THE UNIVERSITY
of EDINBURGH

Investigating Differential T cell Polarization in the Two Pathological Forms of Sheep Paratuberculosis

Louise Nicol

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of Edinburgh

2016

Declaration of Originality

I declare that all work included in this thesis is my own work, except where otherwise stated. No part of this work has been submitted for any other degree of professional qualification.

Louise Nicol, 2016

Acknowledgements

First of all I would like to thank my supervisors Professor John Hopkins, Dr. Bob Dalziel and Dr. Craig Watkins. John has provided endless support and enthusiasm throughout my PhD and taught me the invaluable lesson of 'belt, braces and self-supporting trousers'. Bob and Craig both provided excellent scientific support and much needed 'chin up' encouragement during difficult times in my research. I'd like to thank Dr. Anton Gossner for his help and advice during my time at Roslin. Thanks to Hazel Wilkie and Francesca Chionini who have also contributed to this study.

A huge thank you to all of the friendly lab rats (you know who you are); you made working in the lab good fun.

Finally, I'd like to thank my nearest and dearest who supported me every step of the way and saw me through the stressful times; for that I am immensely grateful.

Abstract

Paratuberculosis is a chronic enteropathy of ruminants that presents as two distinct disease forms in sheep; paucibacillary (or tuberculoid) and multibacillary (or lepromatous) disease. The immunopathology of paucibacillary and multibacillary sheep paratuberculosis has been linked to inflammatory Th1/Th17 cell and Th2/macrophage responses respectively. IL23 and IL25 are key to the development of these responses by interaction with their complex receptors, IL23R/IL12RB1 and IL17RA/IL17RB. Furthermore, the polarization of T cells and the development of appropriate immune responses is controlled by the master regulator transcription factor; T-bet, GATA3, RORγt and RORα. In humans, variations in the structure, sequence and/or expression of the genes encoding these proteins have been implicated in the different pathological forms of tuberculosis and leprosy, and gastrointestinal inflammatory disorders such as Crohn's disease. In the current study, sequencing has identified multiple transcript variants of sheep IL23R, IL12RB1 and IL17RB and a single IL17RA transcript. RT-qPCR assays were developed for the cytokine receptor variants identified in this study and known transcript variants of the transcription factor genes. Expression levels were compared in the ileo cecal lymph node of paucibacillary or multibacillary paratuberculosis diseased sheep. Of the cytokine receptors; the IL12RB1v3 variant, which lacks the receptor activation motif, was differentially expressed and was significantly increased in multibacillary disease; this may contribute to high Th2 responses. Full length IL17RB was differentially expressed and was significantly increased in multibacillary pathology, which may also contribute to Th2 polarization. IL17RA was significantly increased in paucibacillary disease. The contrast between the IL17RA and IL17RB results may indicate that, in addition to Th1 cells, Th17 T cells are also involved in paucibacillary pathology. Of the transcription factor transcripts; full length TBX21 (T-bet) was differentially expressed and was significantly increased in paucibacillary disease; this may explain increased Th1 responses in these sheep. Full length GATA3 was significantly increased in paucibacillary compared to multibacillary sheep, suggesting a loss of Th2 responses in late-stage multibacillary pathology.

RORAv1 variant was differentially expressed and was significantly increased in paucibacillary pathology, indicating a role of Th17 T cells in paucibacillary pathology.

Contents

Abbreviations

Ab	Antibody
ALL	Acute lymphoblastic leukaemia
aa	Amino acid
APC	Antigen presenting cell
AS	Alternative splicing
bp	Base pair
C3b	Complement component 3b
CCL	Chemokine ligand
CD	Crohn's disease
CSF	Colony stimulating factor
CXCL	C-X-C motif chemokine
DC	Dendritic cell
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
ESE	Exonix splicing enhancers
ESS	Exonic splicing silencers
EST	Expressed sequence tag
FB	Fibronectin
Foxp3	Forkhead box p3
GATA3	GATA-binding protein 3
gDNA	Genomic deoxyribonucleic acid
GWAS	Genome wide association studies
H&E	Haematoxylin and eosin
HEK	Human embryo kidney
IBD	Inflammatory bowel disease
IFNγ	Interferon gamma

Ig	Immunoglobulin
IGF	Insulin-like growth factor
IL	Interleukin
ILC	Innate lymphoid cell
IS1626	Insertion sequence 1626
IS900	Insertion sequence 900
ISE	Intronic splicing enhancers
ISS	Intronic splicing silencers
JD	Johne's disease
LAM	lipoarabinomannan
L-lep	Lepromatous leprosy
M. bovis	*Mycobacterium bovis*
M. leprae	*Mycobacterium leprae*
MAA	*Mycobacterium avium*subsp. *avium*
Man-LAM	mannosylatedlipoarabinomannan
MAP	*Mycobacterium avium* subspecies *paratuberculosis*
MAV	*Mycobacterium aviumsubspecies*
MBP	Mannose binding protein
MCP	Monocyte chemoattractant protein
MHC	Major histocompatibility complex
mRNA	Messenger ribonucleic acid
MSMD	Mendelian susceptibility to mycobacterial disease
MTB	*Mycobacterium tuberculosis*
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next generation sequencing
NK	Natural killer
NOD	Nucleotide-binding oligomerization domain
NRAMP	Natural resistance-associated macrophage protein

paraTB	Paratuberculosis
PBMC	peripheral blood mononuclear cell
PPD	Purified protein derivatives
PRR	Pattern recognition receptor
R	Receptor
RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid sequencing
RT-PCR	Reverse transcription polymerase chain reaction
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SNP	Single nucleotide polymorphism
STAT	Signal transducer and activator of transcription
TB	Tuberculosis
T-bet	T-box transcription factor
TGFβ	Transforming growth factor beta
Th	T-helper
T-lep	Tuberculoid leprosy
TLR	Toll-like receptor
TNF-α	Tumour necrosis factor alpha
T-reg	T-regulatory
TSLP	Thymic stromal lymphopoietin
UC	Ulcerative colitis
UTR	Untranslated region
ZN	Ziehl-Neelsen

List of figures

List of tables

Chapter 1

Introduction

Metazoans are, constitutively and transiently, colonized by a variety of microorganisms that engage in mutualistic interactions with their host. However, this relationship can also be antagonistic and the nature of such interactions is often poorly understood. Recent studies have begun to elucidate the receptor and signalling pathways involved in the host response to both commensal and pathogenic microbes (Iwasaki and Medzhitov, 2010). The immune system uses these sensing pathways to induce necessary defence mechanisms to maintain host-microbe homeostasis. Vertebrates possess two types of immunity: innate and adaptive, that work together to protect the host. The innate immune system is the non-specific first line of defence and is genetically programmed to detect shared features of invading microbes. Cells of the innate immune system include dendritic cells/macrophages, epithelial cells and neutrophils, eosinophils and basophils. In contrast, the adaptive immune system is highly specific and is composed of T and B lymphocytes. The adaptive immune system employs antigen receptors which provide the host with immune 'memory'; allowing the induction of an appropriate immune response in the event of subsequent exposures to specific antigens. In the past, innate and adaptive immune systems have been viewed as two distinct organisations providing responses at separate stages of antigen exposure; however as research continues, it is becoming increasingly evident that each response requires the other to function and that these responses operate simultaneously to deliver effective host protection. The innate immune system can be described as the foundations upon which the adaptive immune system develops (Chaffey, 2003). Given that cells of the innate system are not always able to eliminate foreign organisms; cells of the adaptive immune response have evolved to provide a more complete means of defence. The successful control of pathogenic infection requires efficient communication between cells of both innate and adaptive immune systems. Adaptive immune responses are optimised by the activities of the innate immune system and efficient adaptive immunity is required to protect the host against persistent pathogens. This project is focused on components involved in the induction and maintenance of an adaptive T cell immune response to mycobacterial infection in sheep.

1.1 Paratuberculosis

Paratuberculosis (paraTB or Johne's disease - JD) is an infectious, granulomatous disorder that causes chronic inflammation of the gastrointestinal tract of ruminants. Infection occurs predominantly in cattle and sheep; although can occur in goats, deer and wild ruminants and results in wasting and loss of condition (Machackova *et al.*, 2004). *Mycobacterium avium* subspecies *paratuberculosis* (MAP) is the causative agent of paraTB. This slow growing, acid fast, bacterial infection is responsible for extensive economic losses worldwide related to shorter lifespan of infected animals, increased replacement costs and loss of productivity (Ott *et al.*, 1999). Animals become infected at a young age with most infections occurring during the neonatal stage (Seitz *et al.*, 1989). The main route of disease transmission is faecal-oral, including ingestion of bacteria from infected teats, water, feed or pasture which has been contaminated from the diarrhoea of infected animals (Manning and Collins, 2001). Whitlock and Buergelt first classified the infection status of cattle into 4 stages based on the severity of clinical signs (Whitlock and Buergelt, 1996):

1. Silent stage represents young animals under the age of 2 that have no clinical symptoms of disease. These animals shed bacteria at low levels, difficult to detect in faeces, but have MAP present in the gut tissues.
2. Subclinical stage where adult animals are infected but show no clinical symptoms. Antibodies and immune mediated cell responses can be detected during this phase of infection. Cattle with subclinical infection often develop problems with mastitis and infertility (Fawcett *et al.*, 1995).
3. Clinical paraTB infection resulting in weight loss and diarrhoea. Faecal culture of MAP gives positive results.
4. Advanced clinical paraTB characterised by persistent diarrhoea, severe weight loss, oedema of the gut/throat. These animals are sent for emergency slaughter or will die of dehydration.

Susceptibility of animals to infection with MAP is largely age-dependant with the greatest risk period being during the first few months of life. Animals develop increased tolerance to infection from 4 months to 1 year old, by which time they

achieve a level of resistance comparable to an adult (Arsenault *et al.*, 2014). The higher susceptibility to infection seen in younger animals is thought to be due to the general immaturity of the innate and adaptive immune systems. As a result of transmission the majority of animals within a herd that has a case of infection will, in turn, become affected. Information on the host immune response in sheep paraTB is limited; immunological data is predominantly derived from studies of bovine paraTB where disease typically progresses from a subclinical to paucibacillary phase, to an eventual multibacillary endpoint. Although there have been instances of progression of disease noted in sheep (Dennis *et al.*, 2011), paucibacillary and multibacillary forms of paraTB in sheep mostly appear to be distinct end points, separate from one another. Both paucibacillary/multibacillary pathologies are equally fatal in sheep; however there is little evidence to conclude that all asymptomatic animals (infected but present no clinical pathology) will succumb to the disease (Smeed *et al.*, 2007).

1.2 Economic impact of paratuberculosis

There is limited information on the prevalence of ovine paraTB; it is estimated that approximately 0.9-1.3% farmed sheep herds in New Zealand (Morris *et al.*, 2006) and 5-15% of sheep herds in Australia become infected per annum (Reddacliff *et al.*, 2006). Between 1993 and 2002 there were approximately 465 recorded cases of cattle and sheep paraTB in Scotland (Stewart *et al.*, 2005); however this figure reflects only reported cases. Not all cases are reported to Health Protection Scotland as Johne's disease is not currently classified as a notifiable disease. Accurate determination of paraTB prevalence is difficult due to the limitations of available diagnostic techniques (Gillan *et al.*, 2010). In 2010 a DEFRA report estimated the annual UK cost of paraTB at ~£30min all species (Geraghty *et al.*, 2014.DEFRA, 2010). Efforts to control paraTB through improving animal management has had limited success. The ability of the pathogen to persist in the environment and the lack of a reliable detection method of infected animals make traditional approaches to control the spread of infection largely ineffective. Prevalence of the disease is increasing throughout most parts of the world, particularly in temperate climates with frequent rainfall (Pavlik *et al.*, 2000). Animals that are housed in tight confinement show a higher incidence of disease as it has been shown that restrictive conditions contribute to the spread of disease within, and between, herds (Chiodini *et al.*, 1984). The lack of a uniformly defined prevalence estimate hinders the comparison of prevalence across countries and regions. Nielsen and Toft reviewed the prevalence of paraTB in farmed animals in Europe and found that the true prevalence among cattle appeared to be approximately 20%, but further highlighted that there is a need for well-designed method of determining the prevalence of MAP infections within herds and between herds (Nielsen and Toft, 2009.Nielsen *et al.*, 2000).

1.3 Symptoms and Pathology

ParaTB is characterised by granulomatous lesions of the gut. These lesion are formed as immune barriers, functioning to limit infection and control disease progression. Although animals are infected at a young age, the chronic nature of the disease entails that clinical signs may not be visible for a significant period of time, up to 3-5 years post infection (Chiodini *et al.*, 1984). Clinical signs of MAP infection include weight loss; oedema of the abdomen caused by thickening of the abdominal wall and necrosis of lymph glands; diarrhoea (most common in cattle); and eventual death (Ayele *et al.*, 2001). These are caused by inflammation within the intestine resulting in a loss of gut barrier function; lower absorbance of water and nutrients in the intestine contributes to pathology with animals developing anorexia. Three forms of the disease have been described in sheep based on gross gut pathology; with approximately only 30% of infected animals becoming clinically affected. Most animals (~70%) remain asymptomatic, in that they are infected but show no pathology and develop no clinical disease (Smeed *et al.*, 2008.Clarke, 1997). The pathological forms are unambiguously differentiated into multibacillary (lepromatous) and paucibacillary (tuberculoid) forms. Post mortem examination of paucibacillary infected animals describes a corrugated thickening of the intestine with loss of convoluted villi structure. Further examination by histopathology reveals a lymphocytic infiltrate with very low/no mycobacteria present, observed by Ziehl-Neelsen staining. Characteristically, paucibacillary lesions consist of lymphocyte infiltration with very few macrophages. In contrast, multibacillary infected animals have a large macrophage infiltrate; these macrophages are infected with intracellular MAP and so many mycobacteria are present in the intestinal tissues (Clarke, 1997). Again, thickening of the gut wall is noted with loss of villi integrity, this can be seen in figure 1.1. Given that paraTB presents as a spectrum of pathology, similar to that seen in human leprosy, a wide range of histopathology has been observed, including the development of lesions in the Peyer's patches (Corpa *et al.*, 2000).

Similarities have been noted between the pathology of paraTB and the pathology of the major human diseases mycobacterial diseases, tuberculosis (TB) and leprosy. It is yet unknown what causes each of the disease outcomes; however as with TB and

leprosy, a host genetic susceptibility has been suggested; this is described in greater detail in section 1.10.

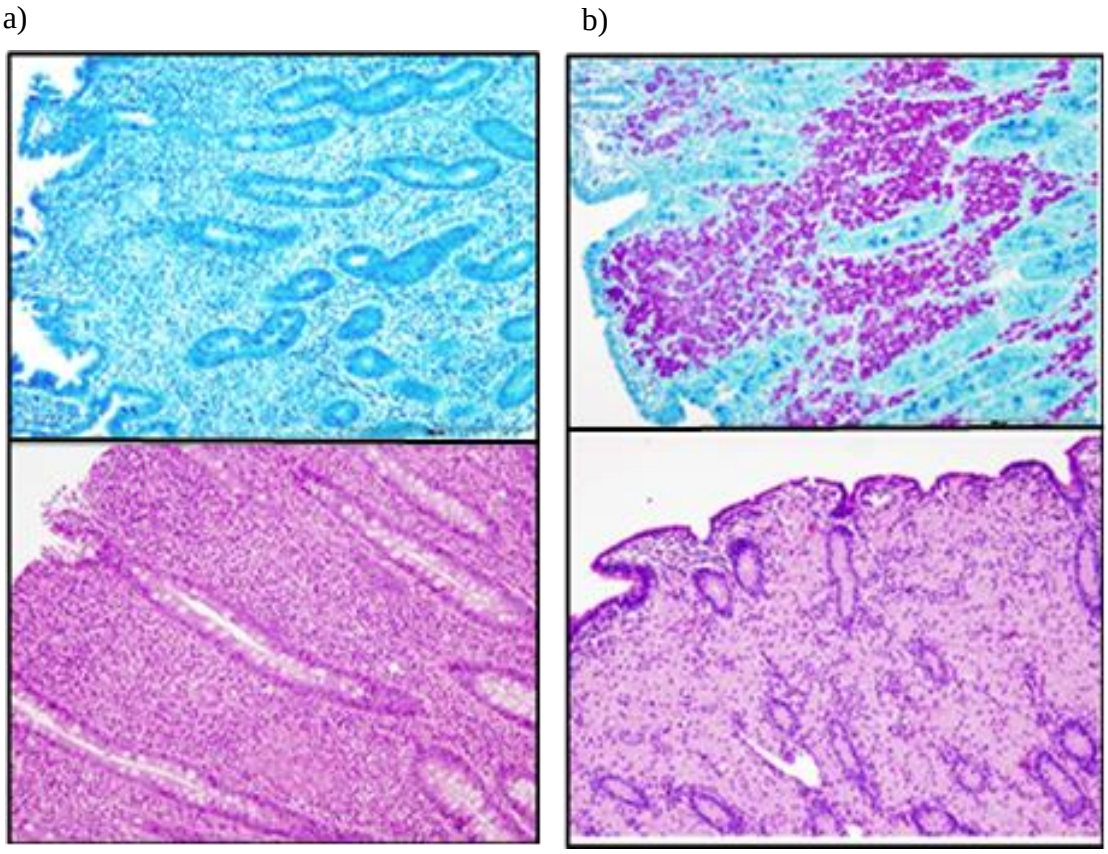


Figure 1.1 Representative sections of ileum showing a) paucibacillary and b) multibacillary MAP infection. Histology sections stained using Zeihl-Neelsen (top) and haematoxylin and eosin (bottom) methods. (Smeed *et al.*, 2007).

1.4 Diagnosis and Detection of MAP infection

A major difficulty when trying to diagnose paraTB is the identification of subclinical cases. Infected animals can show no symptoms of disease for up to 5 years after infection, by this time clinical symptoms are manifested and the shedding of bacteria will already have contaminated the environment. The most common way of identifying shedding of MAP bacteria by infected animals is faecal culture; however MAP has proven to be extremely difficult and slow to culture in the laboratory, especially when present in low abundance (Manning and Collins, 2001). MAP requires the addition of mycobactin, a growth factor derived from mycobacteria, to culture medium in order to grow and can take up to 16 weeks to be visibly detected in medium (Behr and Collins, 2010).Current tests are not highly sensitive or specific and so development of new tests is required.

1.4.1 Detection of immune response

Current bovine tuberculin skin tests detect delayed hypersensitivity and have been used for the diagnosis of bovine TB for many years. These tests are based on the measurement of a delayed-type hypersensitivity (DTH) to intradermally injected tuberculins. The single intradermal cervical comparative test (SICCT) consists of a comparative analysis of immune reaction to *Mycobacterium bovis* and *Mycobacterium avium* subsp. *avium* purified protein derivatives (PPD). When the skin thickness (due to DTH) to *M. Bovis* is 2mm greater than of the response to *M. avium* the animal is considered a TB reactor. Similarly, tests involving intradermal inoculation of MAP extract (johnin) have been used to detect MAP infection; thickness of the skin at the site of injection is monitored. An increase in thickness of >4mm within 72hours is considered as a positive result for MAP infection (Strain *et al.*, 2011.Strain, 2011). However, this test shows a lack of specificity due to cross-reactivity with both *M. bovis* and *M. avium avium*, is highly subjective and provides no information on the infection status of the animal (asymptomatic or pathology case).

Enzyme-linked Immunosorbent Assays (ELISA) are commonly used test for screening large herds/flocks of animals. Animals that have a positive ELISA for paraTB are confirmed by faecal culture. The sensitivity of these tests compared with faecal culture varies widely; from 13.7%, in which case animals are described as low shedders, to 100%; considered high shedders of MAP (Eda *et al.*, 2006). An IFNγ ELISA (Bovigam CSL) is used to detect the increase in IFNγ production in pre-clinical animals infected with MAP; in vitro IFNγ production is assessed following incubation of whole blood with MAP antigens. However responses to this test are time-dependant and in animals where an IFNγ response is short lived, the peak of IFNγ production may not be detected. An evaluation of five antibody detection tests based on ELISA technology for the diagnosis of bovine paraTB found that results were variable between the tests, highlighting the issue that antibody based diagnostics are not always reliable (Collins *et al.*, 2005). ELISA tests used to detect antibody levels can detect immune reactions to MAP specific proteins; however this test does not distinguish between multibacillary and paucibacillary disease forms and can often detect MAP antibodies in both forms of disease; rendering this test ineffective as a way of differentiating multibacillary and paucibacillary infection (Gillan *et al.*, 2010). Research continues to identify new proteins in MAP that will have higher sensitivity than tests currently available (Collins, 2005. Fernández, 2012).

1.4.2 Detection of MAP DNA

The MAP bacterium genome was sequenced in 2004 and it was found to share common genetic homogeneity with members of *Mycobacterium avium* subspecies (MAV) (Thorel *et al.*, 1990); however the genotype of MAP can be specifically distinguished from other mycobacteria by the presence of the insertion sequence *IS900* (Green *et al.*, 1989.Bull *et al.*, 2000). Amplification of this *IS900* by PCR/qPCR methods has been widely used as a diagnostic tool to detect MAP in infected animals and humans (Autschbach *et al.*, 2005). Selection of adequate primers for *IS900* detection is essential for amplification of this element. Similar sequences, '*IS900*-like insertions', have been identified in other mycobacteria; although these sequences are not identical to *IS900*, they are known to share

sequence identity often leading to complications when designing efficient detection assays specific to MAP. For instance, the *IS1626* sequence of *M. avium* shares approximately 82% identity with *IS900*, other elements present in MAP are also present in *M. intracellulare;* therefore are non-specific and unable to be used as a diagnostic tool for the detection of MAP (Puyang *et al.*, 1999). *IS900* PCR is considered one of the most reliable detection methods when confirming MAP infection (Sidoti *et al.*, 2011).

1.5 Treatment and control of MAP infection

Treatment and control of paraTB is extremely difficult and can be very costly. A potential way of reducing the occurrence of animals shedding MAP is by vaccination. The immune response elicited by the vaccine aims to reduce bacterial shedding from infected animals. The Gudair^{TM} vaccine is a whole cell preparation of heat killed MAP which elicits both humoral and cellular immune responses; however such responses can also be detrimental; approximately 50% of vaccinated animals will develop granulomatous lesions at the site of injection (Eppleston and Windsor, 2007). Cattle vaccinated with Gudair^{TM} generate a response that may interfere with the current *M. bovis* SICCT, rendering this test ineffective and leading to an increase incidence of false positive *M. bovis* reactors (Eppleston and Windsor, 2007). Current vaccines used to prevent paraTB are not DIVA (differentiating infected from vaccinated) and so often interfere with serological tests for both MAP and *M. bovis* infections (Bastida and Juste, 2011). Legislation in many countries states that animals that test SICCT positive for tuberculosis are banned from international trade and should be slaughtered; unless it can be proven that the positive result is a consequence of prior vaccination and not tuberculosis infection (Bastida and Juste, 2011). Furthermore, there have been cases of vaccinated sheep that have gone on to develop paraTB of the multibacillary form; continuing to contribute to the spread of paraTB (Reddacliff *et al.*, 2006). These are some of the hurdles affecting MAP vaccine approval by many regulatory and agricultural authorities throughout the world. Aside from these complications vaccination is expensive and the cost of MAP in the UK, especially with low milk prices, is not high enough to provide an incentive for farmers to prevent paraTB infection. Farmers do not find vaccination to

be economical and so most animals displaying symptoms of paraTB are culled (Bastida and Juste, 2011). More recently DNA vaccines have been used in mice and sheep but not in cattle. These consist of mammalian expression vectors containing MAP genes; however further investigation is required to determine whether these vaccine techniques will be successful (Velaz-Faircloth *et al.*, 1999 Sechi, 2006).

The development of anti-mycobacterium drugs has proven to be ineffective. MAP was shown to be highly resistant *in vivo* to standard anti-tuberculosis drugs; even though these drugs have been shown to reduce infection in *M. tuberculosis* infected humans (Ayele *et al.*, 2001).

The most cost effective practice to control the spread of paraTB is the careful maintenance of farmed environments. Preventing the introduction of infection into a new herd through methods such as:

- Ensuring new stock are tested for the presence of MAP.
- Managing risks when transporting animals at exhibitions through separation and hygiene.
- Managing water flow from neighbouring land.
- Managing disposal of manure.
- Culling of any infected animals immediately.
- Decontamination of areas where animals are kept in close confinement on a regular basis i.e. lambing/calving barns.

