML run from which results were taken. The lnL column shows the log-likelihood of the given model. The LRT Result column shows the result of LRTs (if any) for the given model. The Parameter Estimates column shows the parameter estimates of each given model for the current dataset. The Positive Selection column indicates whether positive selection was predicted under the given model. Finally, the Positively Selected Sites column shows the number of positively selected sites (if any).

Tree	Model	p	ω (t=0)	lnL	LRT Result	Parameter Estimates	Positive Selection	Positively Selected Sites (P(ω>1) > 0.5)
Bovidae	modelA	3	0	-104509.6131	m1Neutral, modelAnull	p0=0.89482 p1=0.10518 p2=0.00000 p3=0.00000 w0=0.06834 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-104509.6131	N/A	p0=0.89482 p1=0.10518 p2=0.00000 p3=0.00000 w0=0.06834 w1=1.00000 w2=1.00000	Not Allowed	
Ovis	modelA	3	2	-104481.4119	modelA	p0=0.89328 p1=0.10457 p2=0.00192 p3=0.00022 w0=0.06814 w1=1.00000 w2=546.75144	Yes	32 BEB sites
	modelAnull	3	1	-104498.3404	N/A	p0=0.83385 p1=0.09749 p2=0.06147 p3=0.00719 w0=0.06785 w1=1.00000 w2=1.00000	Not Allowed	
Scottish Blackface	modelA	3	1	-104509.5076	m1Neutral, modelAnull	p0=0.86305 p1=0.10144 p2=0.03177 p3=0.00373 w0=0.06833 w1=1.00000 w2=1.52637	No	
	modelAnull	3	1	-104509.5079	N/A	p0=0.84563 p1=0.09940 p2=0.04919 p3=0.00578 w0=0.06833 w1=1.00000 w2=1.00000	Not Allowed	
Sheep reference	modelA	3	10	-104500.0374	modelA	p0=0.89440 p1=0.10480 p2=0.00072 p3=0.00008 w0=0.06832 w1=1.00000 w2=999.00000	Yes	1 BEB site
	modelAnull	3	1	-104509.6131	N/A	p0=0.89482 p1=0.10518 p2=0.00001 p3=0.00000 w0=0.06834 w1=1.00000 w2=1.00000	Not Allowed	
HighFEC	modelA	3	2	-104508.9427	m1Neutral, modelAnull	p0=0.79176 p1=0.09307 p2=0.10306 p3=0.01211 w0=0.06830 w1=1.00000 w2=1.00000	No	
	modelAnull	3	1	-104508.9427	N/A	p0=0.79177 p1=0.09307 p2=0.10305 p3=0.01211 w0=0.06830 w1=1.00000 w2=1.00000	Not Allowed	
LowFEC	modelA	3	10	-104507.8718	m1Neutral, modelAnull	p0=0.86723 p1=0.10192 p2=0.02761 p3=0.00325 w0=0.06827 w1=1.00000 w2=5.71753	No	
	modelAnull	3	1	-104508.0279	N/A	p0=0.75196 p1=0.08838 p2=0.14287 p3=0.01679 w0=0.06827 w1=1.00000 w2=1.00000	Not Allowed	

Table 4.11: Results of lineage-site selective pressure analysis on OSBPL5. The P column shows the number of free parameters in the ω distribution that are estimated under the given model. The ω (t=0) column shows the initial ω value used in the CodeML run from which results were taken. The lnL column shows the log-likelihood of the given model. The LRT Result column shows the result of LRTs (if any) for the given model. The Parameter Estimates column shows the parameter estimates of each given model for the current dataset. The Positive Selection column indicates whether positive selection was predicted under the given model. Finally, the Positively Selected Sites column shows the number of positively selected sites (if any).

Tree	Model	p	ω (t=0)	lnL	LRT Result	Parameter Estimates	Positive Selection	Positively Selected Sites (P(ω>1) > 0.5)
Bovidae	modelA	3	2	-11909.224667	modelA	p0=0.93557 p1=0.06046 p2=0.00374 p3=0.00024 ω0=0.06534 ω1=1.00000 ω2=68.56723	Yes	7 BEB sites
	modelAnull	3	1	-11912.089275	N/A	p0=0.93454 p1=0.06021 p2=0.00493 p3=0.00032 ω0=0.06550 ω1=1.00000 ω2=1.00000	Not Allowed	
Ovis	modelA	3	1	-11797.067977	modelA	p0=0.82269 p1=0.06463 p2=0.10448 p3=0.00821 ω0=0.05542 ω1=1.00000 w2=741.93558	Yes	64 BEB sites
	modelAnull	3	1	-11854.819238	N/A	p0=0.75231 p1=0.05578 p2=0.17865 p3=0.01325 ω0=0.05446 ω1=1.00000 ω2=1.00000	Not Allowed	
Scottish Blackface	modelA	3	2	-11910.114944	m1Neutral, modelAnull	p0=0.00000 p1=0.00000 p2=0.93958 p3=0.06042 ω0=0.06559 ω1=1.00000 ω2=999.00000	No	
	modelAnull	3	1	-11910.362899	N/A	p0=0.00000 p1=0.00000 p2=0.93955 p3=0.06045 ω0=0.06560 ω1=1.00000 ω2=1.00000	Not Allowed	
Sheep reference	modelA	3	10	-11883.188439	modelA	p0=0.89917 p1=0.06184 p2=0.03648 p3=0.00251 ω0=0.06212 ω1=1.00000 ω2=999.00000	Yes	15 BEB sites
	modelAnull	3	1	-11905.325252	N/A	p0=0.40682 p1=0.02694 p2=0.53107 p3=0.03517 ω0=0.06398 ω1=1.00000 ω2=1.00000	Not Allowed	
HighFEC	modelA	3	0	-11912.341025	m1Neutral, modelAnull	p0=0.93938 p1=0.06061 p2=0.00000 p3=0.00000 ω0=0.06593 ω1=1.00000 ω2=1.00000	No	
	modelAnull	3	1	-11912.342204	N/A	p0=0.93900 p1=0.06059 p2=0.00039 p3=0.00002 ω0=0.06593 ω1=1.00000 ω2=1.00000	Not Allowed	
LowFEC	modelA	3	0	-11907.702012	modelA	p0=0.93467 p1=0.06050 p2=0.00454 p3=0.00029 ω0=0.06555 ω1=1.00000 ω2=999.00000	Yes	2 BEB sites
	modelAnull	3	1	-11909.683135	N/A	p0=0.00002 p1=0.00000 p2=0.93913 p3=0.06085 ω0=0.06556 ω1=1.00000 ω2=1.00000	Not Allowed	

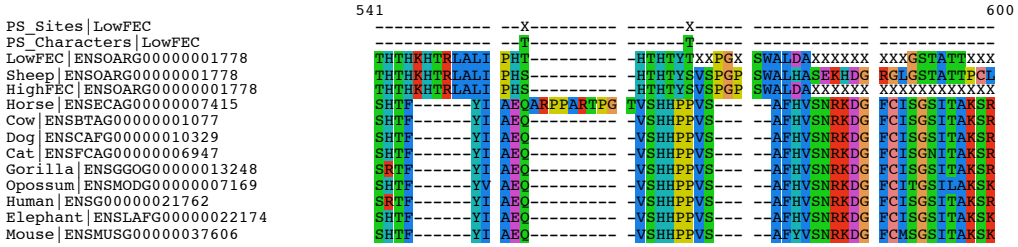


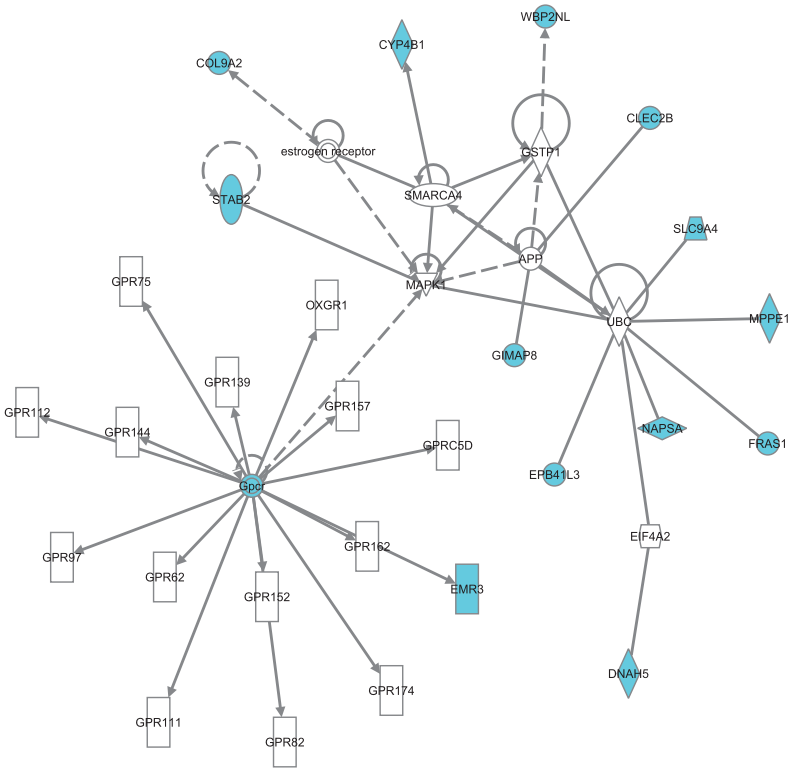
Figure 4.14: Selected region of OSBPL5 multiple sequence alignment. Amino acid sites predicted to be under positive selection in LowFEC Scottish Blackface animals by model A are marked with an X.

4.3.3.4 Pathway analysis of genes with lineage-specific positive selection

Of the 31 genes that showed evidence of lineage-specific positive selection, one to one human homologs could be found for 23 (Table **??**). The top networks from the IPA analysis were "*Cancer, Cell-To-Cell Signalling and Interaction, Nervous System Development and Function*" (score: 33) and "*Gene Expression, Carbohydrate Metabolism, Lipid Metabolism*" (score: 18; Figure 4.15).

A) Network 1

Cancer, Cell-To-Cell Signaling and Interaction, Nervous System Development and Function



B) Network 2

Gene Expression, Carbohydrate Metabolism, Lipid Metabolism

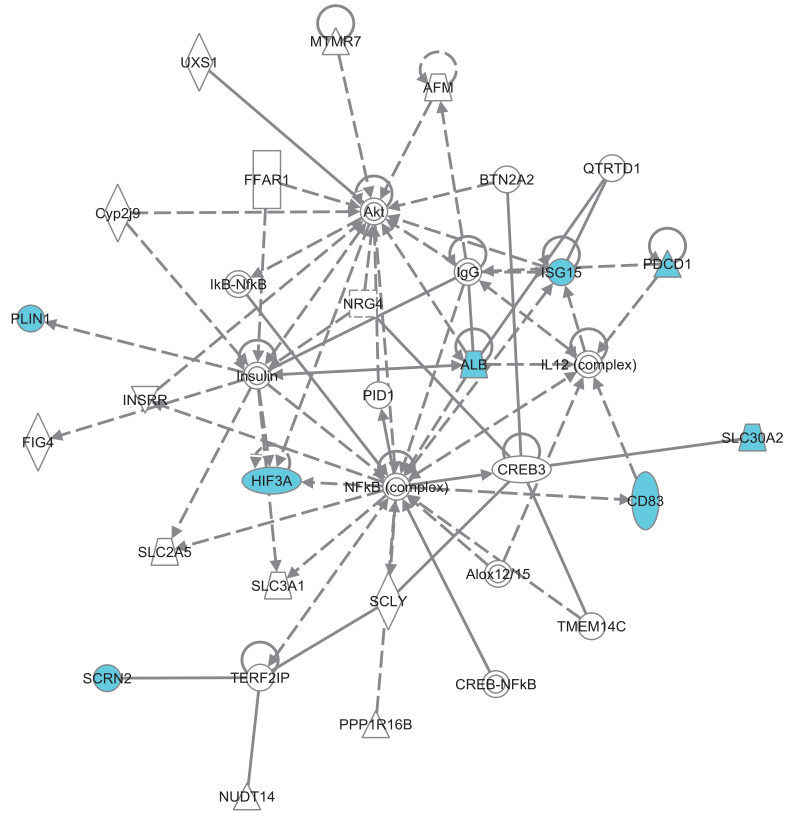


Figure 4.15: Top networks from IPA analysis of genes showing evidence of lineage-specific positive selection. Genes showing lineage-specific positive selection are coloured blue.

The top diseases and disorders associated with the genes showing lineage-specific positive selection included the inflammatory response, dermatological diseases and conditions and connective tissue disorders (Table 4.12). The most significant of these was the inflammatory response, which was associated with 6 of the genes showing positive selection. Inflammation plays an important role in both the innate and adaptive immune response to gastrointestinal parasites, with inflammation and immunity inextricably linked (Sutherland and Scott, 2009).

Table 4.12: Top diseases and bio functions analysis from IPA of genes showing lineage-specific positive selection.

Diseases and disorders	p-value	# Molecules
Inflammatory Response	1.24E-04 - 4.89E-02	6
Cancer	2.17E-04 - 4.60E-02	14
Dermatological Diseases and Conditions	2.17E-04 - 3.08E-02	10
Connective Tissue Disorders	1.20E-03 - 2.04E-02	6
Developmental Disorder	1.20E-03 - 4.74E-02	7

It must be acknowledged that these results are biased as the subset of genes analysed were differentially expressed in sheep with divergent responses to gastrointestinal nematode infection (Chapter 3). Despite this limitation, IPA analysis results suggest that the observed positive selection may be a result of selective pressure placed on sheep and cattle by pathogenic infection.

4.3.3.5 Limitations of the analysis

While care was taken to address known limitations in selective pressure analyses, such as sensitivity to taxa number and short sequence length, there are a number of factors that can result in false positives. Alignment errors may lead to high false positive numbers in the lineage-site test (Fletcher and Yang, 2010), and accurate multiple sequence alignment is essential for detecting positive selection (Redelings, 2014). While PRANK (Loytynoja and Goldman, 2008) has a low false positive rate in comparison to alignments from other popular software such as MUSCLE (Edgar, 2004), MAFFT (Katoh et al., 2002), and ClustalW (Larkin et al., 2007), it is by no means a perfect tool (Fletcher and Yang, 2010).

Evidence of positive selection may therefore not always be indicative of a protein functional shift. Biased gene conversion (Galtier and Duret, 2007), relaxation of functional constraint (Hughes and Friedman, 2004) and recombination events can confound the identification of positively selected amino acid sites, however CodeML is robust to lower levels of recombination (Anisimova et al., 2003). Low effective population size (Ne) may also make it difficult to detect adaptive molecular evolution in some species (Gossmann et al., 2012).

Resistance to gastrointestinal nematodes in sheep is understood to be a complex polygenic trait, controlled by many genes of relatively small effect (Crawford et al., 2006; Kemper

et al., 2009). In humans widespread signals of polygenic selection have been observed in immune-response related pathways (Daub et al., 2013), however selection may be acting on many genes at a time and modifying allele frequencies only slightly, making it difficult to detect signals of adaptation (Pritchard et al., 2010).

4.3.4 Conclusions

Domestication of livestock has resulted in lineage-specific changes in genes involved in the inflammatory response and connective tissue disorders. As these genes have been shown to be differentially expressed in sheep with divergent responses to gastrointestinal nematode infection, the positive selection observed may be a result of selective pressure placed on sheep as a result of GIN infection.

Chapter 5

Associations between polymorphisms in the genome and nematode resistance in Scottish Blackface lambs

5.1 Introduction

Current strategies for the control of gastrointestinal nematodes, based almost entirely on the frequent use of anthelmintics, are increasingly regarded as unsustainable due to the emergence of nematode resistance (Kaplan, 2004). Resistance has been reported to the benzimidazole (BZ) and levamisole (LV) drenches on Irish farms (Patten et al., 2007), leaving only macrocyclic lactones (ML) and the prescription-only monepantel (AD) and the derquantel (SI)/macrocyclic lactones combination Startect to treat drug resistant nematodes. Additionally, there is increasing consumer resistance to the use of systemic chemicals, such as anthelmintics, in food production.

Direct measurement of parasite burden requires slaughter of the animal, therefore indirect phenotypic markers such as faecal egg count (FEC) are routinely used for selection of resistant animals for breeding. FEC has been shown to be moderately heritable, with a wide variability among individuals (Bishop and Morris, 2007; Safari et al., 2005). Therefore selective breeding for resistance is an alternative, sustainable method of nematode control. Rapid genetic progress has been demonstrated in selective breeding programmes in research flocks (Windon, 1990; Cummins et al., 1991; Woolaston et al., 1991; Morris et al., 1997, 2000, 2005; Greeff and Karlsson, 2006), and commercial breeding programmes are currently implemented in Australia (www.sheepgenetics.org.au) and New Zealand (www.sil.co.nz) where resistant animals are identified by low FEC. Recording of this phenotype in Ireland (www.sheep.ie) however is only just beginning to be carried out routinely.

Studies using microsatellite-based linkage analysis (LA) led to the identification of multiple regions of the genome associated with GIN resistance (e.g. Beh et al., 2002; Crawford et al., 2006; Davies et al., 2006), however the identification of candidate genes through QTL mapping has proven difficult, as QTL often span millions of base pairs and contain hundreds of potential candidate genes. Recent advances in high throughput sequencing and genotyping technologies provide new opportunities to understand the host response to GIN at the molecular level, and identify polymorphisms conferring GIN resistance. Genome-wide association studies utilise the information provided by a large number of markers spread evenly throughout the genome to detect variants associated with a trait. While these variants may not be responsible for the observed phenotype, they may be in linkage disequilibrium (LD) with the causative mutation and thus can be used for selection.

With the advent of the Illumina® Ovine SNP50 BeadChip (www.sheephapmap.org) in 2009, microsatellites have largely been replaced with SNP-based genome-wide association studies (GWAS) (e.g. Kemper et al., 2011; Sallé et al., 2012; Riggio et al., 2013, 2014). The information provided by the OvineSNP50 BeadChip has been used across multiple independent populations to identify genomic regions that are associated with GIN resistance (Kemper et al., 2011; Sallé, 2012; Pickering, 2013; Riggio et al., 2013). A subsequent

meta-analysis of three independent populations, including those used by Sallé et al. (2012) and Riggio et al. (2013), successfully identified regions in common between the three populations (Riggio et al., 2014). This analysis revealed regions that had not previously been identified in the separate analyses, indicating the potential power of a meta-analyses.

If SNPs are to be utilised in selection programmes, however, the associations must be independently validated. For this reason the significant QTL and genomic regions from the original studies were genotyped in the Scottish Blackface population used in this study. Additionally, results from the RNA-Seq analysis (Chapter 3) provided a list of candidate genes for GIN resistance. Discovery of the causal mutation, or linked polymorphisms, within these genes would not only help in understanding the biological pathways underlying the trait, but also allow estimation of the proportion of variation in FEC explained by the SNP. For this reason markers in these candidate genes were also genotyped in the Scottish Blackface population used in this study.

5.2 Materials and Methods

5.2.1 Animals

All lambs were sourced from the purebred Scottish Blackface flock at the Teagasc Hill Sheep Farm in Leenane, Co. Mayo, over a 2 year period (2010 - 2011; details in Chapter 2). Flock FEC was monitored weekly from early June, in both 2010 and 2011, when lambs were approximately 8 weeks of age, using the FECPAK system (www.fecpak.co.nz). Once flock FEC reached approximately 600 EPG lambs were individually sampled twice (FEC1A and FEC1B), with samples A and B taken 1 week apart. Following FEC1B sampling, lambs were dosed with a macrocyclic lactone (ML; Oramec, Merial Animal Health Ltd) in accordance with manufacturer's recommendations. Flock FEC was then monitored weekly (FECPAK), and once flock FEC again reached approximately 600 EPG the sampling protocol was repeated to yield FEC2A and FEC2B. Individual FEC was determined using the modified McMaster method (Ministry of Agriculture Fisheries and Food, 1986). Faecal egg counts were distinguished as *Nematodirus* and 'other Trichostrongyles'. Body weight was taken at approximately 14 weeks of age for all animals, and again at approximately 20 weeks of age for the 2011-born animals. At the last faecal sampling (FEC2B) blood samples were collected from each lamb by jugular venipuncture into aseptic vacutainers for DNA extraction (green vacutainer; lithium heparin).

5.2.2 DNA extraction

Genomic DNA was extracted from 9 mL of whole sheep blood using the high salt method, as previously described (Montgomery and Sise, 1990). Briefly, red blood cells were lysed in 2 volumes of red blood cell lysing solution (RBCLS; Appendix 5.1) for 5 min, during which time a change in colour was observed. Tubes were then centrifuged at 1,100 x g for 10 min to pellet the white blood cells. All centrifugation steps were performed at room temperature. To ensure the complete lysis of red blood cells, 10 mL RBCLS was added to the pellet and shaken to mix. Pure white blood cells were then harvested by centrifugation at 800 x g for 5 min. Pelleted white blood cells were washed in Tris buffered saline and centrifuged at 800 x g for 5 min. Washed white blood cells were resuspended by vigorous vortexing in 3 mL of TE buffer before the addition of 3 mL of freshly made Proteinase K-EDTA-SDS (Appendix 5.1) solution while gently swirling the tubes. The tubes were incubated at 50°C in a water bath for a minimum of 2 h, or overnight. The digested sample was extracted with 3 mL of saturated NaCl, shaken vigorously for 1 min, and centrifuged at 1,100 x g for 10 min. The supernatant was carefully decanted into falcon tubes containing 20 ml of 95% Analar ethanol to precipitate the DNA, and the tubes inverted gently, before the DNA was spooled out using a sealed Pasteur pipette. DNA was washed in 70% Analar ethanol and

dried overnight before resuspension in 1 mL of sterile TE. DNA was cleaned using a DNA blood mini kit (Qiagen, Germany).

In some cases whole blood was frozen, which results in the lysis of the white blood cells. For these samples genomic DNA was extracted from 400 uL of sheep blood using the Maxwell® 16 Research System (Promega, UK), as per the manufacturer's recommendations.

5.2.3 Selection of SNP

SNPs for association studies were identified from two sources:

1. Candidate genes chosen because they were differentially expressed (DE) between HighFEC and LowFEC Scottish Blackface (Chapter 3),
2. Candidate genomic regions chosen because they were associated with GIN resistance in a previous study (Riggio et al., 2014).

SNP in genes differentially expressed between HighFEC and LowFEC animals

A total of 41 genes differentially expressed between HighFEC and LowFEC animals were identified (Chapter 3). Eight genes were on scaffolds or contigs, so their genomic location was unknown. These were discarded from further analysis. Genotyping of SNP within the remaining genes (Table 3.6 & Table 3.7) was undertaken in order to determine if polymorphisms within the genes were associated with resistance. The SNPs examined (RNA-Seq cohort) were chosen from those on the Ovine Infinium® HD SNP BeadChip (Table 5.1). Two genes did not contain any SNP and were also discarded.

Table 5.1: Genomic location of SNP in RNA-Seq cohort. SNP were in coding (C) or non-coding (NC) regions.

Gene	Name	Chr	Gene start	Gene end	Gene length	Total SNP	Genotyped	C	NC
ENSOARG00000020224	*COL9A2*	1	14,676,505	14,694,890	18,385	8	2	1	1
ENSOARG00000006960	*FCRL1*	1	106,671,532	106,687,229	15,697	12	3	3	0
ENSOARG00000019308	*BTLA*	1	175,586,573	175,619,634	33,061	14	3	3	0
ENSOARG00000020373	*MFI2*	1	189,724,320	189,750,098	25,778	10	3	0	3
ENSOARG00000020386	*APOD*	1	190,322,087	190,334,583	12,496	2	2	0	2
ENSOARG00000010719	*LPL*	2	45,652,122	45,675,906	23,784	7	3	2	1
ENSOARG00000005490	*SLC30A2*	2	239,927,723	239,938,030	10,307	6	3	3	0
ENSOARG00000021044	*CLEC2B*	3	204,722,589	204,738,493	15,904	3	3	0	3
ENSOARG00000018806	*WBP2NL*	3	217,007,512	217,026,812	19,300	4	2	1	1
ENSOARG00000013341	-	5	11,795,634	12,084,118	288,484	65	5	3	2
ENSOARG00000011529	*ZFR2*	5	17,716,762	17,739,259	22,497	13	3	2	1
ENSOARG00000002342	*SH3RF2*	5	54,627,464	54,758,880	131,416	32	5	1	4
ENSOARG00000013782	*ALB*	6	88,136,611	88,159,187	22,576	8	3	1	2
ENSOARG00000016543	*CXCL9*	6	90,526,788	90,531,488	4,700	2	2	1	1
ENSOARG00000018254	*FRAS1*	6	92,393,951	92,908,337	514,386	164	7	7	0
ENSOARG00000000928	*ECT2L*	8	63,609,664	63,674,228	64,564	19	4	2	2
ENSOARG00000004253	-	11	13,939,705	13,941,126	1,421	1	1	0	1
ENSOARG00000007871	*SCRN2*	11	38,175,720	38,178,899	3,179	2	1	1	0
ENSOARG00000012882	*CACNB2*	13	31,828,331	32,071,214	242,883	58	5	0	5
ENSOARG00000005312	*ZNF461*	14	46,051,699	46,072,267	20,568	6	3	3	0
ENSOARG00000002371	-	14	59,902,860	59,913,735	10,875	5	2	1	1
ENSOARG00000001898	*FOLR4*	15	844,049	848,328	4,279	3	3	3	0
ENSOARG00000007987	*GZMK*	16	24,143,311	24,153,334	10,023	6	3	2	1
ENSOARG00000013498	*DNAH5*	16	59,094,860	59,358,685	263,825	89	6	6	0
ENSOARG00000005549	-	17	6,548,267	6,553,023	4,756	1	1	0	1
ENSOARG00000009194	*ATP10A*	18	1,810,732	1,994,082	183,350	50	5	3	2
ENSOARG00000002035	*WARS*	18	63,910,328	63,930,861	20,533	5	3	1	2
ENSOARG00000009143	-	18	68,547,748	68,551,257	3,509	5	2	1	1
ENSOARG00000001778	-	21	47,746,255	47,802,320	56,065	12	4	1	3
ENSOARG00000001718	*MPPE1*	23	43,090,283	43,103,534	13,251	3	3	1	2
ENSOARG00000002175	*DNASE1*	24	3,019,422	3,021,959	2,537	3	3	2	1

SNP were prioritised for genotying using the following criteria: 1) at least one SNP in each gene of interest, 2) SNP in a coding region of the gene of interest, and 2) SNP polymorphic in the Scottish Blackface RNA-Seq data set. Of the 100 SNP submitted for plex design, 93 were subsequently genotyped (Sequenom GmbH, Germany).

GWAS from European sheep populations

A joint genome-wide association study (GWAS) of three European sheep populations, Scottish Blackface (Riggio et al., 2013), Romane x Martinique BlackBelly (Sallé et al., 2012) and Sarda x Laucaune (AGRIS, Sardinia) was carried out (Riggio et al., 2014). Average animal FEC (*Nematodirus* and other Trichostrongyles combined) was used for each animal as the phenotypic measure of GIN resistance. After quality control, 4,123 animals and 38,991 SNPs were available for analysis. Analyses were performed using the regional heritability mapping (RHM) approach (Nagamine et al., 2012). This led to the identification of genome-wide significant regions on OAR4, 12, 14, 19 and 20. Several other regions (on OAR1, 3, 4, 5, 7, 12, 19, 20 and 24) were significant at the suggestive level (Riggio et al., 2014): 6 of these regions, on OAR3, 4, 5, 7, 12 and 14 (Table 5.2), were chosen for

genotyping in the Scottish Blackface population used in this study (Validation cohort). The regions on OAR 13 and 21 (Table 5.2) were significant only in the Romane X Blackbelly sheep population (Sallé et al., 2012), yet they were included due to their high significance in that population.

Table 5.2: Genomic location or regions chosen for SNP validation genotyping. Regions were chosen from a joint GWAS of three European sheep populations (Riggio et al., 2014).

Chromosome	Region	SNP
OAR 3	87.2 - 92.2 Mb	113
OAR 4	6.7 - 11.7 Mb	109
OAR 5	85.0 - 91.0 Mb	113
OAR 7	50.4 - 55.5 Mb	116
OAR 12	45.0 - 57.9 Mb	262
OAR 13	69.0 - 73.0 Mb	101
OAR 14	45.3 - 50.4 Mb	108
OAR 21	36.2 - 39.2 Mb	10

All regions covered approximately 5 Mb (Table 5.2), with the exception of the region on OAR12, which spanned two extended regions (13 Mb). The region on OAR21 encompassed a potential causative mutation in the pepsinogen gene. Of the 932 SNPs, 588 were sourced from the Illumina® Ovine SNP50 BeadChip, and 344 from the Ovine Infinium® HD SNP BeadChip.

5.2.4 Genotyping

The 93 SNPs in candidate genes (RNA-Seq cohort) were genotyped in 253 Scottish Blackface animals by Sequenom GmbH (Germany) using the MassARRAY® system. The 932 SNPs from the candidate genomic regions (Validation cohort) were genotyped in a subset (202) of these animals by LGC Genomics (Germany), using KASP™ technology. The subset of 202 animals were selected based on availability of DNA. These animals were selected from both 2010 and 2011, and included both male and female lambs.

As a measure of quality control SNPs with a minor allele frequency <0.02 and/or a call rate <90% were removed from analysis. Deviation from Hardy Weinberg equilibrium was not used as a criterion for excluding SNPs. Individual animals with low (<95%) call rates were also removed. After these quality control measures 84 markers and 237 animals remained from the RNA-Seq cohort, and 867 markers and 202 animals from the Validation cohort.