These can all be effective strategies in reducing MAP prevalence(Weber, 2006). However it must be noted when implementing these practices that MAP has the ability to persist in soil and aquatic environments for upwards of a year. The bacterium is relatively susceptible to high pH in soil, longevity is reduced in areas of persistent urine and faeces exposure but this does not stop the spread of infection (Lovell *et al.*, 1944). These factors attribute to the challenges faced in controlling the spread of infection once an animal/environment has become infected with MAP.

1.6 Transmission of MAP

The factors that influence the transmission of MAP between susceptible hosts continue to be defined. Control of the disease has focussed on removing those in the relatively advanced stages of disease which contribute to the spread of infection to younger animals by shedding MAP in their faeces. Pre-natal infection has been described whereby calves have acquired infection *in utero*; however the widely known infection of newborn animals occurs by oral ingestion of the pathogen (Collins, 1994). Contamination of teats and the presence of MAP in the colostrum and milk result in the ingestion of large doses of bacteria by suckling neonates (Behr and Collins, 2010). Contamination of pasture, water and feed also contribute to the spread of infection in animals born from paraTB free dams (Kaevska *et al.*, 2014). As the main site of infection is the lower region of the small intestine and corresponding lymph nodes, paraTB is largely excreted via the faeces of infected animals. Due to faecal shedding, the environment of infected herd/flock is often contaminated (Bannantine and Bermudez, 2013). High shedders are, of course, responsible for high rates of transmission. Whittington et al found that excretion of MAP by multibacillary infected Merino sheep occurred daily, showing that environmental contamination can be continuous on farms with endemic ovine paraTB (Whittington *et al.*, 2005). It has been proposed that MAP can be transmitted during breeding as the bacteria can be cultured from the semen of infected rams and bulls (Larsen *et al.*, 1981). Viable mycobacteria has previously been isolated from the reproductive organs of infected animals; however intrauterine transmission has not been regarded as the most important route of transmission when considering methods of control (Ayele *et al.*, 2001).

Similar to bovine tuberculosis, transmission of MAP between livestock and wildlife species that share the same habitat has been reported (Nugent *et al.*, 2011). Infected wildlife have been detected in areas of significant prevalence of infected domestic ruminants (Greig *et al.*, 1999). This adds to the difficulty of MAP eradication from domesticated livestock grazing areas as infected wildlife, such as rabbits and deer, tend to share grazing ground.

Following oral ingestion, MAP localises in the mucosa of the small intestine and associated lymph nodes. As the disease advances infection disseminates in organs distant from the primary, gastrointestinal site of infection via lymphatic vessels and blood.

1.7 Intestinal translocation

Once animals are infected with MAP, initially the mucosa of the small intestine and associated lymph nodes are targeted. The mechanisms by which MAP is able to bind to, and translocate across, the intestinal wall is a matter of ongoing research and debate (Coussens *et al.*, 2010). It is thought that the binding of fibronectin (FB) plays a key role during mycobacteria invasion. MAP expresses a homologue to the fibronectin attachment protein which enables the bacteria to bind to FB(Bannantine and Bermudez, 2013). The bound FB is then able to bind to integrins on the surface of M-cells and translocate through the Peyer's patches into the gut (Secott *et al.*, 2004). M-cells are a proposed target cell of MAP as they express integrins on their luminal faces unlike other intestinal epithelial cells; β1 integrins have been identified as the receptor for FB-opsonized mycobacteria *in vitro* and these integrins have been found in high density on the surface of M-cells (Secott *et al.*, 2004.Bermudez *et al.*, 2010). Other studies have shown that MAP has the ability to attach to and cross the intestinal epithelium independently of Peyers patches, suggesting a role for enterocytes in MAP invasion (Bermudez et al, 2010). Enterocytes do not express the integrins required for FB binding and so other, yet unknown, receptors may be involved. Once MAP has crossed the intestinal epithelium, the bacterium is phagocytosed by macrophages; furthermore opsonisation of MAP by antibody and complement C3b can occur at this stage, increasing its uptake by macrophages (Tessema *et al.*, 2001). The mechanism of host invasion likely contributes to the ability of MAP to propagate in a variety of host species; given that this bacterium can only persist, and not grow, outside the host, invasion and establishment within a variety of host species is crucial to its survival.

1.8 Host immune response to MAP infection

As previously mentioned, there are fewer studies investigating paraTB infection in sheep than cattle; however the paucibacillary and multibacillary pathological forms of disease in sheep are more easily defined (Clarke, 1997). These forms are more difficult to differentiate in other species and are largely time dependent whereas in sheep definition of the two endpoint pathological forms can be achieved based on examination of gross pathology and characteristics of host immune response. In many cases this makes research on paraTB infection less confusing in sheep.

1.8.1 Macrophage response

Typical bacterial infections involve the phagocytosis of the bacteria following recognition by cells of the innate immune system, including macrophages, and dendritic cells. Once internalised in the cell, lysosomes fuse with cytoplasmic phagosomes to form phagolysosomes, where the components of the bacteria are digested by hydrolytic enzymes. Bacterial peptides are then presented on the cell surface by MHC class II molecules and subsequently initiate the activation of an appropriate adaptive immune response. During paraTB infection the bacteria multiply within macrophages in the terminal section of the small intestine when phagocytosed; macrophages in turn proliferate leading to more cells harbouring the bacterium (Secott *et al.*, 2004, Woo, 2008). Macrophages play a conflicting role in this disease whereby they are the primary site of MAP replication but are also important in the elimination of the bacteria by phagocytosis and stimulation of adaptive immune responses. The pathogenesis of MAP at a cellular level, within macrophages, is poorly understood. Macrophages possess several receptors that are involved in the uptake of mycobacteria. Ferwerda *et al.* investigated the recognition of MAP by PRRs and found that macrophages from TLR2 and TLR4 knockout mice produced less immune stimulating cytokines compared with control mice when infected with MAP. It was also shown that NOD2 knock-out human embryo kidney (HEK) cells but not NOD1 knock-out HEK cells responded to stimulation with MAP; these results indicate that TLR2, TLR4 and NOD2 molecules are involved in

pattern recognition of MAP by the innate immune system (Ferwerda *et al.*, 2007). Other studies have shown CD11/CD18, CD14, complement receptors and mannose receptors as key initiators of phagocytosis in mycobacteria infections, including MAP infection (Souza *et al.*, 2007.Woo and Czuprynski, 2008). During most bacterial infections, macrophages ingest the bacterial organisms via receptor mediated phagocytosis, internalisation leads to the formation of phagolysosomes which degrade the foreign bacteria by producing reactive oxygen and lysosomal enzymes (Hostetter *et al.*, 2003). However, mycobacterial/MAP infection differs to that of other bacterial infections in that once MAP is phagocytosed, persistent infection is often established within the host by altering the functioning and gene expression of macrophages. These alterations lead to failure of some phagosomes containing MAP to progress through the endocytotic pathway by blocking vesicle maturation and, in doing so, avoiding phagosome-lysosome fusion (Tooker and Coussens, 2004). Souza et al investigated the effects of mannosylated-lipoarabinomannan (Man-LAM), a mannose-capped lipoglycan cell wall component of MAP which has been identified as one of the ligands that modulates macrophage function, on survival of the bacteria when cultured with bovine macrophages. This study also looked at the effects of Man-LAM on IL10, IL12p40 and TNFα expression by bovine macrophages during initiation of an adaptive immune response and found that Man-LAM induces rapid expression of IL10 as well as a transient expression of TNFα and IL12p40. It was also shown that MAP-derived Man-LAM had the capacity to inhibit maturation and acidification of phagosomes and was able to attenuate bacteria killing in macrophages when co-cultured with *Mycobacterium avium* subsp. *avium* (MAA) reducing bacterial killing to approx. 8%; whereas in previous studies, incubation of macrophages with MAA alone resulted in killing of 40% to 50% of bacteria (Souza *et al.*, 2013). This is similarly seen in *M. tuberculosis* infection whereby many macrophages containing mycobacteria display a lack of maturation and phagolysosome formation (Clemens, 1996).

Cell proliferation gives rise to extensive numbers of internalised MAP, leading to infiltration of the intestinal submucosa by macrophages, APCs and immune cells; this results in inflamed intestinal tissue and eventual reduction in gut absorption and protective function. As with other pathogenic infections, macrophages that have

successfully phagocytosed the organism migrate to the local lymph nodes where they present antigen to resident T cells and B cells; however research has found that macrophages persistently infected with MAP are less efficient at T cell stimulation (Coussens, 2001). Thus suggesting that the stimulation of T cells may be restricted to macrophages that are successful in degrading/processing MAP or perhaps the other innate cell types, such as dendritic cells (DC) , play a crucial role in antigen presentation to T cells during MAP infection (Coussens, 2004).

1.8.2 B-cell and antibody response

Antibodies to MAP produced by the host have proven to be ineffective at controlling the bacterial infection in many species. Antibodies to the MAP cell wall constituent lipoarabinomannan (LAM) have been identified in cattle and appear at approx. 130 days post infection where infection has already established and these antibodies were not effective at creating protective response (Koo *et al.*, 2004). Antibody can be detected in sheep from around 90 days post infection (Kurade *et al.*, 2004); however Begara-McGorum *et al*. found that in lambs with experimentally infected ParaTB, B cell numbers in the mesenteric lymph node were significantly reduced; this was also seen in jejunal and ileal Peyer's patches, resulting in high T:B cell ratios for infected lambs compared to uninfected controls, 8 weeks after infection (Begara-McGorum *et al.*, 1998).

One of the challenges faced when measuring antibody response is that of detection method. A study in cattle stated that the commercial ELISA test could not detect antibody until the animal is over 3 years of age (Ayele *et al.*, 2001); whereas another study stated that the commercial ELISA was unable to detect antibody responses in cattle that had already shown to be seropositive through different methods (Waters *et al.*, 2003).

During late stages of infection, antibody levels correlate with severity of disease and bacterial load (Burrells *et al.*, 1998a), this is similar to what is known in human tuberculosis infection (Neill *et al.*, 2005). Both cattle and deer show a strong correlation between antibody levels and disease severity (Souza *et al.*, 2006, {Mackintosh, 2007). Although sheep produce antibody, there is no evidence to

confirm this is protective to the animal. High antibody levels in sheep have mostly been correlated with the multibacillary form of pathology; it is thought that paucibacillary animals produce antibody at lower levels; however it is difficult to distinguish between the two forms using antibody data (Kurade *et al.*, 2004, Burrells, 1998).

1.8.3 CD4+ T cell response and polarisation

Macrophages, DC's and other APCs provide a first line of host defence against intracellular pathogens including mycobacteria. The differentiation of the pathological forms of sheep paraTB, multibacillary and paucibacillary, have been associated with differential T cell activation (Smeed *et al.*, 2007). Macrophages are key players in the initiation of immune response to MAP; these cells are infected and subsequently activate T cell responses; influencing development of differential T cell subsets. Once an infection has taken place the host response is primarily regulated by T cells. Cytokines secreted during innate responses can create environments that influence T helper (Th) cell differentiation into specific polarised subsets (Luckheeram *et al.*, 2012). These subsets determine the immune response made to a particular type of antigen. Although numerous subsets continue to be identified; four main Th subsets have been characterised and studied in depth; Th1, Th2, Th17 and T-regulatory cells (Treg), each expressing a unique set of transcription factors and characterised by distinct cytokine profiles, shown in figure1.2 (Zhu and Paul, 2010). Transcription factors function to regulate gene expression in specific cell subsets and determine the cytokine milieu produced by cells, as well as receptor molecules expressed on that cell. The differentiation of CD4+ T cells into Th cell subsets is coordinated by cytokines and transcription factors and determines the outcome of immune response from each subset to different types of pathogens (Zhu and Paul, 2010). Furthermore, cytokine receptors are crucial in the mediation of cytokine driven Th cell activity. These receptors are expressed by activated T cells in response to a specific cytokine environment (Drachman and Kaushansky, 1995). Protection against mycobacterial disease is dependent on T cell immunity (Orme *et al.*, 1993) and this is the case in paraTB infection (Coussens, 2004). Insufficient or inappropriate control and/or regulation of Th cell response results in immune

mediated pathology, including that seen in paraTB. In both sheep and cattle, γδ T cells are increased early at the site of infection (Begara-McGorum *et al.*, 1998 Beard, 2000) with CD4+ and CD8+ T cells increased later in infection. In sheep the T cell response is polarised and results in the observed differences in pathology (Smeed *et al.*, 2007). Once sheep are infected with MAP, approximately 30% of the symptomatic animals are affected by the paucibacillary form of paraTB; a tuberculoid state characterised by low levels of bacteria with γδ and CD4+ T cell infiltration of the gut. In this form of the disease MAP cannot often be seen in tissues by Ziehl-Neelsen (ZN) staining, used for staining of acid-fast mycobacteria, nor can be detected by bacterial culture. Initially, studies described the development of a predominantly Th1/Th17 cell response, characterised by release of pro-inflammatory cytokines including IFNγ, IL1α, IL12, IL17A, IL17F and IL2, in paucibacillary infected animals (Clarke, 1997, Begara-McGorum, 1998, Burrells, 1998 . Robinson *et al.*, 2011). However; some studies have described a mixed Th1/Th17 inflammatory and Th2 response in paucibacillary infection and a predominant Th2 antibody response in multibacillary pathology (Coussens, 2001.Coussens, 2004).

Interleukin-12 (IL12) secreted by cells of the innate system following pathogen recognition induce the development of a Th1 cell population. IL12 is secreted by activated APCs; this cytokine then, in turn, induces IFNγ production by Th1 cells and natural killer cells (NK) (Macatonia *et al.*, 1995). Cells involved in the Th1 phenotype predominantly provide the host with protective immunity to intracellular bacterial and viral infections. Type-1 Th cells are regulated by the transcription factors signal transducer and activator of transcription (STAT) 4 and T-box transcription factor (T-bet) (Zhu and Paul, 2010). T-bet is activated when IL12 is produced by dendritic cells and interacts with the IL12 receptor, also in response to IFNγ production by NK cells. Cytokines produced by Th1 cells activate phagocytic cells and favour the production of opsonising and complement-fixing antibodies, IgG1 and IgG3. This response provides optimum protection against intracellular bacterial and viral infections (Romagnani, 1999). Inadequate regulation of this response leads to chronic inflammation and granuloma formation (Crane and Forrester, 2005). Host immune responses to mycobacterial infection depend on the integrity of the type-1 cytokine pathway. It is well established that IFNγ is required

to control mycobacteria infection in a number of species (Cooper *et al.,* 1993.Flynn *et al.,* 1993). In human tuberculosis infection CD4+ T cells are critical in the regulation of infection, with CD8+ cells contributing to this control by producing IFNγ (North and Jung, 2004). This response aims to control infection by forming granulomatous lesions that control bacterial growth and can, in some cases of tuberculosis, clear the bacteria (Saunders and Cooper, 2000, Barnes, 1993). NK cells produce IFNγ in both human and bovine tuberculosis, resulting in increased mycobacterial killing by activated macrophages (Denis *et al.,* 2007). Furthermore, mice deficient in IL12 show greater susceptibility to mycobacterium-induced diseases, more so in the absence of the IL12p40 subunit than of the IL12p35 subunit (Cooper *et al.,* 2002). This may be related to the involvement of IL12p40 as a subunit of the heterodimeric cytokine IL23, this is described further in Chapter 3. Up-regulation of these cytokines and a cell mediated response is appropriate in the control of mycobacteria replication; however failure to correctly regulate the production of these cytokines leads to pathology caused by chronic granulomatous inflammation (Saunders and Cooper, 2000). This is typical of the immune response and pathology often described in the paucibacillary form of MAP infection (Clarke, 1997. Smeed *et al.,* 2010).

In contrast to Th1, Th2 differentiation is activated in the presence of IL4, IL13, IL25 and IL33. IL4 have been shown to be produced by naive T cells following activation by DCs (Noben-Trauth *et al.,* 2002). Another theory is the involvement of innate lymphoid cells (ILCs); a family of haematopoietic effector that have morphological characteristics of lymphoid cells but lack rearranged antigen receptors; a defining feature of adaptive immune cells. These cells share common functional and phenotypic features and are classified into three broad groups (summarised in table 1.1). Furthermore, it is increasingly recognized that ILCs share conserved transcriptional mechanisms with CD4+ T cells as well as mirroring their Th cell counterparts in terms of cytokine production; thus ILC1s express T-bet, ILC2s express GATA3 and ILC3s express RORγt transcription factors that induce gene expression within these cells. Since ILC2s secrete the Th2 associated cytokines, IL5 and IL13, in response to stimulation by I-25, IL33 and thymic stromal lymphopoietin (TSLP) produced by mast cells, epithelial cells and macrophages (Lund *et al.,*

2013.Wolterink *et al.*, 2012.Barlow *et al.*, 2013), these cells may also produce IL4; however this is yet to be confirmed. IL4 induced STAT6 up-regulates the master transcription factor of Th2 cells GATA3 (GATA-binding protein 3), resulting in the initiation of gene transcription and secretion of IL4, IL5, IL13 and IL25 from Th2 cells. These cytokines stimulate IgA and IgE production by B-cells as well as promoting eosinophil differentiation and activation (Van der Pouw Kraan *et al.*, 1998). Type-2 T cell immunity is typically seen during host response to helminthic infection and allergens; insufficient regulation of the Th2 phenotype has been identified in allergic reactions (Deo *et al.*, 2010). Secretion of IL25 plays a key role in the maintenance of the Th2 phenotype (Wang *et al.*, 2007). This cytokine is secreted by epithelial cells at mucosal surfaces in response to pathogen stimulation and aids the differentiation of naive T cells to Th2 cells. Furthermore this cytokine provides a feedback loop that up-regulates the production of IL4; continuing to stimulate the differentiation of naive T cells into Th2 cells (Wang *et al.*, 2007). Expression of IL4 has been shown to inhibit Th1cell development and so directs cells towards type-2 immunity (Wurtz *et al.*, 2004). During multibacillary (lepromatous) paraTB; high levels of acid fast bacilli in the gut tissues has been noted, with macrophage and B cell infiltration at the site of infection (Clarke, 1997). Multibacillary diseased animals can present with high levels of IgG plasma antibody and very low levels of IFNγ production; however this is not seen in all cases and antibody levels are often unreliable (Burrells *et al.*, 1998b, Gillan, 2010.Collins *et al.*, 2005). Disease occurs as a results of the formation of extensive and severe lesions that occur as a result of autoimmune reactions (Gillan *et al.*, 2010.Perez *et al.*, 1999). These lesions contain few CD4+ T cells and an increased number CD8+ T cells which may attribute to the low levels of IFNγ found in these animals (Little *et al.*, 1996, Burrells, 1999). Multibacillary animals are typically associated with Th2 polarised immunity, showing up-regulation of Th2 associated responses and a lack of cytokines and IFNγ production found in Th1/17 cell responses. This results in a detrimental lack of macrophage activation and, although B cells are present and an antibody response is elicited it is unable to control the bacterial burden, thus the observed multibacillary pathology and eventual dissemination of infection (Navarro *et al.*, 1998).

Table 1.1 Innate lymphoid cell (ILC) families.

ILC NAME	FUNCTION	CYTOKINES PRODUCED	STIMULATING CYTOKINES
NK cells /ILC1/cytotoxic ILCs	Innate immunity against viral infection, tumour immunosurveillance	IFNγ	IL-18, IL12 and IL-15
RORγt+ ILCs/ILC3	Innate immunity against bacteria, lymphoid tissue formation and repair.	IL17A and IL-22	IL-1B and IL23
Type 2 ILCs/ILC2	Innate immunity against extracellular parasites	IL-5 and IL-13	IL-25 and IL-33

In the presence of IL6, TGFβ and IL23 produced by macrophages and DCs, a Th17 cell population is produced. Th17 cells excrete IL17A, IL17F and IL22 (Korn *et al.*, 2009). These cytokines are responsible for inducing inflammation, primarily by the stimulation of non-professional immune cells, including fibroblasts and epithelial cells, and macrophages to produce granulocyte colony-stimulating factor (CSF) and chemokines that subsequently recruit polymorphonuclear leukocytes, including neutrophils (Korn *et al.*, 2009). Macrophage recruitment leads to an increase in IL6 secretion, further stimulating the differentiation of naïve T cells to Th17 cells (Kimura and Kishimoto, 2010). This cell subset has been identified in various auto-inflammatory diseases and was originally thought to be part of the Th1 immune response (Damsker *et al.*, 2010). However, both IL4 and IFNγ have been shown to inhibit the development of the Th17 cell subset (Zhu *et al.*, 2010).

As noted above, regulation occurs throughout T-helper cell immune responses with each cell subset inhibiting/influencing development of other Th cell subsets. Whilst cytokines and transcription factors produced by each subset can redirect Th

responses to a less polarised phenotype, all of these Th cell groups are also regulated by further subset of T cells known as adaptive T-regulatory (T-reg) cells (Zhu *et al.*, 2010). These cells exert suppression of other cell subsets by the production of suppressive cytokines TGFβ and IL10. (Romagnani, 2006). Human and mouse models of tuberculosis infection have shown an increased number of T-reg cells accumulating at the site of infection, preventing migration of immune cells to the site, ultimately reducing inflammation (Josefowicz *et al.*, 2012). Furthermore, T-regs in mice and humans can inhibit the production of IFNγ in response to tuberculosis infection (Chen *et al.*, 2007, Kursar, 2007) and have therefore been suggested as key influencers in tuberculosis pathogenesis. MAP pathogenesis in cattle differs to sheep in that a gradual progression from paucibacillary to multibacillary phenotype is observed - the trigger of transition from a Th1 to Th2 response is currently unknown (Coussens, 2004); however it has been suggested that this shift in immune response in bovine paraTB may be due to a population of MAP-reactive T-reg cells developing during subclinical, limiting Th1 responses to MAP antigens later in infection (Roussey *et al.*, 2014).

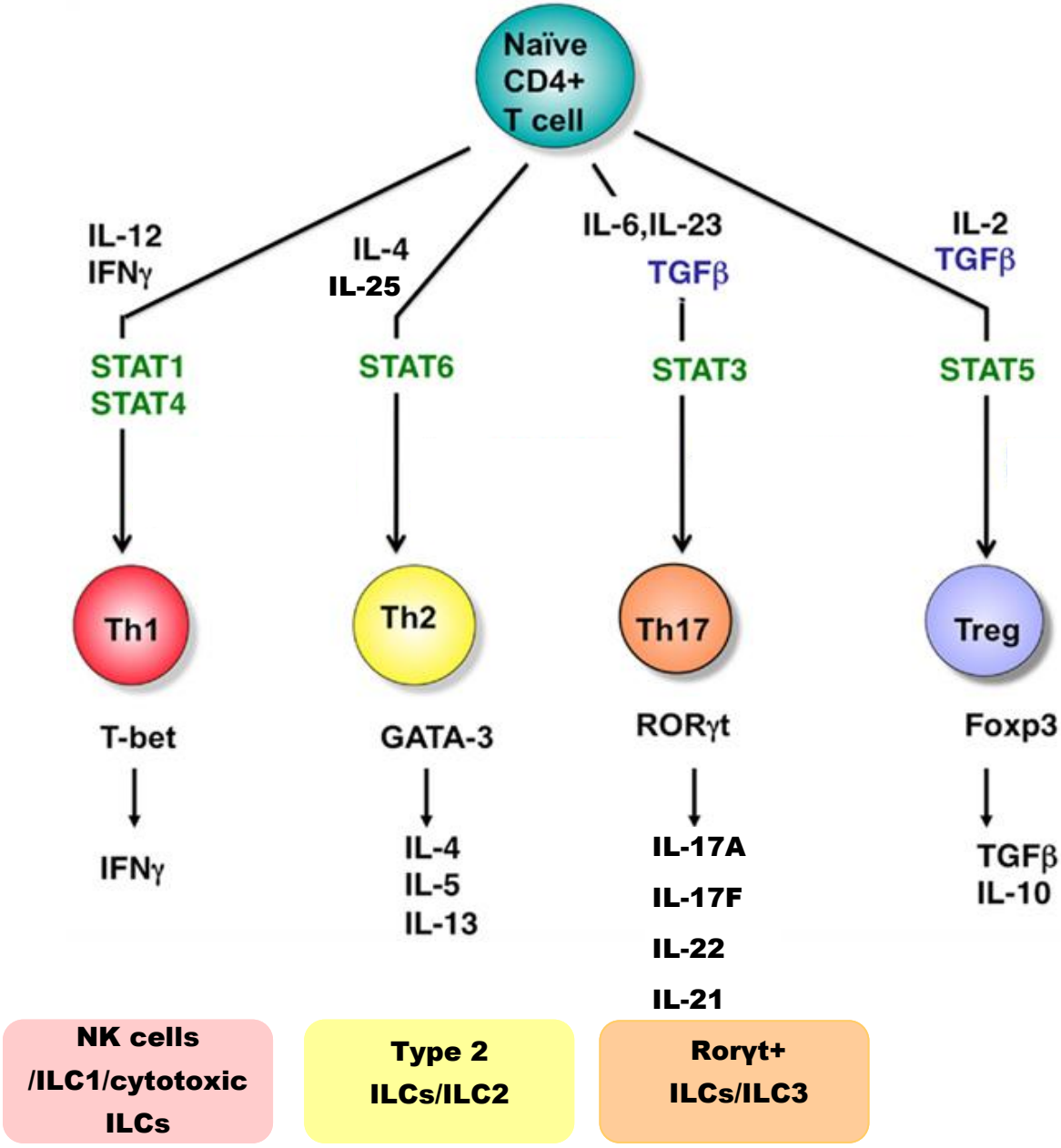


Figure1.2 T-helper cell differentiation pathways and associated ILC subsets. Adapted from (Yoshimura *et al.*, 2012).

1.9 Genetic and genomic studies of MAP infection

1.9.1 Host gene expression studies

Profiling of host gene expression during MAP infection has been the focus of numerous studies and such research is ongoing (Coussens, 2004.Coussens *et al.*, 2004.Lee *et al.*, 2001.Smeed *et al.*, 2010.Coussens *et al.*, 2012). It is thought that dysregulation of host gene expression contributes to the ineffective regulation of MAP-induced immune responses and subsequent pathology (Smeed *et al.*, 2008). Investigation of gene expression at different stages of infection has provided insight into the biological pathways that are involved in immune responses to paraTB (Tanaka *et al.*, 2005). Mostly, studies investigating the genetic basis of paraTB have been carried out in cattle with fewer studies looking at gene expression profiles in sheep. This may be due to the assumption that sheep profiles would be similar to that of cattle; Smeed *et al*. described the differential cytokine gene expression profiles in pathological forms of sheep paraTB. Here it was reported that paucibacillary and multibacillary infection is associated with the differential expression of IFNγ and IL10 respectively. Increased levels of pro-inflammatory cytokines including IL1β, IL18 and TNFα in both diseased groups is indicative of persistent inflammatory lesions (Smeed *et al.*, 2008). In studies of cytokine gene expression in response to MAP infection in cattle, Lee *et al*. and Coussens *et al* have described comparable expression patterns of these cytokines in MAP infected calves compared with controls, with up-regulation of IFNγ and TNFα in infected animals (Coussens *et al.*, 2004.Lee *et al.*, 2001). A further study in sheep found that TNFα, IL1β and IL6 was significantly increased in infected ileum tissues when compared to uninfected controls (Alzuherri *et al.*, 1997). Sweeney *et al* also found that progression of paraTB to clinical stages in cattle was associated with reduced expression of IFNγ at the site of infection and suggested that if the immune system could be manipulated to increase IFNγ in the later stages of infection, resistance in cattle might be enhanced (Sweeney *et al.*, 1998).

1.9.2 Host genomic studies

With the advance of next generation sequencing (NGS) technologies, microarray analysis, RNA-seq and genome wide association studies (GWAS) have also been utilised in the identification of genes which may play a role in the severity and/or progression of paraTB. Coussens *et al*. investigated the differences in gene expression profiles of peripheral blood mononuclear cells (PBMC) from subclinical and clinical paraTB positive cattle following *in vitro* stimulation with MAP. The expression profiles were compared using a bovine-specific cDNA microarray platform with 721 expressed sequence tags (EST) representing known genes. It was shown that stimulation of PBMCs from clinically infected cows with MAP decreased expression of 83 genes, including genes encoding fibroblast growth factor and Lyn B protein kinase, and only 8 genes were recorded to have a modest up-regulation. In contrast, stimulated PBMCs from subclinically infected cattle resulted in an up-regulation of 71 genes and only 16 genes were down-regulated, concluding that the gene expression profiles from clinically and subclinically infected cattle are notably different following exposure to MAP (Coussens *et al.,* 2002). Microarray analysis of ileal gene expression in sheep with clinical paraTB identified significant differential expression of 36 genes; variation in cellular composition of multibacillary and paucibacillary infected sheep were reflected in difference of gene expression between the two groups. The chemokine CXCL10 was significantly raised in paucibacillary animals; this gene is associated with chemo-attraction of type 1 T cells and would explain T cell infiltration common in paucibacillary infection. IGF2R plays a key role in the activation of TGFβ, both genes were found to be up-regulated in multibacillary animals (Smeed *et al.,* 2010).

Discussions from such studies also highlight the limitations of NGS when trying to identify genetic variations that are of consequence to phenotype. Recently, *Zare et al.* investigated susceptibility to paraTB infection in Jersey cattle, the second most common breed of dairy cattle in America, using the bovine SNP50 BeadChip platform. Sample groups consisted of 'cases' that tested positive for MAP infection by ELISA and faecal culture, and controls that were free of MAP in results from both tests, from an initial pool of ~5000 cattle. Comparative analyses of results from the

SNP50 platform found that, although many SNPs were identified in both groups, the significance of polymorphisms were relatively low. This was explained as the complexity of genetic architecture associated with susceptibility to paraTB. NGS studies have the power to identify a vast array of genetic information and identify numerous variants; however the ability to identify which of these variants has an influential effect on disease phenotype and those which are of biological relevance in a whole animal system is limited. In order to overcome this limitation, larger scale studies are required to identify the most frequent and significant variants/SNPs present between infected and non-infected groups (Zare *et al.*, 2014). In the case of paraTB, the number of samples required from naturally infected animals for such a study is often difficult to obtain; therefore smaller scale studies are used as an initial identifier of potentially relevant genes/variants.

1.10 MAP and human inflammatory diseases

The cause of the differential pathology of paraTB is currently unknown; however as previously noted, several studies have suggested that the host genetics play a key role in the clinical outcome of mycobacterial infection (Fortin *et al.*, 2007. Coussens, 2002). ParaTB pathology is associated with differential T cell activation in sheep, with genes that control this activation playing a key role in the pathological outcome of infection. Similarities in gene expression patterns, immune responses and pathology is noted when comparing human diseases such as TB, leprosy and inflammatory bowel disease to paraTB (Neurath, 2014, Sartor, 2005. Ashenafi, 2014). These similarities are not always surprising given the presence of mycobacterial causative agents in TB, leprosy and paraTB. These human diseases are also associated with Th cell differentiation and it is thought that genetic and transcriptional variants of the genes that influence T cell differentiation in human diseases may also play a role in the pathological outcomes of paraTB. This section describes TB, leprosy and Chrone's disease in humans and highlights the genes that impact differential T cell activation in these diseases.

1.10.1 Tuberculosis

Tuberculosis is a major health issue throughout the world causing more deaths than that of any other single infectious agent disease. A report was released by the World Health Organisation stating the TB situation in 2013. Approximately 9 million cases of ill health caused by TB were reported with 1.5 million of those resulting in death from the disease. Similar patterns are seen in paraTB whereby many animals are infected but do not become diseased. Although TB incidence is falling globally, reported at a rate of around 1.5% per year between 2000 and 2013, this infectious agent is still of concern (www.who.int). *Mycobacterium tuberculosis* is the causative agent of human tuberculosis; an aerobic, intracellular pathogen which infects lung tissue rich in oxygen. From the site of infection the bacterium can migrate throughout the body, primarily disseminating across the lungs and regional lymph nodes, via the blood or lymphatics. Similar to MAP, many individuals infected with TB do not show clinical signs and remain asymptomatic (Wang, 1999). Similarities are also noted in the immune responses to TB and paraTB infection; *M. tuberculosis*is recognised by antigen receptors on the surface of macrophages and is phagocytosed, this contact between host and pathogen triggers the host' adaptive immune response and triggers differential T cell polarization (Ernst, 1998). In active (pathological) TB, a Th1 cell mediated immune response develops with an influx of lymphocytes and activated macrophages to the site of infection, leading to the formation of granulomatous lesions; similar to that seen in paraTB pathology. Although the cell mediated response is necessary for the control of *M. tuberculosis* infection; it is also responsible for much of the pathology associated with TB (O'Garra *et al.*, 2013). Host antibody mediated immune responses provide poor protection due to the intracellular nature of *M. tuberculosis* and its ability to evade recognition whilst residing within macrophages. Although highly infectious, only 10% of the infected human population progress to active tuberculosis (Moller and Hoal, 2010)**.** The defining factors which lead to pathology in some infected humans whilst others remain asymptomatic are unknown (O'Garra *et al.*, 2013); as seen with paraTB. Genetic predisposition is thought to play a key role in the determination of symptomatic and asymptomatic, as well as tuberculoid/lepromatous, outcomes

during TB infection in humans (Abel and Casanova, 2000). Individuals susceptible to TB infection are often diagnosed with a rare human syndrome known as Mendelian susceptibility to mycobacterial disease (MSMD) (Qu *et al.*, 2011). Most commonly affected by TB and salmonella infections, these patients have an increase susceptibility to even non-pathogenic mycobacteria. Genetic mutations in the genes of the IL12/IL23/IFNγ axis including *IL12RB1*, *IFNGR1* and *IFNGR2* have been identified in MSMD patients (Schurr, 2011). A functional 874A>T SNP found in *IFNG* gene was associated with TB infection, this variant reduces IFNγ induction due to inferior binding afforded to NFκB (Filipe-Santos *et al.*, 2006). It is assumed that many contributing genetic factors are associated with TB/MSMD susceptibility, many of which are associated with IFNγ inducing pathways (Bellamy, 1998, Al-Muhsen, 2008); however the specific functional involvement of many of the variants identified in these genes is yet unclear. IL12RB1 is crucial in the induction and stabilization of efficient Th1 responses. This cytokine receptor is required for IL12/IL23 signalling and subsequent Th1/Th17 activities; including induction of T-bet, a transcription factor that functions to stabilise Th1 cell activity by initiating transcription of the *IFN*γ gene (described in more detail in chapter 6). Robertson *et al*. have described a splice variant of the *IL12RB1* gene that occurs in response to TB infection. The novel isoform of human *IL12RB1* lacks a transmembrane domain and was found to increase DC migration and activation of TB specific T cells (Robinson *et al.*, 2010). Another study found that genetic variants of IL12RB1 confer genetic susceptibility to TB and contribute to severity of disease in a Japanese population (Kusuhara *et al.*, 2007).

1.10.2 Leprosy

Leprosy, a chronic inflammatory condition caused by infection with *Mycobacterium leprae* (*M. leprae*), shows a wide spectrum of clinical features. Differential pathology is associated with differential T-cell activation, similar to that seen in paraTB, with developing symptoms of granulomata formation occurring in the respiratory tract, skin and nerves. At one end of the spectrum is the tuberculoid form (T-lep) and at the other end is lepromatous leprosy (L-lep). Similarly to paucibacillary form of paraTB, T-lep is characterised by predominance of CD4+ T cells, high levels of IFNγ and Th1 cell mediated immune responses stimulated by *M. leprae*; patients affected with the T-lep form of disease show higher resistance to infection with less severe clinical manifestations and is often described as 'self-curing' (Moubasher 1998. Modlin, 1994). Patients infected with L-lep form of disease are comparable to that of multibacillary paraTB infected animals; characterised by a predominant CD8+ T cell infiltration and Th2 cytokine production. High numbers of infected macrophages and low Th1/cell mediated immune responses mean that those affected by L-lep form of disease are unable to contain the infection and disease progression occurs (Modlin, 1994). Studies investigating the factors associated with susceptibility and resistance of host species to *M. leprae* have identified candidate genes, many of which contain polymorphisms, conferring susceptibility/protection to leprosy. Most of the genes identified are regulators of immunological events involved in *M. leprae*. A toll-like receptor 2 (TLR2) mutation resulting in a Arg>Trp substitute in a conserved region of the protein was found to be present in L-lep patients but not T-lep patients or control subjects, indicating a role for this mutation in susceptibility to L-lep (Kang and Chae, 2001). Smeed *et al* found that *TLR2* expression was greatly up-regulated in multibacillary lesions from paraTB infected sheep when compared with paucibacillary infected and uninfected control animals; suggesting that TLR2 induced signalling pathways are important in lepromatous pathogenesis (Smeed *et al.*, 2010). Many of the genes associated with the differential pathologies in leprosy are regulators of the Th1/Th2 cell paradigm and some of these genes confer susceptibility to disease (Fitness *et al.*, 2002).

1.10.3 Crohn's Disease

The clinical outcomes of MAP infection have been associated with differential T cell activation and is similar to that seen in human intestinal pathologies including Crohn's disease (CD) and ulcerative colitis (UC) (Kobayashi *et al.*, 2008.Grant, 2005).

CD and UC are the principle types of inflammatory bowel disease (IBD) in humans and are of unknown aetiology. CD is driven by the production of IL12 and IFNγ in a predominant Th1 associated response, whereas UC is driven by the production of IL13 in a predominantly Th2 associated immune response; immune dysregulation and uncontrolled production of pro-inflammatory cytokines leads to inflammation of the gastrointestinal tract in both CD and UC (Lennard-Jones, 1989.Bouma and Strober, 2003). IL25 is a potent activator of Th2 responses in the intestine; blocking IL25 signalling protects against gut inflammation in a mouse model of colitis by reducing IL13 production (Camelo *et al.*, 2012). It has also been shown that IL23 drives T cell mediated intestinal inflammation and promotes inflammatory Th17 activity (Yen *et al.*, 2006).

The pathology and immunology of the IBD disease forms are comparable to that of paraTB (Davis and Madsen-Bouterse, 2012a). Both paraTB and IBD result in the breakdown of the intestinal barrier, whereby bacterial invasion of gut tissues and chronic inflammation contribute to the persistent continuation of disease (Chacon *et al.*, 2004). Although CD is considered to be an autoimmune disease; increasing evidence suggests that there may be an infectious stimulant (Sartor, 2006). Controversially, many studies have stated that infection with MAP may lead to immune dysregulation and inflammation in Crohn's disease patients. Isolates of MAP found in animals are linked to active paraTB; whereas the isolation of MAP from human specimens is yet to be consistently confirmed as the potential cause of CD (Sartor, 2005).

It is thought that MAP transmits from infected animals to humans via the consumption of milk, dairy products and possibly contaminated water sources as described in figure 1.3. Due to the resilience of MAP in the environment; the

potential of the bacterium to infiltrate water supplies is highly possible (Pierce, 2009). Members of the *M. avium* complex (MAC) have been shown to be resistant to chlorine and other disinfectants (Taylor *et al.*, 2000). Normal water treatment processes involve filtration and disinfection, usually by chlorination; however these processes have little or no effect on mycobacteria (Whittington *et al.*, 2005.Hilborn *et al.*, 2006). MAP can be detected by *IS900* PCR in raw and pasteurised milk (Ellingson *et al.*, 2005.Paolicchi *et al.*, 2012), and commercial cheese samples (Clark *et al.*, 2006). An extensive study carried out from 1990-1994 across England and Wales found that 7% of bottles and cartons of retail pasteurized cow's milk tested positive for *IS900* by PCR (Millar *et al.*, 1996).

Schwartz et al. investigated the presence of MAP in tissue specimens from CD patients and healthy control subjects using short term culture media; 37% of CD samples tested positive whilst only 5.6% of control specimens contained viable MAP (Schwartz *et al.*, 2000). In a similar study, Bull et al. detected MAP in 42% of CD patients and 9% of control subjects using a mycobacterial growth indicator tube; this was confirmed as MAP by *IS900* PCR (Bull *et al.*, 2003). One study used PCR and culture methods to identify MAP in humans with CD, ulcerative colitis (UC) and healthy individuals. Results found that MAP was present in 46% of individuals with CD, 45% with UC and 20% of healthy individuals (Naser *et al.*, 2004). Furthermore, Nazareth *et al.* found the prevalence of MAP detected by *IS900* PCR to be 68% in peripheral blood samples from 202 IBD patients (Nazareth *et al.*, 2015). It is yet unknown whether MAP can influence the development of CD or whether the presence of MAP in IBD patients occurs due to a loss of protective barrier function in these cases, i.e. is a secondary effect. Other studies have used GWAS analyses and have suggested that there is no clear association between MAP and Crohn's disease and that intestinal inflammation is more likely due to an autoimmune dysregulation caused by changes in the microbiome (Chacon *et al.*, 2004. Davis and Madsen-Bouterse, 2012b).

Studies have been carried out to investigate genetic susceptibility to Crohn's disease in humans and have found that genes associated with immune regulation confer susceptibility to IBD. The NOD2 protein (also known as CARD15) is present in the

gut tissues and plays an important role in recognizing bacterial molecules and stimulating immune reactions (Philpott *et al.*, 2014); variants identified in this gene have been associated with susceptibility to CD, although the specific functions of these variants are unclear (Ogura *et al.*, 2001. Hugot *et al.*, 2001). NRAMP1 (encoded by *SLC11A1* gene) plays an important role in innate immunity, preventing bacterial growth within macrophages during early stages of infection. When investigating nine *SLC11A1*variants in biopsy tissues from CD patients, three polymorphisms were significantly associated with CD when compared to healthy controls; speculating that over expression of the mutated alleles in CD patients could lead to increase severity of disease (Gazouli *et al.*, 2008). Similar variation has also been described in paraTB infected cattle; significant association of *SLC11A1* microsatellite frequency was found in paraTB when compared with uninfected controls (Pinedo *et al.*, 2009). The *IL23R* gene has been described as the 'inflammatory bowel disease gene' as this gene has been associated with IBD in several studies. This gene is required for efficient Th1/Th17 cell responses and production of subsequent cytokines via stimulation of these cells by IL12/IL23 (Yen *et al.*, 2006). The role of this gene in human disease is described in greater detail in section 4.1.

Common disease pathways are found in paraTB and IBD and it is suggested that anti-MAP therapies have the potential to be utilised in prevention and/or treatment of CD. The evidence of MAP present in CD patients and comparable pathology supports association rather than causality; therefore further investigation is required before this theory can be accepted.

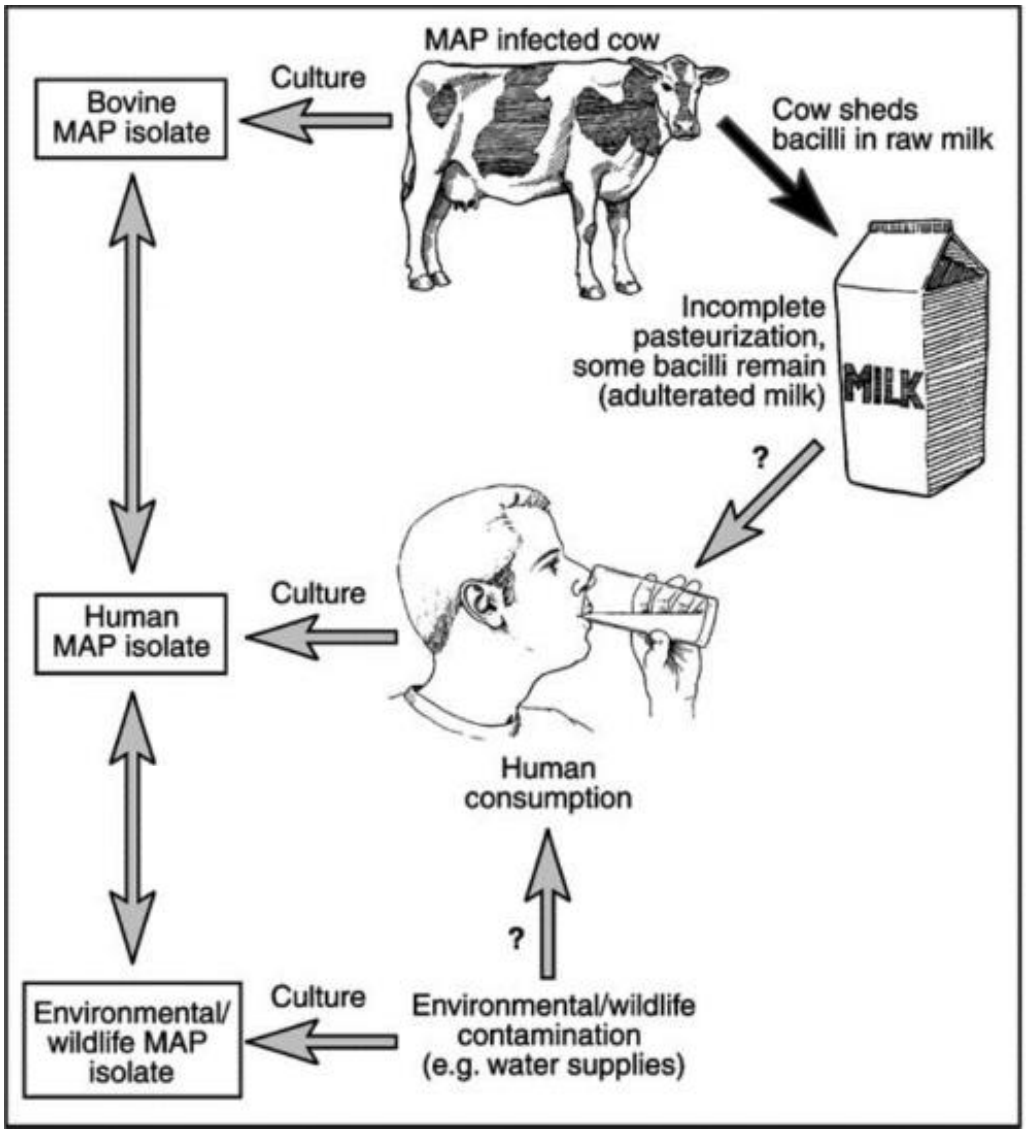


Figure 1.3 Potential sources of MAP transmission from animal to human. Humans can become infected from consumption of milk, dairy products and contaminated water supplies. Image adapted from (Chacon *et al.*, 2004).

There is a recurring link between the differential polarization of T cell subsets and the outcomes of pathology in human inflammatory diseases (Romagnani, 1996). Particularly in relation to genetic variants within genes that influence T cell polarization. Information collected from studies on human diseases has provided a basis for the hypothesis that genes which influence differential T cell polarization are crucial to the differentiation of immune response to paraTB infection in sheep, and determine the pathological outcome of disease.

1.11 mRNA transcription and alternative splicing

The effects that variation within immune regulating genes can have on subsequent immune responses and disease phenotype have been described. This section describes the ways in which this variation can occur and how changes in protein structure, as a result of gene sequence variants, can affect downstream activity.

Before synthesis of particular proteins can occur, an mRNA molecule corresponding to that protein must be synthesised by transcription; this process occurs in the nucleus of cells. In eukaryotes, RNA polymerase II is required for transcription of all genes that encode proteins. In order for RNA polymerase II to imitate transcription at the promoter region of a gene it requires a series of additional proteins, including transcription factors, which function to position the polymerase at a specific DNA template and separate the DNA strands; allowing transcription to begin. Further proteins recruited during initiation of transcription are responsible for histone modification and chromatin-remodelling (Chaffey, 2003). Certain transcription factors function to regulate the rate at which a particular gene is transcribed – described in more detail in chapter 7. Once the RNA polymerase II assembled with the proteins required to initiate transcription, now termed the transcription initiation complex, a series of conformational changes occur to tighten the binding of the polymerase to the DNA template; following which the polymerase disengages from the transcription factors and begins elongation of the mRNA transcript. Once elongation is underway, the transcription factors release from the DNA and can initiate further rounds of transcription.

During the elongation stage of transcription the RNA undergoes three particular processing events; 5’ end capping, intronic splicing (alternative splicing) and cleavage/polyadenylation of the 3’ end. Following these processes the newly synthesised mRNA is transported from the nucleus, through nuclear pore complexes, into the cytosol where subsequent protein translation is initiated (Chaffey, 2003).