5.2.5 Statistical analysis

Prior to analysis FEC A and B (taken one week apart) from each round of natural infection (FEC1 and FEC2) were averaged, with *Nematodirus* and other Trichostrongyles counts

analysed separately. FEC values were then log transformed (ln(X+25)) to stabilise the variance. Body weight (14 weeks and 20 weeks) was adjusted for age using the following formula

$$wtXa = \frac{weight}{age(days)} \times (X \times 7)$$

Where *X* is either 14 or 20. A variance component analysis was carried out in ASReml (v3.0; Gilmour et al., 2009), using the available pedigree information, to estimate heritability of each trait (Appendix 5.2). All animals for which phenotypic data was available were used (n = 261). A model was fitted with sex (male or female) and grazing group (Lowland10, Lowland11 or Hill11) as fixed effects for FEC traits (*Nematodirus* and other Trichostrongyles), and sex, grazing group and dam age (2 years or older than 2 years) as fixed effects for weight traits (wt14a and wt20a). Animal was fitted as random. Other fixed effects (birth type and rearing type) were examined but were not significant and so were discarded from the model. Phenotypic variance and heritability were also calculated using ASReml

$$Phenotype(P) = Genotype(G) + Enviroment(E)$$

Phenotypic variance is defined as additive genetic variance (animal) + residual variance

$$Var(P) = Var(A) + Var(R)$$

The additive genetic variance as a proportion of the phenotypic variance is known as narrow-sense heritability, and is defined as

$$h^2 = \frac{Var(A)}{Var(P)}$$

Association analyses were performed using the GenABEL package (Aulchenko et al., 2007) in the R environment (http://www.r-project.org), as per the method of Riggio et al. (2013) (Appendix 5.3). To account for relatedness, the variance-covariance matrix was estimated from the genomic kinship matrix, constructed using pair-wise average Identities-by-State (IBS). These were calculated for all samples on the basis of all autosomal SNPs. Population stratification was observed using a multidimensional scaling plot of kinship distance based on IBS. A linear mixed (polygenic) model was estimated based on the variance-covariance matrix and the kinship matrix, with sex (male or female) and grazing group (Lowland10, Lowland11 or Hill11) as fixed effects. From this heritability was also calculated.

Association was then tested using an mmscore function (Chen and Abecasis, 2007) on the residuals, which have been corrected for relatedness, and therefore should be independent of pedigree or prior selection. After Bonferroni correction the significance level thresholds for

the RNA-Seq cohort were P < 5.88 x 10^{-4} and P < 1.18 x 10^{-2} for genome-wide significance (P < 0.05) and suggestive significance (that is, one false positive per genome scan), respectively. The corresponding significance level thresholds for the Validation cohort were P < 5.78 x 10^{-5} and P < 1.15 x 10^{-3} respectively. To account for any remaining population substructure not accounted for by the genomic kinship matrix, P-values were corrected for the genomic inflation factor λ (Chen and Abecasis, 2007).

As mmscore SNP effects are biased downwards, a subset of SNPs (n = 5) with the lowest corrected P-values was taken and effects re-estimated in ASReml (Gilmour et al., 2009). This was done by fitting the SNPs one at a time as fixed effects in the previously described model (Appendix 5.4.1). The predicted trait values (Appendix 5.4.2) were converted to genotypic values as follows (Appendix 5.4.3):

$$a = \frac{(AA - BB)}{2}$$

$$d = AB - \frac{AA + BB}{2}$$

The proportion of V_A due to each SNP was estimated as

$$\frac{2pq(a + d(q - p))^2}{V_A}$$

where AA, BB and AB are the predicted trait values for each genotype class, p and q are the allelic frequencies at the SNP locus, a is the genotypic value of the best homozygote, d is the deviation due to dominance and V_A is the total additive genetic variance of the trait obtained when no SNP fixed effects are included in the model (Falconer and Mackay, 1996). Standard errors of a and d were constructed from the variance–covariance matrix of the predicted genotype classes, as were the SED for pairwise contrasts between the SNP genotype classes.

5.3 Results and discussion

5.3.1 Heritability estimation

The association between selected SNPs and *Nematodirus* and other Trichostrongyles FEC in an Irish Scottish Blackface population was examined in this study. Summary statistics for FEC traits are reported in Table 5.3. As expected, *Nematodirus* egg counts were considerably lower than the other Trichostrongyles egg counts. The average other Trichostrongyles FEC was similar for FEC1 and FEC2, however the range was greater from the first natural infection (Table 5.3).

Table 5.3: Descriptive statistics for faecal egg count (FEC) traits of all Scottish Blackface lambs. FEC were reported as *Nematodirus* and "other Trichostrongyles".

Average age (range) in weeks	Trait	Animals	Mean (Range)	Mean ± SD of transformed data
17 (11 - 23)	Other Trichostrongyles FEC1	260	859 (0 - 4,835)	6.29 ± 1.16
	Nematodirus FEC1	260	62 (0 - 500)	4.09 ± 0.86
26 (19 - 35)	Other Trichostrongyles FEC2	260	825 (25 - 3,225)	6.46 ± 0.82
	Nematodirus FEC2	260	24 (0 - 250)	3.62 ± 0.65

Heritability estimates using the pedigree-based approach (ASReml) and the genomic relationship matrix from GenABEL are reported in Table 5.4. The average heritability estimate for FEC in the Scottish Blackface population used in this study is within the range reported in the literature. As expected, the heritability estimates using IBS relationships were lower than the pedigree-based estimates. Despite the low number of markers used in both cohorts, average heritability for FEC was estimated to be 0.08 for the RNA-Seq cohort, and 0.14 for the Validation cohort. This indicates that the overall panel of SNPs used explained a proportion of the variation, and may therefore have the potential to be used as markers for GIN resistance.

Heritability estimates were larger for FEC1 for both *Nematodirus* and other Trichostrongyles measurements (Table 5.4), however the differences were not significant due to the large standard deviations. This may be due to the reduced range in FEC observed during the second infection (Table 5.3), particularly in the Lowland11 grazing group (Figure 5.1). The Validation SNP cohort did not explain any more of the heritability in body weight at 14 weeks of age than the RNA-Seq SNPs. For body weight at 20 weeks of age data was only available from the 2011-born animals (Table 5.4). Heritability estimates for this trait using both the pedigree-based approach (ASReml) and the genomic relationship matrix from GenABEL were larger than those for 14 weeks of age.

Table 5.4: Heritability estimates for faecal egg count (FEC) and weight traits. Heritability (h^2) estimated using pedigree (ASREML) and Identities-By-State (IBS; GenABEL) relationships for *Nematodirus* (Nem) and other Trichostrongyles (OT) FEC1 (11 - 23 weeks) and FEC2 (19 - 35 weeks) and weight at 14 (wt14a) and 20 (wt20a) weeks of age.

		FEC1_lnOT	**FEC2_lnOT**	**FEC1_lnNem**	**FEC2_lnNem**	**wt14a**	**wt20a**
N		259	260	259	260	253	167
Pedigree	ID (additive variance)	0.48	0.04	0.32	0.09	3.94	8.43
	Variance (residual)	0.57	0.46	0.20	0.26	5.46	3.54
	Phenotypic variance (± s.d.)	1.04 (± 0.11)	0.50 (± 0.04)	0.51 (± 0.06)	0.35 (± 0.03)	9.41 (± 1.05)	11.98 (± 2.09)
	Heritability (± s.d.)	0.46 (± 0.21)	0.07 (± 0.10)	0.62 (± 0.21)	0.27 (± 0.16)	0.42 (± 0.23)	0.70 (± 0.39)
	Average heritability	0.27		0.45		NA	
IBS[1]	RNA-Seq cohort h^2	0.10	0.06	0.10	0.09	0.12	0.15
	Average h^2	0.08		0.10		NA	
	Validation cohort h^2	0.22	0.08	0.22	0.14	0.12	0.26
	Average h^2	0.14		0.18		NA	

[1]GenABEL does not currently report standard errors of heritability estimates.

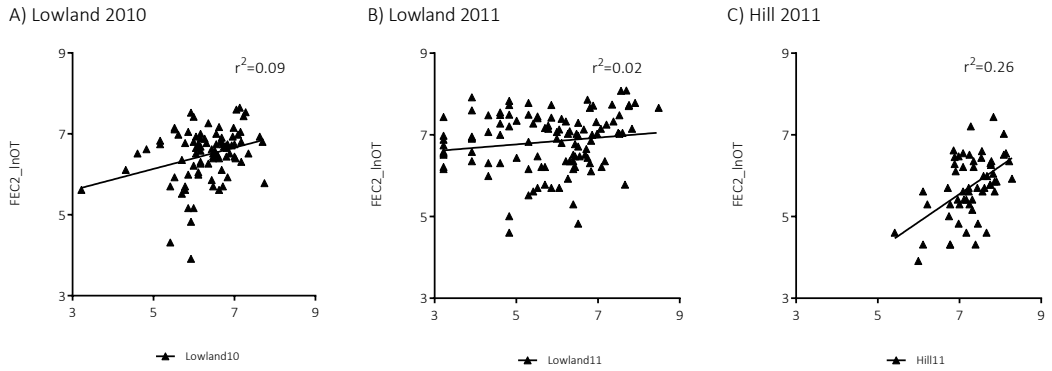


Figure 5.1: Correlation between log-transformed faecal egg count (FEC) values. Correlation between log-transformed FEC values from two natural infections (FEC1 and FEC2) in the three grazing groups contained within the Scottish Blackface population (Lowland10, Lowland11 and Hill11).

5.3.2 Population stratification

Multi-dimensional scaling (MDS) plots of kinship distance based on IBS showed no population stratification of the animals genotyped with the RNA-Seq SNP cohort (Figure 5.2A). The results from the animals genotyped with the Validation SNP cohort showed some genetic sub-structure, with three distinct groups (Figure 5.2B).

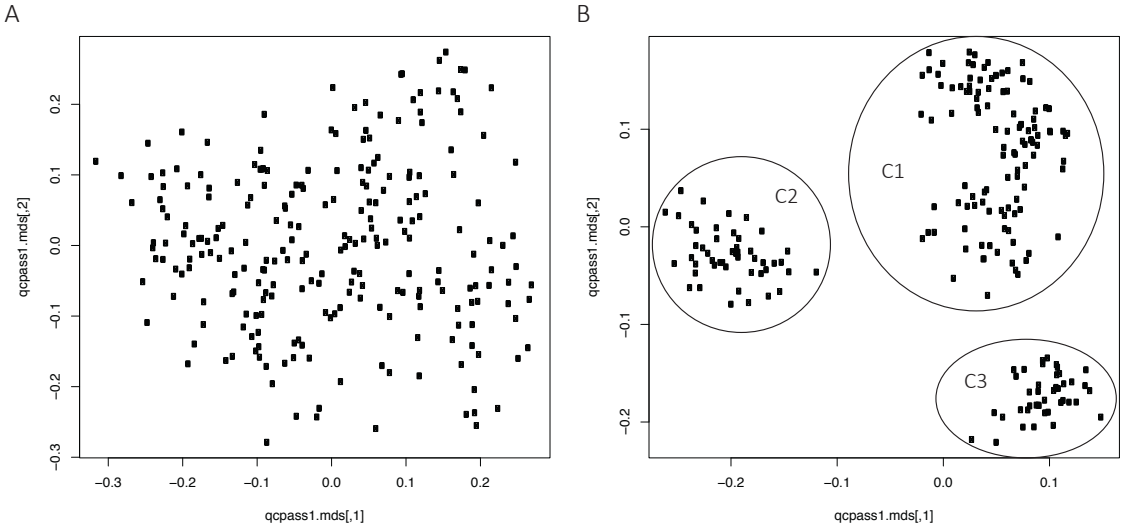


Figure 5.2: Population stratification shown by genotyped SNP. Multi-dimensional scaling (MDS) plots for animals genotyped with RNA-Seq cohort (A) and Validation cohort (B) SNPs. Animals in B are clustered into three groups (C1, C2 and C3).

Upon examination of the animals contained within each group it was revealed that each contained unique sires, with no sire present in more than one group aside from ram 0794334, who had one of his progeny in cluster 1 (Table 5.5) with the remaining 42 progeny in cluster 3. The three distinct groups observed may therefore be a result of haplotypes present in these sires within the regions chosen for genotyping.

Table 5.5: Sire information from the animals used in this study. All animals were purchased for use on the Teagasc Hill Sheep Farm in Leenane, County Mayo. C1, C2 and C3 refer to the number of progeny from each sire represented in each of these groups on the MDS plot for animals genotyped with Validation SNP cohort (Figure 5.2B).

Ram ID	Purchased	Date	Original owner	C1	C2	C3
0492728	Maam X Mart	12/10/08	Tom Whelan, Barrnahowna, Tourmakeady	11	-	-
0792730	Maam X Mart	12/10/08	Tom Egan, Finny Clonbur, Co. Galway	4	-	-
0794368	Maam X Mart	4/10/09	Brendan Varley, Glentrague, Clonbur, Co. Galway	15	-	-
0594333	On farm	2011	Brian Reilly, Glenacolly, Leenane, Co. Mayo	39	-	-
0792729	Maam X Mart	12/10/08	Tom Egan, Finny Clonbur, Co. Galway	7	-	-
0994335	Maam X Mart	10/10/11	Danny Fadian	34	-	-
0894367	On farm	-	Brian Reilly, Glenacolly, Leenane, Co. Mayo	-	49	-
0794334	Maam X Mart	10/10/11	Patrick Henagehen	1	-	42

5.3.3 Association analysis

Association analysis identified five suggestive SNPs (one for other Trichostrongyles FEC2, two for *Nematodirus* FEC1, and one each for weight at 14 and 20 weeks of age) from the RNA-Seq cohort (Figure 5.3B & Figure 5.4A), and one suggestive SNP (*Nematodirus* FEC1) from the Validation cohort (Figure 5.6A). Manhattan plots displaying the association analysis results and the corresponding Q–Q plots of observed P-values against expected P-values for other Trichostrongyles FEC (Figure 5.3 and Figure 5.5) and *Nematodirus* FEC (Figure 5.4 and Figure 5.6) are shown below. The results show a range of significance for individual SNPs, however no individual SNP reached the significance level after Bonferroni correction.

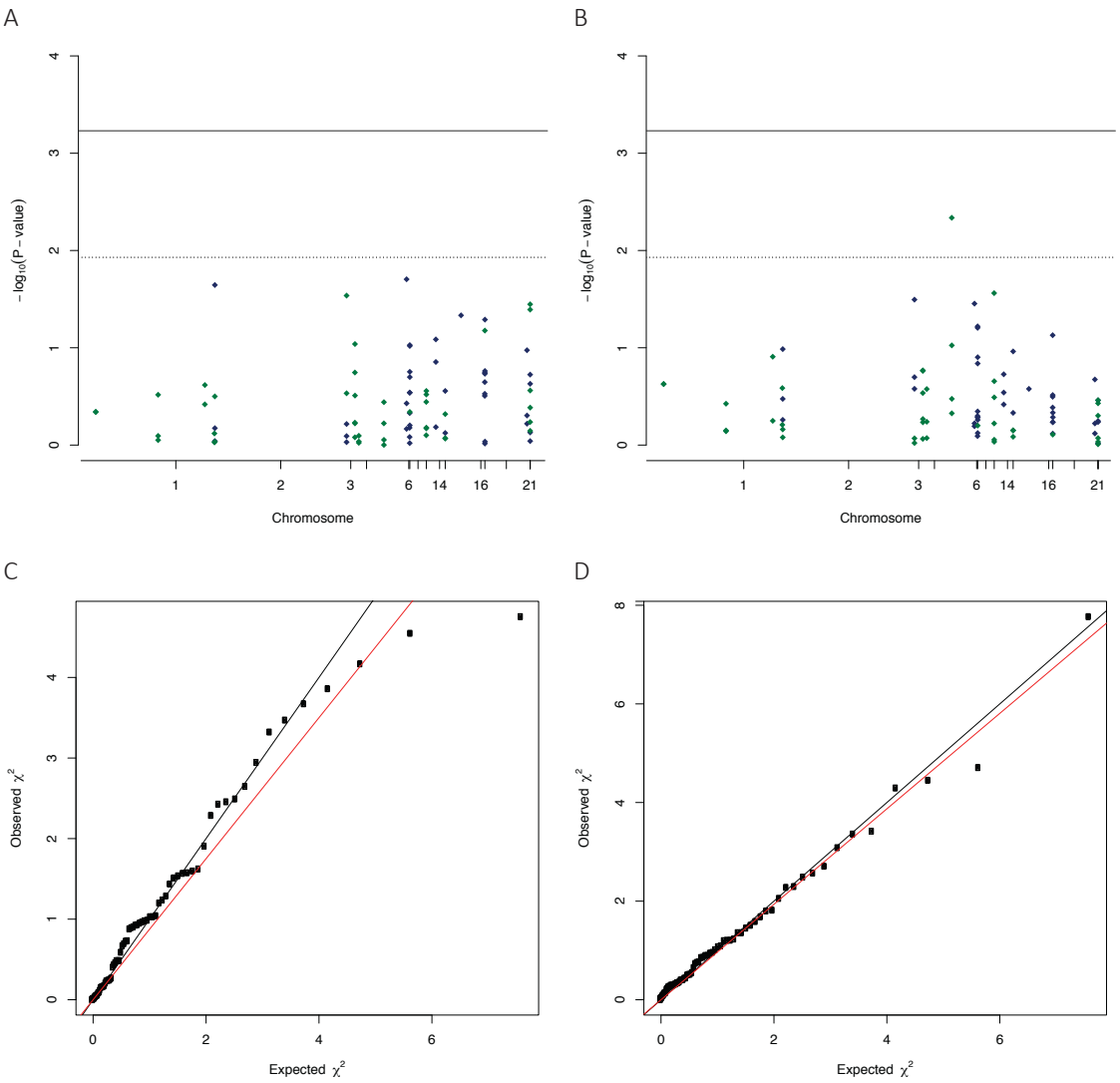


Figure 5.3: RNA-Seq SNP cohort association analysis results for other Trichostrongyles faecal egg count (FEC) traits. Manhattan plot displaying the RNA-Seq cohort association analysis results (-$\log_{10}$(P) of the corresponding Pc1df, P-values corrected for the genomic inflation factor λ) and Q–Q plot (below) of observed P-values against the expected P-values for other Trichostrongyles FEC1 (a, c) and FEC2 (b, d). Genome-wide P<0.05 (solid line) and suggestive (dashed line) thresholds are also shown.

The RNA-Seq cohort SNP OAR5_54653150 reached genome-wide suggestive significance for other Trichostrongyles FEC2 (Figure 5.3B). This SNP is an intronic variant within the SH3 domain containing ring finger 2 gene (*SH3RF2*). SH3RF2 acts as a scaffold protein and expression leads to the activation of the JNK pathway and to nuclear translocation of NF-$\varkappa$B. (Tapon et al., 1998). The JNK pathway is activated primarily by cytokines and exposure to environmental stress (Weston and Davis, 2007), and the nuclear factor NF-$\varkappa$B pathway is considered a prototypical pro-inflammatory signalling pathway (Lawrence, 2009). A deletion that removes all but the first exon is associated with high growth in chickens, with *SH3RF2* expressed in the hypothalamus of low growth animals but not in high growth individuals (**?**). The animals in the high growth line have also been reported to have a genetic defect in hypothalamic appetite regulation (Dunnington and Siegel, 1996).

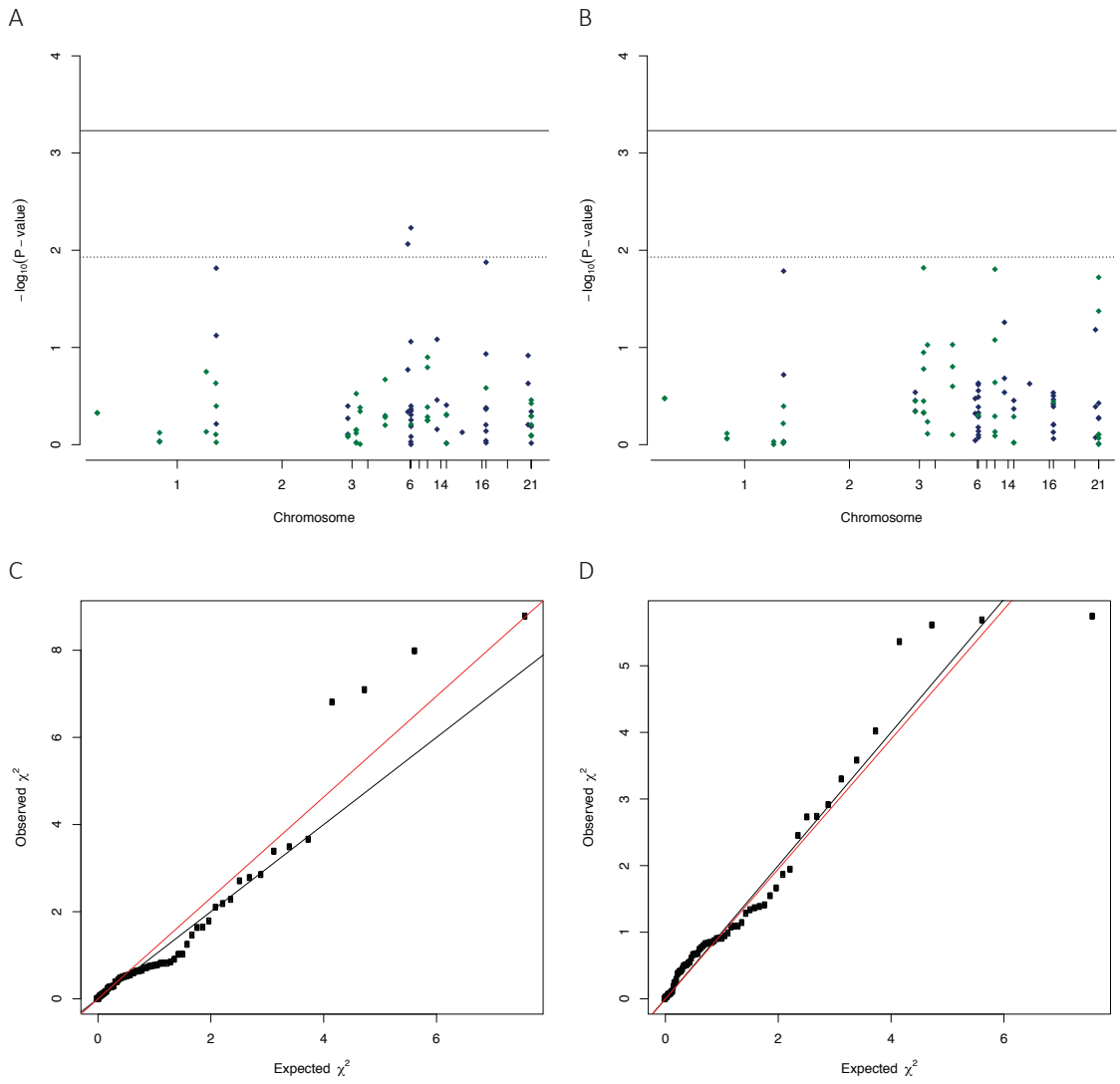


Figure 5.4: RNA-Seq SNP cohort association analysis results for *Nematodirus* faecal egg count (FEC) traits. Manhattan plot displaying the RNA-Seq cohort association analysis results (-$\log_{10}$(P) of the corresponding Pc1df, P-values corrected for the genomic inflation factor λ) and Q–Q plot (below) of observed P-values against the expected P-values for *Nematodirus* FEC1 (a, c) and FEC2 (b, d). Genome-wide P<0.05 (solid line) and suggestive (dashed line) thresholds are also shown.

Two RNA-Seq cohort SNPs reached a suggestive level of significance for *Nematodirus* FEC1 (Figure 5.4A). The first, located within the Fraser syndrome 1 gene (*FRAS1*), codes for a missense variant that results in a change from isoleucine to a valine at codon 1128. This change is classified as 'tolerated' by the SIFT algorithm (score=1) (Kumar et al., 2009). *FRAS1* appears to function in the regulation of epidermal-basement membrane adhesion and organogenesis during development in both humans and mice (McGregor et al., 2003; Short et al., 2007). The second SNP is an intronic variant contained within the albumin (*ALB*) gene. Serum albumin, the main protein of plasma, is a carrier protein for steroids, fatty acids, and thyroid hormones, and functions as a regulator of the colloidal osmotic pressure of blood. *ALB* has been associated with GIN infection in four separate studies (Keane et al., 2006; Knight et al., 2010; Nagaraj et al., 2012; Pemberton et al., 2012). It has

been hypothesised that albumin may be constitutively released into the gastric mucus, and may therefore play an innate protective role (Pemberton et al., 2012).

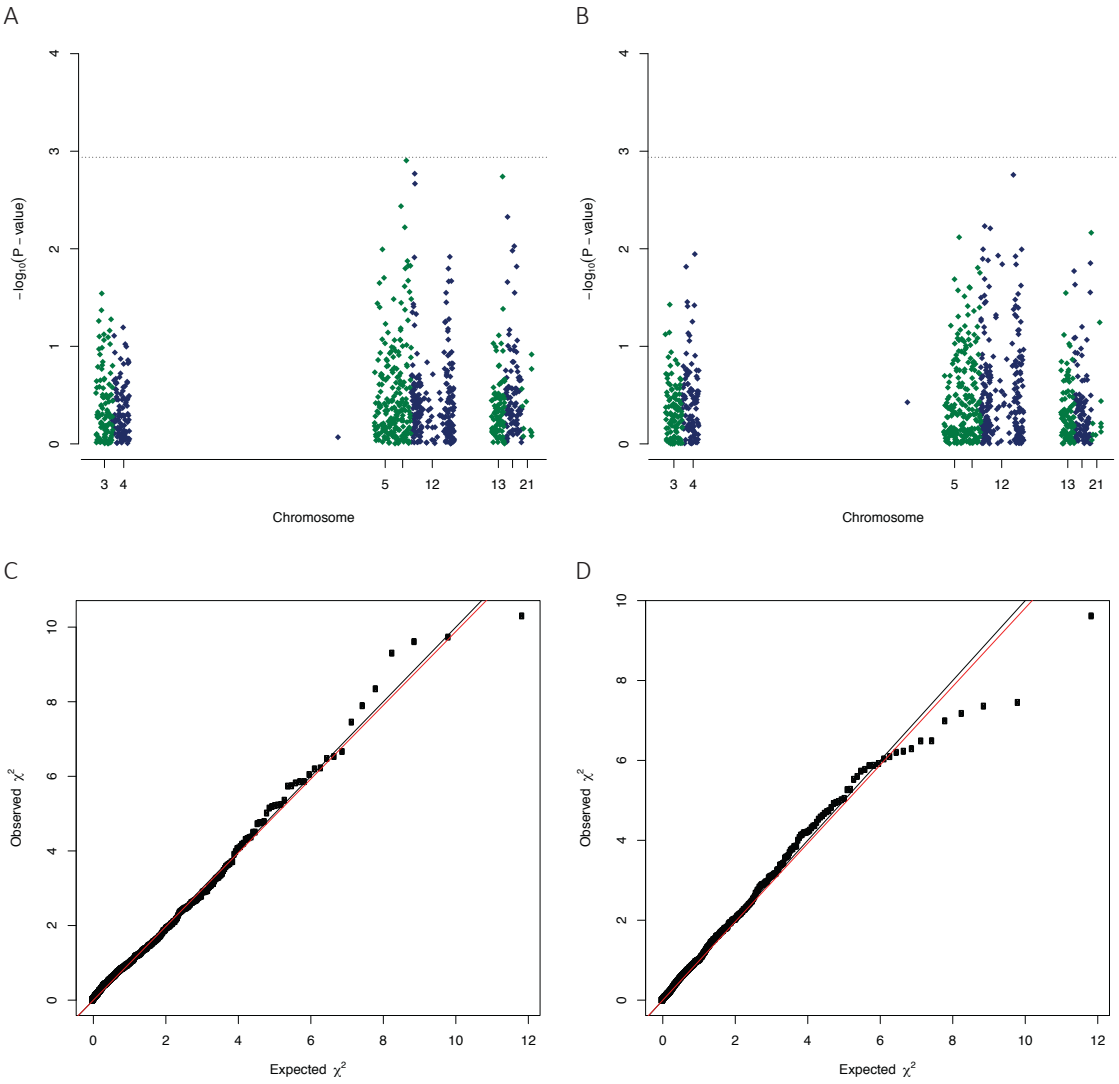


Figure 5.5: Validation SNP cohort association analysis results for other Trichostrongyles faecal egg count (FEC) traits. Manhattan plot displaying the Validation cohort association analysis results (-$\log_{10}$(P) of the corresponding Pc1df, P-values corrected for the genomic inflation factor λ) and Q–Q plot (below) of observed P-values against the expected P-values for other Trichostrongyles FEC1 (a, c) and FEC2 (b, d). Genome-wide suggestive (dashed line) thresholds are also shown.

None of the Validation cohort SNPs within the QTL for pepsinogen variation under infection (Sallé et al., 2012) were significant, although one SNP approached suggestive significance (corrected P-value of 0.007) for other Trichostrongyles FEC1 (Figure 5.5). Pepsinogen level was not found to be significantly different between selected HighFEC and LowFEC animals from this population given a controlled challenge of *T. circumcincta* (Chapter 2), and it is therefore not surprising that SNPs within this QTL were not highly significant.