Alternative splicing (AS) is a crucial mechanism for gene regulation and the generation of proteomic diversity; leading to synthesis of different protein products that function in a vast number of cellular processes including cell growth, cell

differentiation and cell death. Splicing is carried out by the spliceosome; a large structure which functions to accurately recognise the splice sites of a gene and catalyse the splicing reaction. Pre-mRNA splicing involves the removal of introns from the mRNA precursor. Decisions of which sequences of a gene are to be included in the final mRNA transcript rely on the ability of the splicesome to recognise specific splice sites located in the intron-exon boundaries of the pre-mRNA (Martinez and Lynch, 2013). Selection of exons to be included in the mRNA transcript depends on the positioning of the *cis*-regulatory elements - DNA sequences in the vicinity of a gene that are required for gene expression via complex networks of protein-DNA and protein-protein interactions. *Cis*-elements include sequences within exons such as exonic splicing enhancers (ESEs) and exonic splicing silencers (ESSs) as well as intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs). AS can be regulated at different steps of the spliceosome assembly by different splicing factors, most of which rely on *cis*-acting elements (Chen and Manley, 2009). AS refers to any variable of transcript that results from the inclusion/exclusion of an intron, exon or portion-thereof from a given gene. The main forms of variation that occur as a result of alternative splicing are shown in figure 1.4. The most common form is cassette-type alternative exon, also known as exon skipping. Intron retention in human transcripts is usually found in the untranslated regions (UTR) and has been associated with weaker splice sites and the regulation of *cis*-regulatory elements (Sakabe and de Souza, 2007). Differential splicing in the 5' or 3' UTR can alter the structure or presence of regulatory motifs including miRNA binding motifs and promoter binding sites, resulting in changes to mRNA stability, expression and/or translation; thus in turn has a profound effect on the expression pattern of resulting proteins (Martinez and Lynch, 2013). Alternative selection of exon splice sites often results in changes to protein structure which in turn can lead to changes in protein function. Protein changes also occur as a result of mutually exclusive alternative exon selection (Pohl *et al.*, 2013). Conservation of specific alternative splicing patterns within a gene indicates evidence of biological function, as non-functional isoforms are more likely to be subject to negative selection (Keren *et al.*, 2010). In particular, it has been shown that alternative

splicing plays a crucial role in the function and mediation of T cell and adaptive immune responses (Martinez and Lynch, 2013).

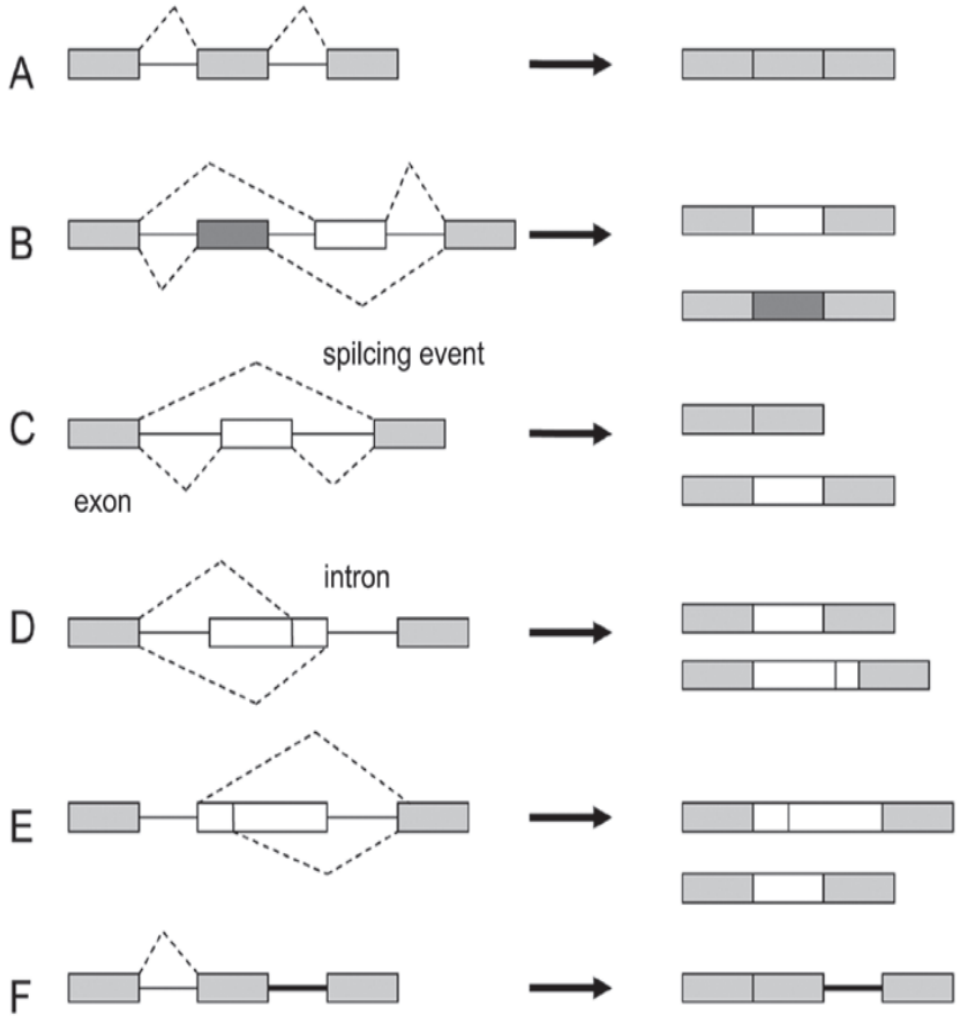


Figure 1.4 Main forms of alternative splicing events. A) constitutive splicing; B) mutually exclusive exons; C) cassette alternative exons; D) alternative 3' splice site; E) alternative 5' splice site; and F) intron retention. (Wang *et al.*, 2015).

The adaptive immune response has the ability to respond efficiently to an array of stimuli including pathogens, allergens and tumours; and utilises AS mechanisms to create multiple and distinct mRNA transcripts with diverse functions. AS isoforms occur in a tissue specific or stimuli dependant manner and function to regulate immune responses; common AS genes include cytokines and cytokine receptors (Wang *et al.*, 2015). Alternative splicing of cytokines and associated signal transduction molecules results in altered structure and function these molecules and can have a detrimental effect on subsequent downstream signalling and immune outcomes (Atamas, 1997). Splice variants often result in functional antagonists of their wild-type cytokine or can be expressed as soluble or membrane bound proteins. Examples include; an identified splice variant of IL10, IL10δ3 that was observed in leukemic cells from children with relapsed acute lymphoblastic leukaemia (ALL). Expression of this variant was found to correlate with a significantly better response to chemotherapy and clinical outcome of disease in children, indicating that the spliced isoform of IL10 may have a critical role in modulating the function of wild-type IL10 (Wu *et al.,* 2005). The IL4 receptor alpha subunit (IL4Rα) exists in two forms; one membrane bound and one soluble form which lacks the exons for the transmembrane and intracellular regions. It has been shown in mice that the soluble form can inhibit IL4-mediated cell proliferation and upregulation of IL5 by T cells; furthermore levels of IL-4, in the presence of IL4Rα isoform, was altered in asthmatic patients when compared to healthy controls (Atamas *et al.*, 1996. Kruse *et al.*, 1999. Bergin *et al.*, 2006).

These studies are just two examples of a vast array of identified AS variants in cytokine and cytokine receptor genes; and highlight the profound effects that AS variants can have on immune activity and outcomes of clinical disease. There is little known of the presence of transcript variants in MAP infected animals or the association of AS variants with the pathological outcomes of disease, particularly within genes that influence the differential Th cell polarisation. This project has investigated the presence of AS transcript variants in genes that influence Th cell polarization, and have been associated with Th polarization in human diseases, in the two pathological forms of sheep paraTB; including cytokine, cytokine receptor and transcription factor genes.

1.12 Aims and Hypothesis

T cell polarizing cytokines determine the differentiation of Th cell mediated immune responses and differential expression of these cytokines are associated with the different immune responses observed in pathological, multibacillary and paucibacillary, forms of paraTB. This study has primarily focussed on the IL23 and IL25 (also known as IL17E) cytokines and their receptor complexes; IL23R/IL12RB1 and IL17RB/IL17RA respectively. These cytokines have been implicated in other species, as key players in the polarization of T cell differentiation, driving adaptive Th1/Th17 and Th2 immune responses. These cytokines and receptors are associated with disease characterised by gut pathology and/or inflammation including TB, leprosy and IBD (Dubinsky, 2007.Kobayashi *et al.*, 2008). It is therefore suggested that these genes may influence the T cell immune responses that contribute to the pathological outcomes of paraTB in sheep. In order to obtain a more in depth view of the Th cell differentiation pathways, this project has also investigated the transcription factors that function to initiate expression of genes that are associated with Th cell profiles.

The aims of this project were to:

- Identify genetic variants in the IL23/IL25 cytokine receptor genes by obtaining and aligning gene sequences from sheep infected with the multibacillary and paucibacillary forms of paraTB, using reverse transcription PCR and cloning methods. Sequences were analysed for differential exon usage that may play a role in the biology of disease.

- Investigate expression levels of full length gene transcripts and variant transcripts of the IL23 and IL25 cytokine and cytokine receptor genes using real-time PCR analysis and determine whether differential expression of these genes, and/or variants therein, are correlated with the pathological outcomes of sheep paraTB.

- Investigate expression levels of full length gene transcripts and variant transcripts of transcription factors associated with Th1, Th17 and Th2 cell differentiation, using real-time PCR analysis and determine whether differential expression of these genes, and/or variants therein, are correlated with the pathological outcomes of sheep paraTB.

By investigating the above points, this project will contribute to the overall understanding of how the sheep immune system responds to MAP infection. Information obtained from the chosen genes aims to determine whether these genes are related to the clinical outcome of paraTB.

The hypothesis tested was that differential gene expression and/or genetic variants of IL23 and IL25 cytokines, their receptors and transcription factors associated with Th1, Th2 and Th17 cell subsets contribute to the differential polarization of the T cell immune responses seen in the two pathological disease outcome of MAP infection in sheep.

Chapter 2

Materials and Methods

2.1 Experimental animals and tissues

Animals were outbred Blackface female sheep with naturally acquired MAP infection. All sheep were between the ages of two and five years and were assessed for the presence of gastrointestinal strongyle nematode infection; details of the ID number, age, breed and faecal egg count (FEC) of these sheep are shown in table 2.1. Animals were euthanized and diagnosis of paraTB was confirmed by physical examination at post-mortem, haematoxylin & eosin and Ziehl-Neelsen histopathology and the presence of MAP specific insertion sequence *IS900*, by PCR using extracted genomic DNA (gDNA). There were six animals in each of the three groups; multibacillary (M), paucibacillary (P) and uninfected controls (C). Ileocecal lymph nodes (ILN) were removed immediately after post-mortem and approximately ~0.5g was placed into five volumes of RNAlater (Ambion, UK) that were then incubated at 4oC overnight before storing at -80oC. All animal experiments were in accordance with, and approved by, the Moredun Research Institute Experiments and Ethics Committee and conducted under the Animals (Scientific Procedures) Act 1986 project Licence.

2.2 *IS900* PCR

Genomic DNA was extracted from approximately 20mg of ileocecal lymph node tissue from each animal sample under sterile conditions. The Wizard® SV Genomic DNA Purification System (Promega Ltd) was used to extract DNA from samples as stated in manufacturer instructions. Concentrations of each DNA sample were measured using a NanoDrop ND-1000 spectrophotometer (Labtech International Ltd). For*IS900* PCR detection of MAP DNA, primers were used as stated by Eishi *et al*. and Bauerfeind *et al.*(Bauerfeind *et al.*, 1996. Eishi *et al.*, 2002). All PCR reactions were performed using FastStartTaq DNA Polymerase (Roche Diagnostics Ltd.) as stated in manufacturer instructions, using 500ng of gDNA and 0.2μM of each primer. The PCR cycle consisted of an initial denaturation step of 5 min at 95°C, followed by 35 cycles of 15 sec at 95°C, 15 sec at 55°C, 58°C, 60°C and 62°C, and 30 sec at 72°C, and a final elongation step of 72°C for 10 minutes. PCR products

were visualised on a 2% agarose gel before being purified, cloned and sequenced (see chapter 2 sections 2.4.5 – 2.6.5).

Sheep of similar age that presented no signs of clinical pathology and were negative for *IS900* were considered uninfected controls (*IS900* PCR/analysis carried out by collaborator Dr. Anton Gossner).

Table 2.1 Details of ID number, age, breed and FEC of sheep used in this study.

Sheep ID	Breed	Age (years)	FEC* (eggs/g)
SH139	Blackface	3	<250
SH140	Blackface	3	<250
SH146	Blackface X Bleu du Maine	2.5	<750 >250
SH204	Blackface	4.5	<750 >250
SH190	Blackface	3	>750
SH199	Blackface	2	<250
SH107	Texel X Blackface	2.5	<750 >250
SH147	Blackface X	2	<250
SH155	Blackface	3	<750 >250
SH160	Blackface X	3	<250
SH188	Blackface X Bleu du Maine	4	>750
SH205	Blackface	4	<750 >250
K207	Blackface	2.5	<250
K208	Blackface	2.5	<250
K213	Blackface	2.5	<250
K224	Blackface	2.5	<250
K227	Blackface	2.5	<250
K229	Blackface	2.5	<250

*FEC counts and categorisation from SCOPS manual (Abbot *et al.*, 2009)

2.3 RNA extraction and RNA quality analysis

2.3.1 RNA extraction

RNA was extracted from lymph nodes using the RiboPure™ RNA extraction kit (Ambion, UK), according to the manufacturer's instructions. Lymph node, stored at -80^{o}C in RNAlater™, was thawed and a piece of tissue was cut (20mg). Tissue samples were homogenised in 1ml of TriZol and proteins were precipitated with 200µl chloroform for 5 min at room temperature. Samples were centrifuged at 12,000 x g for 15 min at 4^{o}C before transferring the aqueous phase (top layer), containing RNA, to a fresh microcentrifuge tube. One sample volume of 70% ethanol was added to each reaction to precipitate the RNA. RNA was bound to spin column membrane as stated in manufacturer's protocol.

2.3.2 DNase treatment

DNase treatment was carried out in order to prevent genomic contamination of RNA samples. According to the On-column PureLink®DNase treatment protocol, 80µl DNase mastermix was prepared for each sample. Each mastermix contained 8µl 10x DNase I Reaction Buffer, 10µl resuspended DNase and 62µl RNase free water. Following binding of RNA to the spin column, 80µl DNase mastermix was added and columns incubated at room temperature for 15 min. Following 15 min incubation column membranes were washed with buffer I and II as stated in manufacturers protocol. RNA samples were eluted in 60ul RNase-free water.

2.3.3 Assessment of RNA quality and concentration

The concentration of RNA in each sample was measured using the Nanodrop 1000 spectrophotometer (Nanodrop Technologies, Inc). RNA quality was assessed using Agilent®Bioanalyzer with a RNA 6000 Nano LabChip kit (Agilent Technologies). This provides information of RNA integrity based on degradation level. This is measured on a scale of 1-10 for all RNA samples and each sample should contain a score of at least 7 RNA integrity number (RIN). RNA concentrations (ng/µl) and RIN for each sample are shown in table 2.2.

Sample ID	RNA Concentration (ng/µl)	RIN
SH139	1204.1	7.7
SH140	539.8	7.4
SH146	1165.7	8
SH204	2308.4	8.6
SH190	2344.6	8.3
SH199	1736.4	8.6
SH107	1348.6	8.4
SH147	1355.8	7.9
SH155	2004.9	7.9
SH160	1529.8	7.9
SH188	1064.7	7.5
SH205	2240.7	8.7
K207	1858.9	8.8
K208	1344.3	8.8
K213	454.9	8.7
K224	1998.2	8.5
K227	1945.6	8.8
K229	1750.4	9

2.4 Reverse transcription polymerase chain reaction (RT-PCR)

2.4.1 SuperScript™ II RT cDNA synthesis

Complementary first-strand DNA was synthesised with SuperScript™ II RT using oligo-dT (Invitrogen, UK cat no. 10777-019). Each reaction contained 1µg total RNA, 1µl oligo-dT, 1µl 10nM dNTP mix, made up to 12µl with deionised water. Reactions were heated to 65ºC for 5 min then chilled on ice before adding 4µl 5xFirst-Strand buffer, 2µl 0.1M DTT, then incubated at 42ºC for 2 min. 1µl of SuperScript™ II was added before final incubations of 42ºC for 50 min and 70ºC for 15 min to inactivate the reaction. cDNA samples were either diluted 1:10 in deionised water or stored as undiluted samples at -20ºC to be used as template for amplification in RT-PCR and standard curve analysis in real-time quantitative PCR (RT-qPCR).

2.4.2 RT-PCR Primer Design

Primers were designed for genes of interest based on ovine (predicted) sequences using Primer-BLAST, available from National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/tools/primer-blast), and NetPrimer (http://www.premierbiosoft.com/netprimer/index.html) software. Primers were selected using the following criteria; GC content between 40-65%, primer length of 19-25 base pairs, minimal secondary structure and self-complementarity. For genes greater than 800bp, overlapping sections were amplified individually using several primer pairs. Once gene section consensus gave sequence for the full length of the gene, primers were designed to obtain the full gene sequence in a single amplification. Primer sequences for full length gene amplification are shown in table 2.3.

2.4.3 Primer optimisation

Each primer set was optimised for use in RT-PCR using FastStart Taq DNA polymerase (Roche Applied Sciences) or Platinum Taq DNA polymerase High Fidelity (Invitrogen by Life Technologies) with reactions prepared as stated in

manufacturer manuals. Nested RT-PCR replaced required volume of DNA template with 1µl of PCR product from first-round amplification. Primers were optimised for annealing temperature (using a range of 45-65ºC) and elongation time period, calculated to 1min/kb.

Table 2.3 Detail of primers used to obtain full gene sequence. * depicts nested primer sets.

Gene	Primer	Sequence (5'-3')	Tm (°C)	Product size (bp)
IL23R	For	AAATTGACAGGCAGCAGTGAGG	55	2030
	Rev	TCTTCAGATTTCAAGGCAGGTTCT		
	For*	AAGGAAGACGTGAGGCCGATAC	62	2000
	Rev*	TTTCAAGGCAGGTTCTGACTACAC		
IL12RB1	For	ACCGGCAGCACAGAAGTTCA	55	2382
	Rev	GCCCTCTCTGAACCTCACTG		
IL17RB	For	AATAAGAGCGCGCGGCCCGAAG	55	1630
	Rev	ACTTCAGATTCTCACAGACACTT		
	For*	CGGCCCGAAGCCGATCCGGA	60	1575
	Rev*	GGGTCAGAAGGCTTTAGGG		
IL17RA	For	TTCCCGTGGTTCACATCGAG	52	539
	Rev	CAGGGTGACCGTGATGTGG		
	3' RACE	ACCGTCCACCACCTGCCTAAGCCCA	60	
	3' RACE*	ACATCACCGTGGAGACCCTTGAGG	68	
	5' RACE	TCCATGGGGTGAAGCTCAGCCGCAG	60	
	5'RACE*	CGTCAGGGATGGGCTTAGGCAGGT	68	
	5' UTR For	CTGTCTGCAAATACAAGGTCCTG	52	399
	5' UTR Rev	TCCACTCGATGTGAACCACG		

2.4.4 Agarose gel electrophoresis and PCR product purification

Agarose gels were prepared to visualise DNA fragments from PCR products and plasmids containing DNA inserts. PCR products were visualised on 1% gels. Gels were prepared by adding agarose (Sigma-Adrich) to 1x Bionic Buffer (Sigma-Aldrich) (1% = 1g:100ml) and heating until fully dissolved. GelRed nucleic acid gel stain (Cambridge Bioscience) was added to the dissolved agarose (1µl/100ml) to visualise PCR products. 2% agarose gels were adjusted accordingly. Electrophoresis was run at a constant 130V until DNA separation was achieved. Gels were viewed/images captured using an ultraviolet transilluminator with FluorChem™ imaging software. PCR products were separated by 1% agarose gel electrophoresis at 130v. Primers which gave a band of expected size were regarded as useable. PCR DNA products were then purified using either; Qiagen MinElute PCR purification Kit or excised from the agarose gel and purified using Qiagen MinElute Gel Extraction Kit (Qiagen, Crawley, UK). Protocols were followed as stated in Qiagen 'bench-top' user manual, centrifuge method. PCR products were eluted into 10µl EB buffer and analysed on a Nanodrop1000 spectrophotometer to determine DNA concentration.

2.5 Cloning and sequencing of PCR product

2.5.1 pGEM®-T Easy Vector ligation

Purified PCR product was ligated into Promega pGEM®-T Easy Vector (Promega, UK) according to manufacturer's protocol. Each reaction contained; required volume of PCR product to give a 1:5 DNA to vector ratio, 5µl 2xrapid ligation buffer, 0.5µl pGEM®-T easy vector (25ng), 1µl T4 DNA ligase (3 Weiss units/µl), made up to 10µl with deionised water. Ligations were either refrigerated overnight or incubated at room temperature for 1 hr before transformation into competent cells.

2.5.2 Transformation into JM109 cells

Ligated pGEM®-T easy vector samples were transformed into JM109 High Efficiency Competent cells (Promega, UK) by adding 2-10µl of ligated reaction to 20µl of cells and chilling on ice for 20 min. Samples were heat shocked for 45 sec at 42^{o}C before returning to ice for a further 2 min. Cell transformations were suspended in 980µl room temperature SOC (super optimal broth with catabolite repression) medium and incubated at 37^{o}C shaking for 1.5hr. Reactions were plated onto LB/Amp/IPTG/XGAL plates and incubated at 37^{o}C overnight. This method was used to allow for blue/white selection of colonies. The presence of white colonies indicates uptake of vector/DNA insert. Single white colonies were picked and transferred to 3ml LB broth and cultured overnight at 37^{o}C.

2.5.3 Purification of plasmid DNA

Plasmid DNA was purified from cloned JM109 cells using either Invitrogen PureLink Quick Plasmid Miniprep Kit or Promega Wizard® *Plus* SV Minipreps DNA Purification System. 1.5ml of overnight culture was lysed by the addition of reagents provided by manufacturer and spun at >12,000 x g to clear the lysate. Supernatant was then transferred to a separation column with a silica membrane designed to bind DNA. Following wash steps stated in manufacturer's manual, DNA

was eluted into 75µl of deionised water and analysed using a NanoDrop1000 spectrophotometer (Nanodrop Technologies) to determine DNA concentration.

2.5.4 Restriction enzyme digestion

Plasmids were digested using EcoRI High Fidelity restriction enzyme with CutSmart buffer (New England Biolabs) to check for the correct size of DNA insert within the plasmid. Restriction digest reactions contained: 2µl x10 restriction enzyme buffer, 0.5µl *Eco*RI restriction enzyme, required volume of plasmid DNA (~200ng), made up to 20µl with deionised water. Reactions were incubated at 37^{o}C for 1hr and visualised on a 1% agarose gel.

2.5.5 Sequencing of DNA

Plasmid DNA samples were sequenced on both strands using SP6 promoter (ATTTAGGTGACACTATAG)and T7 promoter (TAATACGACTCACTATAGGG) primers (Promega) in separate reactions with BigDye® Terminator v3.1 Cycle Sequencing Kit with protocol followed as stated in manufacturer's manual (Applied Biosystems, Warrington, UK). Each reaction contained; 0.5µl BigDye®, 1.75µl 5X sequencing buffer, 3.2µl 1µM T7/SP6 primer, required volume of plasmid DNA template (200ng), made up to 10µl with deionised water. Reactions were cycled as follows: 96^{o}C for 1 min, then 25 cycles of 96^{o}C for 10 sec, 50^{o}C for 5 sec, and 60oC for 4 min, then held at 4^{o}C. Reactions were sent to The GenePool, Edinburgh Genomics (University of Edinburgh) for sequencing.

2.5.6 Analysis of gene sequences

At least 2 clones of each amplified PCR product were sequenced three times, in two directions (i.e 2 forward, 1 reverse or 2 reverse, 1 forward direction) from 2 separate RT-cDNA synthesis reactions (see section 2.2.1), from multibacillary and paucibacillary sheep. Sequences from individual clones were compared to sheep genome assembly OARv3.1
(http://www.livestockgenomics.csiro.au/sheep/oar3.1.php/) by NCBI-BLAST

(http://blast.ncbi.nlm.nih.gov/Blast.) to ensure the correct region had been amplified. Sequences were then compiled using CLC Workbench v6.7.1 (www.clcbio.com) producing a consensus sequence for each sample. Consensus sequences were aligned using ClustalW v2.1 (www.ebi.ac.uk). Differences between consensus sequences were considered to be transcript/splice variants where the same variation was consistent in three or more consensus sequences. Sequences were translated from nucleotide to protein sequence by using ExPASy TRANSLATE tool (Bioinformatics Resource Portal. Swiss Institute of Bioinformatics. http://web.expasy.org/translate).

2.6 Reverse transcription real-time Polymerase Chain Reaction

RT-qPCR experiments were carried out in the Corbett Rotor-Gene 6 real-time PCR machine. SYBR Green FastStart Roche Taq DNA polymerase was used for all reactions.

2.6.1 RT-qPCR primer design and assay optimisation

Primers were designed based on sequence obtained from reverse transcription PCR method. Primer-BLAST and NetPrimer software was used to design primers specific to each gene. Primers were selected using the following criteria; GC content between 40-65%, primer length of 21-27 base pairs, minimal secondary structure and self-complementarity, single product size of between 75-220bp as shown in figure 2.1. Where possible, primers were located in different exons or in exon-exon boundaries. PCR products were sequenced to confirm product specificity. Primer designed to identify gene variants were designed with the variant region within the primer sequence. Sequences for each primer are shown in Table 2.4. RT-qPCR assays were optimised for template concentration using template serial dilutions of cDNA from paraTB infected animals (paucibacillary and multibacillary) and non-infected animals, primer concentration and annealing temperature. Assays were carried out using FastStart Universal SYBR Green Master (ROX) (Roche Applied Sciences). Each reaction contained 7.5µl SYBR green master, required volume of primer and template, made up to 15µl with deionised water.

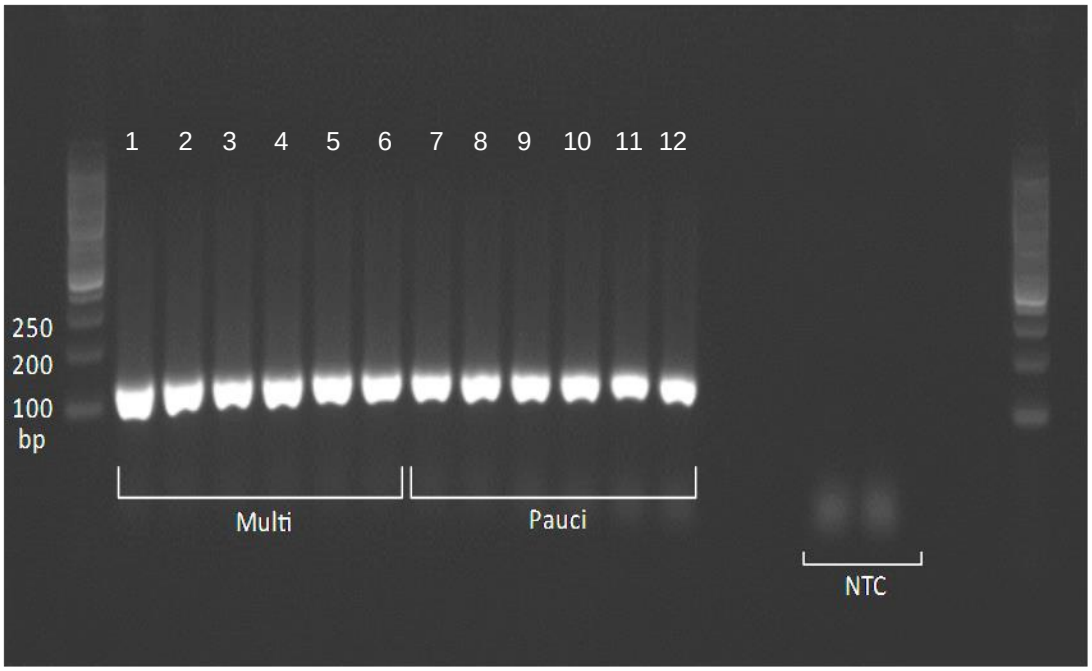


Figure 2.1 Representative gel electrophoresis image of RT-qPCR assay product (*IL17RBv3* - 122bp). Lanes 1-12; SH139, SH140, SH146, SH204, SH190, SH199, SH107, SH147, SH155, SH160, SH188 and SH205. No PCR product present in no template control (NTC). Amplified PCR products were sequenced to confirm specificity.

2.6.2 Selection of reference genes

Template serial dilutions, from paraTB infected animal cDNA, were used for reference genes SDHA and YWHAZ to provide control genes for future experiments. These genes were selected as they are constitutively expressed in the sheep samples and expression levels of these genes covered the range of expression of genes of interest (20-30 cycles). Each reaction contained 0.45µl, 0.5µl and 0.5µl of each reference gene primer (10mM), respectively, and annealing temperature of 62ºC was used for all genes.

2.6.3 Generation of standard curves& positive controls

RT-qPCR assays were optimised by generating a standard curve from a cDNA dilution series; this was carried out for all genes of interest. Each dilution point should have an equal difference between cycles to give acceptable efficiency. Standard curves with efficiency between 0.9 and 1.1, an M value (measurement of the gradient of the slope) of around -3.3 from at least four dilution points, R^2 values of as close to 1 as possible and a single melt peak, were considered acceptable. Figure 2.2 shows a representative image of an optimised assay. Assays where gene expression was late (>30 cycles), standard curves of acceptable efficiency were unable to be generated. These gene were considered too lowly expressed for accurate measurement in the samples used for this study. Positive control assays were carried out for these genes by adding PCR product (2-5ng) that is known to contain the gene region being amplified to the RT-qPCR reaction; this ensured that the assay was valid when a measurable level of the gene was present.

2.6.4 Data analysis & statistics

Relative gene expression levels were calculated using GenEx (MultiD Analyses AB) using the comparative 2-(ΔΔCq) method (Livak and Schmittgen, 2001). Data was collected from three separate biological replicates (3x RT reactions) per animal and two technical replicates for each animal. All genes were normalized to the mean of SDHA and YWHAZ reference genes. Fold changes were calculated from ΔCq values

using GenEx; the difference between group means for each gene was analysed by one-way ANOVA (with Tukey's pairwise comparison) with a significance threshold of $p \leq 0.05$.

A)

B)

C)

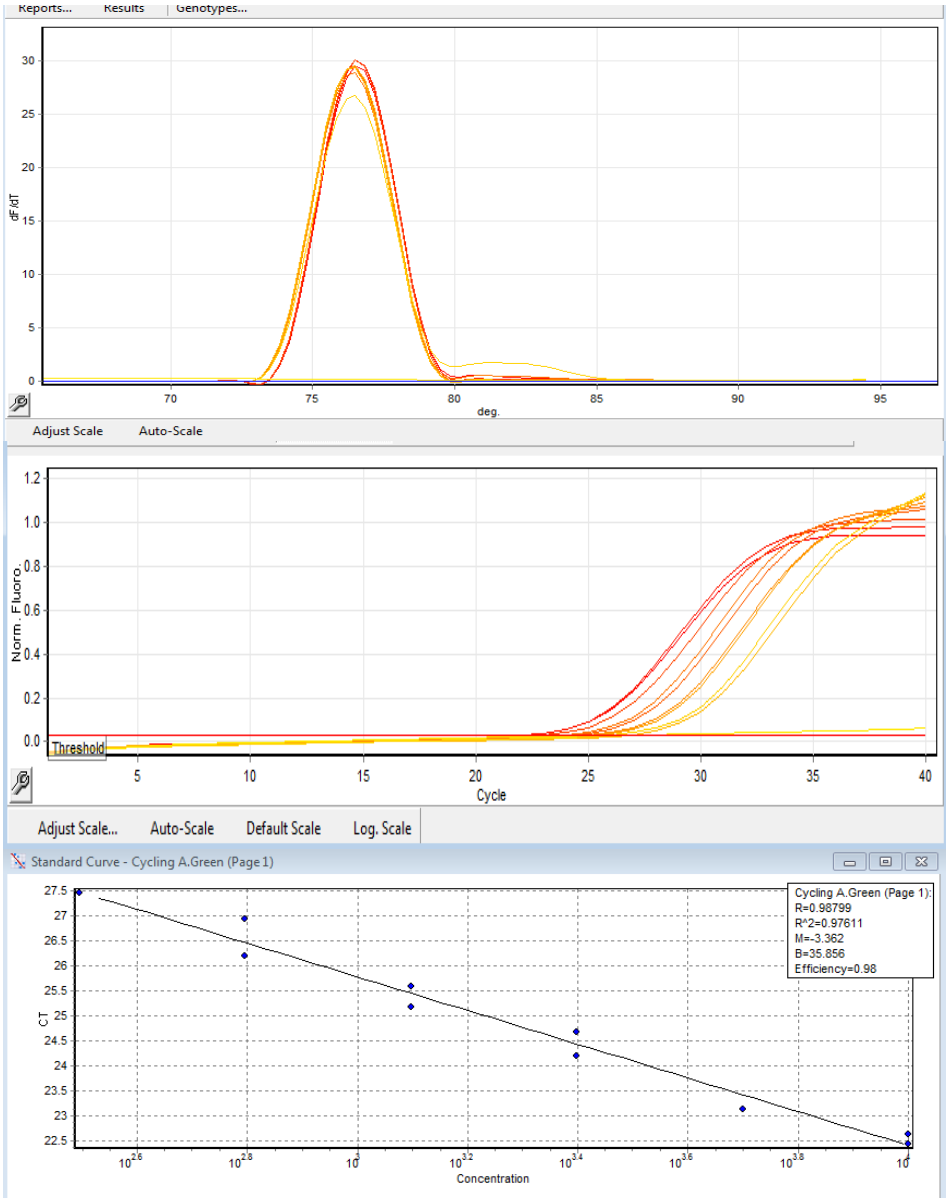


Figure 2.2 Generation of cDNA standard curves for RT-qPCR. A) single melt peak indicating single product amplified B) serial dilution standard curve with equal cycles between dilution points C) Efficiency calculated using at least four dilution points to measure the M value.

Table 2.4 RT-qPCR Primer details.

Gene	Primer	Sequence (5'-3')	Tm (°C)	Product size (bp)
IL23R	For	GGTGTACGTGAAGAGTTTAGAGACAG	62	131
	Rev	CAGTACATTTGAAGCTTGGACCC		
IL23Rv1	For	ACAGGTCGTAAGAGAGCGACAAAG	58	120
	Rev	AATCCCTCCATGACACCAACTGAA		
IL23Rv2	For	CATAGACACAAATTTAGAGACAGAA	60	129
	Rev	CAGTACATTTGAAGCTTGGACCC		
IL23Rv3	For	ACTTGGAACGTAAGCTAAATTCACT	58	85
	Rev	AATTCTGACTGTTGCTCATATGTAA		
IL23Rv4	For	TCCTGAAATAGACAACGGGAAAC	58	79
	Rev	AGAACTGACAGCATAACAGCAAA		
IL23Rv5	For	ACACCTTACTTCTGAATTAAAAGAAG	58	195
	Rev	CTGGGAGAATGATTTCTATCTCTG		
IL12RB1	For	AGGATTCGGTGTCTGTGGACTGGA	61	110
	Rev	CAGACGCTTGGTTGCTGTCCTCCT		
IL12RB1v1	For	TATCCGTGCTCCACAATGTCACTCT	60	150
	Rev	GCAAAGCTGCTCTTGGACATCTTGA		
IL12RB1v2	For	CTGCTCTTCTCAAGGACCGGATGGG	60	82
	Rev	TTCAGGGGGGATGCACAAGAGGAGAA		
IL12RB1v3	For	CTGAGCTGGAGCAGCCCTGTGTGCAT	62	120
	Rev	GCTGCTCATGTAAGGCCACCTGCC		
IL12RB1v4	For	CACCATACCCAGATGCAGACTCA	61	187
	Rev	CGCAGAGCCTGTGGCAAAGA		
IL17RB	For	CGGGCAGATGTGGACGTTTTC	60	130
	Rev	TGAAGTTCACAGCCATGGAGGG		
IL17RB v1	For	GATCATGAGAGAAGTGTGCTGGG	58	118
	Rev	GAAGTAACAAACGTGCCTCCAGA		
IL17RB v2	For	CGGGCAGATGTGGACGTTTTC	56	129
	Rev	TGAAGTTCACAGCCATGGAGGG		

IL17RBv3	For	TAACTCGAACTTCTGTGGTGGTCC	58	122
	Rev	GGTGCAGTGGGTCTTCAGTGAG		
IL17RA	For	CACACTGAGGCATCATCACAAGC	60	144
	Rev	GAAGTTTCTGGATTGGTGGTTTGG		
TBX21	For	CCTGTTGTGGTCCAAGTTC	60	120
	Rev	CGGTAATGGCTGGTGGGCTC		
GATA3	For	CCACAAGATGAACGGACAG	62	129
	Rev	GGCATTTCTTCTCCACAGAGTCGT		
GATA3v1	For	GCGAGATCCAGCACAGGCC	60	130
	Rev	GTTCTGTCCGTTCATCTTGTGG		
RORC2	For	CGCTGTGCCCACCGACTCACCGAG	62	136
	Rev	TGACCAGCACCACTTCCATG		
RORC2v1	For	TTGAAGGCTGCAGTGAACTC	60	116
	Rev	GGCATTGATGAGCACGAGG		
RORAv1	For	GTGATCGCAGCGATGAAAGC	60	101
	Rev	CCTTGCAGCCTTCACATGTAATG		
RORAv2	For	AGTTCTCTGCTGCAGTTGCTAAC	62	124
	Rev	GACGAGCTCATGGGCAAGG		
RORAv3	For	TGCGCAGACAGAGCTATTCC	58	112
	Rev	TGGGTCTTCTTTGTTACTGAGATACC		
RORAv4	For	GCAGCTTTCTTCTGGTGTCGTC	62	113
	Rev	TGGGTCTTCTTTGTTACTGAGATACC		
RORAv5	For	GGCTTTCTGTGGATGGGATC	63	133
	Rev	TGGGTCTTCTTTGTTACTGAGATACC		
YWHAZ	For	TGTAGGAGCCCGTAGGTCATC	60	101
	Rev	TCTCTCTGTATTCTCGAGCCATC		
SDHA	For	ACCTGATGCTTTGTGCTCTGC	62	126
	Rev	CCTGGATGGGCTTGGAGTAA		

Chapter 3

Definition of pathological forms of paratuberculosis

3.1 Introduction

MAP infection in sheep can give rise to two distinct pathological forms of disease. Approximately 30% of symptomatic animals are affected by the paucibacillary form; a tuberculoid state characterised by low bacteria levels and T cell infiltration of the gut causing pathology by chronic granulomatous inflammation. The other 70% of symptomatic cases have the multibacillary (lepromatous) form of disease; showing high levels of acid fast bacilli in the gut tissues, with macrophage and B-cell infiltration at the site of infection, leading to the formation of intestinal lesions (Smeed *et al.*, 2008.Clarke, 1997).

Previous studies have described the histopathology of intestinal samples from paucibacillary and multibacillary affected animals and have defined characteristics of the two pathologies of paraTB in sheep; multibacillary tissues contain numerous macrophages infected with acid-fast organisms, whereas paucibacillary tissues have very few macrophages, lymphocyte infiltration and very few or no acid-fast organisms found (Clark *et al.*, 2010.Corpa *et al.*, 2000). Previously, there have been very few studies that have investigated the presence of MAP using histology of lymph node tissue samples.

Results from physical examination of gut tissues at post-mortem, IS*900*PCR (carried out by collaborators Dr. Craig Watkins, Moredun Research Institute and Dr. Anton Gossner, Roslin Institute), and microscopy of histological section slides of the ileum and lymph node tissues collected from each animal at post-mortem, were combined in order to confirm the presence of para TB and distinguish MAP-infection status of each animal.

The aim of this work was to identify and characterise the paraTB status of each animal to be used in this study as either multibacillary, paucibacillary or control (uninfected). This was done by combining data from post-mortem examination, histopathology examination and *IS900* PCR to define the pathological status of disease.

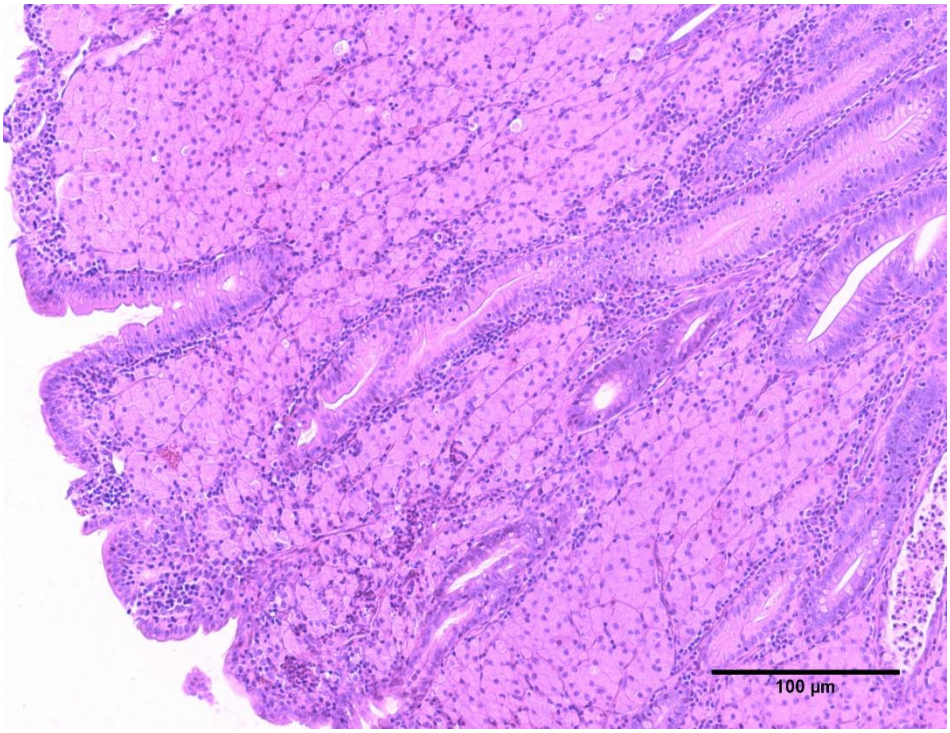
3.2 Histology and ZN analysis

Ileum and ileocecal lymph node tissue samples from suspected MAP infected sheep were stained by ZN or H&E methods (carried out by collaborator Dr. Francesca Chionini, Moredun Research Institute). Images of each stained section were taken using a light microscope (NikonNi2-Zeiss-105) to contribute to the final confirmation of multibacillary or paucibacillary infection status of each animal. Representative images of H&E and ZN stained ileum and lymph node tissue sections from suspected multibacillary/paucibacillary infected sheep and subclinical samples, to provide comparison, are shown.

The images shown in figure 3.1 and 3.2 indicate multibacillary infection given the presence of bacteria in both the ileum and lymph node tissues, identified by ZN staining (purple stain) characteristic of a multibacillary infection. The presence of MAP was confirmed in these tissues by identification of the specific *IS900* through RT-PCR (see figure 3.6). In contrast, the ZN stained sections shown in figures 3.3 and 3.4 showed no bacteria present in either the ileum or lymph node tissues, however a positive *IS900* RT-PCR result shows that MAP is present in these tissues; these samples are defined as a characteristic paucibacillary infection status.

Inflammation of the ileum can be observed in histology samples from both multibacillary and paucibacillary infected animals. Loss of villi structure is noted when comparing these samples to representative images of the ileum of sheep that are not infected with MAP (figure 3.5). Previous work carried out in our lab has identified the cell types causing this inflammation as predominantly macrophages and B-cells in multibacillary samples, and CD4+ T cells in paucibacillary samples (Smeed *et al.*, 2010).

A)



B)

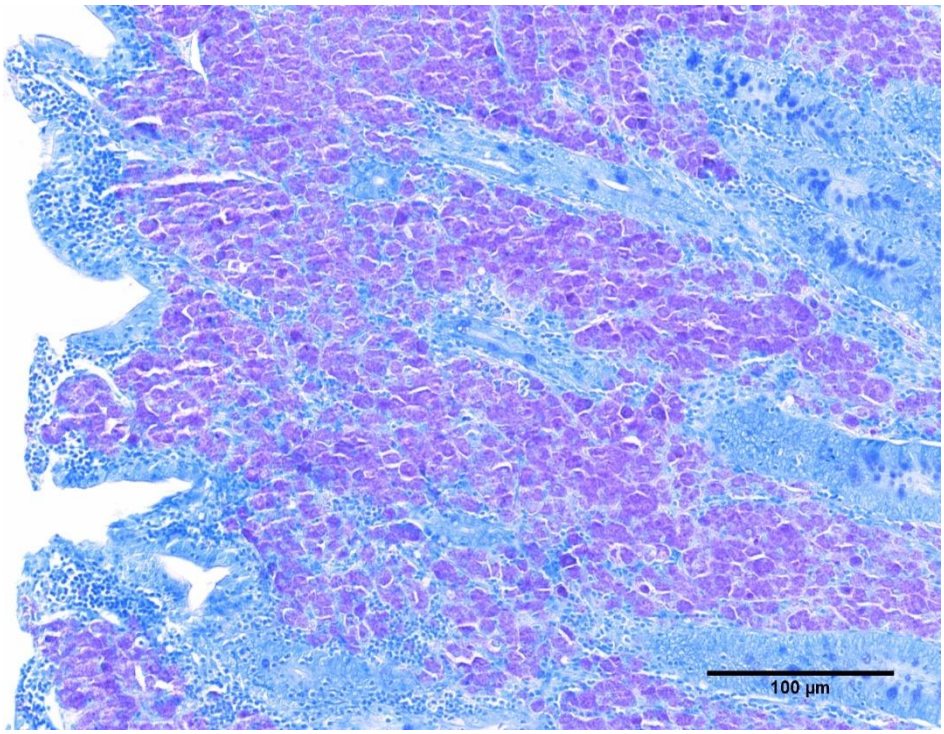
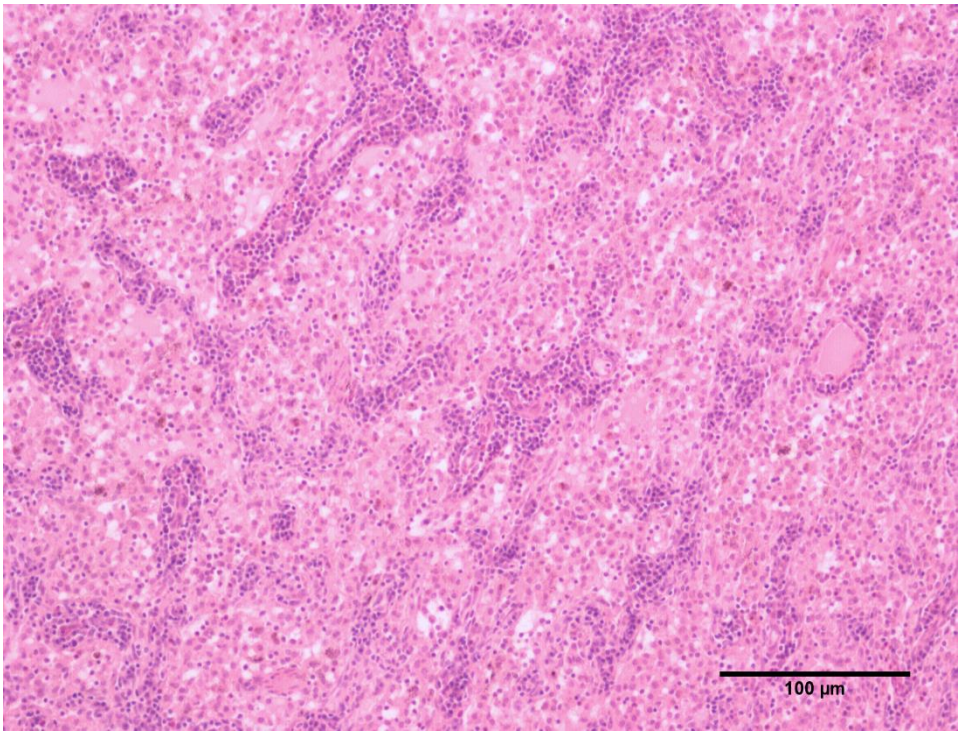


Figure 3.1 Representative histology images of a multibacillary infected sheep. Ileum tissues are shown with both H&E A) and ZN B) staining. Microscopy carried out at X10magnification (NikonNi2-Zeiss-105).