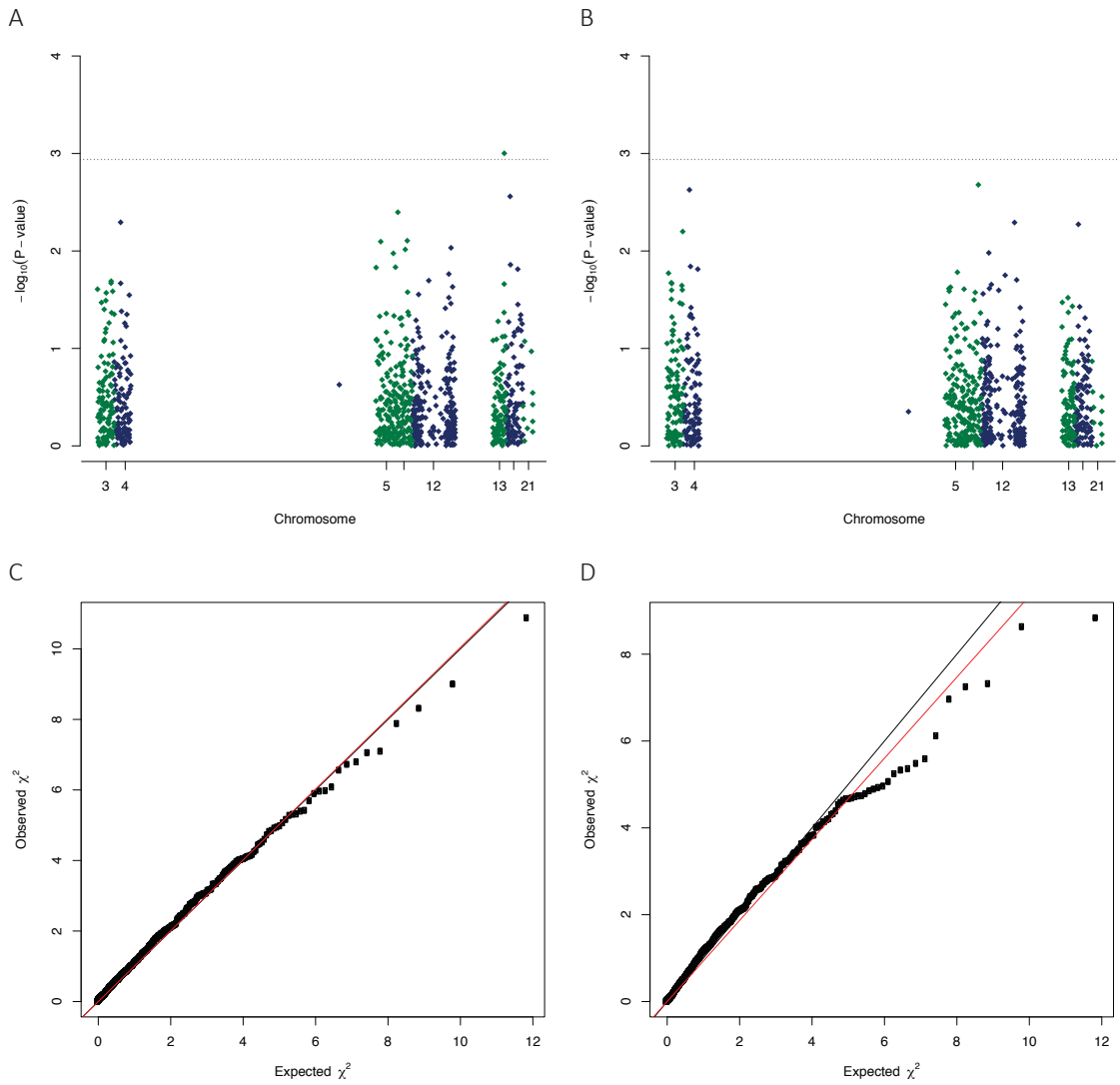


Figure 5.6: Validation SNP cohort association analysis results for *Nematodirus* faecal egg count (FEC) traits. Manhattan plot displaying the Validation cohort association analysis results ($-\log_{10}$(P) of the corresponding Pc1df, P-values corrected for the genomic inflation factor λ) and Q–Q plot (below) of observed P-values against the expected P-values for *Nematodirus* FEC1 (a, c) and FEC2 (b, d). Genome-wide suggestive (dashed line) thresholds are also shown.

The region on OAR13 covered by the Validation SNP cohort was shown to be highly significantly associated with FEC in Romane × Martinik Black Belly backcross lambs after a secondary challenge with *Haemonchus contortus* (Sallé et al., 2012). The observed effects could not be matched to any known functional candidate genes within the QTL, located between 70.1 and 77.8 Mb, however a SNP in this region (s09683.1) is also associated with *Nematodirus* FEC1 in the present analysis. There were no suggestive SNPs in the regions identified from the Scottish Blackface GWAS (Riggio et al., 2013), highlighting the differences between populations of animals of the same breed.

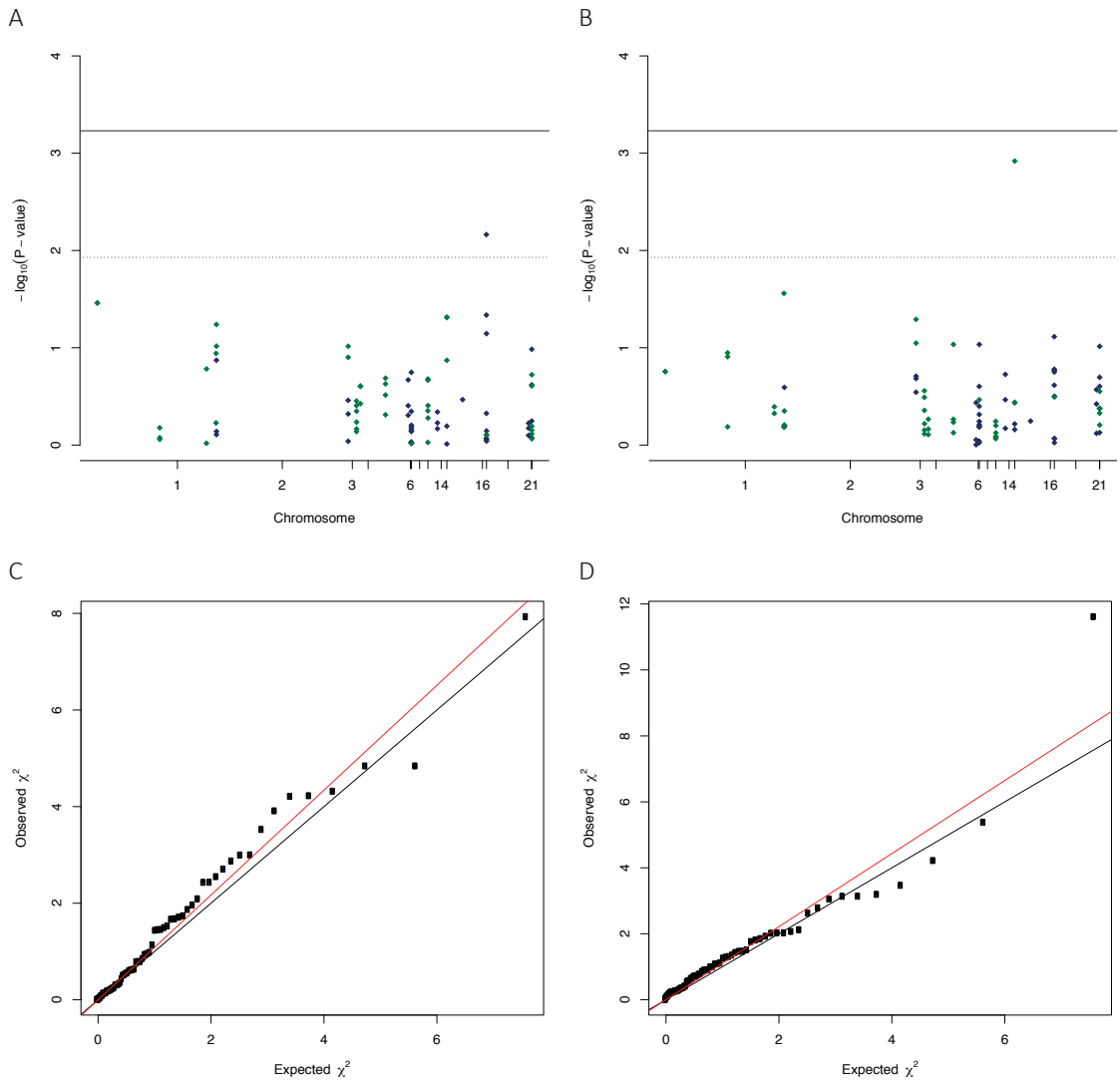


Figure 5.7: RNA-Seq SNP cohort association analysis results for body weight traits. Manhattan plot displaying the RNA-Seq cohort association analysis results (-$\log_{10}$(P) of the corresponding Pc1df, P-values corrected for the genomic inflation factor λ) and Q–Q plot (below) of observed P-values against the expected P-values for body weight at 14 weeks (wt14a; a, c) and 20 weeks (wt20a; b, d) of age. Genome-wide suggestive (dashed line) thresholds are also shown.

Two RNA-Seq cohort SNPs were suggestively associated with weight traits (Figure 5.7A and B). The missense variant OAR16_59155635 (arginine changed to histidine at amino acid 1078) within *DNAH5* (dynein heavy chain 5, axonema), suggestively associated with weight at 14 weeks of age, is classified as 'tolerated' by the SIFT algorithm (score=0.29). Dyneins are microtubule-associated motor protein complexes, and are required for structural and functional integrity of cilia and flagella (Ibañez-Tallon et al., 2003). DNAH5 has been shown to be decreased in the intestine of pigs on a protein restricted diet compared to controls, indicating that DNAH5 may play a role in nutrition absorption (Ren et al., 2014).

Infection with *T. circumcincta* has been shown to reduce growth rate in young lambs (Coop et al., 1982; Bishop et al., 1996; Bouix et al., 1998; Bishop and Stear, 2000a, 2001), with animals with lower FEC growing more quickly (Coop et al., 1982, 1985). The nutritional

status of the host during infection is important, with the provision of additional protein to growing sheep during infection resulting in enhanced immunity to GIN (Brunsdon, 1964; Coop et al., 1995; Van Houtert and Sykes, 1996). The presence of SNPs in genes involved in appetite and metabolism may therefore be associated with the ability of animals to combat GIN infection through increased ingestion of food - and thus of protein. Infection induces protein deficiency by increasing the demand for amino acids in the alimentary tract while reducing supply through depression of appetite (Sykes and Coop, 2001). *T. colubriformis*-infected sheep given a choice between two feeds that differ in their protein concentration are able to modify their diet selection in order to meet the increased protein requirements resulting from the infection (Kyriazakis et al., 1994).

The SNP OAR15_846268, associated with weight at 20 weeks of age, is also a tolerated (SIFT score=0.9) missense variant (lysine to gluatmic acid at amino acid 30). This SNP lies within *FOLR4* (folate receptor 4, delta). T_{reg} cells have been shown to constitutively express high amounts of folate receptor 4, which is essential for their maintenance (Yamaguchi et al., 2007). T_{reg} are a subpopulation of $CD4^{+}$T cells that modulate the immune system through production of the immunosuppressive cytokines IL-10 and TGF-β. They are an important “self-check” in the immune system, and have been shown to be induced and expanded during helminth infection (Allen and Maizels, 2011).

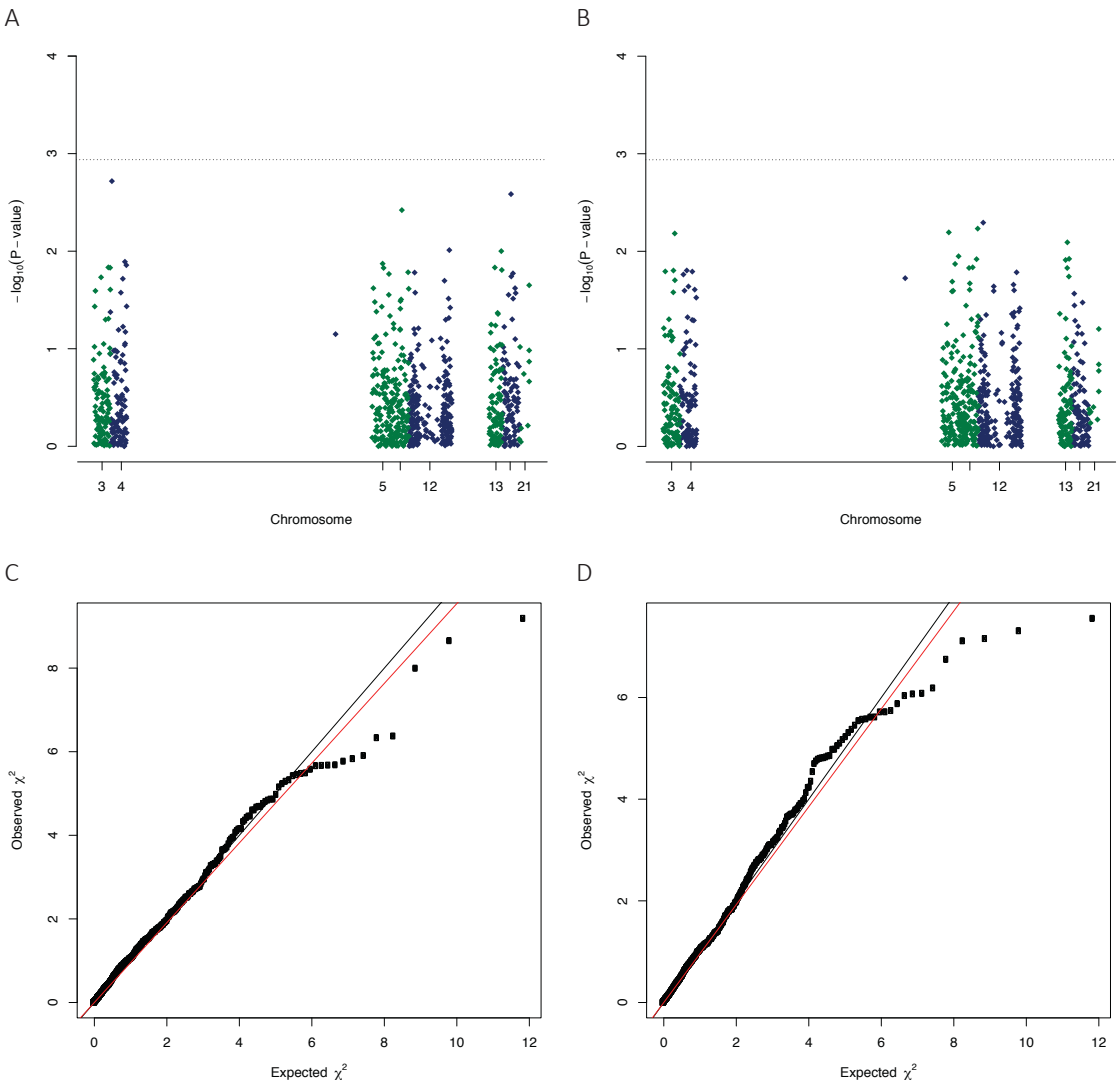


Figure 5.8: Validation SNP cohort association analysis results for body weight traits. Manhattan plot displaying the Validation cohort association analysis results (-$\log_{10}$(P) of the corresponding Pc1df, P-values corrected for the genomic inflation factor λ) and Q–Q plot (below) of observed P-values against the expected P-values for body weight at 14 weeks (wt14a; a, c) and 20 weeks (wt20a; b, d) of age. Genome-wide suggestive (dashed line) thresholds are also shown.

No Validation cohort SNPs reached a suggestive level of significance (Figure 5.8). A summary of the 5 most significant SNPs for each trait, their significance level (Pc1df, GenABEL P-value corrected for the genomic inflation factor λ) and ASReml p-value for the RNA-Seq and Validation SNP cohorts are reported in Table 5.6 and Table 5.7, respectively.

Table 5.6: Top 5 RNA-Seq cohort SNPs for each FEC trait. SNP are sorted by significance level (Pc1df, GenABEL P-value corrected for the genomic inflation factor λ). Chromosome number and position are described within the SNP name. SNPs are either coding (C) or non-coding (NC). P-values for each SNP were also re-estimated in ASReml (P-value) with each SNP as a fixed effect.

Trait	SNP	A1	A2	N	Pc1df	Sig.	Sug.	Coding	Gene	P-value
FEC1_lnOT	OAR2_45658663	C	T	235	0.0204			C	*LPL*	0.0220
	OAR6_88151153	G	T	234	0.0258			N	*ALB*	0.3050
	OAR3_204733979	G	A	236	0.0296			N	*CLEC2B*	0.2170
	OAR21_47759570	T	C	236	0.0332			N		0.1210
	OAR21_47759083	C	T	236	0.0388			N		0.1460
FEC2_lnOT	OAR5_54653150	C	T	237	0.0046		Y	N	*SH3RF2*	0.0080
	OAR11_38176374	C	G	237	0.0274			C	*SCRN2*	0.0880
	OAR2_239931409	T	A	237	0.0320			C	*SLC30A2*	0.0310
	OAR6_88151153	G	T	235	0.0352			N	*ALB*	0.0340
	OAR6_92741314	T	C	237	0.0602			C	*FRAS1*	0.0450
FEC1_lnNem	OAR6_92746652	A	G	235	0.0059		Y	C	*FRAS1*	0.1100
	OAR6_88151153	G	T	234	0.0088		Y	N	*ALB*	0.0850
	OAR18_1929542	A	G	235	0.0133			N	*ATP10A*	0.0760
	OAR2_45653400	G	T	236	0.0151			N	*LPL*	0.2630
	OAR2_45675893	A	G	235	0.0752			C	*LPL*	0.2210
FEC2_lnNem	OAR5_11926149	T	C	217	0.0152			N		0.0280
	OAR13_32030417	G	A	234	0.0157			N	*CACNB2*	0.0660
	OAR2_45653400	G	T	237	0.0164			N	*LPL*	0.1310
	OAR21_47759570	T	C	237	0.0190			N		0.0560
	OAR21_47759083	C	T	237	0.0422			N		0.1120
wt14a	OAR16_59155635	A	G	232	0.0069		Y	C	*DNAH5*	0.002
	OAR1_14688520	C	T	232	0.0347			C	*COL9A2*	0.012
	OAR1_14691551	G	A	232	0.0347			N	*COL9A2*	0.012
	OAR16_59153719	T	C	232	0.0461			C	*DNAH5*	0.016
	OAR15_847289	G	A	232	0.0485			C	*FOLR4*	0.006
wt20a	OAR15_846268	A	G	156	0.0012		Y	C	*FOLR4*	0.005
	OAR1_189741476	A	G	156	0.0276			N	*MFI2*	0.065
	OAR3_204733979	G	A	156	0.0510			N	*CLEC2B*	0.183
	OAR16_59117728	A	G	156	0.0769			C	*DNAH5*	0.024
	OAR3_204731481	C	T	147	0.0896			N	*CLEC2B*	0.125

Table 5.7: Top 5 Validation cohort SNPs for each FEC trait. SNP are sorted by significance level (Pc1df, GenABEL P-value corrected for the genomic inflation factor λ). P-values for each SNP were also re-estimated in ASReml (P-value) with each SNP as a fixed effect.

Trait	SNP	Chr	Position	A1	A2	N	Pc1df	Sig.	Sug.	P-value
FEC1_lnOT	OAR7_53958778	7	53958778	G	A	199	0.0013			0.006
	s00660_1	13	72309372	A	G	200	0.0017			0.007
	OAR12_46119459	12	46119459	G	T	197	0.0021			0.003
	OAR12_46045439	12	46045439	G	A	198	0.0022			0.004
	OAR7_57813736_1	7	52367152	T	C	196	0.0034			0.024
FEC2_lnOT	s02002_1	12	54784267	C	T	198	0.0017			<0.001
	OAR12_46212305	12	46212305	A	G	201	0.0059			0.003
	s53649_1	12	47882527	C	A	201	0.0062			0.003
	s54156_1	21	36240663	C	T	201	0.0069			0.002
	OAR5_89535202	5	89535202	C	T	199	0.0076			0.008
FEC1_lnNem	s09683_1	13	72474617	A	G	200	0.0010		Y	0.013
	s75832_1	14	46590917	C	A	198	0.0027			0.010
	OAR7_51056435	7	51056435	A	G	199	0.0040			0.018
	OAR4_8687452_1	4	8658001	C	T	199	0.0050			0.031
	OAR7_53867898	7	53867898	C	T	198	0.0079			0.034
FEC2_lnNem	OAR7_60254585_1	7	54439369	C	T	201	0.0021			0.009
	OAR4_8579647_1	4	8554085	A	G	197	0.0024			0.007
	OAR12_60974719_1	12	54847345	C	T	200	0.0051			0.019
	OAR14_48676584_1	14	46319049	A	G	200	0.0053			0.017
	OAR3_97664476_1	3	91999429	G	C	202	0.0063			0.028
wt14a	OAR4_6967496_1	4	7211794	T	C	195	0.0019			0.007
	OAR14_47918448	14	47918448	G	A	198	0.0026			0.013
	OAR7_58944510_1	7	53341820	G	A	197	0.0038			0.037
	OAR12_57110529	12	57110529	A	G	197	0.0097			0.098
	s43133_1	13	72638850	T	C	198	0.0100			0.078
wt20a	OAR12_46430191	12	46430191	A	G	156	0.0051			0.021
	s63393_1	7	55204866	G	A	156	0.0059			0.014
	OAR5_87108294	5	87108294	T	G	156	0.0064			0.003
	OAR3_96268784_1	3	90654749	C	T	156	0.0066			0.038
	OAR13_76998194_1	13	71580447	A	G	156	0.0081			0.017

Resistance to gastrointestinal nematodes in sheep is understood to be a complex polygenic trait, controlled by many genes of relatively small effect (Crawford et al., 2006; Kemper et al., 2009). To this end, single markers have been estimated to explain a maximum of 0.48% or 0.08% of the phenotypic variance in FEC following challenge with either *T. colubriformis* or *H. contortus* respectively (Kemper et al., 2011). Single markers were estimated to explain a maximum of 0.50% of the additive genetic variance in this study (Table 5.8). The RNA-Seq SNP cohort accounted for a larger proportion of the additive genetic variance (Table 5.8). This is expected, as they were chosen to be polymorphic within the population, and are from genes that were differentially expressed between HighFEC and LowFEC animals.

The differences between the P-values using the mmscore function in GenABEL and those from the re-estimation in ASReml may be attributable to lack of power. Sample size is very important for GWAS, particularly when looking for variants with small effect sizes (Spencer et al., 2009), which is likely to be why no SNPs reached statistical significance. Using the

Table 5.8: Summary of SNP association analysis results. Genotype means (0 and 2 represent homozygotes and 1 heterozygotes), the genotypic value of the best homozygote (a), dominance deviation (d), and the percent of additive genetic variance that each SNP accounts for (V_A%) are shown.

Trait		**FEC2_lnOT**	**FEC1_lnNem**	**FEC1_lnNem**	**wt14a**	**wt20a**	**FEC1_lnNem**
Cohort		RNA-Seq	RNA-Seq	RNA-Seq	RNA-Seq	RNA-Seq	Validation
SNP		**OAR5_54653150**	**OAR6_92746652**	**OAR6_88151153**	**OAR16_59155635**	**OAR15_846268**	**s09683_1**
Chr		5	6	6	16	15	13
Location		54,653,150	92,746,652	88,151,153	59,155,635	846,268	72,474,617
Gene		*SH3RF2*	*FRAS*	*ALB*	*DNAH5*	*FOLR4*	N/A
Genotype	*0 (±s.e.)*	6.12 (0.13)	3.69 (0.22)	3.47 (0.30)	20.8 (0.82)	26.1 (0.55)	3.95 (0.13)
	1 (±s.e.)	6.26 (0.07)	4.01 (0.14)	4.00 (0.15)	21.3 (0.47)	27.4 (0.56)	4.19 (0.12)
	2 (±s.e.)	6.52 (0.07)	4.11 (0.14)	4.11 (0.13)	22.7 (0.41)	29.0 (1.06)	4.50 (0.20)
Genetic effect	*a (±s.e.)*	0.19 (0.07)	0.20 (0.09)	0.31 (0.14)	0.94 (0.40)	1.47 (0.52)	0.27 (0.10)
	d (±s.e.)	-0.05 (0.09)	0.11 (0.11)	0.20 (0.15)	-0.4 (0.51)	-0.1 (0.62)	-0.03 (0.10)
	V_A%	0.50	0.05	0.08	0.08	0.10	0.10

same approach as the present study, Riggio et al. (2013) did an association analysis using 752 Scottish Blackface lambs. Using this much larger population, only two significant SNPs were found, one for *Nematodirus* FEC and one for other Trichostrongyles FEC at 16 weeks, highlighting the polygenic nature of the trait and the large number of animals required for association studies.

5.3.4 Conclusions

The results from this study support the use of a panel of SNPs rather than individual SNPs for predicting nematode resistance, in agreement with the complexity of this polygenic trait. While a small subset of the SNPs examined in this study had been shown to be significantly associated with gastrointestinal nematode resistance in other populations, including animals of the same breed, there were no significant associations found between individual SNPs and FEC in the Scottish Blackface population used in the association analysis. A number of suggestive associations were detected however. The observed results highlight the shared and unique mechanisms of resistance to gastrointestinal nematodes within and between populations. Overall, the panel of SNPs as a whole was useful in explaining a proportion of the observed heritability of the trait.

Chapter 6

General Discussion

Gastrointestinal nematodes are a serious cause of morbidity and mortality in grazing ruminants. Infected lambs have a reduced ability to absorb nutrients from the gastrointestinal tract resulting in ill-thrift and occasional death. Sub-clinical infection adds to the production losses, in the form of reduced growth rate and light, under-finished carcasses. However, lambs must be challenged with infective L3 in order to allow the development of acquired immunity. Effective nematode control strategies must therefore limit host exposure to nematodes to a level that allows the development of immunity while not compromising performance.

Anthelmintic drenching has been the method of choice for nematode control for the last 50 years. However, consumer concerns about food products from animals subjected to chemical treatment, combined with the inevitable evolution of anthelmintic resistant nematodes means alternative, sustainable methods of nematode control are required. A sustainable method of nematode control is to select for genetically resistant individuals. Selection on phenotype (faecal egg count, FEC) requires detailed trait measurement. This is time-consuming, unappealing and expensive. Selecting resistant animals would be simplified if animals could be selected by genotype; this could also accelerate genetic gain. A detailed understanding of the genes and mechanisms involved in protective immunity and the factors that regulate this response would also aid the development of effective and sustainable nematode control methods, such as immunomodulatory anthelmintics. The aim of this study was to identify genes and biological processes mediating the response to nematode infection.

In Ireland, two lowland sheep breeds have been extensively studied with respect to nematode resistance and it has been demonstrated that the Texel breed is more resistant to nematode infection than the Suffolk breed (Hanrahan and Crowley, 1999; Sayers et al., 2005a; Good et al., 2006; Sayers et al., 2008; Hassan et al., 2011a). However, there have been no studies to date in Ireland on breeds that occupy hill and marginal land. Extensive studies in Scotland have shown substantial genetic variation among Scottish Blackface lambs in both FEC and in worm length (Stear et al., 1999b). This breed is commonly found on hill country in Ireland, and is also an important source of crossbred ewes for the lowland sheep sector. For these reasons the Scottish Blackface breed was the focus of this study.

Reduced FEC in Scottish Blackface lambs is primarily a result of reduced fecundity of adult *T. circumcincta* females

This is the first study in Ireland to confirm within-breed differences in the ability of Scottish Blackface lambs to resist gastrointestinal nematode infection (Chapter 2). After a controlled challenge with *Teladorsagia circumcincta*, LowFEC animals displayed lower FEC throughout the course of infection, validating FEC measurements from two independent natural infections as a method of identifying resistant and susceptible animals. Worm fecundity was significantly lower in the LowFEC (resistant) animals, with shorter, less fecund adult

females. While there were no significant differences in worm burden between the HighFEC and LowFEC animals, there was a correlation between number of mature females and FEC on day of slaughter. There was some evidence that the HighFEC animals tended towards a higher worm burden in both the 2010W cohort animals slaughtered at 14 days post infection, and the 2011E animals slaughtered at 71 days post infection, although this was not statistically significant. FEC in the 2011E cohort peaked at 56 days post infection, and had begun to decline by 71 dpi. This could be a consequence of a reduction in worm number or in the number of eggs produced per worm. Therefore slaughter at 71 days post infection may have been too late to fully capture the peak values for the number of adult worms. Taken together, these results indicate that adult worm burden may also play a role in reduced FEC in resistant Scottish Blackface lambs, although further experiments would be required to validate this. In agreement with our study, studies in Scotland indicate that variability in FEC in Scottish Blackface lambs is largely due to differences in average worm fecundity (Stear et al., 1996), although adult worm burden was also a contributing factor (Stear et al., 1995b). Results from the study by Stear et al. (1996) indicates that the acquired immune response in Scottish Blackface lambs develops in two stages. Firstly, lambs develop specific local IgA responses that regulate worm development and therefore fecundity, decreasing egg production. Secondly, in association with the production of globule leukocytes in the abomasal mucosa, an effective hypersensitivity response is developed that regulates worm burden, in conjunction with decreased fecundity of the worms that do establish (Seaton et al., 1989; Stear et al., 1995b, 1996).