A)



B)

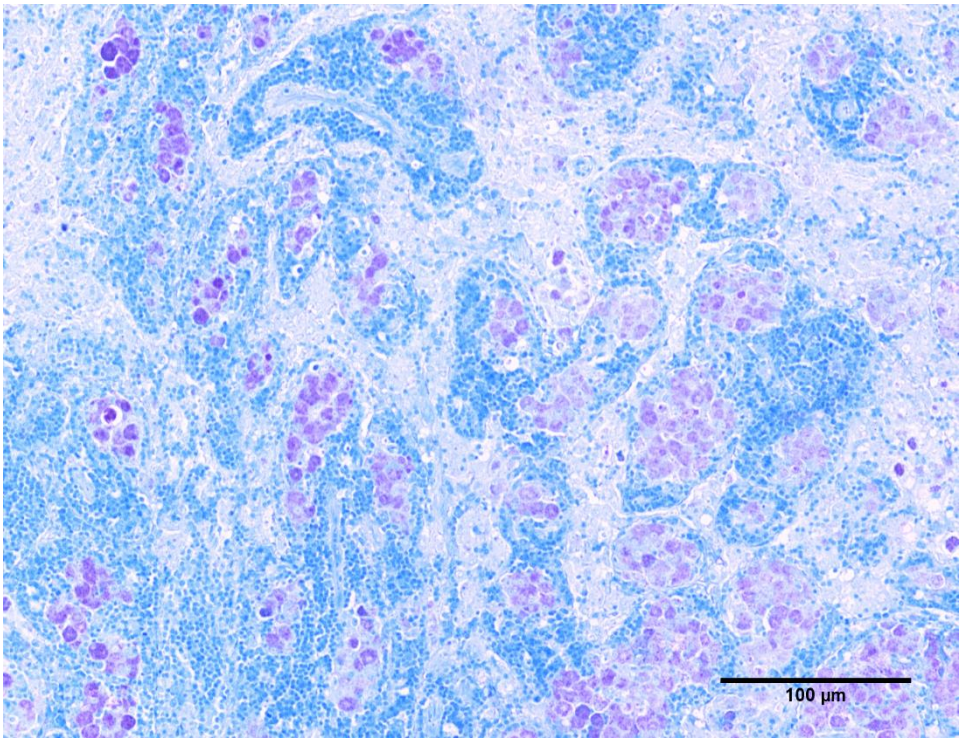
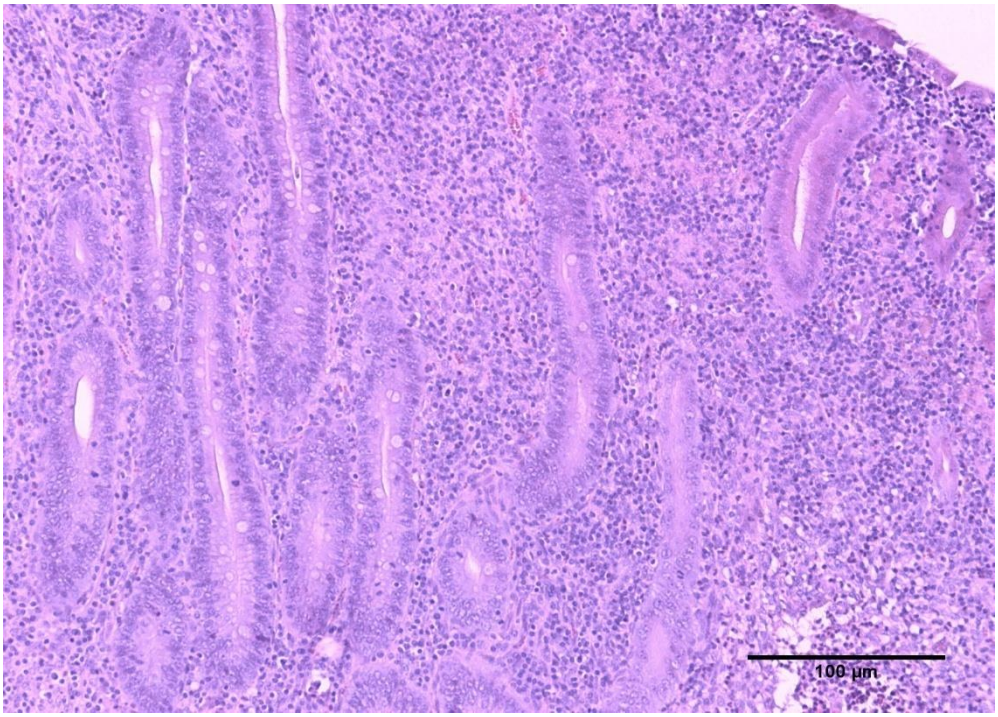


Figure 3.2 Representative histology images of a multibacillary infected sheep. Mesenteric lymph node tissues are shown with both H&E A) and ZN B) staining. Microscopy carried out at X10magnification (NikonNi2-Zeiss-105).

A)



B)

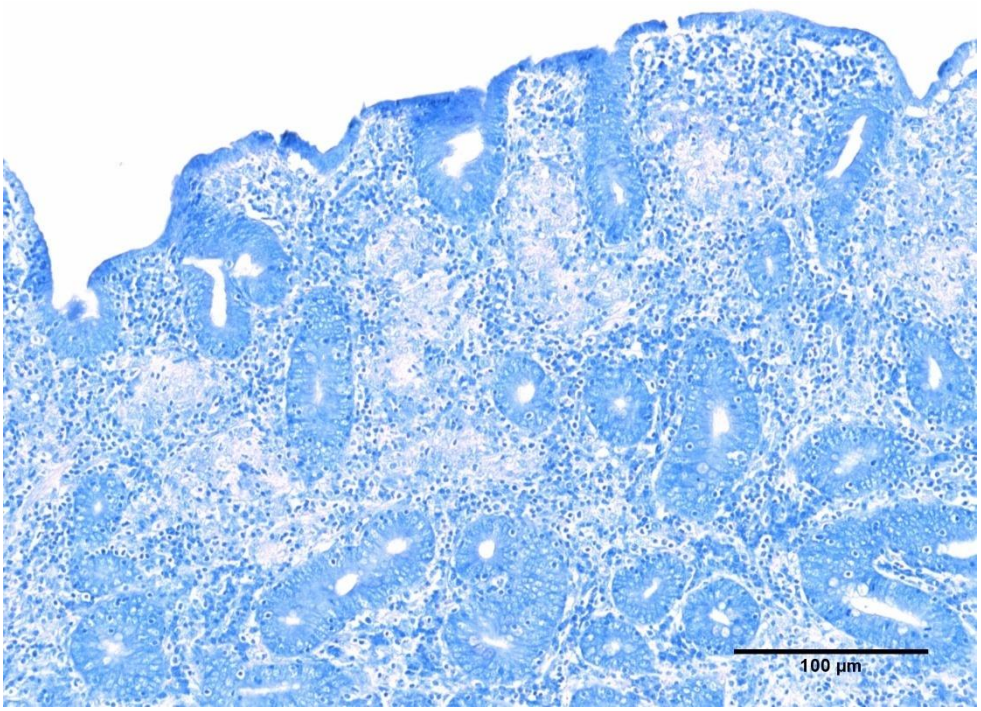
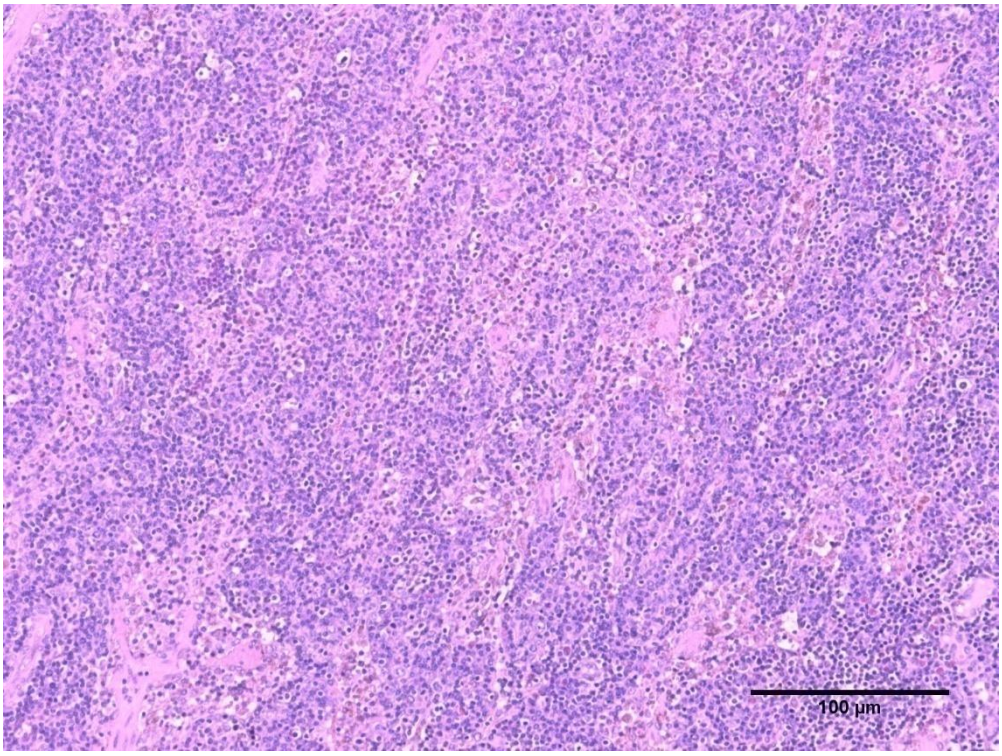


Figure 3.3 Representative histology images of a paucibacillary infected sheep. Ileum tissues are shown with both H&E A) and ZN B) staining. Microscopy carried out at X10magnification (NikonNi2-Zeiss-105).

A)



B)

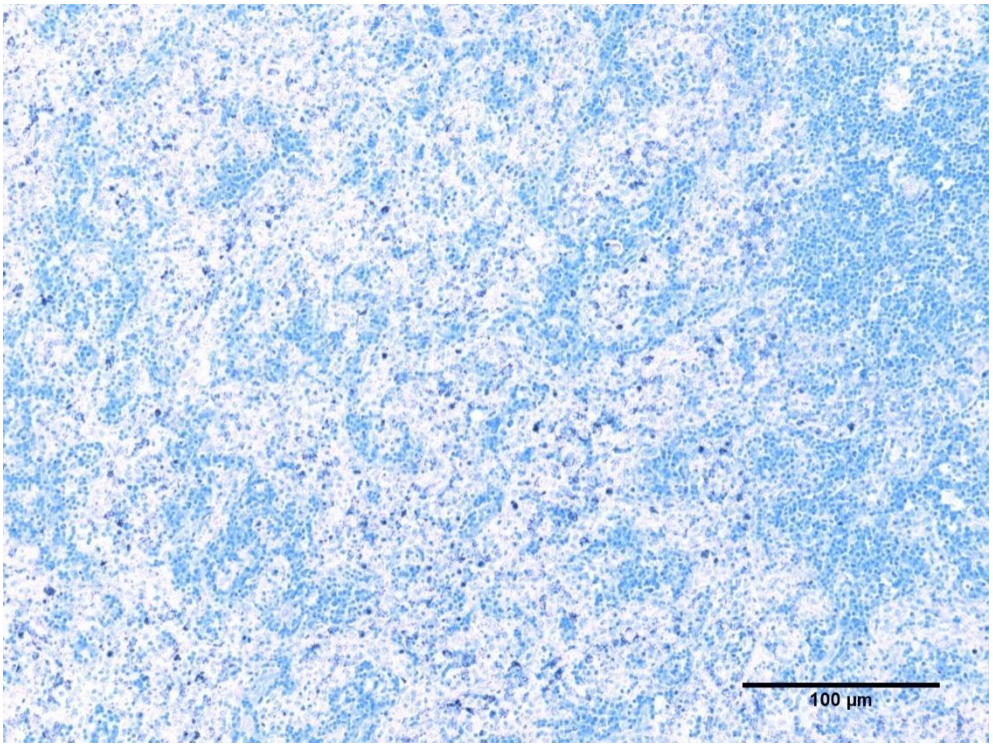
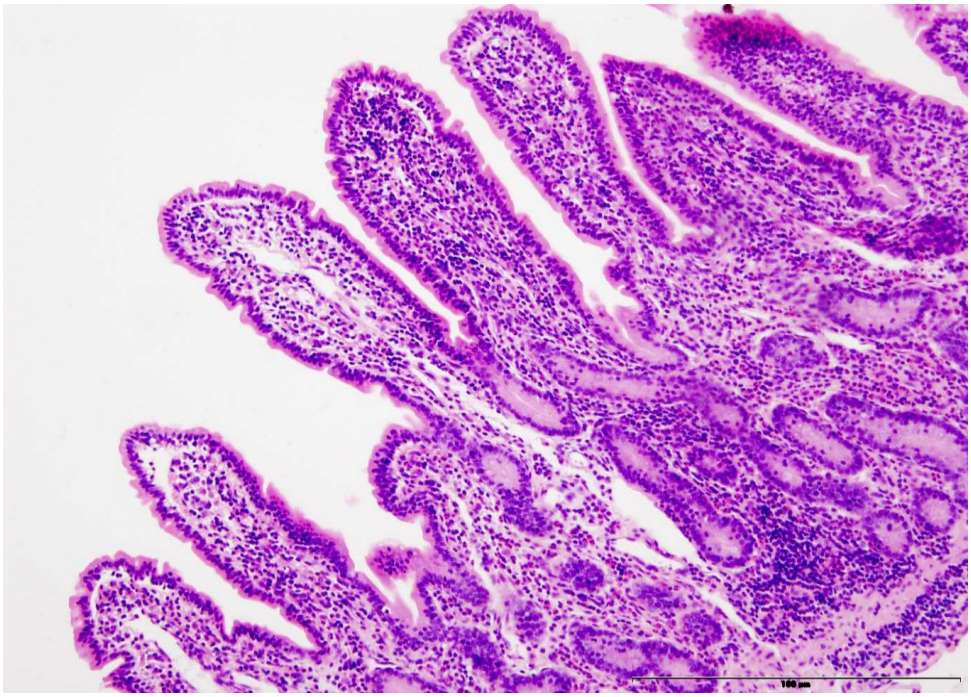


Figure 3.4 Representative histology images of a paucibacillary infected sheep. Mesenteric lymph node tissues are shown with both H&E A) and ZN B) staining. Microscopy carried out at X10magnification (NikonNi2-Zeiss-105).

A)



B)

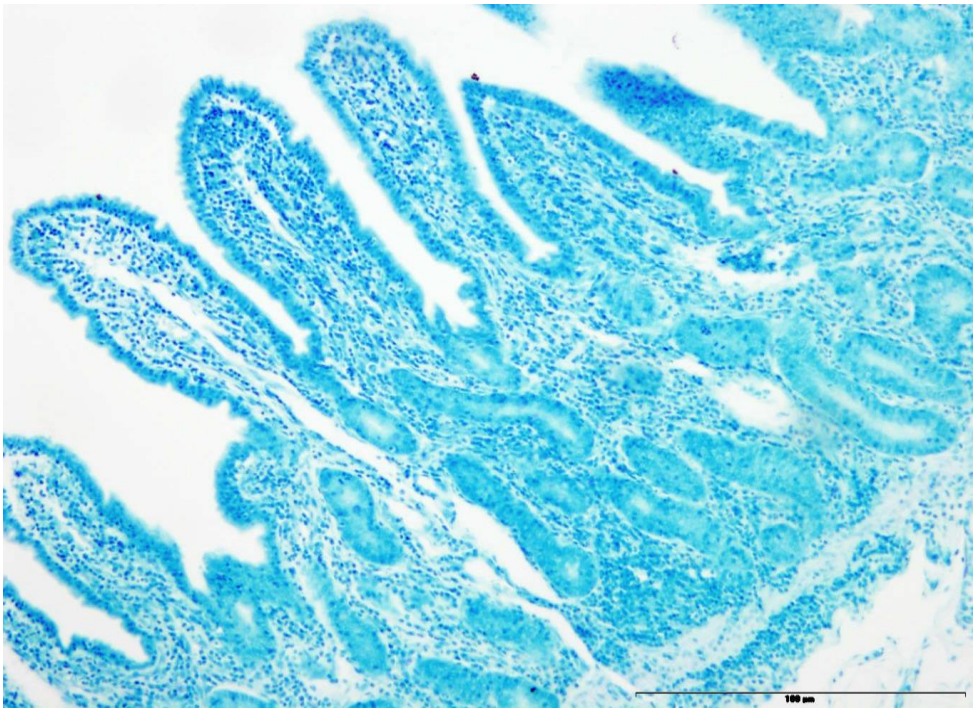
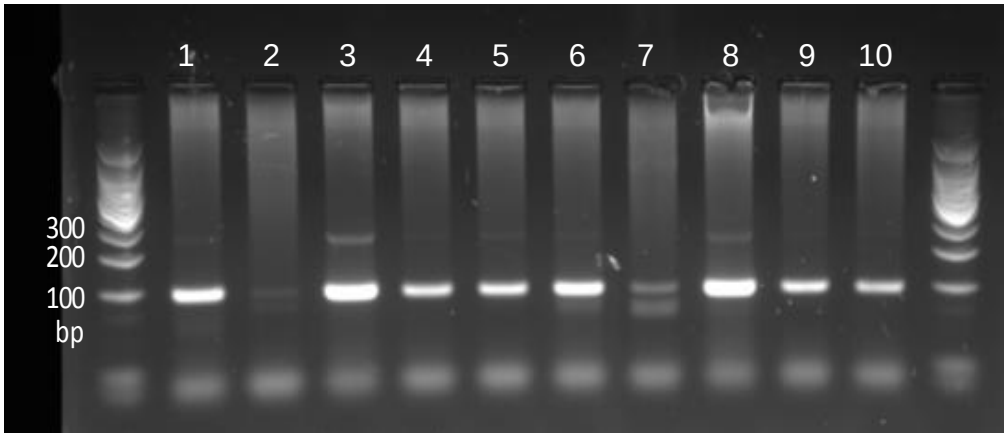


Figure 3.5 Representative histology images of healthy sheep. Ileum tissues are shown with both H&E A) and ZN B) staining. Microscopy carried out at X10 magnification (Smeed *et al.*, 2007).

3.3 Detection of *IS900* by PCR

IS900 PCR was carried out using gDNA from ICLN of each animal as described in section 2.1. Ileocecal lymph node (ICLN) was used as this is the draining lymph node from the site of infection. The presence of IS*900* amplicons, confirmed by sequencing of expected size PCR products analysed on 2% agarose gel (figure 3.6) from either primer sets were considered a positive identification of MAP in the sample/sheep. Two separate primer sets were used to provide confidence of results. The absence of amplicon from all PCR reactions in a sample was considered a negative result, indicating no MAP infection present. This study found all multibacillary and paucibacillary classified sample to be positive and uninfected control samples to be negative for the presence of *IS900*; results of the *IS900* PCR analysis are shown in table 3.1.

A)



B)

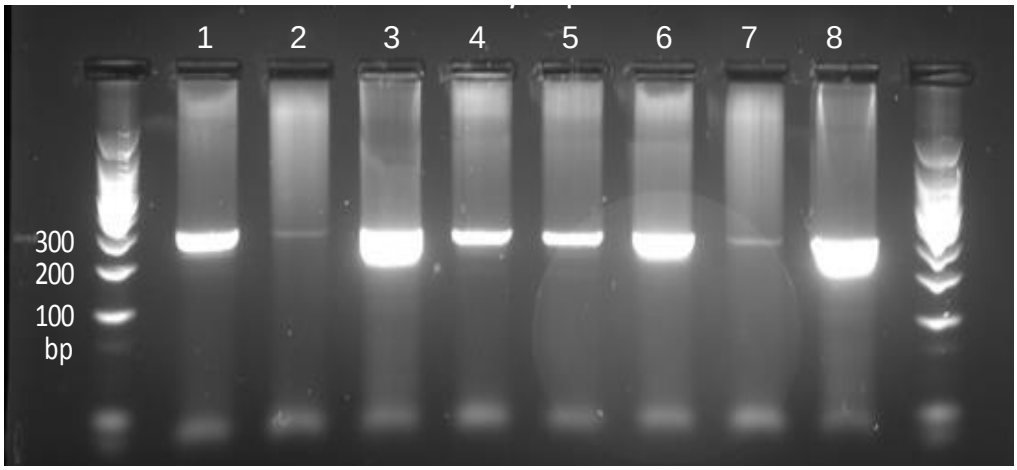


Figure 3.6 Representative gel image of *IS900* PCR product identified in ICLN of paratuberculosis infected samples using A) primers designed by (Eishi *et al.*, 2002). Expected product size of 99bp; SH167 (1-2), SH166 (3-4), SH140 (5-6), SH139 (7-8) and SH199 (9-10), and B) primers designed by (Bauerfeind *et al.*, 1996). Expected product size of 314bp; SH167 (1-2), SH166 (3-4), SH140 (5-6) and SH139 (7-8) (Images provided by Dr. Anton Gossner).

3.4 Classification of sheep samples

Information collected from post mortem examination, histopathology of ileum and lymph node tissues and *IS900* PCR was collated in order to define each animal as either multibacillary, paucibacillary infection status or uninfected control. Results of infection status for each animal used in this study are shown in table 3.1.

Table 3.1 MAP infection status of each sheep used in this study.

Sheep ID	*IS900* +/-	ZN: bacteria	Lesion Grade*	Classification
SH139	+	+	Severe	Multi
SH140	+	+	Severe	Multi
SH146	+	+	Severe	Multi
SH204	+	+	Severe	Multi
SH190	+	+	Severe	Multi
SH199	+	+	Severe	Multi
SH107	+	-	Mild	Pauci
SH147	+	-	Mild	Pauci
SH155	+	-	Mild	Pauci
SH160	+	-	Mild	Pauci
SH188	+	-	Moderate	Pauci
SH205	+	-	Mild	Pauci
K207	-	N/A	N/A	Control
K208	-	N/A	N/A	Control
K213	-	N/A	N/A	Control
K224	-	N/A	N/A	Control
K227	-	N/A	N/A	Control
K229	-	N/A	N/A	Control

*****Lesion grade: severity based on both total numbers of epithelioid macrophages and leukocyte distribution patterns of the terminal ileum (Nicol *et al.*, 2016).

3.5 Conclusions

- The presence of MAP bacterium can be seen by ZN staining in the lymph node and ileum tissues of multibacillary infected animals; however this is not seen in paucibacillary tissues.
- Inflammation of the ileum tissue, with loss of villi integrity, is noted in both paucibacillary and multibacillary animals when compared to subclinical animals; as shown by ZN and H&E tissue staining.
- *IS900* PCR has identified the presence of MAP bacterium in both multibacillary and paucibacillary infected animals. Control animals are *IS900* negative.
- The paraTB infection status of each sheep used in this study have been defined as either multibacillary or paucibacillary based on the combined data collected from port-mortem examination, histology of lymph node and ileum tissues and *IS900* PCR.

Chapter 4

IL23 axis genes in sheep paratuberculosis

4.1 Introduction

The interleukin 23 (IL23) cytokine is a novel member of the IL12 heterodimeric cytokine family that is created by the binding of IL12p40 to a p19 (IL23A) subunit, the former of which is shared with IL12. Initially it was thought that IL23stimulated Th1 immune responses along with IL12; however, although IL23 and IL12 exhibit some similar biological activities, IL12 functions in the differentiation of naïve T cells into Th1 lymphocytes and initiates subsequent IFNγ production, whereas IL23 has been shown to mediate pro-inflammatory responses via the activation of Th17 lymphocytes and subsequent production of Th17 cytokines including IL17A and IL17F (Kan *et al.*, 2008); creating a link between Th1/Th17 polarization (Kobayashi *et al.*, 2008).

Mice deficient in IL12 receptor subunit (IL12p35), IFNγ, IFNyR or STAT1, which are all critical molecules in the IL12/IFNγ induction of Th1 mediated responses, exhibit an increased severity of diseases including collagen induced arthritis (CIA), IBD and experimental autoimmune encephalomyelitis (EAE); indicating that the IL12/IFNγ pathway is not solely responsible for inflammation during autoimmune disease (Hunter, 2005.McKenzie *et al.*, 2006). In addition, treatment with mAbs against IL12p40 suppresses disease development in both human and animal disease models (Mannon *et al.*, 2004). Since the IL12p40 subunit is shared with IL12 and IL23, and anti-p40 mAbs inhibit both cytokines, involvement of IL23 in the induction of auto-inflammatory responses is suggested (Iwakura and Ishigame, 2006). In the presence of environmental danger signals and antigen recognition, antigen presenting macrophages and dendritic cells produce IL23 (Andersson *et al.*, 2004.Langrish *et al.*, 2005). During Th17 cell development, naive CD4+ T cells are activated by IL6 produced by dendritic cells which acts in conjunction with TGFβ to induce expression of the retinoic orphan receptor RORγt and up-regulate expression of IL23 receptor (IL23R) on the surface of T cells, thus making them competent for Th17 cytokine production via Signal Transducer and Activator of Transcription 3 (STAT3) signalling (Weaver et al 2007). The IL23R consists of IL12RB1 and IL23R subunits and this receptor complex is highly expressed on activated/memory T cells. Binding of IL23 and IL23R results in stimulation of memory T cells; IL23 binding to

IL23R expressed on Th17 cells triggers the activation of Janus kinase 2 (JAK2) and STAT3; resulting in phosphorylation and subsequent dimerization of these molecules (Lankford and Frucht, 2003). This in turn stimulates downstream expression of Th17 cytokines as depicted in figure 4.1. IL17A and IL17F are pro-inflammatory cytokines and are the best characterised cytokines of the IL17 family (Gu *et al.*, 2013). IL17A displays greater strength of signalling in terms of initiating downstream gene expression in comparison to IL17F; however both cytokines have been shown to stimulate endothelial cells, macrophages, epithelial cells and fibroblasts to produce inflammatory mediators including IL6, TNFα, IL1β and chemokines, leading to inflammation if not regulated (Gaffen 2009; Pappu et al 2011.).

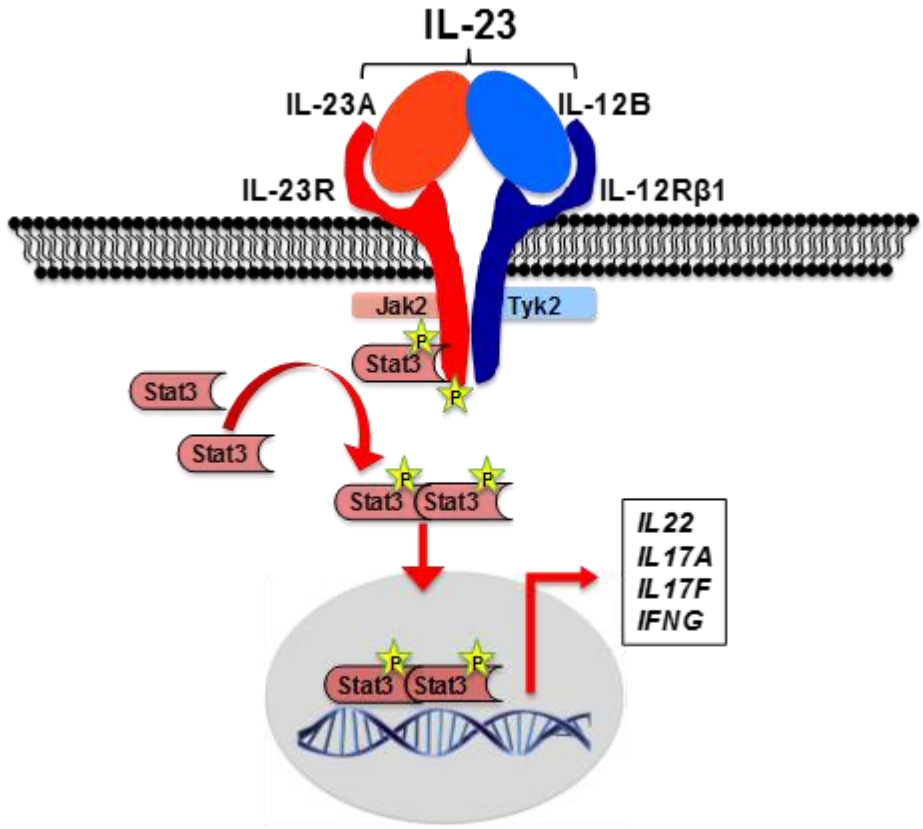


Figure 4.1 Binding of IL23 cytokine to IL23R. Phosphorylation of STAT3 molecules induces activation of STAT3 and initiates downstream expression of Th17 cytokines. Image adapted from Nature (Di Cesare *et al.*, 2009).

The IL23R is a type-1 cytokine receptor containing the highly conserved WSxWS motif, where x can be any amino acid; which is essential for receptor activation (Dagil et al. 2012). IL23R expression has been detected on macrophages and dendritic cells at low levels, which are consequently IL23 responsive (Langrish et al. 2004). The importance of IL23R signalling in controlling immunity through Th17 cell activity may explain why the *IL23R* gene has been implicated in susceptibility to, and pathogenesis of, several autoimmune and inflammatory diseases including rheumatoid arthritis, psoriasis, ankylosing spondylitis and multiple sclerosis (Abraham & Cho., 2009). Furthermore, IL23 plays a key role in the induction of intestinal inflammation. Th1 and Th17 immune responses are regulated by IL23 in ulcerative colitis (UC) and Crohn's disease (CD) patients. IL23R was up-regulated in CD4+ T cells from both CD and UC patients, whereas IL17A production was significantly increased by IL23 in UC, but not CD patients (Kobayashi *et al.*, 2008). Yen *et al.* have also shown that IL23 promotes T cell mediated chronic inflammation of the gut by up-regulating IL17 production (Yen *et al.*, 2006). At the gene level, variants of the *IL23R* gene have been identified and linked to susceptibility in IBD patients (Duerr *et al.*, 2006). In particular, the exon 9 deletion (Δ9) of IL23R in humans has been identified in, and associated with, CD patients when compared to healthy controls, however it is yet to be determined whether this variant plays a role in initiation of pathology in CD (Gallagher *et al.*, 2012). A GWAS study identified *IL23R* as one of the genetic factors linked to CD. One missense polymorphism within *IL23R*, the R318Q SNP, was found to be significantly higher amongst healthy controls than in diseased patients, suggesting a protective role for this allele from immune mediated chronic inflammation (Di Meglio *et al.*, 2011). This SNP consisting in a guanine (G) to adenine (A) substitution at DNA level results in an arginine (R) to glutamine (Q) substitution within the cytoplasmic domain of IL23R (Duerr *et al.*, 2006). A functional study of the R318Q SNP found that IL23-induced Th17 cell effector function was impaired in carriers of the protective allele, resulting in reduced STAT3 phosphorylation and IL17A production (Di Meglio *et al.*, 2011). These findings support a critical role for IL23/IL23R signalling in the generation of pathogenic Th17 responses. The IL12RB1 subunit of IL23R has also been implicated in human disease, including susceptibility to mycobacterial and *Salmonella*

infections. A lack of IL12RB1 chain expression was identified in individuals presenting with severe idiopathic mycobacterial and *Salmonella* infections (de Jong *et al.*, 1998). Sequence analysis revealed genetic mutations resulting in premature stop codons in the extracellular domain of IL12RB1. A detrimental lack of IL23/IL12 signalling with reduced IFNγ, IL17A and IL17F production resulted in human immunodeficiency and susceptibility to intracellular bacterial infection (de Jong *et al.*, 1998). Other studies have shown complete IL12RB1 deficiency to be the most frequently known genetic aetiology of the syndrome of Mendelian susceptibility to mycobacterial disease (MSMD). Reduced IL12RB1 signalling results in a lack of STAT3 phosphorylation and downstream induction of gene expression (Fieschi *et al.*, 2004. Miller and Robinson, 2012). The studies described highlight the importance of IL12Rβ1/IL23R signalling in the induction of both Th1 and Th17 cell responses to control autoimmune and intracellular bacterial diseases.

When investigating the role of *IL23* in paraTB, Gossner *et al*. described expression levels of the IL23 cytokine subunit *IL23A* in four sheep groups; uninfected controls, asymptomatic, paucibacillary and multibacillary. It was shown that *IL23A* transcript expression in ileal mucosa from paucibacillary infected animals were significantly higher than that of the other three groups; furthermore *IL23A* expression levels were higher at the site of infection, the ileal mucosa, than in lymph node tissues. This study also identified an allelic variant present in the sheep *IL23A* gene. The allelic variant was identified when measuring expression levels of *IL23A* in two distinct gastrointestinal disease infections in sheep; pataTB and *T. Circumcincta* (Gossner *et al.*, 2012b). Expression of the p19 subunit (*IL23A*) by mature monocyte-derived dendritic cells is enhanced by peptidoglycan, a major constituent of gram positive bacterial cell walls, but not by lipopolysaccharide derived from gram negative bacteria (Cummings *et al.*, 2007), which may explain increased expression of IL23 in gram positive MAP infected host.

In this chapter the sequencing of full gene transcripts and transcript variants of the IL23 cytokine receptors, *IL23R* and *IL12RB1*, as well as sequence of the *IL23A* cytokine subunit, is described in the two pathological forms of sheep paraTB. This

chapter also quantifies gene expression levels of *IL23A* and the two IL23 receptor genes, including identified transcript variants, in paraTB infected and control sheep.

4.2 Ovine *IL23A* gene

4.2.1 Sequencing of *IL23A* allelic variants in the two pathological forms of sheep paraTB

In order to investigate the presence of the known ovine *IL23A* allelic variants, present in exon 1 of sheep *IL23A*, in the two study groups of paraTB infected animals, RT-PCR was performed to amplify partial *IL23A* sequence as described by Gossner and colleagues (Gossner *et al.*, 2012b). Three clones from each animal were sequenced to give a representative sample of transcripts within each animal and experimental group. Consensus sequence of *IL23A* clones from both paucibacillary and multibacillary groups gave identical sequence. This *IL23A* sequence (IL23A1) was aligned with ovine *IL23A* sequence containing allelic variant found in two separate studies; GenBank FN822243.2 (IL23A2), GenBank NM_001185122.2 (IL23A3) and sheep genome sequence (OARv3.1). Both nucleotide and protein sequence alignments are shown in figure 4.2. There were no allelic variants found in either the multibacillary or paucibacillary MAP infected sheep in this study.

A)

```
IL23A2          GGAGCCAGCCAGATTGGAGAGGCAGGCAACGAGATGCTGGGGAACAGAGTTGTGATGCTG 60
IL23A3          GGAGCCAGCCAGATTGGAGAGGCAGGCAACGAGATGCTGGGGAACAGAGTTGTGATGCTG 60
OARv3.1         GGAGCCAGCCAGATTGGAGAGGCAGGCAACGAGATGCTGGGGAACAGAGTTGT------G 54
IL23A1          GGAGCCAGCCAGATTGGAGAGGCAGGCAACGAGATGCTGGGGAACAGAGTTGT------G 54
                ************************************************     *

IL23A2          CTGCTACTGCTACTGCTGCCCTGGACAGCTCAGGGCCGGGCTGTGTCAGAGGACAGCAGC 120
IL23A3          CTGCTACTGCTACTGCTGCCCTGGACAGCTCAGGGCCGGGCTGTGTCAGAGGACAGCAGC 120
OARv3.1         CTGCTACTGCTACTGCTGCCCTGGACAGCTCAGGGCCGGGCTGTGTCAGAGGACAGCAGC 114
IL23A1          CTGCTACTGCTACTGCTGCCCTGGACAGCTCAGGGCCGGGCTGTGTCAGAGGACAGCAGC 114
                ************************************************************

IL23A2          CCTGCTTGGGCTCGGGGCCAACAGCTCTCACAGCAACTCTGCATGCTAGCCTGGAGTGCA 180
IL23A3          CCTGCTTGGGCTCGGGGCCAACAGCTCTCACAGCAACTCTGCATGCTAGCCTGGAGTGCA 180
OARv3.1         CCTGCTTGGGCTCGGGGCCAACAGCTCTCACAGCAACTCTGCATGCTAGCCTGGAGTGCA 174
IL23A1          CCTGCTTGGGCTCGGGGCCAACAGCTCTCACAGCAACTCTGCATGCTAGCCTGGAGTGCA 174
                ************************************************************

IL23A2          CACCTACCAATGGGACATGTGGATCTACCAAGAGAAGAAGGAGGTGATGAGACTACAGAT 240
IL23A3          CACCTACCAATGGGACATGTGGATCTACCAAGAGAAGAAGGAGGTGATGAGACTACAGAT 240
OARv3.1         CACCTACCAATGGGACATGTGGATCTACCAAGAGAAGAAGGAGGA--------------- 219
IL23A1          CACCTACCAATGGGACATGTGGATCTACCAAGAGAAGAAGGAGGA--------------- 219
                ******************************************:
```

B)

```
OARv3.1         MLGNRVV--LLLLLLPWTAQGRAVSEDSSPAWARGQQLSQQLCMLAWSAHLPMGHVDLPR 58
IL23A1          MLGNRVV--LLLLLLPWTAQGRAVSEDSSPAWARGQQLSQQLCMLAWSAHLPMGHVDLPR 58
IL23A3          MLGNRVVMLLLLLLLPWTAQGRAVSEDSSPAWARGQQLSQQLCMLAWSAHLPMGHVDLPR 60
IL23A2          MLGNRVVMLLLLLLLPWTAQGRAVSEDSSPAWARGQQLSQQLCMLAWSAHLPMGHVDLPR 60
                *******  **************************************************
```

Figure 4.2 Sequence alignment of *IL23A* alleles, Nucleotide A) and amino acid B) and OARV3.1 genome sequence. Previously identified allelic variant is highlighted in yellow.

4.2.2 Gene expression levels of *IL23A*

The IL23 cytokine is a heterodimeric ligand composed of IL23p19 (IL23A) and IL12p40 (IL12B) subunits. Gene expression levels of *IL23A* within ICLN were measured by RT-qPCR; and found that *IL23A* could only be identified at >30cyles and therefore was too low for accurate measurement. Positive control for this assay was carried out using cDNA from paraTB infected animals spiked with *IL23A* PCR product as described in section 2.6.3. The positive control confirms that this assay is optimised to measure *IL23A* when then gene is present in greater abundance, see figure 4.3.

A)

B)

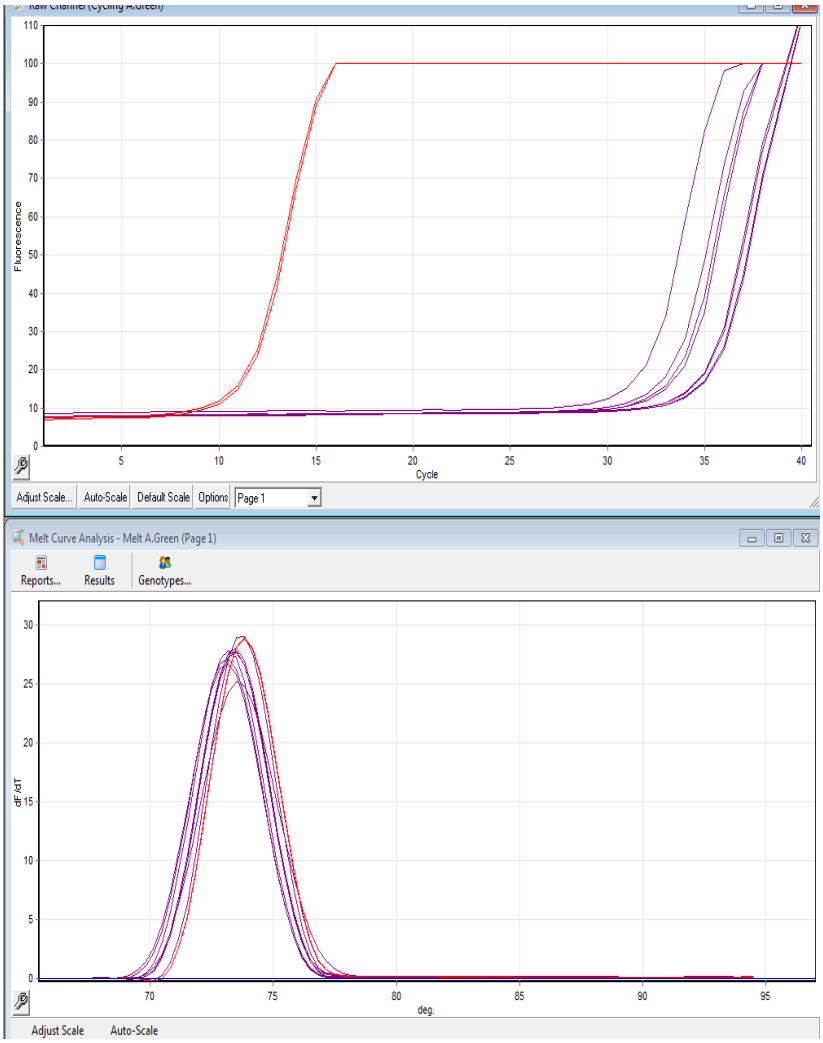


Figure 4.3 Positive control RT-qPCR assay for *IL23A*. Cycling A) and melt curve B) analysis of pooled cDNA from paraTB infected and control animals spiked with *IL23A* PCR product, shown in red. Cycling and melt curve analysis of dilution series (1:4-1:64) pooled cDNA from paraTB infected animals, shown in purple.

4.3 Ovine *IL23R* gene

4.3.1 Identifying *IL23R* transcript variants

To identify alternatively spliced transcript variants present in the sheep *IL23R* gene, full length *IL23R* was amplified by RT-PCR using primers described in table 2.3. At least six clones from each animal were sequenced to give a representative sample of transcripts present within individual animals and within experimental groups. Each clone was aligned against sheep genome assembly OARv3.1 and differences in sequence that were present in 2 or more clones from different animals were considered transcript variants. Alternatively spliced transcript variants were identified in several clones from both paucibacillary and multibacillary affected groups as summarised below. Gene maps of variants identified are depicted in figure 4.4. *IL23R* sequence and *IL23R* transcript variant sequences were aligned using ClustalW v2.1 (as described in Chapter2) and are shown in figure 4.5. The ovine *IL23R* gene is found on chromosome 1 and consists of 11 exons. Expression patterns of each variant was measured using RT-qPCR.

Cloning and sequencing identified 5 splice variants of ovine *IL23R*. The *IL23R* is encoded on the plus strand of chromosome 1 (OARv3.1: NC_019458.1). Differences between full length *IL23R* (LN868336) and identified transcript variants are described below and are detailed in figures 4.5, 4.6 and 4.7;

IL23Rv1 (LN868337) has an insertion of 29bp (bp79 – 108) in the 5' UTR (Chr1: 42,464,555 – 42,464,584) which results in no change to coding region and the same 627 aa predicted amino acid sequence as full length IL23R and would presumably function as full length *IL23R*.

IL23Rv2 (LN868338) has a deletion of 21bp (bp530-551) at the 5' end of exon 4 (Chr1: 42,476, 073 – 42,476,094) resulting in a 7 amino acid (YVVYVKS) deletion at position 150 and a 2bp insertion (bp857-859) resulting in change of amino acid sequence from position 259 consequently leading to a premature stop codon at position 287 within the protein; it is predicted that this is a truncated protein.

IL23Rv3 (LN868339) has an insertion of 28bp (bp878-906) at the 3' end of exon 6 (Chr1: 42,492,950 – 42,492,977) resulting in premature stop codon and a protein of 268 amino acids.

IL23Rv4(LN868340) has the deletion of exon 8 (Chr1: 42,512,417 – 42,512,506) and encodes a predicted protein of 597 amino acids as a consequence of a 30 amino acid (PQVTMKSFQHDTQNSGLLIASIFKKHLTSD) deletion at position 319. This deletion causes truncation of the extracellular region; even though this protein still contains the conserved motifs. The deletion of exon 8 may lead to changes during receptor-ligand binding.

IL23Rv5 (LN868341) has the deletion of exon 9 (Chr1: 42,515,314 – 42,515,419) resulting in a premature stop codon at the 5' end of exon 10. This is position 356 within the protein. This protein is comprised of a truncated extracellular region that still contains the WSxWS motif but no transmembrane domain. Therefore this variant may encode a soluble form of the IL23R.

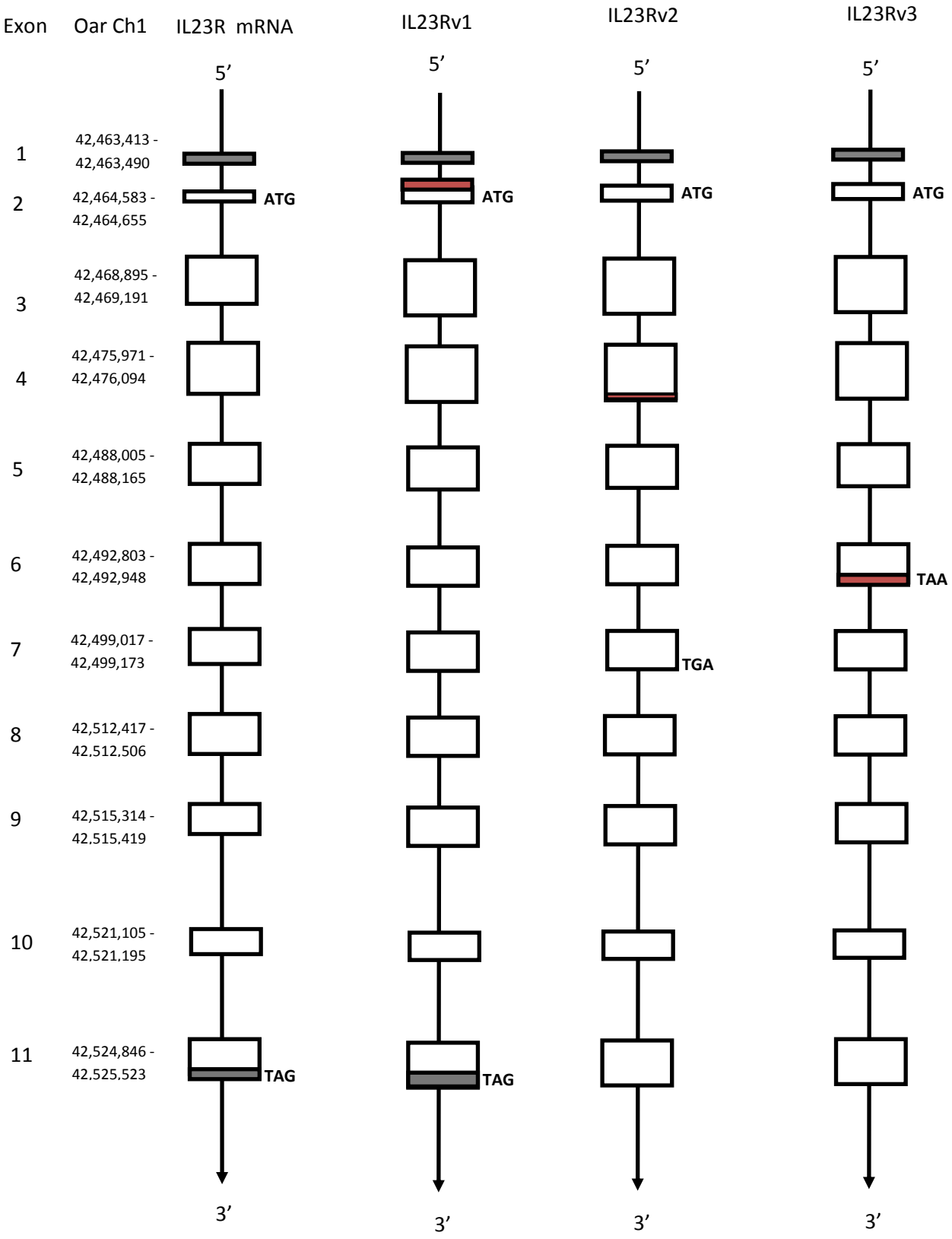


Figure 4.4 Schematic diagrams of the ovine *IL23R* and splice variants *(IL23Rv1-3*) identified in this study. Boxes represent individual exons with chromosome location shown. Grey boxes indicate 5'and 3'UTRs with variant regions highlighted in red. The 'ATG' is initial start codon and stop codons are shown for each transcript.