These results are in contrast to the mechanisms of resistance found in in the two other breeds studied extensively in Ireland, Texel and Suffolk, where differences in FEC are largely a result of variation in worm burden rather than variation in average worm fecundity (Good et al., 2006). These differences in FEC are observed irrespective of age, suggesting that the increased resistance to gastrointestinal nematodes in Texel lambs may be innate. This could be a result of either the innate immune response, or innate characteristics of the breed, that may include smooth muscle function (Vallance et al., 1997; Diez-Tascon et al., 2005) or other innate defences (Douch et al., 1984). Transcriptional profiling of lambs from both breeds identified a panel of genes expressed in the absence of infection that were mostly related to the innate immune response (Ahmed, 2013). Future work could include examination of the innate and acquired immune response in all three breeds, examining the mechanisms of resistance over time in Texel, Suffolk and resistant and susceptible Scottish Blackface. The animal selection model validated in Chapter 2 required exposing the animals to gastrointestinal nematodes, and selecting for resistance or susceptibility based on FEC. Examining the innate immune response within breed requires selection of known resistant and susceptible naïve animal. This is not currently possible in Ireland, however genetic selection for health traits, such as parasite resistance, is on the verge of implementation in Ireland (www.sheep.ie). Phenotype recording for FEC commenced in 2013, and estimated breeding values (EBVs) for parasite resistance should be available for

some sires in 2015. This would allow monitoring of the response to infection in resistant and susceptible individuals that had not had prior exposure to gastrointestinal nematodes.

A significant difference in worm burden was observed between the 2010W and 2011W cohorts. This is most likely the result of the use of older and potentially less infective larvae for the infection of the 2011W cohort. In particular the low level of larvae in the HighFEC animals at 7 dpi was a cause for concern. The difference in worm burden between the two cohorts meant that results from both years were analysed separately. The 2011W cohort had consistently lower plasma pepsinogen, and serum and mucosa anti-*T. circumcincta* IgA than the 2010W cohort, reflective of the lower worm burden.

While pepsinogen levels have previously been reported to be an indicator of ostertagiosis in lambs (Lawton et al., 1996; Balic et al., 2000b; Davies et al., 2005), other studies of within-breed differences of gastrointestinal nematode resistance have found that it is not significantly different between lambs with consistently high or low FEC (Stear et al., 1995a). Pepsinogen levels increased post-infection in all cohorts, indicating that abomasal damage did occur, however there was no effect of phenotype (HighFEC or LowFEC). Basophils and eosinophils produce cytokines that can enhance the protective immune response against gastrointestinal nematode infection (Anthony et al., 2007), however it does not appear that levels of these cells circulating in the blood are able to differentiate HighFEC and LowFEC animals. Previous work has found an association between the number of peripheral eosinophils and both FEC (Stear et al., 2002) and worm burden (Beraldi et al., 2008) in Scottish Blackface lambs, with resistant (low FEC) animals having higher circulating eosinophils. This is in contrast to our study, where none of the haematology parameters taken at the time of sampling were associated with FEC2. While both peripheral eosinophil counts and plasma pepsinogen concentrations, in conjunction with FEC, have been postulated to be an effective tool in selecting resistant Scottish Blackface lambs (Stear et al., 1995c), under a controlled challenge they did not differentiate animals with high and low FEC in this study. Peripheral eosinophil concentrations were higher in LowFEC animals over the course of infection in the 2011E cohort although this difference was not statistically significant. This could be due to low power, which could be resolved in future by increasing the number of animals studied per group. The relationship between peripheral blood eosinophilia and tissue eosinophilia is reasonably weak, with only a proportion of eosinophils found in the blood moving into the abomasal mucosa (Henderson and Stear, 2006). The lack of a statistically significant difference between HighFEC and LowFEC lambs in peripheral eosinophilia may therefore not be reflective of eosinophil levels in the abomasal mucosa, the site of infection. Unfortunately the number of eosinophils in abomasal tissue was not measured in any of our cohorts. Future work should therefore consider examining the role of tissue eosinophilia in GIN resistance of Irish sheep populations.

In sheep the majority of plasma IgA, the isotype closely associated with intestinal mucosal immune responses, derives from the mucosal surfaces of the gastrointestinal tract (Sheldrake et al., 1984), and association between levels of plasma and mucosal IgA is a lot stronger than

that of eosinophils (Henderson and Stear, 2006). IgA produced in the gastrointestinal tract can either bind to parasites and parasite secretions, or be transported into the bloodstream. Increased levels of IgA has been positively associated with resistance to *T. circumcincta*, regulating both worm length and fecundity (Smith et al., 1985; Stear et al., 2004; Strain et al., 2002; Stear et al., 1995b, 1999b,c; Strain and Stear, 1999; Halliday et al., 2007). This resistance is regulated through suppressed parasite growth, development and fecundity, and mediated by IgA activity against 4^{th}-stage larvae. While the antigen used in this study was generated from L3 larvae, it has been reported that there is a correlation ($r = 0.68$; $P < 0.001$) between the IgA response to both L3 and L4 antigen (Stear et al., 1995b). Despite this, future work could examine the IgA response to antigen from 4^{th}-stage larvae, as this is strongly correlated with worm fecundity (Stear et al., 1995b). Results from the 2011E cohort show LowFEC animals had significantly higher levels of serum anti-*T. circumcincta* IgA throughout the infection. The 2010W cohort LowFEC animals also had numerically higher levels of serum IgA at day 7, although this was not statistically significant. This may be a result of only having 5 lambs per group in the 2010W and 2011W cohorts, compared with 10 in the 2011E cohort, resulting in reduced statistical power. The resistant animals therefore have larger quantities of unbound IgA entering the bloodstream, particularly at day 7. This could potentially be due to excess IgA production, or a result of excess free IgA due to low worm numbers or reduced worm length. A commercial antibody test, CarLa, targets the carbohydrate larval antigen of *T. colubriformis*. Antibodies to CarLA have been shown to be higher in resistant lambs (Harrison et al., 2008; Shaw et al., 2012). In our study the serum IgA response to CarLA was found to mirror that of the response to L3 *T. circumcincta* antigen, despite the fact that the antigens in both tests were derived from different nematode species. This indicates that the CarLA test could potentially be used as a proxy test for anti-*T. circumcincta* antibodies if *T. circumcincta* antigen could not be sourced.

While the model for selection of resistant and susceptible animals was validated in female lambs, the animals used to define the acute response to infection were male. This was unavoidable, due to the number of lambs present on the Teagasc Hill Sheep Farm, however this must be kept in mind when interpreting the results. Sex has been shown to affect host resistance to *T. circumcincta* infection, with females having lower FEC (Barger, 1993; Stear et al., 1996; Bouix et al., 1998), which may be a result of higher IgA activity (Strain et al., 2002). Future work could examine the differences between male and female lambs selected for high or low FEC, and determine the differences and similarities between the mechanisms of resistance in each sex.

Rapid effective immune response observed in the abomasal lymph node of LowFEC Scottish Blackface in response to infection with *T. circumcincta*

Following on from the identification of Scottish Blackface lambs with divergent phenotypes for GIN resistance, we characterised the transcriptome of the abomasal lymph node following a controlled challenge with *T. circumcincta*. The aim of Chapter 3 was to identify genes and biological processes associated with the host response to GIN in resistant and susceptible individuals. While this experiment was designed so that all gene expression data could be analysed together, fitting year as a variable, the observed differences in worm burden combined with the separation of samples by year on the MDS plot meant that this was not viable. This had the consequence of lower numbers of animals per group than originally intended when determining the acute response to infection. This may have reduced our power to detect differentially expressed genes. A number of genes were found to be differentially expressed between both HighFEC and LowFEC animals, and between 7 and 14 days post infection. Despite the difference in worm burden between the two cohorts, examining upstream regulators revealed pathways and biological mechanisms in common between the HighFEC and LowFEC animals. At 7 dpi genes downstream from the regulators *PPARG*, *IFNG* and *IL4* were activated in LowFEC animals, whereas by day 14 genes downstream of these regulators were upregulated in HighFEC animals. Interferon gamma (*IFN*γ) is secreted by T_H1 lymphocytes, and plays a critical role in regulating the type 1 versus type 2 immune responses in vertebrates (Wakelin, 1996), whereas *IL4* is typically associated with a T_H2-type response (Anthony et al., 2007). The differentially expressed genes, along with the results from the upstream regulator analysis, indicate that it appears to be the differential interplay between T_H1/T_H2 genes that controls the response to gastrointestinal nematodes in resistant compared to susceptible animals. This is in agreement with previous studies using Texel and Suffolk animals (Hassan et al., 2011b; Ahmed, 2013). The LowFEC (resistant) Scottish Blackface lambs are generating an immune response to *T. circumcincta* at 7 days post infection, whereas in their HighFEC counterparts this response is delayed until ~14 days post infection. The immune response generated by the LowFEC animals may therefore be influencing the larval stages of *T. circumcincta*. This response, in conjunction with the IgA response, could result in shorter, less fecund adults in LowFEC animals, as reported in Chapter 2.

Many of these genes were found to be in agreement with other studies on resistance to gastrointestinal nematodes in sheep, including those investigating the transcriptome (Ahmed, 2013; Gossner et al., 2013) and the proteome (Nagaraj et al., 2012). These results are in agreement with the systems genetics study undertaken by Sayre and Harris (2012), which combined data from multiple QTL and gene expression studies. These authors discovered common pathways between genes in QTL associated with genetic resistance to internal parasites. This suggests that there may be some pathways in common to GIN resistance over multiple breeds of sheep and species of gastrointestinal nematode, despite the different mechanisms of resistance between breeds already discussed.

The number of genes differentially expressed for each comparison was relatively low when compared to other RNA-Seq studies of the same tissue (Pemberton et al., 2011; Ahmed, 2013; Gossner et al., 2013). While Pemberton et al. (2011) (and subsequently Gossner et al., 2013) looked at variation in the transcriptome within Scottish Blackface lambs, the study population consisted of 789 individuals bred over a 3-year period (Davies et al., 2006). The animals used in this study, however, came from of a population of 258 individuals bred over a 2-year period, and therefore animals selected within this population may not have been sufficiently divergent to detect a large number of differentially expressed genes. In the future, progeny from tested sires with high and low EBVs for parasite resistance could be utilised, which would ensure the use of animals with divergent phenotypes for resistance to gastrointestinal nematodes. This would also allow an increase in the number of animals used, which would increase the statistical power of the study. Alternatively, there may be variation between animals in how they manifest resistance or susceptibility, which would result in a low number of genes in common between animals within-group. Despite this, there may be pathways in common between the genes that are upregulated in each animal in response to infection. It is for this reason that genes found to be significant using common dispersion estimates were used for network and pathway analysis.

Utilising progeny from sires with high and low EBVs would allow the examination of the development of immunity over time. The selection model used required FEC from two independent natural infections, resulting in the lambs being approximately six to seven months old at slaughter. Using EBVs for selection would allow the response to gastrointestinal nematode infection to be examined in both naïve lambs, and lambs after primary or secondary infection.

Results throughout this thesis, in particular Chapter 3 and Chapter 4, have been reliant on both the sheep genome assembly and annotation quality. The ovine genome assembly (OARv3.1) was produced by the International Sheep Genome Consortium (ISGC) in September 2012, with the Ensembl annotation released in December 2013 (Ensembl release 74). Of the differentially expressed genes identified in this study, approximately 30% were classified as novel protein coding genes in Ensembl. Future releases of the sheep genome will result in the number of functionally annotated genes increasing, ultimately resulting in increased information from the RNA-Seq data. For example, the two most studied regions with regards to resistance to gastrointestinal nematodes in sheep are the region of the Major Histocompatibility Complex (MHC) on chromosome 20, and the region containing the interferon gamma (*IFN*γ) gene on chromosome 3 (Bishop and Morris, 2007). The MHC region is extremely polymorphic (Stear et al., 2005), which makes assembly and annotation difficult. To this end the extensively studied MHC class IIa gene *Ovar-DRB1* (Schwaiger et al., 1995; Sayers et al., 2005a; Stear et al., 2005) is not found in the current Ensembl release (75). Additionally, some genes, such as ENSOARG00000013341, have multiple within-species paralogs (19 in Ensembl release 75). This has implications when mapping RNA-Seq reads to the genome assembly, as only uniquely mapped reads were

kept for read counts.

In this study gene expression was examined in the abomasal lymph node, the site from which the immune response is generated, rather than the abomasal mucosa, the site of interaction between the nematode and the immune response. It was shown in a study of Johne's disease in calves that cytokine levels were generally higher in the mesenteric lymph nodes than in the intestine, and that the cytokines expressed had different profiles depending on the type of tissue examined and the time of sampling following infection (Wu et al., 2007). Both abomasal tissue and abomasal mucosa is available for all animals sequenced in this study. Future work could include sequencing the entire transcriptome, or selected genes through targeted re-sequencing and exome capture (Ng et al., 2009), of the abomasal mucosa. This would allow examination of communication between the abomasal mucosa, the site of infection, and the lymph node, from where the immune response is generated.

While Chapter 3 focussed on the discovery of differentially expressed genes between HighFEC and LowFEC animals, and subsequently the pathways and networks in which these genes are located, RNA-Seq data can also be used for other discovery applications. Novel transcripts, exons, or alternative splicing events can all be detected from RNA-Seq data (Iyer and Chinnaiyan, 2011). One of the genes found to be upregulated in HighFEC animals at 7dpi in 2011 was a novel micro RNA. These small, non-coding RNAs function in transcriptional and post-transcriptional regulation of gene expression (Chen and Rajewsky, 2007) and have been shown to affect inflammatory and immune mediated diseases (O'Connell et al., 2012; Singh et al., 2013). Profiling of micro RNA expression could help to elucidate their role in the response to gastrointestinal nematodes in sheep. Future work could therefore examine the RNA-Seq data for expression and differential expression of a range of non-protein coding RNAs such as micro RNAs, long non-coding RNAs, and short interfering RNAs. The data could also be interrogated to examine novel transcripts and splice variants.

Positive selective pressure in ruminant genomes within genes associated with high and low FEC in Scottish Blackface lambs

The selective pressure caused by gastrointestinal nematodes in ruminants may leave a genetic footprint for directional positive selection, quantifiable by measures of sequence change (as in Morgan et al., 2010). The aim of Chapter 4 was to the take subset of genes, identified as being differentially expressed in the abomasal lymph node of resistant and susceptible Scottish Blackface lambs after infection with the gastrointestinal nematode *T. circumcincta* (Chapter 3), and look for selective pressure variation that may be associated with resistance or susceptibility to gastrointestinal nematodes. Selective pressure variation was examined using a lineage-site specific analysis, focussing on the lineages to Bovidae, *Ovis* (sheep reference and Scottish Blackface), sheep (reference) and the HighFEC and

LowFEC animals. The *Ovis* lineage showed the largest number of gene families with lineage-specific positive selection, whereas no genes were found to be under positive selection uniquely in the Scottish Blackface groups. This analysis was hampered by the use of RNA-Seq reads as a proxy for genomic sequencing, which resulted in no coverage in some regions of the coding sequences being examined. Pathway analysis of the genes under lineage-specific positive selection revealed that these genes were involved in the inflammatory response, dermatological diseases and conditions, and connective tissue disorders. Inflammation plays an important role in both the innate and adaptive immune response to gastrointestinal parasites, with inflammation and immunity inextricably linked (Sutherland and Scott, 2009). As the genes examined have been shown to be differentially expressed in animals with divergent phenotypes for resistance to gastrointestinal nematodes, pathway analysis suggests that the observed positive selection may be a result of selective pressure placed on sheep and cattle by pathogenic infection. While genes that are involved in immunity and have undergone adaptive evolution have been identified in a number of studies in cattle (Lynn et al., 2005; Larson et al., 2006; Babiuk et al., 2007; Freeman et al., 2008; Jann et al., 2008; Takeshima et al., 2009), studies of adaptive evolution in the sheep genome have previously been limited by a lack of sheep coding sequence data and lack of an assembled sheep genome. This chapter therefore represents the first large-scale comparative genomics study of the considerable selective pressure placed on the sheep genome by gastrointestinal nematodes.

Suggestive associations between polymorphisms in the genome and variation in nematode resistance and body weight in Scottish Blackface lambs

Selective pressure can result in polymorphisms that increase fitness, in this case increased resistance to gastrointestinal nematodes. If the selective advantage is great enough these polymorphisms will be swept to fixation in the population, as seen in the species level selective pressure analyses of Chapter 4. Association analyses can be used to identify either causal mutations, or a marker in close linkage disequilibrium (LD) with a causal mutation that can subsequently be used for selection. Markers in the candidate genes identified in the RNA-Seq analysis (Chapter 3) were genotyped in the Scottish Blackface population used in this study. Markers in genomic regions previously identified as associated with nematode resistance in a meta-analysis of three genome-wide association studies (Riggio et al., 2014) were also examined for association in our flock. All lambs born within the Scottish Blackface flock in 2010 and 2011 were genotyped, and an association analysis performed between the genotyped SNPs and FEC (both *Trichostrongyles* and *Nematodirus* FEC1 and FEC2) and weight (14 weeks and 20 weeks) traits. While no SNP reached significance after correction for multiple testing, six SNPs (five from the RNA-Seq cohort and one from the Validation cohort) reached the suggestive level of significance for association with FEC or weight traits. Single markers were estimated to explain a maximum of 0.50% of the additive genetic variance, however the panel of SNPs as whole was useful in explaining

a proportion of the observed heritability of the trait. Animals within and between populations appear to have both shared and unique mechanisms of resistance. These observed results highlight the polygenic nature of resistance to gastrointestinal nematodes. This study was limited by the number of animals available (253) for genotyping. Previous studies have used numbers ranging from 750 to over 1000 individuals (Sallé et al., 2012; Riggio et al., 2013) for the detection of QTL regions. Future work on validating the SNP of interest from this study would require access to a larger population of Scottish Blackface animals, for increased power. As resistance to gastrointestinal nematodes is a complex polygenic trait, controlled by many genes of relatively small effect (Crawford et al., 2006; Kemper et al., 2009), future work on identifying SNP to be used in selection programs will likely focus on utilising a panel of SNP rather than searching for individual causative mutations.

By using the same gene set for both selective pressure analysis and SNP genotyping the commonalities between the two studies can be examined. An intronic variant contained within the albumin (*ALB*) gene was suggestively associated with *Nematodirus* FEC1. The coding sequence of this gene also shows evidence of positive selection in the Bovidae lineage. Albumin is a negative acute phase protein, which has been shown to decline in response to internal challenges, such as infection, inflammation or stress Murata et al. (2004). *ALB* has been associated with GIN infection in four separate studies in sheep (Keane et al., 2006; Knight et al., 2010; Nagaraj et al., 2012; Pemberton et al., 2012). It has been hypothesised that albumin may be constitutively released into the gastric mucus, and may therefore play an innate protective role (Pemberton et al., 2012). The two proteins with the most positively selected sites in the *Ovis* lineage were FRAS1 and DNAH5. A SNP located within the Fraser syndrome 1 gene (*FRAS1*), was suggestively associated with *Nematodirus* FEC1. This SNP codes for a missense variant that results in a change from isoleucine to a valine at codon 1128. While 20 sites in FRAS1 show *Ovis* lineage-specific evidence of positive selection, codon 1128 is not one of them. Positive selection was also observed within DNAH5 (dynein heavy chain 5, axonema). A missense variant within *DNAH5* results in a change from arginine to histidine at amino acid 1078, and is potentially associated with weight at 14 weeks of age. However, this particular SNP is not under positive selection. Despite neither genotyped SNP showing evidence of positive selection, further genotyping could help to elucidate if SNPs within any of the positively selected sites within each gene are more strongly associated with the trait of interest. The evidence of association with traits of interest could be the result of linkage disequilibrium with a causal mutation within the gene.

Conclusions

This study represents the first characterisation of resistance to gastrointestinal nematodes in Scottish Blackface animals in Ireland. Resistance was found to manifest through reduced fecundity in nematodes, although reduced worm burden may also play a role. The anti-

nematode response was mediated, at least in part, by IgA. Transcriptional profiling of the abomasal lymph node during a controlled challenge with *T. circumcincta* indicated that the LowFEC (resistant) Scottish Blackface lambs are generating an immune response to *T. circumcincta* at 7 days post infection, whereas in their HighFEC counterparts this response is delayed until ~14 days post infection. This is in agreement with a previous study of resistant and susceptible breeds in Ireland in which it was found the resistant breed (Texel) generated a more rapid immune response than the susceptible breed (Suffolk). The early immune response generated by the LowFEC animals may therefore be influencing the larval stages of *T. circumcincta*, which could result in the shorter, less fecund adults, as reported in Chapter 2. Selective pressure analysis revealed selection acting on the coding sequence a number of differentially expressed genes in sheep and cattle, potentially as a result of the selective pressure placed on these species by gastrointestinal nematodes. Finally, suggestive associations were found between SNP in differentially expressed genes and FEC traits, however the lack of any genome-wide significant associations may be due to the polygenic nature of the trait.

Bibliography

(2001). Refugia–overlooked as perhaps the most potent factor concerning the development of anthelmintic resistance. *Onderstepoort Journal of Veterinary Research*, 68(1):55–67.

Abbott, K. A., Taylor, M. A., and Stubbings, L. A. (2009). Sustainable Worm Control Strategies for Sheep. Technical Report 3rd Edition, Packington, UK.

Abraham, S. N. and St John, A. L. (2010). Mast cell-orchestrated immunity to pathogens. *Nature Reviews Immunology*, 10(6):440–452.

Ahmed, A. M. (2013). *Immunological and transcriptomic analysis of genetically resistant and susceptible sheep to gastrointestinal nematodes*. PhD thesis, University College Dublin, Dublin, Ireland.

Akey, J. M. (2009). Constructing genomic maps of positive selection in humans: where do we go from here? *Genome Research*, 19(5):711–722.

Albers, G. A. A., Gray, G. D., Piper, L. R., Barker, J. S. F., Jambre, L. F. L., and Barger, I. A. (1987). The genetics of resistance and resilience to *Haemonchus contortus* infection in young merino sheep. *International Journal for Parasitology*, 17(7):1355–1363.

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

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402.

Amarante, A. F., Bricarello, P. A., Rocha, R. A., and Gennari, S. M. (2004). Resistance of Santa Ines, Suffolk and Ile de France sheep to naturally acquired gastrointestinal nematode infections. *Veterinary Parasitology*, 120(1-2):91–106.

Amarante, A. F., Craig, T. M., Ramsey, W. S., Davis, S. K., and Bazer, F. W. (1999). Nematode burdens and cellular responses in the abomasal mucosa and blood of Florida Native, Rambouillet and crossbreed lambs. *Veterinary Parasitology*, 80(4):311–324.

Amarante, A. F. T., Susin, I., Rocha, R. A., Silva, M. V. B., Mendes, C. Q., and Pires, A. V. (2009). Resistance of Santa Ines and crossbred ewes to naturally acquired gastrointestinal nematode infections. *Veterinary Parasitology*, 165(3-4):273–280.

Amills, M., Ramiya, V., Norimine, J., and Lewin, H. A. (1998). The major histocompatibility complex of ruminants. *Revue scientifique et technique (International Office of Epizootics)*, 17(1):108–120.

Anderson, N., Martin, P. J., and Jarrett, R. G. (1991a). Field evaluation of a mixture of albendazole sulphoxide and levamisole against *Ostertagia* and *Trichostrongylus* spp in sheep. *Australian Veterinary Journal*, 68(4):133–136.

Anderson, N., Martin, P. J., and Jarrett, R. G. (1991b). The efficacy of mixtures of albendazole sulphoxide and levamisole against sheep nematodes resistant to benzimidazole and levamisole. *Australian Veterinary Journal*, 68(4):127–132.

Andronicos, N. M., Hunt, P., and Windon, R. (2010). Expression of genes in gastrointestinal and lymphatic tissues during parasite infection in sheep genetically resistant or susceptible to *Trichostrongylus colubriformis* and *Haemonchus contortus*. *International Journal for Parasitology*, 40(4):417–429.

Anisimova, M., Bielawski, J. P., and Yang, Z. (2001). Accuracy and Power of the Likelihood Ratio Test in Detecting Adaptive Molecular Evolution. *Molecular Biology and Evolution*, 18(8):1585–1592.

Anisimova, M., Bielawski, J. P., and Yang, Z. (2002). Accuracy and power of bayes prediction of amino acid sites under positive selection. *Molecular Biology and Evolution*, 19(6):950–958.

Anisimova, M., Bielawski, J. P., and Yang, Z. (2008). Phylogeny-Aware Gap Placement Prevents Errors in Sequence Alignment and Evolutionary Analysis. *Science*, 320(5883):1632–1635.

Anisimova, M., Nielsen, R., and Yang, Z. (2003). Effect of Recombination on the Accuracy of the Likelihood Method for Detecting Positive Selection at Amino Acid Sites. *Genetics*, 164(3):1229–1236.

Anthony, R. M., Rutitzky, L. I., Urban, J. F., Stadecker, M. J., and Gause, W. C. (2007). Protective immune mechanisms in helminth infection. *Nature Reviews Immunology*, 7(12):975–987.

Armour, J., Jarrett, W., and Jennings, F. W. (1966). Experimental *Ostertagia circumcincta* infections in sheep: development and pathogenesis of a single infection. *American Journal of Veterinary Research*, 27(9):1267–1278.

Ashenafi, S., Aderaye, G., Bekele, A., Zewdie, M., Aseffa, G., Hoang, A. T. N., Carow, B., Habtamu, M., Wijkander, M., Rottenberg, M., Aseffa, A., Andersson, J., Svensson, M., and Brighenti, S. (2014). Progression of clinical tuberculosis is associated with a Th2 immune response signature in combination with elevated levels of SOCS3. *Clinical Immunology*, 151(2):84–99.

Athanasiadou, S., Kyriazakis, I., Jackson, F., and Coop, R. L. (2001). Direct anthelmintic effects of condensed tannins towards different gastrointestinal nematodes of sheep: in vitro and in vivo studies. *Veterinary Parasitology*, 99(3):205–219.

Aulchenko, Y. S., Ripke, S., Isaacs, A., and van Duijn, C. M. (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, 23(10):1294–1296.

Babiuk, S., Horseman, B., Zhang, C., Bickis, M., Kusalik, A., Schook, L. B., Abrahamsen, M. S., and Pontarollo, R. (2007). BoLA class I allele diversity and polymorphism in a herd of cattle. *Immunogenetics*, 59(2):167–176.

Baker, R., Watson, T. G., Bisset, S. A., Vlassoff, A., and Douch, P. G. C. (1991). Breeding sheep in New Zealand for resistance to internal parasites, research results and commercial application. In Gray, G. D. and Woolaston, R. R., editors, *Breeding for disease resistance in sheep*, pages 19–32. Australian Wool Corporation, Melbourne, Australia.

Balic, A., Bowles, V. M., Liu, Y. S., and Meeusen, E. N. T. (2003). Local immune responses in sensitized sheep following challenge infection with *Teladorsagia circumcincta*. *Parasite Immunology*, 25(7):375–381.

Balic, A., Bowles, V. M., and Meeusen, E. N. (2000a). Cellular profiles in the abomasal mucosa and lymph node during primary infection with *Haemonchus contortus* in sheep. *Veterinary Immunology and Immunopathology*, 75(1-2):109–120.

Balic, A., Bowles, V. M., and Meeusen, E. N. (2000b). The immunobiology of gastrointestinal nematode infections in ruminants. *Advances in Parasitology*, 45:181–241.

Balic, A., Bowles, V. M., and Meeusen, E. N. T. (2002). Mechanisms of immunity to *Haemonchus contortus* infection in sheep. *Parasite Immunology*, 24(1):39–46.

Balic, A., Cunningham, C. P., and Meeusen, E. N. T. (2006). Eosinophil interactions with *Haemonchus contortus* larvae in the ovine gastrointestinal tract. *Parasite Immunology*, 28(3):107–115.

Banerji, S., Ni, J., Wang, S.-X., Clasper, S., Su, J., Tammi, R., Jones, M., and Jackson, D. G. (1999). LYVE-1, a New Homologue of the CD44 Glycoprotein, Is a Lymph-specific Receptor for Hyaluronan. *The Journal of Cell Biology*, 144(4):789–801.

Bange, F. C., Flohr, T., Buwitt, U., and Bottger, E. C. (1992). An interferon-induced protein with release factor activity is a tryptophanyl-tRNA synthetase. *FEBS Letters*, 300(2):162–166.

Baranov, K. O., Volkova, O. Y., Mechetina, L. V., Chikaev, N. A., Reshetnikova, E. S., Nikulina, G. M., Taranin, A. V., and Najakshin, A. M. (2012). Expression of human B-Cell specific receptor FCRL1 in healthy individuals and in patients with autoimmune diseases. *Molecular Biology*, 46(3):450–456.

Barger, I. A. (1993). Influence of sex and reproductive status on susceptibility of ruminants to nematode parasitism. *International Journal for Parasitology*, 23(4):463–469.

Barnes, E. H., Dobson, R. J., and Barger, I. A. (1995). Worm control and anthelmintic resistance: adventures with a model. *Parasitology Today*, 11(2):56–63.

Barrett, J. C., Hansoul, S., Nicolae, D. L., Cho, J. H., Duerr, R. H., Rioux, J. D., Brant, S. R., Silverberg, M. S., Taylor, K. D., Barmada, M. M., Bitton, A., Dassopoulos, T., Datta, L. W., Green, T., Griffiths, A. M., Kistner, E. O., Murtha, M. T., Regueiro, M. D., Rotter, J. I., Schumm, L. P., Steinhart, A. H., Targan, S. R., Xavier, R. J., Libioulle, C., Sandor, C., Lathrop, M., Belaiche, J., Dewit, O., Gut, I., Heath, S., Laukens, D., Mni, M., Rutgeerts, P., Van Gossum, A., Zelenika, D., Franchimont, D., Hugot, J.-P., de Vos, M., Vermeire, S., Louis, E., Cardon, L. R., Anderson, C. A., Drummond, H., Nimmo, E., Ahmad, T., Prescott, N. J., Onnie, C. M., Fisher, S. A., Marchini, J. L., Ghori, J., Bumpstead, S., Gwilliam, R., Tremelling, M., Deloukas, P., Mansfield, J., Jewell, D., Satsangi, J., Mathew, C. G., Parkes, M., Georges, M., and Daly, M. J. (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genetics*, 40(8):955–962.

Bartram, D. J., Leathwick, D. M., Taylor, M. A., Geurden, T., and Maeder, S. J. (2012). The role of combination anthelmintic formulations in the sustainable control of sheep nematodes. *Veterinary Parasitology*, 186(3-4):151–158.

Beasley, A. M., Kahn, L. P., and Windon, R. G. (2010). The periparturient relaxation of immunity in Merino ewes infected with *Trichostrongylus colubriformis*: Parasitological and immunological responses. *Veterinary Parasitology*, 168(1-2):60–70.

Becker, D., Tetens, J., Brunner, A., Bürstel, D., Ganter, M., Kijas, J. W., Drögemüller, C., and for the International Sheep Genomics Consortium (2010). Microphthalmia in Texel Sheep Is Associated with a Missense Mutation in the Paired-Like Homeodomain 3 PITX3 Gene. *PLoS ONE*, 5(1):e8689.

Beef + Lamb New Zealand (2008). Risk factors for drench resistance in sheep.

Beh, K. J., Callaghan, M. J., Leish, Z., Lenane, I., Windon, R. G., and Maddox, J. F. (2002). A genome scan for quantitative trait loci affecting resistance to *Trichostrongylus colubriformis* in sheep. *Animal Genetics*, 33(2):97.

Benavides, M. V., Weimer, T. A., Borba, M. F. S., Berne, M. E. A., and Sacco, A. M. S. (2002). Association between microsatellite markers of sheep chromosome 5 and faecal egg counts. *Small Ruminant Research*, 46(2–3):97–105.

Benavides, M. V., Weimer, T. A., Borba, M. F. S., Berne, M. E. A., and Sacco, A. M. S. (2009). Genetic analyses of polymorphisms on ovine chromosomes 5 and 20 and their effect on resistance to internal parasites. *Small Ruminant Research*, 83(1-3):67–73.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B Methodological*, 57(1):289–300.

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

Beraldi, D., McRae, A. F., Gratten, J., Pilkington, J. G., Slate, J., Visscher, P. M., and Pemberton, J. M. (2007). Quantitative trait loci (QTL) mapping of resistance to strongyles and coccidia in the free-living Soay sheep (*Ovis aries*). *International Journal for Parasitology*, 37(1):121–129.

Berghen, P., Hilderson, H., Vercruysse, J., and Dorny, P. (1993). Evaluation of pepsinogen, gastrin and antibody response in diagnosing ostertagiasis. *Veterinary Parasitology*, 46:175–195.

Bishop, S. C., Bairden, K., and McKellar, Q. A. (1996). Genetic parameters for faecal egg count following mixed, natural, predominantly *Ostertagia circumcinta* infection and relationships with live weight in young lambs. *Animal Science*, 63:423–428.

Bishop, S. C. and Morris, C. A. (2007). Genetics of disease resistance in sheep and goats. *Small Ruminant Research*, 70(1):48–59.

Bishop, S. C. and Stear, M. J. (1999). Genetic and epidemiological relationships between productivity and disease resistance: gastrointestinal parasite infection in growing lambs. *Animal Science*, 69:515–524.

Bishop, S. C. and Stear, M. J. (2000a). Genetic control of resistance to gastrointestinal parasites in hill sheep. Technical Report Report to MAFF on project LS2204.

Bishop, S. C. and Stear, M. J. (2000b). The use of a gamma-type function to assess the relationship between the number of adult *Teladorsagia circumcincta* and total egg output. *Parasitology*, 121(04):435–440.

Bishop, S. C. and Stear, M. J. (2001). Genetic improvement of resistance to nematodes in terminal sire sheep. Technical Report Report to LINK on project LK0610.

Bisset, S. A., Morris, C. A., McEwan, J. C., and Vlassoff, A. (2001). Breeding sheep in New Zealand that are less reliant on anthelmintics to maintain health and productivity. *New Zealand Veterinary Journal*, 49(6):236–246.

Bisset, S. A., Morris, C. A., Squire, D. R., Hickey, S. M., and Wheeler, M. (1994). Genetics of resilience to nematode parasites in Romney sheep. *New Zealand Journal of Agricultural Research*, 37(4):521–534.

Bisset, S. A., Vlassoff, A., Douch, P. G. C., Jonas, W. E., West, C. J., and Green, R. S. (1996). Nematode burdens and immunological responses following natural challenge in Romney lambs selectively bred for low or high faecal worm egg count. *Veterinary Parasitology*, 61(3–4):249–263.

Blattman, A. N., Kinghorn, B. P., Woolaston, R. R., Gray, G. D., and Beh, K. J. (1993). A search for associations between major histocompatibility complex restriction fragment length polymorphism bands and resistance to *Haemonchus contortus* infection in sheep. *Animal Genetics*, 24(4):277–282.

Bouix, J., Krupinski, J., Rzepecki, R., Nowosad, B., Skrzyzala, I., Roborzynski, M., Fudalewicz-Niemczyk, W., Skalska, M., Malczewski, A., and Gruner, L. (1998). Genetic resistance to gastrointestinal nematode parasites in Polish long-wool sheep. *International Journal for Parasitology*, 28(11):1797–1804.

Breloer, M. and Fleischer, B. (2008). CD83 regulates lymphocyte maturation, activation and homeostasis. *Trends in Immunology*, 29(4):186–194.

Bricarello, Gennari, Vaz, and Echevarria (2004). Worm burden and immunological responses in Corriedale and Crioula Lanada sheep following natural infection with *Haemonchus contortus*. *Small Ruminant Research*, 51(1):9–9.

Bromley, S. K., Mempel, T. R., and Luster, A. D. (2008). Orchestrating the orchestrators: chemokines in control of T cell traffic. *Nature Immunology*, 9(9):970–980.

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., and 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(3):757–773.

Brown, H. D., Matzuk, A. R., Ilves, I. R., Peterson, L. H., Harris, S. A., Sarett, L. H., Egerton, J. R., Yakstis, J. J., Campbell, W. C., and Cuckler, A. C. (1961). Antiparasitic drugs. IV. 2-(4'-thiazolyl)-benzimidazole, a new anthelmintic. *Journal of the American Chemical Society*, 83(7):1764–1765.

Broxmeyer, H. E., Sherry, B., Lu, L., Cooper, S., Oh, K. O., Tekamp-Olson, P., Kwon, B. S., and Cerami, A. (1990). Enhancing and suppressing effects of recombinant murine

macrophage inflammatory proteins on colony formation in vitro by bone marrow myeloid progenitor cells. *Blood*, 76(6):1110–1116.

Brunsdon, R. V. (1964). The effect of nutrition on the establishment and persistence of trichostrongyle infestation. *New Zealand Veterinary Journal*, 12(5):108–111.

Brunsdon, R. V. (1970). Seasonal changes in the level and composition of nematode worm burdens in young sheep. *New Zealand Journal of Agricultural Research*, 13:126–148.

Buddle, B. M., Jowett, G., Green, R. S., Douch, P. G. C., and Risdon, P. L. (1992). Association of blood eosinophilia with the expression of resistance in Romney lambs to nematodes. *International Journal for Parasitology*, 22(7):955–960.

Buitkamp, J., Filmether, P., Stear, M. J., and Epplen, J. T. (1996). Class I and class II major histocompatibility complex alleles are associated with faecal egg counts following natural, predominantly *Ostertagia circumcincta* infection. *Parasitology Research*, 82(8):693–696.

Burgess, C. G. S., Bartley, Y., Redman, E., Skuce, P. J., Nath, M., Whitelaw, F., Tait, A., Gilleard, J. S., and Jackson, F. (2012). A survey of the trichostrongylid nematode species present on UK sheep farms and associated anthelmintic control practices. *Veterinary Parasitology*, 189(2-4):299–307.

Burke, J. M. and Miller, J. E. (2004). Relative resistance to gastrointestinal nematode parasites in Dorper, Katahdin, and St. Croix lambs under conditions encountered in the southeastern region of the United States. *Small Ruminant Research*, 54(1):43–51.

Bystry, R. S., Aluvihare, V., Welch, K. A., Kallikourdis, M., and Betz, A. G. (2001). B cells and professional APCs recruit regulatory T cells *via* CCL4. *Nature Immunology*, 2(12):1126–1132.

Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15):1972–1973.

Cardia, D. F. F., Rocha-Oliveira, R. A., Tsunemi, M. H., and Amarante, A. F. T. (2011). Immune response and performance of growing Santa Ines lambs to artificial *Trichostrongylus colubriformis* infections. *Veterinary Parasitology*, 182(2–4):248–258.

Caunt, C. J. and Keyse, S. M. (2013). Dual-specificity MAP kinase phosphatases (MKPs). *FEBS Journal*, 280(2):489–504.

Chabala, J. C. J., Mrozik, H. H., Tolman, R. L. R., Eskola, P. P., Lusi, A. A., Peterson, L. H. L., Woods, M. F. M., Fisher, M. H. M., Campbell, W. C. W., Egerton, J. R. J., and Ostlind, D. A. D. (1980). Ivermectin, a new broad-spectrum antiparasitic agent. *Journal of Medicinal Chemistry*, 23(10):1134–1136.

Charlier, J., Dorny, P., Levecke, B., Demeler, J., von Samson-Himmelstjerna, G., Hoglund, J., and Vercruysse, J. (2011). Serum pepsinogen levels to monitor gastrointestinal nematode infections in cattle revisited. *Research in Veterinary Science*, 90(3):451–456.

Charon, K. M., Moskwa, B., Rutkowski, R., Gruszczynska, J., and Swiderek, W. (2002). Microsatellite polymorphism in DRB1 gene (MHC class II) and its relation to nematode faecal egg count in Polish Heath sheep. *Journal of Animal and Feed Sciences*, 11:47–58.

Chen, K. and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. *Nature Reviews Genetics*, 8(2):93–103.

Chen, W.-M. and Abecasis, G. R. (2007). Family-based association tests for genomewide association scans. *The American Journal of Human Genetics*, 81(5):913–926.

Clarke, R. A., Burn, A. L., Lenane, I., Windon, R. G., and Beh, K. J. (2001). Molecular analysis and nematode resistance association of a polymorphism at the 5' end of the sheep *IgE* gene. *Veterinary Immunology and Immunopathology*, 79(1–2):15–29.

Coles, G. C. and Roush, R. T. (1992). Slowing the spread of anthelmintic resistant nematodes of sheep and goats in the United Kingdom. *Veterinary Record*, 130(23):505–510.

Coltman, D. W., Wilson, K., Pilkington, J. G., Stear, M. J., and Pemberton, J. M. (2001). A microsatellite polymorphism in the gamma interferon gene is associated with resistance to gastrointestinal nematodes in a naturally-parasitized population of Soay sheep. *Parasitology*, 122(05):571–582.

Conway, D. P. (1964). Variance in the effectiveness of thiabendazole against *Haemonchus contortus* in sheep. *American Journal of Veterinary Research*, 25:844–846.

Coop, R. L., Graham, R. B., Jackson, F., Wright, S. E., and Angus, K. W. (1985). Effect of experimental *Ostertagia circumcincta* infection on the performance of grazing lambs. *Research in Veterinary Science*, 38(3):282–287.

Coop, R. L., Huntley, J. F., and Smith, W. D. (1995). Effect of dietary protein supplementation on the development of immunity to *Ostertagia circumcincta* in growing lambs. *Research in Veterinary Science*, 59(1):24–29.

Coop, R. L. and Kyriazakis, I. (1999). Nutrition-parasite interaction. *Veterinary Parasitology*, 84(3-4):187–204.

Coop, R. L., Sykes, A. R., and Angus, K. W. (1982). The effect of three levels of intake of *Ostertagia circumcincta* larvae on growth rate, food intake and body composition of growing lambs. *The Journal of Agricultural Science*, 98(02):247.

Cooper, D. M., Petchkovsky, D. V., Hackett, T. L., Knight, D. A., and Granville, D. J. (2011). Granzyme K Activates Protease-Activated Receptor-1. *PLoS ONE*, 6(6):e21484.

Cox, L. A., Glenn, J. P., Spradling, K. D., Nijland, M. J., Garcia, R., Nathanielsz, P. W., and Ford, S. P. (2012). A genome resource to address mechanisms of developmental programming: determination of the fetal sheep heart transcriptome. *The Journal of Physiology*, 590(12):2873–2884.

Craig, N. M., Miller, H. R. P., Smith, W. D., and Knight, P. A. (2007). Cytokine expression in naïve and previously infected lambs after challenge with *Teladorsagia circumcincta*. *Veterinary Immunology and Immunopathology*, 120(1-2):47–54.

Crawford, A. M., Dodds, K. G., and McEwan, J. C. (2000). DNA Markers, Genetic Maps and the Identification of QTL: General Principles. In Axford, R. B. S. N. F. O. J., editor, *Breeding for Disease Resistance in Farm Animals*, pages 3–26. CABI Publishing, New York.

Crawford, A. M., Paterson, K. A., Dodds, K. G., Diez-Tascon, C., Williamson, P. A., Roberts Thomson, M., Bisset, S. A., Beattie, A. E., Greer, G. J., Green, R. S., Wheeler, R., Shaw, R. J., Knowler, K., and McEwan, J. C. (2006). Discovery of quantitative trait loci for resistance to parasitic nematode infection in sheep: I. Analysis of outcross pedigrees. *BMC Genomics*, 7:178.

Crofton, H. D. (1965). Ecology and Biological Plasticity of Sheep Nematodes. I. the Effect of Temperature on the Hatching of Eggs of Some Nematode Parasites of Sheep. *The Cornell Veterinarian*, 55:242–250.

Cummins, L. J., Thompson, R. L., Yong, W. K., Riffkin, G. G., Goddard, M. E., Callinan, A., and Saunders, M. J. (1991). Genetics of *Ostertagia* selection lines. In Gray, G. D. and Woolaston, R. R., editors, *Breeding for Disease Resistance in Sheep*, pages 11–18. Australian Wool Corporation, Melbourne, Australia.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*, 27(8):1164–1165.

Daub, J. T., Hofer, T., Cutivet, E., Dupanloup, I., Quintana-Murci, L., Robinson-Rechavi, M., and Excoffier, L. (2013). Evidence for polygenic adaptation to pathogens in the human genome. *Molecular Biology and Evolution*, 30(7):1544–1558.

Davies, G., Stear, M. J., Benothman, M., Abuagob, O., Kerr, A., Mitchell, S., and Bishop, S. C. (2006). Quantitative trait loci associated with parasitic infection in Scottish blackface sheep. *Heredity*, 96(3):252–258.

Davies, G., Stear, M. J., and Bishop, S. C. (2005). Genetic relationships between indicator traits and nematode parasite infection levels in 6-month-old lambs. *Animal Science*, 80(02):143–150.

Davis, R. S., Wang, Y.-H., Kubagawa, H., and Cooper, M. D. (2001). Identification of a family of Fc receptor homologs with preferential B cell expression. *Proceedings of the National Academy of Sciences*, 98(17):9772–9777.

Dawicki, W. and Marshall, J. S. (2007). New and emerging roles for mast cells in host defence. *Current Opinion in Immunology*, 19(1):31–38.

de Veer, M. J., Kemp, J. M., and Meeusen, E. N. T. (2007). The innate host defence against nematode parasites. *Parasite Immunology*, 29(1):1–9.

Debes, G. F., Dahl, M. E., Mahiny, A. J., Bonhagen, K., Campbell, D. J., Siegmund, K., Erb, K. J., Lewis, D. B., Kamradt, T., and Hamann, A. (2006). Chemotactic Responses of IL-4-, IL-10-, and IFN-γ-Producing CD4+ T Cells Depend on Tissue Origin and Microbial Stimulus. *The Journal of Immunology*, 176(1):557–566.

deSchoolmeester, M. L., Martinez-Pomares, L., Gordon, S., and Else, K. J. (2009). The mannose receptor binds *Trichuris muris* excretory/secretory proteins but is not essential for protective immunity. *Immunology*, 126(2):246–255.

Diez, J. J. and Iglesias, P. (2003). The role of the novel adipocyte-derived hormone adiponectin in human disease. *European Journal of Endocrinology*, 148(3):293–300.

Diez-Tascon, C., Keane, O. M., Wilson, T., Zadissa, A., Hyndman, D. L., McEwan, J. C., and Crawford, A. M. (2005). Microarray analysis of selection lines from outbred populations to identify genes involved with nematode parasite resistance in sheep. *Physiological Genomics*, 21(1):59–69.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., Guernec, G., Jagla, B., Jouneau, L., Laloë, D., Le Gall, C., Schaëffer, B., Le Crom, S., Guedj, M., Jaffrézic, F., and French StatOmique Consortium (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in Bioinformatics*, 14(6):671–683.

Dion, C., Carter, C., Hepburn, L., Coadwell, W. J., Morgan, G., Graham, M., Pugh, N., Anderson, G., Butcher, G. W., and Miller, J. R. (2005). Expression of the Ian family of putative GTPases during T cell development and description of an Ian with three sets of GTP/GDP-binding motifs. *International Immunology*, 17(9):1257–1268.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21.

Dobson, R. J., Hosking, B. C., Besier, R. B., Love, S., Larsen, J. W. A., Rolfe, P. F., and Bailey, J. N. (2011). Minimising the development of anthelmintic resistance, and

optimising the use of the novel anthelmintic monepantel, for the sustainable control of nematode parasites in Australian sheep grazing systems. *Australian Veterinary Journal*, 89(5):160–166.

Doligalska, M., Moskwa, B., and Stear, M. J. (1999). Relationships among peripheral eosinophilia, eosinophil peroxidase activity, interleukin-5 concentration and faecal nematode egg count during natural, mixed gastrointestinal nematode infection. *Veterinary Immunology and Immunopathology*, 70(3-4):299–308.

Dominik, S. (2005). Quantitative trait loci for internal nematode resistance in sheep: a review. *Genetics Selection Evolution*, 37 Suppl 1:S83–96.

Dominik, S., Hunt, P. W., McNally, J., Murrell, A., Hall, A., and Purvis, I. W. (2010). Detection of quantitative trait loci for internal parasite resistance in sheep. I. Linkage analysis in a Romney x Merino sheep backcross population. *Parasitology*, 137(8):1275–1282.

Donald, A. D., Morley, F. H. W., Waller, P. J., Axelsen, A., Dobson, R. J., and Donnelly, J. R. (1982). Effects of reproduction, genotype and anthelmintic treatment of ewes on Ostertagia spp. populations. *International Journal for Parasitology*, 12(5):403–411.

Dorny, P., Shaw, D. J., and Vercruysse, J. (1999). The determination at housing of exposure to gastrointestinal nematode infections in first-grazing season calves. *Veterinary Parasitology*, 80(4):325–340.

Douch, P. G., Harrison, G. B., Buchanan, L. L., and Brunsdon, R. V. (1984). Relationship of histamine in tissues and antiparasitic substances in gastrointestinal mucus to the development of resistance to *Trichostrongyle* infections in young sheep. *Veterinary Parasitology*, 16(3-4):273–288.

Douch, P. G. C., Green, R. S., Morris, C. A., McEwan, J. C., and Windon, R. G. (1996). Phenotypic markers for selection of nematode-resistant sheep. *International Journal for Parasitology*, 26(8-9):899–911.

Douch, P. G. C., Green, R. S., and Risdon, P. L. (1994). Antibody responses of sheep to challenge with *Trichostrongylus colubriformis* and the effect of dexamethasone treatment. *International Journal for Parasitology*, 24(7):921–928.

Draber, P., Vonkova, I., Stepanek, O., Hrdinka, M., Kucova, M., Skopcova, T., Otahal, P., Angelisova, P., Horejsi, V., Yeung, M., Weiss, A., and Brdicka, T. (2011). SCIMP, a Transmembrane Adaptor Protein Involved in Major Histocompatibility Complex Class II Signaling. *Molecular and Cellular Biology*, 31(22):4550–4562.

Dubuquoy, L., Dharancy, S., Nutten, S., Pettersson, S., Auwerx, J., and Desreumaux, P. (2002). Role of peroxisome proliferator-activated receptor γ and retinoid X receptor heterodimer in hepatogastroenterological diseases. *The Lancet*, 360(9343):1410–1418.

Dukkipati, V. S., Blair, H. T., Garrick, D. J., and Murray, A. (2006). '*Ovar-Mhc*' – Ovine major histocompatibility complex: Role in genetic resistance to diseases. *New Zealand Veterinary Journal*, 54(4):153–160.

Dunnington, E. A. and Siegel, P. B. (1996). Long-term divergent selection for eight-week body weight in white Plymouth rock chickens. *Poultry science*, 75(10):1168–1179.

Ebstein, F., Kloetzel, P.-M., Krüger, E., and Seifert, U. (2012). Emerging roles of immunoproteasomes beyond MHC class I antigen processing. *Cellular and Molecular Life Sciences*, 69(15):2543–2558.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5):1792–1797.

Engstrom, P. G., Steijger, T., Sipos, B., Grant, G. R., Kahles, A., Consortium, T. R., Ratsch, G., Goldman, N., Hubbard, T. J., Harrow, J., Guigo, R., and Bertone, P. (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. *Nature Methods*, 10(12):1185–1191.

Erives, A. (2011). A model of proto-anti-codon RNA enzymes requiring L-amino acid homochirality. *Journal of Molecular Evolution*, 73(1-2):10–22.

Eysker, M. and Kooyman, F. N. J. (1993). Notes on necropsy and herbage processing techniques for gastrointestinal nematodes of ruminants. *Veterinary Parasitology*, 46(1–4):205–213.

Eysker, M. and Ploeger, H. W. (2000). Value of present diagnostic methods for gastrointestinal nematode infections in ruminants. *Parasitology*, 120(07):109–119.

Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Longman, Harlow, Essex, England, 4th ed. edition.

Fan, R., Xie, J., Bai, J., Wang, H., Tian, X., Bai, R., Jia, X., Yang, L., Song, Y., Herrid, M., Gao, W., He, X., Yao, J., Smith, G. W., and Dong, C. (2013). Skin transcriptome profiles associated with coat color in sheep. *BMC Genomics*, 14(1):389.

Fielding, P. E. and Fielding, C. J. (1980). A cholesteryl ester transfer complex in human plasma. *Proceedings of the National Academy of Sciences*, 77(6):3327–3330.

Fife, B. T. and Pauken, K. E. (2011). The role of the PD-1 pathway in autoimmunity and peripheral tolerance. *Annals of the New York Academy of Sciences*, 1217(1):45–59.

Filali, H., Martin-Burriel, I., Harders, F., Varona, L., Hedman, C., Mediano, D., Monzon, M., Bossers, A., Badiola, J., and Bolea, R. (2014). Gene expression profiling of mesenteric lymph nodes from sheep with natural scrapie. *BMC Genomics*, 15(1):59.

Flajnik, M. F. and Kasahara, M. (2001). Comparative Genomics of the MHC: Glimpses into the Evolution of the Adaptive Immune System. *Immunity*, 15(3):351–362.

Fletcher, W. and Yang, Z. (2010). The Effect of Insertions, Deletions, and Alignment Errors on the Branch-Site Test of Positive Selection. *Molecular Biology and Evolution*, 27(10):2257–2267.

Flicek, P., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Girón, C. G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juettemann, T., Kähäri, A. K., Keenan, S., Kulesha, E., Martin, F. J., Maurel, T., McLaren, W. M., Murphy, D. N., Nag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H. S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S. J., Vullo, A., Wilder, S. P., Wilson, M., Zadissa, A., Aken, B. L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T. J. P., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D. R., and Searle, S. M. J. (2014). Ensembl 2014. *Nucleic Acids Research*, 42(D1):D749–D755.

Foley, C. (2014). *Divergent Endometrial Inflammation, Calcium Signalling and Systemic Immune Response in Cows with Healthy and Inflamed Endometria Early Postpartum*. PhD thesis, Trinity College Dublin.

Freeman, A. R., Lynn, D. J., Murray, C., and Bradley, D. G. (2008). Detecting the effects of selection at the population level in six bovine immune genes. *BMC Genetics*, 9:62.

Fuhrman, J. A. and Piessens, W. F. (1985). Chitin synthesis and sheath morphogenesis in *Brugia malayi* microfilariae. *Molecular and Biochemical Parasitology*, 17(1):93–104.

Galtier, N. and Duret, L. (2007). Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. *Trends in Genetics*, 23(6):273–277.

Gao, J., Liu, K., Liu, H., Blair, H. T., Li, G., Chen, C., Tan, P., and Ma, R. Z. (2010). A complete DNA sequence map of the ovine Major Histocompatibility Complex. *BMC Genomics*, 11(1):466.

Gawenis, L. R., Greeb, J. M., Prasad, V., Grisham, C., Sanford, L. P., Doetschman, T., Andringa, A., Miller, M. L., and Shull, G. E. (2005). Impaired Gastric Acid Secretion in Mice with a Targeted Disruption of the NHE4 Na+/H+ Exchanger. *Journal of Biological Chemistry*, 280(13):12781–12789.

Geijtenbeek, T. B. H. and Gringhuis, S. I. (2009). Signalling through C-type lectin receptors: shaping immune responses. *Nature Reviews Immunology*, 9(7):465–479.

Gilfillan, A. M. and Tkaczyk, C. (2006). Integrated signalling pathways for mast-cell activation. *Nature Reviews Immunology*, 6(3):218–230.

Gilmour, A. R., Gogel, B. J., Cullis, B. R., and Thompson, R. (2009). ASReml User Guide Release 3.0. pages 1–398.

Glass, E. J. (2012). The molecular pathways underlying host resistance and tolerance to pathogens. *Frontiers in Genetics*, 3:263.

Goddard, M. E., Wray, N. R., Verbyla, K., and Visscher, P. M. (2009). Estimating Effects and Making Predictions from Genome-Wide Marker Data. *Statistical Science*, 24(4):517–529.

Goldfinch, G. M., Smith, W. D., Imrie, L., McLean, K., Inglis, N. F., and Pemberton, A. D. (2008). The proteome of gastric lymph in normal and nematode infected sheep. *Proteomics*, 8(9):1909–1918.

Goldman, N. and Yang, Z. (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular Biology and Evolution*, 11(5):725–736.

Gonzalez, L. C., Loyet, K. M., Calemine-Fenaux, J., Chauhan, V., Wranik, B., Ouyang, W., and Eaton, D. L. (2005). A coreceptor interaction between the CD28 and TNF receptor family members B and T lymphocyte attenuator and herpesvirus entry mediator. *Proceedings of the National Academy of Sciences*, 102(4):1116–1121.

Good, B., Grennan, E. J., Crowley, B. A., and Hanrahan, J. P. (2001). *Grazing system and anthelmintic treatment of ewes on parasite challenge and lamb growth*. End of Project Reports: Sheep Series No. 17. Project No. 4471. Teagasc Sheep Research Centre, Athenry.

Good, B., Hanrahan, J. P., Crowley, B. A., and Mulcahy, G. (2006). Texel sheep are more resistant to natural nematode challenge than Suffolk sheep based on faecal egg count and nematode burden. *Veterinary Parasitology*, 136(3–4):317–327.

Good, B., Hanrahan, J. P., de Waal, D. T., Patten, T., Kinsella, A., and Lynch, C. O. (2012). Anthelmintic-resistant nematodes in Irish commercial sheep flocks- the state of play. *Irish Veterinary Journal*, 65(1):21–21.

Gossmann, T. I., Keightley, P. D., and Eyre-Walker, A. (2012). The Effect of Variation in the Effective Population Size on the Rate of Adaptive Molecular Evolution in Eukaryotes. *Genome Biology and Evolution*, 4(5):658–667.

Gossner, A. G., Venturina, V. M., Shaw, D. J., Pemberton, J. M., and Hopkins, J. (2012). Relationship between susceptibility of Blackface sheep to *Teladorsagia circumcincta* infection and an inflammatory mucosal T cell response. *Veterinary Research*, 43(1):26.

Gossner, A. G., Wilkie, H., Joshi, A., and Hopkins, J. (2013). Exploring the abomasal lymph node transcriptome for genes associated with resistance to the sheep nematode *Teladorsagia circumcincta*. *Veterinary Research*, 44(1):68.

Gratchev, A., Schmuttermaier, C., Mamidi, S., Gooi, L., Goerdt, S., and Kzhyshkowska, J. (2008). Expression of Osteoarthritis Marker YKL-39 is Stimulated by Transforming Growth Factor Beta (TGF-beta) and IL-4 in Differentiating Macrophages. *Biomarker Insights*, 3:39–44.

Greeff, J. C. and Karlsson, L. (2006). Breeding for worm resistance–whole farm benefits. In Cronje, P. B. and Maxwell, D., editors, *Wool meets Meat - Tools for a modern sheep enterprise*, pages 102–108. The Australian Sheep Industry CRC, Orange, Australia.

Greenberg, A. S., Egan, J. J., Wek, S. A., Garty, N. B., Blanchette-Mackie, E. J., and Londos, C. (1991). Perilipin, a major hormonally regulated adipocyte-specific phosphoprotein associated with the periphery of lipid storage droplets. *Journal of Biological Chemistry*, 266(17):11341–11346.

Grenfell, B. T., Wilson, K., Isham, V. S., Boyd, H. E., and Dietz, K. (1995). Modelling patterns of parasite aggregation in natural populations: trichostrongylid nematode-ruminant interactions as a case study. *Parasitology*, 111 Suppl:S135–51.

Gruner, L., Aumont, G., Getachew, T., Brunel, J. C., Pery, C., Cognié, Y., and Guérin, Y. (2003). Experimental infection of Black Belly and INRA 401 straight and crossbred sheep with trichostrongyle nematode parasites. *Veterinary Parasitology*, 116(3):239–249.

Gulland, F. M. and Fox, M. (1992). Epidemiology of nematode infections of Soay sheep (*Ovis aries* L.) on St Kilda. *Parasitology*, 105 (Pt 3):481–492.

Gutierrez-Gil, B., Perez, J., Alvarez, L., Martínez-Valladares, M., de la Fuente, L.-F., Bayon, Y., Meana, A., Primitivo, F. S., Rojo-Vazquez, F.-A., and Arranz, J.-J. (2009a). Quantitative trait loci for resistance to trichostrongylid infection in Spanish Churra sheep. *Genetics Selection Evolution*, 41(1):46.

Gutierrez-Gil, B., Pérez, J., de la Fuente, L. F., Meana, A., Martínez-Valladares, M., San Primitivo, F., Rojo-Vázquez, F. A., and Arranz, J. J. (2009b). Genetic parameters for resistance to trichostrongylid infection in dairy sheep. *Animal*, 4(04):505.

Haile, A., Tembely, S., Anindo, D. O., Mukasa-Mugerwa, E., Rege, J. E. O., Yami, A., and Baker, R. (2002). Effects of breed and dietary protein supplementation on the responses to gastrointestinal nematode infections in Ethiopian sheep. *Small Ruminant Research*, 44(3):247–261.

Halliday, A. M., Routledge, C. M., Smith, S. K., Matthews, J. B., and Smith, W. D. (2007). Parasite loss and inhibited development of *Teladorsagia circumcincta* in relation to the kinetics of the local IgA response in sheep. *Parasite Immunology*, 29(8):425–434.

Han, G., Ye, M., Zhou, H., Jiang, X., Feng, S., Jiang, X., Tian, R., Wan, D., Zou, H., and Gu, J. (2008). Large-scale phosphoproteome analysis of human liver tissue by enrichment

and fractionation of phosphopeptides with strong anion exchange chromatography. *Proteomics*, 8(7):1346–1361.

Hanrahan, J. P. and Crowley, B. A. (1999). Evidence for breed differences in resistance to nematode infection. In *Sheep and Goat Commission of European Association for Animal Production (50th Annual Meeting)*, Zurich, Switzerland.

Hansen, J. D., Vojtech, L. N., and Laing, K. J. (2011). Sensing disease and danger: a survey of vertebrate PRRs and their origins. *Developmental and Comparative Immunology*, 35(9):886–897.

Harriman, G. R., Kunimoto, D. Y., Elliott, J. F., Paetkau, V., and Strober, W. (1988). The role of IL-5 in IgA B cell differentiation. *The Journal of Immunology*, 140(9):3033–3039.

Harrison, G. B. L., Pulford, H. D., Doolin, E. E., Pernthaner, A., Shoemaker, C. B., and Hein, W. R. (2008). Antibodies to surface epitopes of the carbohydrate larval antigen CarLA are associated with passive protection in strongylid nematode challenge infections. *Parasite Immunology*, 30(11-12):577–584.

Harrison, G. B. L., Pulford, H. D., Hein, W. R., Barber, T. K., Shaw, R. J., McNeill, M., Wakefield, S. J., and Shoemaker, C. B. (2003a). Immune rejection of *Trichostrongylus colubriformis* in sheep; a possible role for intestinal mucus antibody against an L3-specific surface antigen. *Parasite Immunology*, 25(1):45–53.

Harrison, G. B. L., Pulford, H. D., Hein, W. R., Severn, W. B., and Shoemaker, C. B. (2003b). Characterization of a 35-kDa carbohydrate larval antigen (CarLA) from T*richostrongylus colubriformis*; a potential target for host immunity. *Parasite Immunology*, 25(2):79–86.

Hassan, M., Good, B., Hanrahan, J. P., Campion, D., Sayers, G., Mulcahy, G., and Sweeney, T. (2011a). The dynamic influence of the *DRB1*1101* allele on the resistance of sheep to experimental Teladorsagia circumcincta infection. *Veterinary Research*, 42(1):46.

Hassan, M., Hanrahan, J. P., Good, B., Mulcahy, G., and Sweeney, T. (2011b). A differential interplay between the expression of Th1/Th2/Treg related cytokine genes in *Teladorsagia circumcincta* infected *DRB1*1101* carrier lambs. *Veterinary Research*, 42(1):45.

Henderson, N. G. and Stear, M. J. (2006). Eosinophil and IgA responses in sheep infected with *Teladorsagia circumcincta*. *Veterinary Immunology and Immunopathology*, 112(1-2):62–66.

Hendrix, C. M. (1998). Common Laboratory Procedures for Diagnosing Parasites. In *Diagnostic Veterinary Parasitology*, page 321. Mosby, St. Louis, MO, USA.

Hirschowitz, B. I. (1955). Pepsinogen in the blood. *Journal of Laboratory and Clinical Medicine*, 46(4):568–579.

Holden, M., Deng, S., Wojnowski, L., and Kulle, B. (2008). GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies. *Bioinformatics*, 24(23):2784–2785.

Hu, Z.-L., Park, C. A., Wu, X.-L., and Reecy, J. M. (2013). Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research*, 41(Database issue):D871–9.

Huang, J. T., Welch, J. S., Ricote, M., Binder, C. J., Willson, T. M., Kelly, C., Witztum, J. L., Funk, C. D., Conrad, D., and Glass, C. K. (1999). Interleukin-4-dependent production of PPAR-gamma ligands in macrophages by 12/15-lipoxygenase. *Nature*, 400(6742):378–382.

Hughes, A. and Friedman, R. (2004). Recent Mammalian Gene Duplications: Robust Search for Functionally Divergent Gene Pairs. *Journal of Molecular Evolution*, 59(1):114–120.

Hughes, A. L. and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335(6186):167–170.

Hughes, A. L. and Yeager, M. (1998). Natural selection at major histocompatibility complex loci of vertebrates. *Annual Review of Genetics*, 32(1):415–435.

Hunt, P. W., McEwan, J. C., and Miller, J. E. (2008). Future perspectives for the implementation of genetic markers for parasite resistance in sheep. *Tropical Biomedicine*, 25(1 Suppl):18–33.

Huntley, J. F., Gibson, S., Brown, D., Smith, W. D., Jackson, F., and Miller, H. R. (1987). Systemic release of a mast cell proteinase following nematode infections in sheep. *Parasite Immunology*, 9(5):603–614.

Huntley, J. F., Redmond, J., Welfare, W., Brennan, G., Jackson, F., Kooyman, F., and Vervelde, L. (2001). Studies on the immunoglobulin E responses to *Teladorsagia circumcincta* in sheep: purification of a major high molecular weight allergen. *Parasite Immunology*, 23(5):227–235.

Hurst, L. D. (2009). Fundamental concepts in genetics: Genetics and the understanding of selection. *Nature Reviews Genetics*, 10(2):83–93.

Ibañez-Tallon, I., Heintz, N., and Omran, H. (2003). To beat or not to beat: roles of cilia in development and disease. *Human Molecular Genetics*, 12(suppl 1):R27–R35.

Ihara, E., Beck, P. L., Chappellaz, M., Wong, J., Medlicott, S. A., and MacDonald, J. A. (2009). Mitogen-Activated Protein Kinase Pathways Contribute to Hypercontractility and

Increased Ca2+ Sensitization in Murine Experimental Colitis. *Molecular Pharmacology*, 75(5):1031–1041.

Ingham, A., Menzies, M., Hunt, P., Reverter, A., Windon, R., and Andronicos, N. M. (2011). Divergent ghrelin expression patterns in sheep genetically resistant or susceptible to gastrointestinal nematodes. *Veterinary Parasitology*, 181(2–4):194–202.

Ingham, A., Reverter, A., Windon, R., Hunt, P., and Menzies, M. (2008). Gastrointestinal nematode challenge induces some conserved gene expression changes in the gut mucosa of genetically resistant sheep. *International Journal for Parasitology*, 38(3-4):431–442.

Ito, M., Nagasawa, M., Hara, T., Ide, T., and Murakami, K. (2010). Differential roles of CIDEA and CIDEC in insulin-induced anti-apoptosis and lipid droplet formation in human adipocytes. *Journal of Lipid Research*, 51(7):1676–1684.

Iyer, M. K. and Chinnaiyan, A. M. (2011). RNA-Seq unleashed. *Nature Biotechnology*, 29(7):599–600.

Jackson, F., Bartley, D., Bartley, Y., and Kenyon, F. (2009). Worm control in sheep in the future. *Small Ruminant Research*, 86(1-3):40–45.

Jackson, F., Greer, A. W., Huntley, J., Bartley, D. J., Stanley, A., Stenhouse, L., Stankiewicz, M., and Sykes, A. R. (2004). Studies using *Teladorsagia circumcincta* in an in vitro direct challenge method using abomasal tissue explants. *Veterinary Parasitology*, 124(1-2):73–89.

Jager, M., Ott, C.-E., Grunhagen, J., Hecht, J., Schell, H., Mundlos, S., Duda, G., Robinson, P., and Lienau, J. (2011). Composite transcriptome assembly of RNA-seq data in a sheep model for delayed bone healing. *BMC Genomics*, 12(1):158.

Jann, O., Werling, D., Chang, J.-S., Haig, D., and Glass, E. (2008). Molecular evolution of bovine Toll-like receptor 2 suggests substitutions of functional relevance. *BMC Evolutionary Biology*, 8(1):288.

Janssen, M., Weimann, C., Gauly, M., and Erhardt, G. (2002). Associations between infections with *Haemonchus contortus* and genetic markers on ovine chromosome 20. In *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, pages 13–11, Montpellier, France.

Jeimy, S. B., Tasneem, S., Cramer, E. M., and Hayward, C. P. M. (2008). Multimerin 1. *Platelets*, 19(2):83–95.

Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J. F., Faraut, T., Wu, C., Muzny, D. M., Li, Y., Zhang, W., Stanton, J.-A., Brauning, R., Barris, W. C., Hourlier, T., Aken, B. L., Searle, S. M. J., Adelson, D. L., Bian, C., Cam, G. R., Chen, Y., Cheng, S., DeSilva, U.,

Dixen, K., Dong, Y., Fan, G., Franklin, I. R., Fu, S., Fuentes-Utrilla, P., Guan, R., Highland, M. A., Holder, M. E., Huang, G., Ingham, A. B., Jhangiani, S. N., Kalra, D., Kovar, C. L., Lee, S. L., Liu, W., Liu, X., Lu, C., Lv, T., Mathew, T., McWilliam, S., Menzies, M., Pan, S., Robelin, D., Servin, B., Townley, D., Wang, W., Wei, B., White, S. N., Yang, X., Ye, C., Yue, Y., Zeng, P., Zhou, Q., Hansen, J. B., Kristiansen, K., Gibbs, R. A., Flicek, P., Warkup, C. C., Jones, H. E., Oddy, V. H., Nicholas, F. W., McEwan, J. C., Kijas, J. W., Wang, J., Worley, K. C., Archibald, A. L., Cockett, N., Xu, X., Wang, W., and Dalrymple, B. P. (2014). The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*, 344(6188):1168–1173.

Johnston, S. E., McEwan, J. C., Pickering, N. K., Kijas, J. W., Beraldi, D., Pilkington, J. G., Pemberton, J. M., and Slate, J. (2011). Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, 20(12):2555–2566.

Jung, M.-Y., Park, S.-Y., and Kim, I.-S. (2007). Stabilin-2 is involved in lymphocyte adhesion to the hepatic sinusoidal endothelium via the interaction with αMβ2 integrin. *Journal of Leukocyte Biology*, 82(5):1156–1165.

Kadarmideen, H. N., Watson-Haigh, N. S., and Andronicos, N. M. (2011). Systems biology of ovine intestinal parasite resistance: disease gene modules and biomarkers. *Molecular BioSystems*, 7(1):235–246.

Kahn, L. P., Knox, M. R., Gray, G. D., Lea, J. M., and Walkden-Brown, S. W. (2003). Enhancing immunity to nematode parasites in single-bearing Merino ewes through nutrition and genetic selection. *Veterinary Parasitology*, 112(3):211–225.

Kaminsky, R., Ducray, P., Jung, M., Clover, R., Rufener, L., Bouvier, J., Weber, S. S., Wenger, A., Wieland-Berghausen, S., Goebel, T., Gauvry, N., Pautrat, F., Skripsky, T., Froelich, O., Komoin-Oka, C., Westlund, B., Sluder, A., and Mäser, P. (2008). A new class of anthelmintics effective against drug-resistant nematodes. *Nature*, 452(7184):176–180.

Kane, M., Yadav, S. S., Bitzegeio, J., Kutluay, S. B., Zang, T., Wilson, S. J., Schoggins, J. W., Rice, C. M., Yamashita, M., Hatziioannou, T., and Bieniasz, P. D. (2013). MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature*, 502(7472):563–566.

Kanobana, K., Ploeger, H. W., and Vervelde, L. (2002). Immune expulsion of the trichostrongylid *Cooperia oncophora* is associated with increased eosinophilia and mucosal IgA. *International Journal for Parasitology*, 32(11):1389–1398.

Kaplan, R. M. (2004). Drug resistance in nematodes of veterinary importance: a status report. *Trends in Parasitology*, 20(10):477–481.

Kates, K. C. K., Colglazier, M. L. M., Enzie, F. D. F., Lindahl, I. L. I., and Samuelson, G. G. (1971). Comparative activity of thiabendazole, levamisole, and parbendazole against natural infections of helminths in sheep. *Parasitology Research*, 57(2):356–362.

Katoh, K., Misawa, K., Kuma, K. i., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14):3059–3066.

Keane, O. M., Dodds, K. G., Crawford, A. M., and McEwan, J. C. (2007). Transcriptional profiling of *Ovis aries* identifies *Ovar-DQA1* allele frequency differences between nematode-resistant and susceptible selection lines. *Physiological Genomics*, 30(3):253–261.

Keane, O. M., Zadissa, A., Wilson, T., Hyndman, D. L., Greer, G. J., McCulloch, A. F., Crawford, A. M., and McEwan, J. C. (2006). Gene expression profiling of naïve sheep genetically resistant and susceptible to gastrointestinal nematodes. *BMC Genomics*, 7:42.

Keller, P., Petrie, J. T., De Rose, P., Gerin, I., Wright, W. S., Chiang, S.-H., Nielsen, A. R., Fischer, C. P., Pedersen, B. K., and MacDougald, O. A. (2008). Fat-specific Protein 27 Regulates Storage of Triacylglycerol. *Journal of Biological Chemistry*, 283(21):14355–14365.

Kelly, P., Good, B., Hanrahan, J. P., Fitzpatrick, R., and de Waal, T. (2009). Screening for the presence of nematophagous fungi collected from Irish sheep pastures. *Veterinary Parasitology*, 165(3-4):345–349.

Kemper, K. E., Elwin, R. L., Bishop, S. C., Goddard, M. E., and Woolaston, R. R. (2009). *Haemonchus contortus* and *Trichostrongylus colubriformis* did not adapt to long-term exposure to sheep that were genetically resistant or susceptible to nematode infections. *International Journal for Parasitology*, 39(5):607–614.

Kemper, K. E., Emery, D. L., Bishop, S. C., Oddy, H., Hayes, B. J., Dominik, S., Henshall, J. M., and Goddard, M. E. (2011). The distribution of SNP marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections. *Genetics Research*, 93(03):203–219.

Kemper, K. E., Palmer, D. G., Liu, S. M., Greeff, J. C., Bishop, S. C., and Karlsson, L. J. E. (2010). Reduction of faecal worm egg count, worm numbers and worm fecundity in sheep selected for worm resistance following artificial infection with *Teladorsagia circumcincta* and *Trichostrongylus colubriformis*. *Veterinary Parasitology*, 171(3–4):238–246.

Kerboeuf, D. (1985). Winter survival of trichostrongyle larvae: a study using tracer lambs. *Research in Veterinary Science*, 38(3):364–367.

Kettle, P. R., Vlassoff, A., Ayling, J. M., McMurtry, L. W., Smith, S. J., and Watson, A. J. (1982). A survey of nematode control measures used by sheep farmers and of anthelmintic resistance on their farms. Part 2: South Island excluding the Nelson region. *New Zealand Veterinary Journal*, 30(6):79–81.

Kettle, P. R., Vlassoff, A., Lukies, J. M., Ayling, J. M., and McMurtry, L. W. (1981). A survey of nematode control measures used by sheep farmers and of anthelmintic resistance on their farms. Part 1. North Island and the Nelson region of South Island. *New Zealand Veterinary Journal*, 29(5):81–83.

Kijas, J. W., Lenstra, J. A., Hayes, B. J., Boitard, S., Porto Neto, L. R., San Cristobal, M., Servin, B., McCulloch, R., Whan, V., Gietzen, K., Paiva, S., Barendse, W., Ciani, E., Raadsma, H., McEwan, J. C., Dalrymple, B. P., and other members of the International Sheep Genomics Consortium (2012). Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. *PLoS Biology*, 10(2):e1001258.

Kijas, J. W., Townley, D., Dalrymple, B. P., Heaton, M. P., Maddox, J. F., McGrath, A., Wilson, P., Ingersoll, R. G., McCulloch, R., McWilliam, S., Tang, D., McEwan, J. C., Cockett, N., Oddy, V. H., Nicholas, F. W., and Raadsma, H. (2009). A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLoS ONE*, 4(3):e4668.

Kimura, M. (1984). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kitaura, M., Suzuki, N., Imai, T., Takagi, S., Suzuki, R., Nakajima, T., Hirai, K., Nomiyama, H., and Yoshie, O. (1999). Molecular cloning of a novel human CC chemokine (Eotaxin-3) that is a functional ligand of CC chemokine receptor 3. *Journal of Biological Chemistry*, 274(39):27975–27980.

Knight, J., Hein, W., and Pernthaner, A. (2010). The gastrointestinal nematode *Trichostrongylus colubriformis* down-regulates immune gene expression in migratory cells in afferent lymph. *BMC Immunology*, 11(1):51.

Knight, P., Griffith, S., Pemberton, A., Pate, J., Guarneri, L., Anderson, K., Talbot, R., Smith, S., Waddington, D., Fell, M., Archibald, A., Burgess, S., Smith, W., Miller, H., and Morrison, W. (2011). Novel gene expression responses in the ovine abomasal mucosa to infection with the gastric nematode *Teladorsagia circumcincta*. *Veterinary Research*, 42(1):78.

Knight, P. A., Pate, J., Smith, W. D., and Miller, H. R. (2007). An ovine chitinase-like molecule, chitinase-3 like-1 (YKL-40), is upregulated in the abomasum in response to challenge with the gastrointestinal nematode, *Teladorsagia circumcincta*. *Veterinary Immunology and Immunopathology*, 120(1-2):55–60.

Kooyman, F., Schallig, H., MA, V. L., Mackellar, A., Huntley, J., Cornelissen, A., and Vervelde, L. (2000). Protection in lambs vaccinated with *Haemonchus contortus* antigens

is age related, and correlates with IgE rather than IgG1 antibody. *Parasite Immunology*, 22(1):13–20.

Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V. K. (2009). IL-17 and Th17 Cells. *Annual Review of Immunology*, 27(1):485–517.

Kumánovics, A., Takada, T., and Lindahl, K. F. (2003). Genomic Organization of the Mammalian MHC. *Annual Review of Immunology*, 21(1):629–657.

Kumar, P., Henikoff, S., and Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, 4(7):1073–1081.

Kyriazakis, I., Oldham, J. D., Coop, R. L., and Jackson, F. (1994). The effect of subclinical intestinal nematode infection on the diet selection of growing sheep. *British Journal of Nutrition*, 72(5):665–677.

Lacroux, C., Nguyen, T. H. C., Andreoletti, O., Prevot, F., Grisez, C., Bergeaud, J. P., Gruner, L., Brunel, J.-C., Francois, D., Dorchies, P., and Jacquiet, P. (2006). *Haemonchus contortus* (Nematoda: Trichostrongylidae) infection in lambs elicits an unequivocal Th2 immune response. *Veterinary Research*, 37(4):607–622.

Lanigan, T. M., Liu, A., Huang, Y. Z., Mei, L., Margolis, B., and Guan, K.-L. (2003). Human homologue of *Drosophila* CNK interacts with Ras effector proteins Raf and Rlf. *The FASEB Journal*, 17(14):2048–2060.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948.

Larsen, M. (2000). Prospects for controlling animal parasitic nematodes by predacious micro fungi. *Parasitology*, 120(07):121–131.

Larson, J. H., Marron, B. M., Beever, J. E., Roe, B. A., and Lewin, H. A. (2006). Genomic organization and evolution of the ULBP genes in cattle. *BMC Genomics*, 7:227.

Lawrence, T. (2009). The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harbor Perspectives in Biology*, 1(6):a001651.

Lawrie, D. S., Messer, P. W., Hershberg, R., and Petrov, D. A. (2013). Strong Purifying Selection at Synonymous Sites in *D. melanogaster*. *PLoS Genetics*, 9(5):e1003527.

Lawton, D. E. B., Reynolds, G. W., Hodgkinson, S. M., Pomroy, W. E., and Simpson, H. V. (1996). Infection of sheep with adult and larval *Ostertagia circumcincta*: Effects on abomasal pH and serum gastrin and pepsinogen. *International Journal for Parasitology*, 26(10):1063–1074.

Leathwick, D. M., Waghorn, T. S., Miller, C. M., and Candy, P. M. (2012). Managing anthelmintic resistance–use of a combination anthelmintic and leaving some lambs untreated to slow the development of resistance to ivermectin. *Veterinary Parasitology*, 187(1-2):285–294.

Leathwick, D. M. D. (2012). Modelling the benefits of a new class of anthelmintic in combination. *Veterinary Parasitology*, 186(1-2):93–100.

Leathwick, D. M. D., Hosking, B. C. B., Bisset, S. A. S., and McKay, C. H. C. (2009). Managing anthelmintic resistance: is it feasible in New Zealand to delay the emergence of resistance to a new anthelmintic class? *New Zealand Veterinary Journal*, 57(4):181–192.

Lebre, M. C., Burwell, T., Vieira, P. L., Lora, J., Coyle, A. J., Kapsenberg, M. L., Clausen, B. E., and De Jong, E. C. (2005). Differential expression of inflammatory chemokines by Th1- and Th2-cell promoting dendritic cells: A role for different mature dendritic cell populations in attracting appropriate effector cells to peripheral sites of inflammation. *Immunology and Cell Biology*, 83(5):525–535.

Lecker, S. H., Goldberg, A. L., and Mitch, W. E. (2006). Protein Degradation by the Ubiquitin–Proteasome Pathway in Normal and Disease States. *Journal of the American Society of Nephrology*, 17(7):1807–1819.

Lee, C. G., Da Silva, C. A., Dela Cruz, C. S., Ahangari, F., Ma, B., Kang, M.-J., He, C.-H., Takyar, S., and Elias, J. A. (2011a). Role of chitin and chitinase/chitinase-like proteins in inflammation, tissue remodeling, and injury. *Annual Review of Physiology*, 73:479–501.

Lee, C. Y., Munyard, K. A., Gregg, K., Wetherall, J. D., Stear, M. J., and Groth, D. M. (2011b). The Influence of MHC and Immunoglobulins A and E on Host Resistance to Gastrointestinal Nematodes in Sheep. *Journal of Parasitology Research*, 2011(1):1–11.

Legge, T. (1996). The beginning of caprine domestication. In Harris, D. R., editor, *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, pages 238–262. Smithsonian Institution Press, New York.

Letterio, J. J. and Roberts, A. B. (1998). Regulation of immune responses by TGF-β*. *Annual Review of Immunology*, 16(1):137–161.

Leu, C.-M., Davis, R. S., Gartland, L. A., Fine, W. D., and Cooper, M. D. (2005). FcRH1: an activation coreceptor on human B cells. *Blood*, 105(3):1121–1126.

Levasseur, A., Gouret, P., Lesage-Meessen, L., Asther, M., Asther, M., Record, E., and Pontarotti, P. (2006). Tracking the connection between evolutionary and functional shifts using the fungal lipase/feruloyl esterase A family. *BMC Evolutionary Biology*, 6(1):92.

Levi-Acobas, F., Mars, L. T., Orth, A., Bureau, J.-F., and Bonhomme, F. (2009). Adaptive evolution of interferon-γ in Glire lineage and evidence for a recent selective sweep in *Mus. m. domesticus*. *Genes and Immunity*, 10(4):297–308.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16):2078–2079.

Li, R. W., Choudhary, R. K., Capuco, A. V., and Urban, J. J. F. (2012). Exploring the host transcriptome for mechanisms underlying protective immunity and resistance to nematode infections in ruminants. *Veterinary Parasitology*, 190(1-2):1–11.

Li, W. H., Wu, C. I., and Luo, C. C. (1985). A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Molecular Biology and Evolution*, 2(2):150–174.

Little, P. R., Hodge, A., Watson, T. G., Seed, J. A., and Maeder, S. J. (2010). Field efficacy and safety of an oral formulation of the novel combination anthelmintic, derquantel-abamectin, in sheep in New Zealand. *New Zealand Veterinary Journal*, 58(3):121–129.

Little, T. J., Nelson, L., and Hupp, T. (2007). Adaptive Evolution of a Stress Response Protein. *PLoS ONE*, 2(10):e1003.

Lluis, A., Schedel, M., Liu, J., Illi, S., Depner, M., von Mutius, E., Kabesch, M., and Schaub, B. (2011). Asthma-associated polymorphisms in 17q21 influence cord blood ORMDL3 and GSDMA gene expression and IL-17 secretion. *Journal of Allergy and Clinical Immunology*, 127(6):1587–1594.e6.

Long, M., Betrán, E., Thornton, K., and Wang, W. (2003). The origin of new genes: glimpses from the young and old. *Nature Reviews Genetics*, 4(11):865–875.

Loughran, N. B. (2010). *Evolutionary and paleobiochemical analyses of heme peroxidases*. PhD thesis, Dublin City University, Dublin, Ireland.

Loughran, N. B., Hinde, S., McCormick-Hill, S., Leidal, K. G., Bloomberg, S., Loughran, S. T., O'Connor, B., Ó'Fágáin, C., Nauseef, W. M., and O'Connell, M. J. (2012). Functional Consequence of Positive Selection Revealed through Rational Mutagenesis of Human Myeloperoxidase. *Molecular Biology and Evolution*, 29(8):2039–2046.

Loughran, N. B., O'Connor, B., Ó'Fágáin, C., and O'Connell, M. J. (2008). The phylogeny of the mammalian heme peroxidases and the evolution of their diverse functions. *BMC Evolutionary Biology*, 8:101.

Loytynoja, A. and Goldman, N. (2008). A model of evolution and structure for multiple sequence alignment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1512):3913–3919.

Lynch, M. and Conery, J. S. (2000). The Evolutionary Fate and Consequences of Duplicate Genes. *Science*, 290(5494):1151–1155.

Lynn, D. J., Freeman, A. R., Murray, C., and Bradley, D. G. (2005). A Genomics Approach to the Detection of Positive Selection in Cattle: Adaptive Evolution of the T-Cell and Natural Killer Cell-Surface Protein CD2. *Genetics*, 170(3):1189–1196.

MacKinnon, K. M., Burton, J. L., Zajac, A. M., and Notter, D. R. (2009). Microarray analysis reveals difference in gene expression profiles of hair and wool sheep infected with *Haemonchus contortus*. *Veterinary Immunology and Immunopathology*, 130(3–4):210–220.

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature*, 456(7218):18–21.

Maizels, R. M., Pearce, E. J., Artis, D., Yazdanbakhsh, M., and Wynn, T. A. (2009). Regulation of pathogenesis and immunity in helminth infections. *The Journal of Experimental Medicine*, 206(10):2059–2066.

Maizels, R. M. and Yazdanbakhsh, M. (2003). Immune Regulation by helminth parasites: cellular and molecular mechanisms. *Nature Reviews Immunology*, 3(9):733–744.

Malan, F. S. (1988). Resistance of field strains of *Haemonchus contortus* to ivermectin, closantel, rafoxanide and the benzimidazoles in South Africa. *Veterinary Record*, 123(9):226–228.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E. R., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., Boehnke, M., Clark, A. G., Eichler, E. E., Gibson, G., Haines, J. L., Mackay, T. F. C., McCarroll, S. A., and Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, 461(7265):747–753.

Mansén, A., Guardiola-Diaz, H., Rafter, J., Branting, C., and Gustafsson, J.-Å. (1996). Expression of the Peroxisome Proliferator-Activated Receptor (PPAR) in the Mouse Colonic Mucosa. *Biochemical and Biophysical Research Communications*, 222(3):844–851.

Marshall, K., Maddox, J. F., Lee, S. H., Zhang, Y., Kahn, L., Graser, H. U., Gondro, C., Walkden-Brown, S. W., and Van der Werf, J. H. J. (2009). Genetic mapping of quantitative trait loci for resistance to *Haemonchus contortus* in sheep. *Animal Genetics*, 40(3):262–272.

Marshall, K., Mugambi, J. M., Nagda, S., Sonstegard, T. S., Van Tassell, C. P., Baker, R., and Gibson, J. P. (2012). Quantitative trait loci for resistance to *Haemonchus contortus*

artificial challenge in Red Maasai and Dorper sheep of East Africa. *Animal Genetics*, 44(3):285–295.

Martinez, F. O., Helming, L., Milde, R., Varin, A., Melgert, B. N., Draijer, C., Thomas, B., Fabbri, M., Crawshaw, A., Ho, L. P., Ten Hacken, N. H., Cobos Jiménez, V., Kootstra, N. A., Hamann, J., Greaves, D. R., Locati, M., Mantovani, A., and Gordon, S. (2013). Genetic programs expressed in resting and IL-4 alternatively activated mouse and human macrophages: similarities and differences. *Blood*, 121(9):e57–69.

Martínez-Valladares, M., Vara-Del Río, M. P., Cruz-Rojo, M. A., and Rojo-Vázquez, F. A. (2005). Genetic resistance to *Teladorsagia circumcincta*: IgA and parameters at slaughter in Churra sheep. *Parasite Immunology*, 27(6):213–218.

Matika, O., Nyoni, S., van Wyk, J. B., Erasmus, G. J., and Baker, R. (2003). Resistance of Sabi and Dorper ewes to gastrointestinal nematode infections in an African semi-arid environment. *Small Ruminant Research*, 47(2):95–102.

Matika, O., Pong-Wong, R., Woolliams, J. A., and Bishop, S. C. (2011). Confirmation of two quantitative trait loci regions for nematode resistance in commercial British terminal sire breeds. *Animal*, 5(8):1149–1156.

Maurer, M. and von Stebut, E. (2004). Macrophage inflammatory protein-1. *The International Journal of Biochemistry & Cell Biology*, 36(10):1882–1886.

McAleese, S. M., Pemberton, A. D., McGrath, M. E., Huntley, J. F., and Miller, H. R. (1998). Sheep mast-cell proteinases-1 and -3: cDNA cloning, primary structure and molecular modelling of the enzymes and further studies on substrate specificity. *The Biochemical Journal*, 333 (Pt 3):801–809.

McDonald, J. H. and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, 351(6328):652–654.

McEwan, J. C. and Dodds, K. G. (2009). Confidential report to Ovita Ltd: Breeding for host resistance to nematode parasites using CARLA IgA antibodies in Saliva. Technical report.

McEwan, J. C., Dodds, K. G., and Gemmell, N. J. (2014). Signatures of selection in sheep bred for resistance or susceptibility to gastrointestinal nematodes. *BMC Genomics*, 15(1):637.

McEwan, J. C., Dodds, K. G., Watson, T. G., Greer, G. J., Hosking, B. C., and Douch, P. G. C. (1995). Selection for host resistance to roundworms by the New Zealand sheep breeding industry : the WormFEC service. In *Proceedings of the Association for the Advancement of Animal Breeding and Genetics 11*.

McEwan, J. C., Weston, N. K., Payne, G. M., O'Sullivan, N. H., Auvray, B. A., and Dodds, K. G. (2008). Ovine identification method. Technical report.

McGregor, L., Makela, V., Darling, S. M., Vrontou, S., Chalepakis, G., Roberts, C., Smart, N., Rutland, P., Prescott, N., Hopkins, J., Bentley, E., Shaw, A., Roberts, E., Mueller, R., Jadeja, S., Philip, N., Nelson, J., Francannet, C., Perez-Aytes, A., Megarbane, A., Kerr, B., Wainwright, B., Woolf, A. S., Winter, R. M., and Scambler, P. J. (2003). Fraser syndrome and mouse blebbed phenotype caused by mutations in *FRAS1/Fras1* encoding a putative extracellular matrix protein. *Nature Genetics*, 34(2):203–208.

McKellar, Q. A. (1993). Interactions of *Ostertagia* species with their bovine and ovine hosts. *International Journal for Parasitology*, 23(4):451–462.

McKenna, P. B. (1981). The diagnostic value and interpretation of faecal egg counts in sheep. *New Zealand Veterinary Journal*, 29(8):129–132.

McKenna, P. B. (1990). The use of benzimidazole-levamisole mixtures for the control and prevention of anthelmintic resistance in sheep nematodes: an assessment of their likely effects. *New Zealand Veterinary Journal*, 38(2):45–49.

McKenna, P. B. (2008). Comparison of two worm counting procedures for the enumeration of abomasal and small intestinal nematode parasites of sheep. *Veterinary Parasitology*, 157(3-4):254–259.

McMahon, C., Barley, J. P., Edgar, H. W. J., Ellison, S. E., Hanna, R. E. B., Malone, F. E., Brennan, G. P., and Fairweather, I. (2013a). Anthelmintic resistance in Northern Ireland. II: Variations in nematode control practices between lowland and upland sheep flocks. *Veterinary Parasitology*, 192(1-3):173–182.

McMahon, C., Bartley, D. J., Edgar, H. W. J., Ellison, S. E., Barley, J. P., Malone, F. E., Hanna, R. E. B., Brennan, G. P., and Fairweather, I. (2013b). Anthelmintic resistance in Northern Ireland (I): prevalence of resistance in ovine gastrointestinal nematodes, as determined through faecal egg count reduction testing. *Veterinary Parasitology*, 195(1-2):122–130.

McNeilly, T. N., Devaney, E., and Matthews, J. B. (2009). *Teladorsagia circumcincta* in the sheep abomasum: defining the role of dendritic cells in T cell regulation and protective immunity. *Parasite Immunology*, 31(7):347–356.

Meadows, J. R. S., Cemal, I., Karaca, O., Gootwine, E., and Kijas, J. W. (2007). Five Ovine Mitochondrial Lineages Identified From Sheep Breeds of the Near East. *Genetics*, 175(3):1371–1379.

Meagher, C., Arreaza, G., Peters, A., Strathdee, C. A., Gilbert, P. A., Mi, Q.-S., Santamaria, P., Dekaban, G. A., and Delovitch, T. L. (2007). CCL4 Protects From Type 1 Diabetes by Altering Islet β-Cell–Targeted Inflammatory Responses. *Diabetes*, 56(3):809–817.

Meeusen, E. N. T. and Balic, A. (2000). Do Eosinophils have a Role in the Killing of Helminth Parasites? *Parasitology Today*, 16(3):95–101.

Melén, K., Keskinen, P., Ronni, T., Sareneva, T., Lounatmaa, K., and Julkunen, I. (1996). Human MxB Protein, an Interferon-α-inducible GTPase, Contains a Nuclear Targeting Signal and Is Localized in the Heterochromatin Region beneath the Nuclear Envelope. *Journal of Biological Chemistry*, 271(38):23478–23486.

Menten, P., Wuyts, A., and Van Damme, J. (2002). Macrophage inflammatory protein-1. *Cytokine & Growth Factor Reviews*, 13(6):455–481.

Meraro, D., Gleit-Kielmanowicz, M., Hauser, H., and Levi, B.-Z. (2002). IFN-Stimulated Gene 15 Is Synergistically Activated Through Interactions Between the Myelocyte/Lymphocyte-Specific Transcription Factors, PU.1, IFN Regulatory Factor-8/IFN Consensus Sequence Binding Protein, and IFN Regulatory Factor-4: Characterization of a New Subtype of IFN-Stimulated Response Element. *The Journal of Immunology*, 168(12):6224–6231.

Miao, X. and Luo, Q. (2013). Genome-wide transcriptome analysis between small-tail Han sheep and the Surabaya fur sheep using high-throughput RNA sequencing. *Reproduction*, 145(6):587–596.

Miller, J. E. and Horohov, D. W. (2006). Immunological aspects of nematode parasite control in sheep. *Journal of Animal Science*, 84:E124–E132.

Miller, J. E. J., Bahirathan, M. M., Lemarie, S. L. S., Hembry, F. G. F., Kearney, M. T. M., and Barras, S. R. S. (1998). Epidemiology of gastrointestinal nematode parasitism in Suffolk and Gulf Coast Native sheep with special emphasis on relative susceptibility to *Haemonchus contortus* infection. *Veterinary Parasitology*, 74(1):55–74.

Ministry of Agriculture Fisheries and Food (1986). *Manual of Veterinary Parasitological Laboratory Techniques. Reference book 418*. London.

Miyanokoshi, M., Tanaka, T., Tamai, M., Tagawa, Y.-i., and Wakasugi, K. (2013). Expression of the rodent-specific alternative splice variant of tryptophanyl-tRNA synthetase in murine tissues and cells. *Scientific Reports*, 3:3477.

Miyazono, K., Hellman, U., Wernstedt, C., and Heldin, C. H. (1988). Latent high molecular weight complex of transforming growth factor beta 1. Purification from human platelets and structural characterization. *Journal of Biological Chemistry*, 263(13):6407–6415.

Molan, A. L., Waghorn, G. C., and McNabb, W. C. (2002). Effect of condensed tannins on egg hatching and larval development of *Trichostrongylus colubriformis in vitro*. *Veterinary Record*, 150(3):65–69.

Mömke, S., Kerkmann, A., Wöhlke, A., Ostmeier, M., Hewicker-Trautwein, M., Ganter, M., Kijas, J. W., Distl, O., and for the International Sheep Genomics Consortium

(2011). A Frameshift Mutation within *LAMC2* Is Responsible for Herlitz Type Junctional Epidermolysis Bullosa (HJEB) in Black Headed Mutton Sheep. *PLoS ONE*, 6(5):e18943.

Montgomery, G. W. and Sise, J. A. (1990). Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research*, 33(3):437–441.

Mooney, L., Good, B., Hanrahan, J. P., Mulcahy, G., and de Waal, T. (2009). The comparative efficacy of four anthelmintics against a natural acquired *Fasciola hepatica* infection in hill sheep flock in the west of Ireland. *Veterinary Parasitology*, 164(2–4):201–205.

Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G., and McEwan, J. C. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*, 13(1):10.

Morgan, C. C., Foster, P. G., Webb, A. E., Pisani, D., McInerney, J. O., and O'Connell, M. J. (2013). Heterogeneous Models Place the Root of the Placental Mammal Phylogeny. *Molecular Biology and Evolution*, 30(9):2145–2156.

Morgan, C. C., Loughran, N. B., Walsh, T. A., Harrison, A. J., and O'Connell, M. J. (2010). Positive selection neighboring functionally essential sites and disease-implicated regions of mammalian reproductive proteins. *BMC Evolutionary Biology*, 10:39.

Morgan, E. R., Hosking, B. C., Burston, S., Carder, K. M., Hyslop, A. C., Pritchard, L. J., Whitmarsh, A. K., and Coles, G. C. (2012). A survey of helminth control practices on sheep farms in Great Britain and Ireland. *The Veterinary Journal*, 192(3):390–397.

Mori, Y., Iwasaki, H., Kohno, K., Yoshimoto, G., Kikushige, Y., Okeda, A., Uike, N., Niiro, H., Takenaka, K., Nagafuji, K., Miyamoto, T., Harada, M., Takatsu, K., and Akashi, K. (2009). Identification of the human eosinophil lineage-committed progenitor: revision of phenotypic definition of the human common myeloid progenitor. *The Journal of Experimental Medicine*, 206(1):183–193.

Morris, C. A. (2002). Host genetics and internal parasitism. In *Proceedings of the 32nd Seminar of the Sheep & Cattle Society of the NZ Veterinary Association*, pages 99–104, Wellington, New Zealand.

Morris, C. A. (2009). Review of genetic parameters for disease resistance in sheep in New Zealand and Australia. In *Proceedings of the Association for the Advancement of Animal Breeding and Genetics 18*, pages 263–271, Barossa Valley, Australia.

Morris, C. A., Vlassoff, A., Bisset, S., Baker, R., Watson, T. G., West, C. J., and Wheeler, M. (2000). Continued selection of Romney sheep for resistance or susceptibility to nematode infection: estimates of direct and correlated responses. *Animal Science*, 70:17–27.

Morris, C. A., Vlassoff, A., Bisset, S. A., Baker, R., West, C. J., and Hurford, A. P. (1997). Responses of Romney sheep to selection for resistance or susceptibility to nematode infection. *Animal Science*, 64(2):319–329.

Morris, C. A., Wheeler, M., Watson, T. G., Hosking, B. C., and Leathwick, D. M. (2005). Direct and correlated responses to selection for high or low faecal nematode egg count in Perendale sheep. *New Zealand Journal of Agricultural Research*, 48:1–10.

Mugambi, J. M., Bain, R. K., Wanyangu, S. W., Ihiga, M. A., Duncan, J. L., Murray, M., and Stear, M. J. (1997). Resistance of four sheep breeds to natural and subsequent artificial *Haemonchus contortus* infection. *Veterinary Parasitology*, 69(3-4):265–273.

Mugambi, J. M., Wanyangu, S. W., Bain, R. K., Owango, M. O., Duncan, J. L., and Stear, M. J. (1996). Response of Dorper and red Maasai lambs to trickle *Haemonchus contortus* infections. *Research in Veterinary Science*, 61(3):218–221.

Murata, H., Shimada, N., and Yoshioka, M. (2004). Current research on acute phase proteins in veterinary diagnosis: an overview. *The Veterinary Journal*, 168(1):28–40.

Murphy, K. P., Travers, P., Walport, M., and Janeway, C. (2008). *Janeway's Immunobiology*. Garland Science.

Nagamine, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Rudan, I., Campbell, H., Wilson, J., Wild, S., Hicks, A. A., Pramstaller, P. P., Hastie, N., Wright, A. F., and Haley, C. S. (2012). Localising loci underlying complex trait variation using Regional Genomic Relationship Mapping. *PLoS ONE*, 7(10):e46501.

Nagaraj, S. H., Harsha, H. C., Reverter, A., Colgrave, M. L., Sharma, R., Andronicos, N. M., Hunt, P., Menzies, M., Lees, M. S., Sekhar, N. R., Pandey, A., and Ingham, A. (2012). Proteomic analysis of the abomasal mucosal response following infection by the nematode, *Haemonchus contortus*, in genetically resistant and susceptible sheep. *Journal of Proteomics*, 75(7):2141–2152.

Nair, M. G., Gallagher, I. J., Taylor, M. D., Loke, P., Coulson, P. S., Wilson, R. A., Maizels, R. M., and Allen, J. E. (2005). Chitinase and Fizz Family Members Are a Generalized Feature of Nematode Infection with Selective Upregulation of Ym1 and Fizz1 by Antigen-Presenting Cells. *Infection and Immunity*, 73(1):385–394.

Nakayamada, S., Takahashi, H., Kanno, Y., and O'Shea, J. J. (2012). Helper T cell diversity and plasticity. *Current Opinion in Immunology*, 24(3):297–302.

Nei, M. and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3(5):418–426.

Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E. E., Bamshad, M., Nickerson, D. A., and Shendure, J. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, 461(7261):272–276.

Nicholas, F. W., Windon, R. G., Brown, S. C., and Beh, K. J. (1993). The MHC class II region and resistance to an intestinal parasite in sheep. *Journal of Animal Breeding and Genetics*, 110(1-6):459–472.

Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., Fledel-Alon, A., Tanenbaum, D. M., Civello, D., White, T. J., J Sninsky, J., Adams, M. D., and Cargill, M. (2005). A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees. *PLoS Biology*, 3(6):e170.

Nielsen, R. and Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*, 148(3):929–936.

Nighswonger, A. M., Austin, K. J., Ealy, A. D., Han, C. S., and Hansen, T. R. (2000). Rapid communication: the ovine cDNA encoding interferon-stimulated gene product 17 (ISG17). *Journal of Animal Science*, 78(5):1393–1394.

Nolan, T., Hanrahan, J. P., and O'Malley, L. (2003). Integrated Hill Sheep Production Systems. Technical report, Dublin, Ireland.

Nowosad, B., Gruner, L., Skalska, M., Fudalewicz-Niemczyk, W., Molenda, K., and Kornas, S. (2003). Genetic difference in natural resistance to gastrointestinal nematodes in Polish long-wool, Blackface and Weisses Alpenschaf sheep. *Acta Parasitologica*, 48(2):131–134.

Obbard, D. J., Welch, J. J., Kim, K.-W., and Jiggins, F. M. (2009). Quantifying Adaptive Evolution in the *Drosophila* Immune System. *PLoS Genetics*, 5(10):e1000698.

O'Connell, R. M., Rao, D. S., and Baltimore, D. (2012). microRNA regulation of inflammatory responses. *Annual Review of Immunology*, 30:295–312.

O'Connor, L. J., Walkden-Brown, S. W., and Kahn, L. P. (2006). Ecology of the free-living stages of major trichostrongylid parasites of sheep. *Veterinary Parasitology*, 142(1–2):1–15.

Ohno, S. (1970). *Evolution by gene duplication*. Springer-Verlag, New York.

Ohta, T. (1973). Slightly Deleterious Mutant Substitutions in Evolution. *Nature*, 246(5428):96–98.

O'Loughlin, A., McGee, M., Waters, S., Doyle, S., and Earley, B. (2011). Examination of the bovine leukocyte environment using immunogenetic biomarkers to assess immunocompetence following exposure to weaning stress. *BMC Veterinary Research*, 7(1):45.

Orlowski, J., Kandasamy, R. A., and Shull, G. E. (1992). Molecular cloning of putative members of the Na/H exchanger gene family. cDNA cloning, deduced amino acid sequence, and mRNA tissue expression of the rat Na/H exchanger NHE-1 and two structurally related proteins. *Journal of Biological Chemistry*, 267(13):9331–9339.

Outteridge, P. M., Andersson, L., Douch, P., Green, R. S., Gwakisa, P. S., Hohenhaus, M. A., and Mikko, S. (1996). The PCR typing of *MHC-DRB* genes in the sheep using primers for an intronic microsatellite: Application to nematode parasite resistance. *Immunology and Cell Biology*, 74(4):330–336.

Pacheco, R., Gallart, T., Lluis, C., and Franco, R. (2007). Role of glutamate on T-cell mediated immunity. *Journal of Neuroimmunology*, 185(1-2):9–19.

Paolini, V., Fouraste, I., and Hoste, H. (2004). In vitro effects of three woody plant and sainfoin extracts on 3rd-stage larvae and adult worms of three gastrointestinal nematodes. *Parasitology*, 129(Pt 1):69–77.

Paterson, K. A., McEwan, J. C., Dodds, K. G., and Crawford, A. M. (2001). Fine mapping a locus affecting host resistance to internal parasites of sheep. In *Proceedings of the 14th Conference of the Association for the Advancement of Animal Breeding and Genetics*, pages 91–94.

Paterson, S., Wilson, K., and Pemberton, J. M. (1998). Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries L.*). *Proceedings of the National Academy of Sciences*, 95(7):3714–3719.

Patten, T., Good, B., Hanrahan, J. P., and de Waal, D. T. (2007). A survey on anthelmintic resistance in lowland sheep flocks in Ireland. Technical report.

Patten, T., Good, B., Hanrahan, J. P., Mulcahy, G., and de Waal, T. (2011). Gastrointestinal nematode control practices on lowland sheep farms in Ireland with reference to selection for anthelmintic resistance. *Irish Veterinary Journal*, 64(1):4.

Pemberton, A. D., Brown, J. K., Craig, N. M., Pate, J., McLean, K., Inglis, N. F., Knox, D., and Knight, P. A. (2012). Changes in protein expression in the sheep abomasum following trickle infection with *Teladorsagia circumcincta*. *Parasitology*, 139(03):375–385.

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

Pernthaner, A., Cabaj, W., Stankiewicz, M., Davies, J., and Maass, D. R. (1997). Cytokine mRNA expression and IFN-gamma production of immunised nematode resistant and susceptible lambs against natural poly-generic challenge. *Acta Parasitologica*, 42(3):180–186.

Pernthaner, A., Cole, S.-A., Morrison, L., Green, R., Shaw, R. J., and Hein, W. R. (2006). Cytokine and antibody subclass responses in the intestinal lymph of sheep during repeated experimental infections with the nematode parasite *Trichostrongylus colubriformis*. *Veterinary Immunology and Immunopathology*, 114(1–2):135–148.

Pernthaner, A., Cole, S. A., Morrison, L., and Hein, W. R. (2005). Increased Expression of Interleukin-5 (IL-5), IL-13, and Tumor Necrosis Factor Alpha Genes in Intestinal Lymph Cells of Sheep Selected for Enhanced Resistance to Nematodes during Infection with *Trichostrongylus colubriformis*. *Infection and Immunity*, 73(4):2175–2183.

Peterman, J. H. and Butler, J. E. (1989). Application of theoretical considerations to the analysis of ELISA data. *BioTechniques*, 7(6):608–615.

Pettit, J. J., Jackson, F., Rocchi, M., and Huntley, J. F. (2005). The relationship between responsiveness against gastrointestinal nematodes in lambs and the numbers of circulating IgE-bearing cells. *Veterinary Parasitology*, 134(1–2):131–139.

Pickering, N. K. (2013). *Genetics of flystrike, dagginess and associated traits in New Zealand dual-purpose sheep*. PhD thesis, Massey University, Palmerston North, New Zealand.

Ploeger, H. W., Kloosterman, A., Bargeman, G., von Wuijckhuise, L., and van den Brink, R. (1990a). Milk yield increase after anthelmintic treatment of dairy cattle related to some parameters estimating helminth infection. *Veterinary Parasitology*, 35(1-2):103–116.

Ploeger, H. W., Kloosterman, A., Borgsteede, F. H., and Eysker, M. (1990b). Effect of naturally occurring nematode infections in the first and second grazing season on the growth performance of second-year cattle. *Veterinary Parasitology*, 36(1-2):57–70.

Ploeger, H. W., Kloosterman, A., Eysker, M., Borgsteede, F. H., van Straalen, W., and Verhoeff, J. (1990c). Effect of naturally occurring nematode infections on growth performance of first-season grazing calves. *Veterinary Parasitology*, 35(4):307–322.

Ploeger, H. W., Kloosterman, A., Rietveld, F. W., Berghen, P., Hilderson, H., and Hollanders, W. (1994). Quantitative estimation of the level of exposure to gastrointestinal nematode infection in first-year calves. *Veterinary Parasitology*, 55(4):287–315.

Politz, O., Gratchev, A., McCourt, P. A. G., Schledzewski, K., Guillot, P., Johansson, S., Svineng, G., Franke, P., Kannicht, C., Kzhyshkowska, J., Longati, P., Velten, F. W., Johansson, S., and Goerdt, S. (2002). Stabilin-1 and -2 constitute a novel family of fasciclin-like hyaluronan receptor homologues. *The Biochemical Journal*, 362(Pt 1):155–164.

Prada Jiménez de Cisneros, J., Matthews, L., Mair, C., Stefan, T., and Stear, M. J. (2014). The transfer of IgA from mucus to plasma and the implications for diagnosis and control of nematode infections. *Parasitology*, 141(7):875–879.

Pritchard, J. K., Pickrell, J. K., and Coop, G. (2010). The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, 20(4):R208–15.

Przetak, M. M., Yoast, S., and Schmidt, B. F. (1995). Cloning of cDNA for human granzyme 3. *FEBS Letters*, 364(3):268–271.

Puri, V., Konda, S., Ranjit, S., Aouadi, M., Chawla, A., Chouinard, M., Chakladar, A., and Czech, M. P. (2007). Fat-specific Protein 27, a Novel Lipid Droplet Protein That Enhances Triglyceride Storage. *Journal of Biological Chemistry*, 282(47):34213–34218.

Rainbird, M. A., Macmillan, D., and Meeusen, E. N. (1998). Eosinophil-mediated killing of *Haemonchus contortus* larvae: effect of eosinophil activation and role of antibody, complement and interleukin-5. *Parasite Immunology*, 20(2):93–103.

Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C. E., Socci, N. D., and Betel, D. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biology*, 14(9):R95.

Rattray, P. V. (2003). Helminth Parasites in the New Zealand Meat & Wool Pastoral Industries: A Review of Current Issues. Technical Report Report commissioned by Meat & Wool Innovation Ltd.

Redelings, B. (2014). Erasing Errors due to Alignment Ambiguity When Estimating Positive Selection. *Molecular Biology and Evolution*.

Reich, D. E. and Lander, E. S. (2001). On the allelic spectrum of human disease. *Trends in Genetics*, 17(9):502–510.

Ren, M., Liu, C., Zeng, X., Yue, L., Mao, X., Qiao, S., and Wang, J. (2014). Amino acids modulates the intestinal proteome associated with immune and stress response in weaning pig. *Molecular Biology Reports*, 41(6):3611–3620.

Reynecke, D. P., Waghorn, T. S., Miller, C. M., Vlassoff, A., and Leathwick, D. M. (2011). Dynamics of the free-living stages of sheep intestinal parasites on pasture in the North Island of New Zealand. 2. Weather variables associated with development. *New Zealand Veterinary Journal*, 59(6):287–292.

Riffkin, G. G. and Dobson, C. (1979). Predicting resistance of sheep to *Haemonchus contortus* infections. *Veterinary Parasitology*, 5(4):365–378.

Riggio, V., Matika, O., Pong-Wong, R., Stear, M. J., and Bishop, S. C. (2013). Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. *Heredity*, 110(5):420–429.

Riggio, V., Pong-Wong, R., Sallé, G., Usai, M. G., Casu, S., Moreno, C. R., Matika, O., and Bishop, S. C. (2014). A joint analysis to identify loci underlying variation in nematode resistance in three European sheep populations. *Journal of Animal Breeding and Genetics*, pages 1–11.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140.

Robinson, M. D. and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3):R25.

Roeber, F., Jex, A. R., and Gasser, R. B. (2013). Advances in the diagnosis of key gastrointestinal nematode infections of livestock, with an emphasis on small ruminants. *Biotechnology Advances*, 31(8):1135–1152.

Ross, J. G., Purcell, D. A., Dow, C., and Todd, J. R. (1967). Experimental infections of calves with *Trichostrongylus axei*; the course and development of infection and lesions in low level infections. *Research in Veterinary Science*, 8(2):201–206.

Roux, J., Privman, E., Moretti, S., Daub, J. T., Robinson-Rechavi, M., and Keller, L. (2014). Patterns of positive selection in seven ant genomes. *Molecular Biology and Evolution*, 31(7):1661–1685.

Rowe, A., Gondro, C., Emery, D., and Sangster, N. (2009). Sequential microarray to identify timing of molecular responses to *Haemonchus contortus* infection in sheep. *Veterinary Parasitology*, 161(1-2):76–87.

Rubin, B. Y., Anderson, S. L., Xing, L., Powell, R. J., and Tate, W. P. (1991). Interferon induces tryptophanyl-tRNA synthetase expression in human fibroblasts. *Journal of Biological Chemistry*, 266(36):24245–24248.

Saddiqi, H. A., Sarwar, M., Iqbal, Z., Nisa, M., and Shahzad, M. A. (2012). Markers/parameters for the evaluation of natural resistance status of small ruminants against gastrointestinal nematodes. *Animal*, 6(6):994–1004.

Saeki, N., Kim, D. H., Usui, T., Aoyagi, K., Tatsuta, T., Aoki, K., Yanagihara, K., Tamura, M., Mizushima, H., Sakamoto, H., Ogawa, K., Ohki, M., Shiroishi, T., Yoshida, T., and Sasaki, H. (2007). *GASDERMIN*, suppressed frequently in gastric cancer, is a target of LMO1 in TGF-β-dependent apoptotic signalling. *Oncogene*, 26(45):6488–6498.

Saeki, N., Kuwahara, Y., Sasaki, H., Satoh, H., and Shiroishi, T. (2000). Gasdermin (*Gsdm*) localizing to mouse Chromosome 11 is predominantly expressed in upper gastrointestinal tract but significantly suppressed in human gastric cancer cells. *Mammalian Genome*, 11(9):718–724.

Safari, E., Fogarty, N. M., and Gilmour, A. R. (2005). A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. *Livestock Production Science*, 92(3):271–289.

Sallé, G. (2012). *Functional validation and detection regions of the genome that affect the resistance to gastrointestinal strongyles in sheep*. PhD thesis, University of Toulouse.

Sallé, G., Jacquiet, P., Gruner, L., Cortet, J., Sauvé, C., Prevot, F., Grisez, C., Bergeaud, J. P., Schibler, L., Tircazes, A., Francois, D., Pery, C., Bouvier, F., Thouly, J. C., Brunel, J. C., Legarra, A., Elsen, J. M., Bouix, J., Rupp, R., and Moreno, C. R. (2012). A genome scan for QTL affecting resistance to Haemonchus contortus in sheep. *Journal of Animal Science*, 90(13):4690–4705.

Sangster, N. C., Whitlock, H. V., Russ, I. G., Gunawan, M., Griffin, D. L., and Kelly, J. D. (1979). *Trichostrongylus colubriformis* and *Ostertagia circumcincta* resistant to levamisole, morantel tartrate and thiabendazole: occurrence of field strains. *Research in Veterinary Science*, 27(1):106–110.

Saqui-Salces, M., Dowdle, W. E., Reiter, J. F., and Merchant, J. L. (2012). A high-fat diet regulates gastrin and acid secretion through primary cilia. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 26(8):3127–3139.

Sargison, N. D. (2012). Pharmaceutical treatments of gastrointestinal nematode infections of sheep-Future of anthelmintic drugs. *Veterinary Parasitology*, 189(1):79–84.

Sargison, N. D., Jackson, F., Bartley, D. J., Wilson, D. J., Stenhouse, L. J., and Penny, C. D. (2007). Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of Scotland. *Veterinary Parasitology*, 145(1-2):12–12.

Sasaki, K., Tungtrakoolsub, P., Morozumi, T., Uenishi, H., Kawahara, M., and Watanabe, T. (2014). A single nucleotide polymorphism of porcine MX2 gene provides antiviral activity against vesicular stomatitis virus. *Immunogenetics*, 66(1):25–32.

Sayers, G., Good, B., Hanrahan, J. P., O'Donovan, J., Mulcahy, G., and Sweeney, T. (2008). Breed differences in mucosal and systemic antibody response to nematode infection in sheep: an important role for IgE? *Parasitology*, 135(Pt 1):71–80.

Sayers, G., Good, B., Hanrahan, J. P., Ryan, M., Angles, J. M., and Sweeney, T. (2005a). *Major histocompatibility complex DRB1* gene: its role in nematode resistance in Suffolk and Texel sheep breeds. *Parasitology*, 131(Pt 3):403–409.

Sayers, G., Good, B., Hanrahan, J. P., Ryan, M., and Sweeney, T. (2005b). Intron 1 of the interferon γ gene: Its role in nematode resistance in Suffolk and Texel sheep breeds. *Research in Veterinary Science*, 79(3):191–196.

Sayers, G. and Sweeney, T. (2005). Gastrointestinal nematode infection in sheep - a review of the alternatives to anthelmintics in parasite control. *Animal Health Research Reviews*, 6(2):159–171.

Sayre, B. L. and Harris, G. C. (2012). Systems genetics approach reveals candidate genes for parasite resistance from quantitative trait loci studies in agricultural species. *Animal Genetics*, 43(2):190–198.

Scherf, B. D. (2000). *2000 World Watch List for Domestic Animal Diversity*. Food and Agriculture Organization of the United Nations, Rome, 3rd edition.

Schneider, A., Souvorov, A., Sabath, N., Landan, G., Gonnet, G. H., and Graur, D. (2009). Estimates of positive Darwinian selection are inflated by errors in sequencing, annotation, and alignment. *Genome Biology and Evolution*, 1:114–118.

Scholler, N., Hayden-Ledbetter, M., Hellström, K.-E., Hellström, I., and Ledbetter, J. A. (2001). CD83 Is a Sialic Acid-Binding Ig-Like Lectin (Siglec) Adhesion Receptor that Binds Monocytes and a Subset of Activated CD8+ T Cells. *The Journal of Immunology*, 166(6):3865–3872.

Schrum, S., Probst, P., Fleischer, B., and Zipfel, P. F. (1996). Synthesis of the CC-chemokines MIP-1alpha, MIP-1beta, and RANTES is associated with a type 1 immune response. *The Journal of Immunology*, 157(8):3598–3604.

Schwaiger, F. W., Gostomski, D., Stear, M. J., Duncan, J. L., McKellar, Q. A., Epplen, J. T., and Buitkamp, J. (1995). An Ovine *Major Histocompatibility Complex DRB1 Allele* Is Associated with Low Faecal Egg Counts Following Natural, Predominantly *Ostertagia Circumcincta* Infection. *International Journal for Parasitology*, 25(7):815–822.

Scott, I., Pomroy, W. E., Kenyon, P. R., Smith, G., Adlington, B., and Moss, A. (2013). Lack of efficacy of monepantel against *Teladorsagia circumcincta* and *Trichostrongylus colubriformis*. *Veterinary Parasitology*, 198(1-2):166–171.

Seaton, D. S., Jackson, F., Smith, W. D., and Angus, K. W. (1989). Development of immunity to incoming radiolabelled larvae in lambs continuously infected with *Ostertagia circumcincta*. *Research in Veterinary Science*, 46(2):241–246.

Sedy, J. R., Gavrieli, M., Potter, K. G., Hurchla, M. A., Lindsley, R. C., Hildner, K., Scheu, S., Pfeffer, K., Ware, C. F., Murphy, T. L., and Murphy, K. M. (2005). B and T lymphocyte attenuator regulates T cell activation through interaction with herpesvirus entry mediator. *Nature Immunology*, 6(1):90–98.

Shakya, K. P., Miller, J. E., and Horohov, D. W. (2009). A Th2 type of immune response is associated with increased resistance to *Haemonchus contortus* in naturally infected Gulf Coast Native lambs. *Veterinary Parasitology*, 163(1-2):57–66.

Shariflou, M. R., Wade, C. M., Kijas, J. W., McCulloch, R., Windsor, P. A., Tammen, I., and Nicholas, F. W. (2013). Brachygnathia, cardiomegaly and renal hypoplasia syndrome (BCRHS) in Merino sheep maps to a 1.1-megabase region on ovine chromosome OAR2. *Animal Genetics*, 44(2):231–233.

Shaw, D. J., Vercruysse, J., Claerebout, E., and Dorny, P. (1998a). Gastrointestinal nematode infections of first-grazing season calves in Western Europe: associations between parasitological, physiological and physical factors. *Veterinary Parasitology*, 75(2-3):133–151.

Shaw, D. J., Vercruysse, J., Claerebout, E., and Dorny, P. (1998b). Gastrointestinal nematode infections of first-grazing season calves in Western Europe: general patterns and the effect of chemoprophylaxis. *Veterinary Parasitology*, 75(2-3):115–131.

Shaw, R. J., Gatehouse, T. K., and McNeill, M. M. (1998c). Serum IgE responses during primary and challenge infections of sheep with *Trichostrongylus colubriformis*. *International Journal for Parasitology*, 28(2):293–302.

Shaw, R. J., Morris, C. A., Green, R. S., Wheeler, M., Bisset, S. A., Vlassoff, A., and Douch, P. (1999). Genetic and phenotypic relationships among *Trichostrongylus colubriformis*-specific immunoglobulin E, anti-*Trichostrongylus colubriformis* antibody, immunoglobulin G1, faecal egg count and body weight traits in grazing Romney lambs. *Livestock Production Science*, 58(1):25–32.

Shaw, R. J., Morris, C. A., and Wheeler, M. (2013). Genetic and phenotypic relationships between carbohydrate larval antigen (CarLA) IgA, parasite resistance and productivity in serial samples taken from lambs after weaning. *International Journal for Parasitology*, 43(8):661–667.

Shaw, R. J., Morris, C. A., Wheeler, M., Tate, M., and Sutherland, I. A. (2012). Salivary IgA: A suitable measure of immunity to gastrointestinal nematodes in sheep. *Veterinary Parasitology*, 186(1–2):109–117.

Sheldrake, R. F., Husband, A. J., Watson, D. L., and Cripps, A. W. (1984). Selective transport of serum-derived IgA into mucosal secretions. *The Journal of Immunology*, 132(1):363–368.

Shinkai, A., Yoshisue, H., Koike, M., Shoji, E., Nakagawa, S., Saito, A., Takeda, T., Imabeppu, S., Kato, Y., Hanai, N., Anazawa, H., Kuga, T., and Nishi, T. (1999). A novel human CC chemokine, eotaxin-3, which is expressed in IL-4-stimulated vascular endothelial cells, exhibits potent activity toward eosinophils. *The Journal of Immunology*, 163(3):1602–1610.

Short, K., Wiradjaja, F., and Smyth, I. (2007). Let's stick together: the role of the Fras1 and Frem proteins in epidermal adhesion. *IUBMB life*, 59(7):427–435.

Silva, M. V. B., Sonstegard, T. S., Hanotte, O., Mugambi, J. M., Garcia, J. F., Nagda, S., Gibson, J. P., Iraqi, F. A., McClintock, A. E., Kemp, S. J., Boettcher, P. J., Malek, M., Van Tassell, C. P., and Baker, R. (2012). Identification of quantitative trait loci affecting resistance to gastrointestinal parasites in a double backcross population of Red Maasai and Dorper sheep. *Animal Genetics*, 43(1):63–71.

Singh, R. P., Massachi, I., Manickavel, S., Singh, S., Rao, N. P., Hasan, S., Mc Curdy, D. K., Sharma, S., Wong, D., Hahn, B. H., and Rehimi, H. (2013). The role of miRNA in inflammation and autoimmunity. *Autoimmunity reviews*, 12(12):1160–1165.

Sinski, E., Bairden, K., Duncan, J. L., Eisler, M. C., Holmes, P. H., McKellar, Q. A., Murray, M., and Stear, M. J. (1995). Local and plasma antibody responses to the parasitic larval stages of the abomasal nematode *Ostertagia circumcincta*. *Veterinary Parasitology*, 59(2):107–118.

Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., and Kasprzyk, A. (2009). BioMart – biological queries made easy. *BMC Genomics*, 10(1):22.

Smith, G. (1990). A mathematical model for the evolutions of anthelmintic resistance in a direct life cycle nematode parasite. *International Journal for Parasitology*, 20(7):913–921.

Smith, W. D. (1988). Mechanisms of Immunity to Gastrointestinal Nematodes of Sheep. In *Increasing Small Ruminant Productivity in Semi-arid Areas*, pages 275–286. Springer Netherlands, Dordrecht.

Smith, W. D. (2014). A commercial vaccine for Barber's pole worm – further development. Technical report, Sydney, Australia.

Smith, W. D., Jackson, F., Jackson, E., and Williams, J. (1983). Studies on the local immune response of the lactating ewe infected with *Ostertagia circumcincta*. *Journal of Comparative Pathology*, 93(2):295–305.

Smith, W. D., Jackson, F., Jackson, E., and Williams, J. (1985). Age immunity to *Ostertagia circumcincta*: comparison of the local immune responses of 4 1/2- and 10-month-old lambs. *Journal of Comparative Pathology*, 95(2):235–245.

Southcott, W. H., Major, G. W., and Barger, I. A. (1976). Seasonal pasture contamination and availability of nematodes for grazing sheep. *Crop and Pasture Science*, 27(2):277–286.

Spencer, C. C. A., Su, Z., Donnelly, P., and Marchini, J. L. (2009). Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. *PLoS Genetics*, 5(5):e1000477.

Sréter, T., Kassai, T., and Takács, E. (1994). The heritability and specificity of responsiveness to infection with Haemonchus contortus in sheep. *International Journal for Parasitology*, 24(6):871–876.

Stacey, M., Lin, H.-H., Hilyard, K. L., Gordon, S., and McKnight, A. J. (2001). Human Epidermal Growth Factor (EGF) Module-containing Mucin-like Hormone Receptor 3 Is a New Member of the EGF-TM7 Family That Recognizes a Ligand on Human Macrophages and Activated Neutrophils. *Journal of Biological Chemistry*, 276(22):18863–18870.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313.

Stear, M. J., Bairden, K., Bishop, S. C., Duncan, J. L., Karimi, S. K., McKellar, Q. A., and Murray, M. (1995a). Different patterns of faecal egg output following infection of Scottish Blackface lambs with *Ostertagia circumcincta*. *Veterinary Parasitology*, 59:29–38.

Stear, M. J., Bairden, K., Duncan, J. L., Holmes, P. H., McKellar, Q. A., Park, M., Strain, S., Murray, M., Bishop, S. C., and Gettinby, G. (1997). How hosts control worms. *Nature*, 389(6646):27–27.

Stear, M. J., Bairden, K., Innocent, G. T., Mitchell, S., Strain, S., and Bishop, S. C. (2004). The relationship between IgA activity against 4th-stage larvae and density-dependent effects on the number of 4th-stage larvae of *Teladorsagia circumcincta* in naturally infected sheep. *Parasitology*, 129(03):363–369.

Stear, M. J., Bairden, K., McKeller, Q. A., Scott, I., Strain, S., and Bishop, S. C. (1999a). The relationship between the number and size of nematodes in the abomasum and the concentration of pepsinogen in ovine plasma. *Research in Veterinary Science*, 67(1):89–92.

Stear, M. J. and Bishop, S. C. (1999). The curvilinear relationship between worm length and fecundity of *Teladorsagia circumcincta*. *International Journal for Parasitology*, 29(5):777–780.

Stear, M. J., Bishop, S. C., Doligalska, M., Duncan, J. L., Holmes, P. H., Irvine, J., McCririe, L., McKellar, Q. A., Sinski, E., and Murray, M. A. X. (1995b). Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with *Ostertagia circumcinct*a. *Parasite Immunology*, 17(12):643–652.

Stear, M. J., Bishop, S. C., Duncan, J. L., McKellar, Q. A., and Murray, M. (1995c). The repeatability of faecal egg counts, peripheral eosinophil counts, and plasma pepsinogen concentrations during deliberate infections with *Ostertagia circumcincta*. *International Journal for Parasitology*, 25(3):375–380.

Stear, M. J., Bishop, S. C., Henderson, N. G., and Scott, I. (2003). A key mechanism of pathogenesis in sheep infected with the nematode *Teladorsagia circumcincta*. *Animal Health Research Reviews*, 4(01):45–52.

Stear, M. J., Boag, B., Cattadori, I., and Murphy, L. (2009). Genetic variation in resistance to mixed, predominantly *Teladorsagia circumcincta* nematode infections of sheep: from heritabilities to gene identification. *Parasite Immunology*, 31(5):274–282.

Stear, M. J., Henderson, N. G., Kerr, A., McKellar, Q. A., Mitchell, S., Seeley, C., and Bishop, S. C. (2002). Eosinophilia as a marker of resistance to *Teladorsagia circumcincta* in Scottish Blackface lambs. *Parasitology*, 124(05).

Stear, M. J., Innocent, G. T., and Buitkamp, J. (2005). The evolution and maintenance of polymorphism in the major histocompatibility complex. *Veterinary Immunology and Immunopathology*, 108(1-2):53–57.

Stear, M. J., Park, M., and Bishop, S. C. (1996). The key components of resistance to *Ostertagia circumcincta* in lambs. *Parasitology Today*, 12(11):438–441.

Stear, M. J., Strain, S., and Bishop, S. C. (1999b). How lambs control infection with *Ostertagia circumcincta*. *Veterinary Immunology and Immunopathology*, 72(1-2):213–218.

Stear, M. J., Strain, S., and Bishop, S. C. (1999c). Mechanisms underlying resistance to nematode infection. *International Journal for Parasitology*, 29(1):51–56.

Strain, S., Bishop, S. C., Henderson, N. G., Kerr, A., McKellar, Q. A., Mitchell, S., and Stear, M. J. (2002). The genetic control of IgA activity against *Teladorsagia circumcincta* and its association with parasite resistance in naturally infected sheep. *Parasitology*, 124(05):545–552.

Strain, S. and Stear, M. J. (1999). The recognition of molecules from fourth-stage larvae of *Ostertagia circumcincta* by IgA from infected sheep. *Parasite Immunology*, 21(3):163–168.

Suda, T., Tsunoda, T., Uchida, N., Watanabe, T., Hasegawa, S., Satoh, S., Ohgi, S., Furukawa, Y., Nakamura, Y., and Tahara, H. (2006). Identification of *secernin 1* as a novel immunotherapy target for gastric cancer using the expression profiles of cDNA microarray. *Cancer science*, 97(5):411–419.

Sutherland, I. and Scott, I. (2009). *Gastrointestinal Nematodes of Sheep and Cattle: Biology and Control*. John Wiley & Sons.

Sutherland, I. A., Brown, A. E., Green, R. S., Miller, C. M., and Leathwick, D. M. (1999). The immune response of sheep to larval challenge with *Ostertagia circumcincta* and *O. ostertagi*. *Veterinary Parasitology*, 84(1-2):125–135.

Sutherland, I. A., Damsteegt, A., Miller, C. M., and Leathwick, D. M. (2008). Multiple species of nematodes resistant to ivermectin and a benzimidazole-levamisole combination on a sheep farm in New Zealand. *New Zealand Veterinary Journal*, 56(2):67–70.

Sykes, A. R. and Coop, R. L. (2001). Interactions between nutrition and gastrointestinal parasitism in sheep. *New Zealand Veterinary Journal*, 49(6):222–226.

Sykes, A. R. and Poppi, D. P. (1986). Effects of parasitism on metabolism in sheep. In Ross, A. D., editor, *Control of internal parasites in sheep, Animal Industries Workshop, Lincoln College*, pages 25–36, Lincoln, New Zealand.

Takeshima, S.-n., Sarai, Y., Saitou, N., and Aida, Y. (2009). MHC class II DR classification based on antigen-binding groove natural selection. *Biochemical and Biophysical Research Communications*, 385(2):137–142.

Tapon, N., Nagata, K., Lamarche, N., and Hall, A. (1998). A new rac target POSH is an SH3-containing scaffold protein involved in the JNK and NF-kappaB signalling pathways. *EMBO J*, 17(5):1395–1404.

Taylor, M. A., Coop, R. L., and Wall, R. L. (2007). *Veterinary Parasitology*. Wiley-Blackwell, 3rd edition.

Terefe, G., Lacroux, C., Andreoletti, O., Grisez, C., Prevot, F., Bergeaud, J. P., Penicaud, J., Rouillon, V., Gruner, L., Brunel, J. C., Francois, D., Bouix, J., Dorchies, P., and Jacquiet, P. (2007). Immune response to *Haemonchus contortus* infection in susceptible (INRA 401) and resistant (Barbados Black Belly) breeds of lambs. *Parasite Immunology*, 29(8):415–424.

The Bovine HapMap Consortium, Gibbs, R. A., Taylor, J. F., Van Tassell, C. P., Barendse, W., Eversole, K. A., Gill, C. A., Green, R. D., Hamernik, D. L., Kappes, S. M., Lien, S., Matukumalli, L. K., McEwan, J. C., Nazareth, L. V., Schnabel, R. D., Weinstock, G. M., Wheeler, D. A., Ajmone-Marsan, P., Boettcher, P. J., Caetano, A. R., Garcia, J. F., Hanotte, O., Mariani, P., Skow, L. C., Sonstegard, T. S., Williams, J. L., Diallo, B., Hailemariam, L., Martinez, M. L., Morris, C. A., Silva, L. O., Spelman, R. J., Mulatu, W., Zhao, K., Abbey, C. A., Agaba, M., Araujo, F. R., Bunch, R. J., Burton, J., Gorni, C., Olivier, H., Harrison, B. E., Luff, B., Machado, M. A., Mwakaya, J., Plastow, G., Sim, W., Smith, T., Thomas, M. B., Valentini, A., Williams, P., Womack, J., Woolliams, J. A., Liu, Y., Qin, X., Worley, K. C., Gao, C., Jiang, H., Moore, S. S., Ren, Y., Song, X. Z., Bustamante, C. D., Hernandez, R. D., Muzny, D. M., Patil, S., San Lucas, A., Fu, Q., Kent, M. P., Vega, R., Matukumalli, A., McWilliam, S., Sclep, G., Bryc, K., Choi, J., Gao, H., Grefenstette, J. J., Murdoch, B., Stella, A., Villa-Angulo, R., Wright, M., Aerts, J., Jann, O., Negrini, R., Goddard, M. E., Hayes, B. J., Bradley, D. G., Barbosa da Silva, M., Lau, L. P., Liu, G. E., Lynn, D. J., Panzitta, F., Dodds, K. G., Sequencing, T. B. G., and Consortium, A. (2009).

Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, 324(5926):528–532.

The MHC Sequencing Consortium (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature*, 401(6756):921–923.

Urb, M. and Sheppard, D. C. (2012). The Role of Mast Cells in the Defence against Pathogens. *PLoS Pathogens*, 8(4):e1002619.

Vallance, B. A., Blennerhassett, P. A., and Collins, S. M. (1997). Increased intestinal muscle contractility and worm expulsion in nematode-infected mice. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, 272(2):G321–G327.

Vallender, E. J. and Lahn, B. T. (2004). Positive selection on the human genome. *Human Molecular Genetics*, 13(suppl 2):R245–R254.

van Deventer, S. and Neefjes, J. (2010). The Immunoproteasome Cleans up after Inflammation. *Cell*, 142(4):517–518.

Van Houtert, M. F. J. and Sykes, A. R. (1996). Implications of nutrition for the ability of ruminants to withstand gastrointestinal nematode infections. *International Journal for Parasitology*, 26(11):1151–1167.

Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1(1):1–30.

Vanimisetti, H. B., Greiner, S. P., Zajac, A. M., and Notter, D. R. (2004). Performance of hair sheep composite breeds: Resistance of lambs to *Haemonchus contortus*. *Journal of Animal Science*, 82(2):595–604.

Venturina, V. M., Gossner, A. G., and Hopkins, J. (2013). The immunology and genetics of resistance of sheep to *Teladorsagia circumcincta*. *Veterinary Research Communications*, 37(2):171–181.

Vilà, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., Rice, J. E., Honeycutt, R. L., Crandall, K. A., Lundeberg, J., and Wayne, R. K. (1997). Multiple and ancient origins of the domestic dog. *Science*, 276(5319):1687–1689.

Visscher, P. M., Yang, J., and Goddard, M. E. (2010). A Commentary on 'Common SNPs Explain a Large Proportion of the Heritability for Human Height' by Yang et al. (2010). *Twin Res Hum Genet*, 13(6):517–524.

Vlassoff, A. (1973). Seasonal incidence of infective trichostrongyle larvae on pasture grazed by lambs. *New Zealand journal of experimental agriculture*, 1(3):293–301.

Vlassoff, A., Leathwick, D. M., and Heath, A. C. (2001). The epidemiology of nematode infections of sheep. *New Zealand Veterinary Journal*, 49(6):213–221.

Vlassoff, A. and McKenna, P. B. (1994). Nematode parasites of economic importance in sheep in New Zealand. *New Zealand Journal of Zoology*, 21:1–8.

Voehringer, D. (2013). Protective and pathological roles of mast cells and basophils. *Nature Reviews Immunology*, 13(5):362–375.

Waghorn, T. S., Leathwick, D. M., Rhodes, A. P., Lawrence, K. E., Jackson, R., Pomroy, W. E., West, D. M., and Moffat, J. R. (2006). Prevalence of anthelmintic resistance on sheep farms in New Zealand. *New Zealand Veterinary Journal*, 54(6):271–277.

Wakelin, D. (1996). *Immunity to Parasites*. How Parasitic Infections are Controlled. Cambridge University Press, Cambridge, 2nd edition.

Waller, P. J. and Faedo, M. (1996). The prospects for biological control of the free-living stages of nematode parasites of livestock. *International Journal for Parasitology*, 26(8-9):915–925.

Wang, K., Li, M., and Bucan, M. (2007). Pathway-Based Approaches for Analysis of Genomewide Association Studies. *The American Journal of Human Genetics*, 81(6):1278–1283.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57–63.

Watanabe, N., Gavrieli, M., Sedy, J. R., Yang, J., Fallarino, F., Loftin, S. K., Hurchla, M. A., Zimmerman, N., Sim, J., Zang, X., Murphy, T. L., Russell, J. H., Allison, J. P., and Murphy, K. M. (2003). BTLA is a lymphocyte inhibitory receptor with similarities to CTLA-4 and PD-1. *Nature Immunology*, 4(7):670–679.

Way, G., Morrice, N., Smythe, C., and O'Sullivan, A. J. (2002). Purification and identification of secernin, a novel cytosolic protein that regulates exocytosis in mast cells. *Molecular biology of the cell*, 13(9):3344–3354.

Welch, J. S., Ricote, M., Akiyama, T. E., Gonzalez, F. J., and Glass, C. K. (2003). PPARγ and PPARδ negatively regulate specific subsets of lipopolysaccharide and IFN-γ target genes in macrophages. *Proceedings of the National Academy of Sciences*, 100(11):6712–6717.

Weston, C. R. and Davis, R. J. (2007). The JNK signal transduction pathway. *Cell Regulation*, 19(2):142–149.

Windon, R. G. (1990). Selective breeding for the control of nematodiasis in sheep. *Revue scientifique et technique (International Office of Epizootics)*, 9(2):555–576.

Wion, K. L., Kirchgessner, T. G., Lusis, A. J., Schotz, M. C., and Lawn, R. M. (1987). Human lipoprotein lipase complementary DNA sequence. *Science*, 235(4796):1638–1641.

Woolaston, R. R., Manueli, P., Eady, S. J., Barger, I. A., Le Jambre, L. F., Banks, D. J. D., and Windon, R. G. (1996). The value of circulating eosinophil count as a selection criterion for resistance of sheep to trichostrongyle parasites. *International Journal for Parasitology*, 26(1):123–126.

Woolaston, R. R., Windon, R. G., and Gray, G. D. (1991). Genetic variation in resistance to internal parasites in Armidale experimental flocks. In Gray, G. D. and Woolaston, R. R., editors, *Breeding for Disease Resistance in Sheep*, pages 1–9. Australian Wool Corporation, Melbourne, Australia.

Wrigley, J. J., McArthur, M. M., McKenna, P. B. P., and Mariadass, B. B. (2006). Resistance to a triple combination of broad-spectrum anthelmintics in naturally-acquired *Ostertagia circumcincta* infections in sheep. *New Zealand Veterinary Journal*, 54(1):47–49.

Wu, C. w., Livesey, M., Schmoller, S. K., Manning, E. J. B., Steinberg, H., Davis, W. C., Hamilton, M. J., and Talaat, A. M. (2007). Invasion and Persistence of *Mycobacterium avium* subsp. *paratuberculosis* during Early Stages of Johne's Disease in Calves. *Infection and Immunity*, 75(5):2110–2119.

Yamaguchi, T., Hirota, K., Nagahama, K., Ohkawa, K., Takahashi, T., Nomura, T., and Sakaguchi, S. (2007). Control of Immune Responses by Antigen-Specific Regulatory T Cells Expressing the Folate Receptor. *Immunity*, 27(1):145–159.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., and Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7):565–569.

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences*, 13(5):555–556.

Yang, Z. (2006). *Computational Molecular Evolution*. Oxford University Press.

Yang, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*, 24(8):1586–1591.

Yang, Z. and Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends in Ecology & Evolution*, 15(12):496–503.

Yang, Z. and Nielsen, R. (1998). Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution*, 46(4):409–418.

Yang, Z. and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution*, 19(6):908–917.

Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.-M. K. (2000). Codon-Substitution Models for Heterogeneous Selection Pressure at Amino Acid Sites. *Genetics*, 155(1):431–449.

Yang, Z., Wong, W. S. W., and Nielsen, R. (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, 22(4):1107–1118.

Yokoyama, S., Tada, T., Zhang, H., and Britt, L. (2008). Elucidation of phenotypic adaptations: Molecular analyses of dim-light vision proteins in vertebrates. *Proceedings of the National Academy of Sciences*, 105(36):13480–13485.

Young, A. J. (1999). The physiology of lymphocyte migration through the single lymph node in vivo. *Seminars in Immunology*, 11(2):73–83.

Zhang, F.-R., Huang, W., Chen, S.-M., Sun, L.-D., Liu, H., Li, Y., Cui, Y., Yan, X.-X., Yang, H.-T., Yang, R.-D., Chu, T.-S., Zhang, C., Zhang, L., Han, J.-W., Yu, G.-Q., Quan, C., Yu, Y.-X., Zhang, Z., Shi, B.-Q., Zhang, L.-H., Cheng, H., Wang, C.-Y., Lin, Y., Zheng, H.-F., Fu, X.-A., Zuo, X.-B., Wang, Q., Long, H., Sun, Y.-P., Cheng, Y.-L., Tian, H.-Q., Zhou, F.-S., Liu, H.-X., Lu, W.-S., He, S.-M., Du, W.-L., Shen, M., Jin, Q.-Y., Wang, Y., Low, H.-Q., Erwin, T., Yang, N.-H., Li, J.-Y., Zhao, X., Jiao, Y.-L., Mao, L.-G., Yin, G., Jiang, Z.-X., Wang, X.-D., Yu, J.-P., Hu, Z.-H., Gong, C.-H., Liu, Y.-Q., Liu, R.-Y., Wang, D.-M., Wei, D., Liu, J.-X., Cao, W.-K., Cao, H.-Z., Li, Y.-P., Yan, W.-G., Wei, S.-Y., Wang, K.-J., Hibberd, M. L., Yang, S., Zhang, X.-J., and Liu, J.-J. (2009). Genomewide Association Study of Leprosy. *New England Journal of Medicine*, 361(27):2609–2618.

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S. L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M., Lu, C., Chen, S.-C., Wei, L., Collins-Underwood, J. R., Ma, J., Roberts, K. G., Pounds, S. B., Ulyanov, A., Becksfort, J., Gupta, P., Huether, R., Kriwacki, R. W., Parker, M., McGoldrick, D. J., Zhao, D., Alford, D., Espy, S., Bobba, K. C., Song, G., Pei, D., Cheng, C., Roberts, S., Barbato, M. I., Campana, D., Coustan-Smith, E., Shurtleff, S. A., Raimondi, S. C., Kleppe, M., Cools, J., Shimano, K. A., Hermiston, M. L., Doulatov, S., Eppert, K., Laurenti, E., Notta, F., Dick, J. E., Basso, G., Hunger, S. P., Loh, M. L., Devidas, M., Wood, B., Winter, S., Dunsmore, K. P., Fulton, R. S., Fulton, L. L., Hong, X., Harris, C. C., Dooling, D. J., Ochoa, K., Johnson, K. J., Obenauer, J. C., Evans, W. E., Pui, C.-H., Naeve, C. W., Ley, T. J., Mardis, E. R., Wilson, R. K., Downing, J. R., and Mullighan, C. G. (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature*, 481(7380):157–163.

Zhao, X., Dittmer, K. E., Blair, H. T., Thompson, K. G., Rothschild, M. F., and Garrick, D. J. (2011). A Novel Nonsense Mutation in the *DMP1* Gene Identified by a Genome-Wide Association Study Is Responsible for Inherited Rickets in Corriedale Sheep. *PLoS ONE*, 6(7):e21739.

Zhou, L. J. and Tedder, T. F. (1995). Human blood dendritic cells selectively express CD83, a member of the immunoglobulin superfamily. *The Journal of Immunology*, 154(8):3821–3835.

Zhu, J., Yamane, H., and Paul, W. E. (2010). Differentiation of effector CD4 T cell populations (*). *Annual Review of Immunology*, 28:445–489.