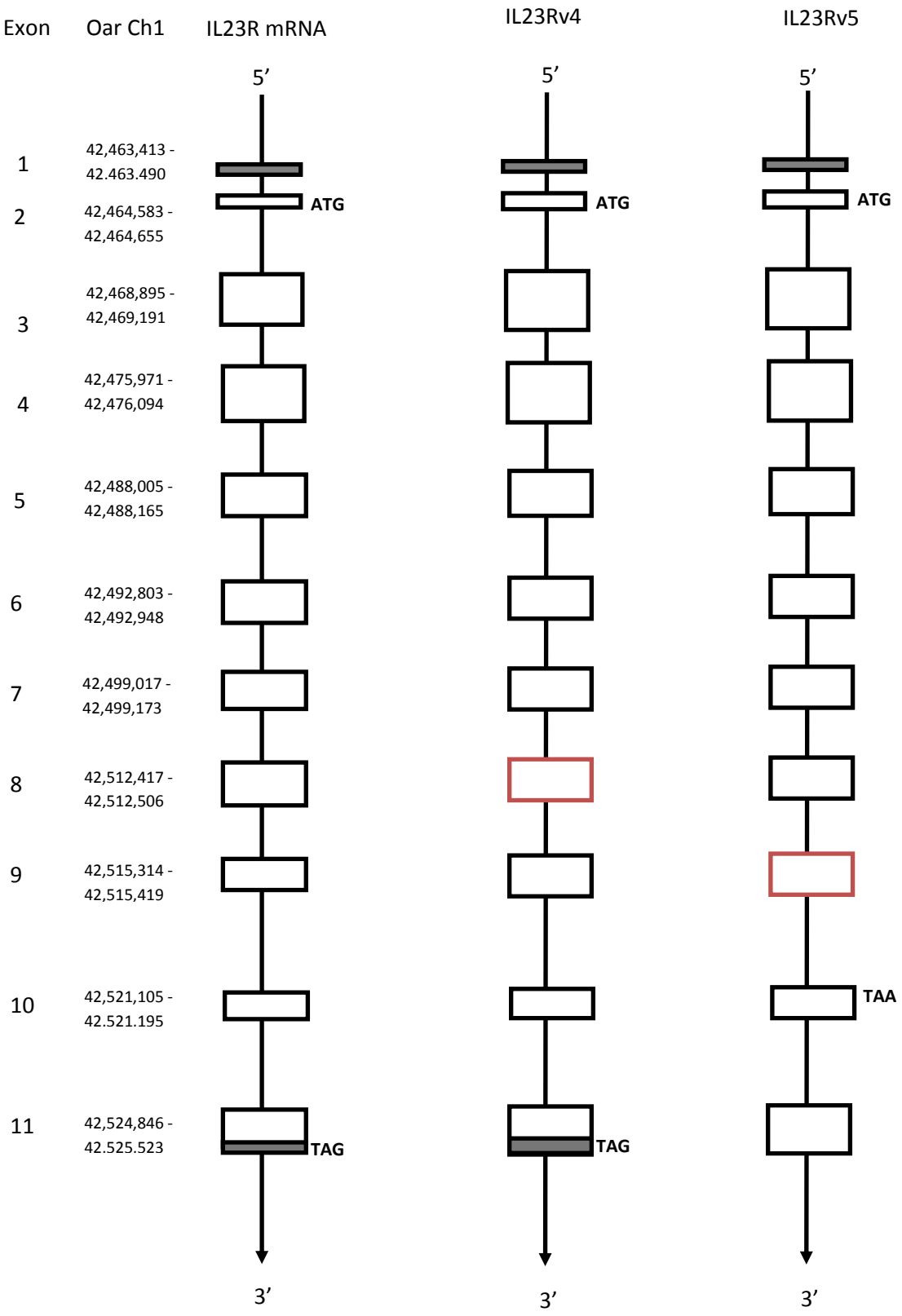


Figure 4.4 cont. Schematic diagrams of the ovine *IL23R* and alternatively spliced transcript variants *(IL23R v4& v5*) identified in this study. Boxes represent individual exons with chromosome location shown. Grey boxes indicate 5' and 3' UTRs with variant regions highlighted in red. The 'ATG' is initial start codon and stop codons are shown for each transcript.

```
V3              TAAGGAAGACGTGAGGCCGATACGAGGGCAGCAGCCTGGCTCTGAGGGGGAATTATGTGC 60
V4              TAAGGAAGACGTGAGGCCGATACGAGGGCAGCAGCCTGGCTCTGAGGGGGAATTATGTGC 60
V2              TAAGGAAGACGTGAGGCCGATACGAGGGCAGCAGCCTGGCTCTGAGGGGGAATTATGTGC 60
V1              TAAGGAAGACGTGAGGCCGATACGAGGGCAGCAGCCTGGCTCTGAGGGGGAATTATGTGC 60
V5              TAAGGAAGACGTGAGGCCGATACGAGGGCAGCAGCCTGGCTCTGAGGGGGAATTATGTGC 60
Full_sequence   TAAGGAAGACGTGAGGCCGATACGAGGGCAGCAGCCTGGCTCTGAGGGGGAATTATGTGC 60
                ************************************************************

V3              TTCAAACAGGTCGTAAGG------------------------------CATGAACCAGGT 90
V4              TTCAAACAGGTCGTAAGG------------------------------CATGAACCAGGT 90
V2              TTCAAACAGGTCGTAAGG------------------------------CATGAACCAGGT 90
V1              TTCAAACAGGTCGTAAGG-AGCGACAAAGAATTTTTTCTGCTTTCAGGCATGAACCAGGT 120
V5              TTCAAACAGGTCGTAAGG------------------------------CATGAACCAGGT 90
Full_sequence   TTCAAACAGGTCGTAAGG------------------------------CATGAACCAGGT 90
                ******************************

V3              CACAATTCATTGGGATGTGGTAATAGCTCTCTACATATTCTTCAGTTGGTGTCATGGAGG 150
V4              CACAATTCATTGGGATGTGGTAATAGCTCTCTACATATTCTTCAGTTGGTGTCATGGAGG 150
V2              CACAATTCATTGGGATGTGGTAATAGCTCTCTACATATTCTTCAGTTGGTGTCATGGAGG 150
V1              CACAATTCATTGGGATGTGGTAATAGCTCTCTACATATTCTTCAGTTGGTGTCATGGAGG 180
V5              CACAATTCATTGGGATGTGGTAATAGCTCTCTACATATTCTTCAGTTGGTGTCATGGAGG 150
Full_sequence   CACAATTCATTGGGATGTGGTAATAGCTCTCTACATATTCTTCAGTTGGTGTCATGGAGG 150
                ************************************************************

V3              GATTACAAATATAAACTGCTCTGGACACATCTGGGTAGAACCTGCCACAATTTTTAAGAT 210
V4              GATTACAAATATAGACTGCTCTGGACACATCTGGGTAGAACCTGCCACAATTTTTAAGAT 210
V2              GATTACAAATATAAACTGCTCTGGACACATCTGGGTAGAACCTGCCACAATTTTTAAGAT 210
V1              GATTACAAATATAAACTGCTCTGGACACATCTGGGTAGAACCTGCCACAATTTTTAAGAT 240
V5              GATTACAAATATAAACTGCTCTGGACACATCTGGGTAGAACCTGCCACAATTTTTAAGAT 210
Full_sequence   GATTACAAATATAAACTGCTCTGGACACATCTGGGTAGAACCTGCCACAATTTTTAAGAT 210
                *************.**********************************************

V3              GGGTATGAATATCTCTATATATTGCCAAGCAGCAATTAGGAACTGCCAACCAAGGAAACT 270
V4              GGGTATGAATATCTCTATATATTGCCAAGCAGCAATTAGGAACTGCCAACCAAGGAAACT 270
V2              GGGTATGAATATCTCTATATATTGCCAAGCAGCAATTAGGAACTGCCAACCAAGGAAACT 270
V1              GGGCATGAATATCTCTATATATTGCCAAGCAGCAATTAGGAACTGCCAACCAAGGAAACT 300
V5              GGGTATGAATATCTCTATATATTGCCAAGCAGCAATTAGGAACTGCCAACCAAGGAAACT 270
Full_sequence   GGGCATGAATATCTCTATATATTGCCAAGCAGCAATTAGGAACTGCCAACCAAGGAAACT 270
                *** ************************************************************

V3              TTATTTTTATAAAAATGGCATCAAAGAAAGATTTCATATCACAAGAATCAATAAAACAAC 330
V4              TTATTTTTATAAAAATGGCATCAAAGAAAGATTTCATATCACAAGAATCAATAAAACAAC 330
V2              TTATTTTTATAAAAATGGCATCAAAGAAAGATTTCATATCACAAGAATCAATAAAACAAC 330
V1              TTATTTTTATAAAAATGGCATCAAAGAAAGATTTCATATCACAAGAATCAATAAAACAAC 360
V5              TTATTTTTATAAAAATGGCATCAAAGAAAGATTTCATATCACAAGAATCAATAAAACAAC 330
Full_sequence   TTATTTTTATAAAAATGGCATCAAAGAAAGATTTCATATCACAAGAATCAATAAAACAAC 330
                ************************************************************

V3              AGCTCGCCTTTGGTATAACAACTTTGTGGAGCCACAAGCCTTTGTGTACTGTACTGCTGA 390
V4              AGCTCGCCTTTGGTATAACAACTTTGTGGAGCCACAAGCCTTTGTGTACTGTACTGCTGA 390
V2              AGCTCGCCTTTGGTATAACAACTTTGTGGAGCCACAAGCCTTTGTGTACTGTACTGCTGA 390
V1              AGCTCGCCTTTGGTATAACAACTTTGTGGAGCCACAAGCCTTTGTGTACTGTACTGCTGA 420
V5              AGCTCGCCTTTGGTATAACAACTTTGTGGAGCCACAAGCCTTTGTGTACTGTACTGCTGA 390
Full_sequence   AGCTCGCCTTTGGTATAACAACTTTGTGGAGCCACAAGCCTTTGTGTACTGTACTGCTGA 390
                ************************************************************

V3              ATGTT-CCAGATATTTTCCAGAGACACTGATTT-GTGGAAAAGACATTTCTTCTGGATAT 448
V4              ATGTT-CCAGATATTTTCCAGAGACACTGATTT-GTGGAAAAGACATTTCTTCTGGATAT 448
V2              ATGTT-CCAGATATTTTCCAGAGACACTGATTT-GTGGAAAAGACATTTCTTCTGGATAT 448
V1              ATGTTCCCAGATATTTTTCAGAGACACTGATTTTGTGGAAAAGACATTTCTTCTGGATAT 480
V5              ATGTT-CCAGATATTTTCCAGAGACACTGATTT-GTGGAAAAGACATTTCTTCTGGATAT 448
Full_sequence   ATGTT-CCAGATATTTTCCAGAGACACTGATTT-GTGGAAAAGACATTTCTTCTGGATAT 448
                ***** *********** *************** **************************

V3              CCACCAGATGTACCTGACAAAGTAACCTGTGTCATTTATGAATATTCAGGCAACATGACT 508
V4              CCACCAGATGTACCTGACAAAGTAACCTGTGTCATTTATGAATATTCAGGCAACATGACT 508
V2              CCACCAGATGTACCTGACAAAGTAACCTGTGTCATTTATGAATATTCAGGCAACATGACT 508
V1              CCACCAGATGTACCTGACAAAGTAACCTGTGTCATTTATGAATATTCAGGCAACATGACT 540
V5              CCACCAGATGTACCTGACAAAGTAACCTGTGTCATTTATGAATATTCAGGCAACATGACT 508
Full_sequence   CCACCAGATGTACCTGACAAAGTAACCTGTGTCATTTATGAATATTCAGGCAACATGACT 508
                ************************************************************

V3              TGTACCTGGAACCCTGGGAGGCCCACCTACATAGACACAAAGTATGTGGTGTACGTGAAG 568
V4              TGTACCTGGAACCCTGGGAGGCCCACCTACATAGACACAAAGTATGTGGTGTACGTGAAG 568
V2              TGTACCTGGAACCCTGGGAGGCCCACCTACATAGACACAAA------------------- 549
V1              TGTACCTGGAACCCTGGGAGGCCCACCTACATAGACACAAAGTATGTGGTGTACGTGAAG 600
V5              TGTACCTGGAACCCTGGGAGGCCCACCTACATAGACACAAAGTATGTGGTGTACGTGAAG 568
Full_sequence   TGTACCTGGAACCCTGGGAGGCCCACCTACATAGACACAAAGTATGTGGTGTACGTGAAG 568
                ****************************************
```

```
V3                AGTTTAGAGACAGAAGAAGAGCAAGAATATCTCACTTCAAGTTACATTAACATCTCCACT 628
V4                AGTTTAGAGACAGAAGAAGAGCAAGAATATCTCACTTCAAGTTACATTAACATCTCCACT 628
V2                  --TTTAGAGACAGAAGAAGAGCAAGAATATCTCACTTCAAGTTACATTAACATCTCCACT 607
V1                AGTTTAGAGACAGAAGAAGAGCAAGAATATCTCACTTCAAGTTACATTAACATCTCCACT 660
V5                AGTTTAGAGACAGAAGAAGAGCAAGAATATCTCACTTCAAGTTACATTAACATCTCCACT 628
Full_sequence     AGTTTAGAGACAGAAGAAGAGCAAGAATATCTCACTTCAAGTTACATTAACATCTCCACT 628
                    ************************************************************

V3                GATTCATTGCAAAAGGGAAAGAAGTATTTGGTTTGGGTCCAAGCTTCAAATGTACTGGGC 688
V4                GATTCATTGCAAAAGGGAAAGAAGTATTTGGTTTGGGTCCAAGCTTCAAATGTACTGGGC 688
V2                GATTCATTGCAAAAGGGAAAGAAGTATTTGGTTTGGGTCCAAGCTTCAAATGTACTGGGC 667
V1                GATTCATTGCAAAAGGGAAAGAAGTATTTGGTTTGGGTCCAAGCTTCAAATGTACTGGGC 720
V5                GATTCATTGCAAAAGGGAAAGAAGTATTTGGTTTGGGTCCAAGCTTCAAATGTACTGGGC 688
Full_sequence     GATTCATTGCAAAAGGGAAAGAAGTATTTGGTTTGGGTCCAAGCTTCAAATGTACTGGGC 688
                  ************************************************************

V3                ACGGAAAAGTCAAAACAACTACAAATTCATCTGGACGATATAGTGATACCTTCTGCATCC 748
V4                ACGGAAAAGTCAAAACAACTACAAATTCATCTGGACGATATAGTGATACCTTCTGCATCC 748
V2                ACGGAAAAGTCAAAACAACTACAAATTCATCTGGACGATATAGTGATACCTTCTGCATCC 727
V1                ACGGAAAAGTCAAAACAACTACAAATTCATCTGGACGATATAGTGATACCTTCTGCATCC 780
V5                ACGGAAAAGTCAAAACAACTACAAATTCATCTGGACGATATAGTGATACCTTCTGCATCC 748
Full_sequence     ACGGAAAAGTCAAAACAACTACAAATTCATCTGGACGATATAGTGATACCTTCTGCATCC 748
                  ************************************************************

V3                ATTATTTCCAGGGCTGAGGATATAAATACTACAGTGTCCAAGACTGTAATCCACTGGAAT 808
V4                ATTATTTCCAGGGCTGAGGATATAAATACTACAGTGTCCAAGACTGTAATCCACTGGAAT 808
V2                ATTATTTCCAGGGCTGAGGATATAAATACTACAGTGTCCAAGACTGTAATCCACTGGAAT 787
V1                ATTATTTCCAGGGCTGAGGATATAAATACTACAGTGTCCAAGACTGTAATCCACTGGAAT 840
V5                ATTATTTCCAGGGCTGAGGATATAAATACTACAGTGTCCAAGACTGTAATCCACTGGAAT 808
Full_sequence     ATTATTTCCAGGGCTGAGGATATAAATACTACAGTGTCCAAGACTGTAATCCACTGGAAT 808
                  ************************************************************

V3                AGTCAAACATCAATTGAAAAAGTTTCCTGTGAAATGAGATACAAAGCTACAACAAACCAA 868
V4                AGTCAAACATCAATTGAAAAAGTTTCCTGTGAAATGAGATACAAAGCTACAACAAACCAA 868
V2                AGTCAAACATCAATTGAAAAAGTTTCCTGTGAAATGAGATACAAAGCTACAACAAACCAA 847
V1                AGTCAAACATCAATTGAAAAAGTTTCCTGCGAAATGAGATACAAAGCTACAACAAACCAA 900
V5                AGTCAAACATCAATTGAAAAAGTTTCCTGCGAAATGAGATACAAAGCTACAACAAACCAA 868
Full_sequence     AGTCAAACATCAATTGAAAAAGTTTCCTGCGAAATGAGATACAAAGCTACAACAAACCAA 868
                  ***************************** ******************************

V3                ACTTGGAACGTAAGCTAAATTCACTTTCCTTTTACATGTTAAAGAATTTGATACCAATTT 928
V4                ACTTGGAACG----------------------------TTAAAGAATTTGATACCAATTT 900
V2                ACTTGGAACACG--------------------------TTAAAGAATTTGATACCAATTT 881
V1                ACTTGGAACG----------------------------TTAAAGAATTTGATACCAATTT 932
V5                ACTTGGAACG----------------------------TTAAAGAATTTGATACCAATTT 900
Full_sequence     ACTTGGAACG----------------------------TTAAAGAATTTGATACCAATTT 900
                  *********.                            ***********************

V3                TACATATGAGCAACAGTCAGAATTCTACTTGCAGCCAAATGCTACATACGTATTTCAAGT 988
V4                TACATATGAGCAACAGTCAGAATTCTACTTGCAGCCAAATGCTACATACGTATTTCAAGT 960
V2                TACATATGAGCAACAGTCAGAATTCTACTTGCAGCCAAATGCTACATACGTATTTCAAGT 941
V1                TACATATGAGCAACAGTCAGAATTCTACTTGCAGCCAAATGCTACATACGTATTTCAAGT 992
V5                TACATATGAGCAACAGTCAGAATTCTACTTGCAGCCAAATGCTACATACGTATTTCAAGT 960
Full_sequence     TACATATGAGCAACAGTCAGAATTCTACTTGCAGCCAAATGCTACATACGTATTTCAAGT 960
                  ************************************************************

V3                GAGATGTCAAGAAACAGGTAAAAGGTACTGGCAGCCCTGGAGTTCACCCTTTTTTCATAA 1048
V4                GAGATGTCAAGAAACAGGTAAAAGGTACTGGCAACCCTGGAGTTCACCCTTTTTTCATAA 1020
V2                GAGATGTCAAGAAACAGGTAAAAGGTACTGGCAACCCTGGAGTTCACCCTTTTTTCATAA 1001
V1                GAGATGTCAAGAAACAGGTAAAAGGTACTGGCAGCCCTGGAGTTCACCCTTTTTTCATAA 1052
V5                GAGATGTCAAGAAACAGGTAAAAGGTACTGGCAGCCCTGGAGTTCACCCTTTTTTCATAA 1020
Full_sequence     GAGATGTCAAGAAACAGGTAAAAGGTACTGGCAGCCCTGGAGTTCACCCTTTTTTCATAA 1020
                  *********************************.**************************


V3                AACTCCTGAAATAGTTCCTCAGGTCACAATGAAATCATTCCAACATGATACTCAGAATTC 1108
V4                AACTCCTGAAATAG-------------------------------------------- 1034
V2                AACTCCTGAAATAGTTCCTCAGGTCACAATGAAATCATTCCAACATGATACTCAGAATTC 1061
V1                AACTCCTGAAATAGTTCCTCAGGTCACAGTGAAATCATTCCAACATGATACTCAGAATTC 1112
V5                AACTCCTGAAATAGTTCCTCAGGTCACAATGAAATCATTCCAACATGATACTCAGAATTC 1080
Full_sequence     AACTCCTGAAATAGTTCCTCAGGTCACAATGAAATCATTCCAACATGATACTCAGAATTC 1080
                  **************

V3                TGGACTTCTAATTGCTTCCATCTTTAAAAAACACCTTACTTCTGACAACGGGAAACAAGA 1168
V4                ----------------------------------------ACAACGGGAAACAAGA 1050
V2                TGGACTTCTAATTGCTTCCATCTTTAAAAAACACCTTACTTCTGACAACGGGAAACAAGA 1121
V1                TGGACTTCTAATTGCTTCCATCTTTAAAAAACACCTTACTTCTGACAACGGGAAACAAGG 1172
V5                TGGACTTCTAATTGCTTCCATCTTTAAAAAACACCTTACTTCT------------------ 1123
Full_sequence     TGGACTTCTAATTGCTTCCATCTTTAAAAAACACCTTACTTCTGACAACGGGAAACAAGA 1140


V3                CATTGGACTTTTATTGGGAATGGTCTTCTTTGCTGTTATGCTGTCAGTTCTATCTTTGTT 1228
V4                CATTGGACTTTTATTGGGAATGGTCTTCTTTGCTGTTATGCTGTCAGTTCTATCTTTGTT 1110
```

```
V2              CATTGGACTTTTATTGGGAATGGTCTTCTTTGCTGTTATGCTGTCAGTTCTATCTTTGTT 1181
V1              CATTGGACTTTTATTGGGAATGGTCTTCTTTGCTGTTATGCTGTCAGTTCTATCTTTGTT 1232
V5              ------------------------------------------------------------
Full_sequence   CATTGGACTTTTATTGGGAATGGTCTTCTTTGCTGTTATGCTGTCAGTTCTATCTTTGTT 1200

V3              TGGGATATTTAACAAATCGCTTCGAACTGGAATTAAAAGAAGAATCTTATTGCTAATACC 1288
V4              TGGGATATTTAACAAATCGCTTCGAACTGGAATTAAAAGAAGAATCTTATTGCTAATACC 1170
V2              TGGGATATTTAACAAATCGCTTCGAACTGGAATTAAAAGAAGAATCTTATTGCTAATACC 1241
V1              TGGGATATTTAACAAATCGCTTCGAACTGGAATTAAAAGAAGAATCTTATTGCTAATACC 1292
V5              ------------------------------GAATTAAAAGAAGAATCTTATTGCTAATACC 1154
Full_sequence   TGGGATATTTAACAAATCGCTTCGAACTGGAATTAAAAGAAGAATCTTATTGCTAATACC 1260
                                               ********************************

V3              AAAATGGCTCTATGAAGATATTCCTAAATTGGAAAACAGTAAAGTTGTGAAAATACTTCA 1348
V4              AAAATGGCTCTATGAAGATATTCCTAAATTGGAAAACAGTAAAGTTGTGAAAATACTTCA 1230
V2              AAAATGGCTCTATGAAGATATTCCTAAATTGGAAAACAGTAAAGTTGTGAAAATACTTCA 1301
V1              AAAATGGCTCTATGAAGATATTCCTAAATTGGAAAACAGTAAAGTTGTGAAAATACTTCA 1352
V5              AAAATGGCTCTATGAAGATATTCCTAAATTGGAAAACAGTAAAGTTGTGAAAATACTTCA 1214
Full_sequence   AAAATGGCTCTATGAAGATATTCCTAAATTGGAAAACAGTAAAGTTGTGAAAATACTTCA 1320
                ************************************************************

V3              GGAAAAAAATGAATTTATGAATAACAATTCCAGTGAACAGGTCCTATATGTTGATCCGGT 1408
V4              GGAAAAAAATGAATTTATGAATAACAATTCCAGTGAACAGGTCCTATATGTTGATCCGGT 1290
V2              GGAAAAAAATGAATTTATGAATAACAATTCCAGTGAACAGGTCCTATATGTTGATCCGGT 1361
V1              GGAAAAAAATGAATTTATGAATAACAATTCCAGTGAACAGGTCCTATATGTTGATCCGGT 1412
V5              GAAAAAAATGAATTTATGAATAACAATTCCAGTGAACAGGTCCTATATGTTGATCCGGT 1274
Full_sequence   GGAAAAAAATGAATTTATGAATAACAATTCCAGTGAACAGGTCCTATATGTTGATCCGGT 1380
                ************************************************************

V3              GATTACAGAGATAGAAATCATTCTCCCAGAAGAAAAACCCATAGGCTACAAGAAAGAAAA 1468
V4              GATTACAGAGATAGAAATCATTCTCCCAGAAGAAAAACCCATAGGCTACAAGAAAGAAAA 1350
V2              GATTACAGAGATAGAAATCATTCTCCCAGAAGAAAAACCCATAGGCTACAAGAAAGAAAA 1421
V1              GATTACAGAGATAGAAATCATTCTCCCAGAAGAAAAACCCATAGGCTACAAGAAAGAAAA 1472
V5              GATTACAGAGATAGAAATCATTCTCCCAGAAGAAAAACCCATAGGCTACAAGAAAGAAAA 1334
Full_sequence   GATTACAGAGATAGAAATCATTCTCCCAGAAGAAAAACCCATAGGCTACAAGAAAGAAAA 1440
                ************************************************************

V3              CAATACAGGATGCCTGGAGAGAAAAGAGAGCCTGGAGAAGTCACTACTCACCAACACTAC 1528
V4              CAATACAGGATGCCTGGAGAGAAAAGAGAGCCTGGAGAAGTCACTACTCACCAACACTAC 1410
V2              CAATACAGGATGCCTGGAGAGAAAAGAGAGCCTGGAGAAGTCACTACTCACCAACACTAC 1481
V1              CCATACAGGATGCCTGGAGAGAAAAGAGAGCCTGGAGAAGTCACTACTCACCAACACTAC 1532
V5              CAATACAGGATGCCTGGAGAGAAAAGAGAGCCTGGAGAAGTCACTACTCACCAACACTAC 1394
Full_sequence   CAATACAGGATGCCTGGAGAGAAAAGAGAGCCTGGAGAAGTCACTACTCACCAACACTAC 1500
                *.************************************************************

V3              AGTTGTATATATTCCTGATCTCAACACTGGGTATAAACCCCAGATTTCAAGTTTTCTCCC 1588
V4              AGTTGTATATATTCCTGATCTCAACACTGGGTATAAACCCCAGATTTCAAGTTTTCTCCC 1470
V2              AGTTGTATATATTCCTGATCTCAACACTGGGTATAAACCCCAGATTTCAAGTTTTCTCCC 1541
V1              AGTTGTATATATTCCTGATCTCAACACTGGGTATAAACCCCAGATTTCAAGTTTTCTCCC 1592
V5              AGTTGTATATATTCCTGATCTCAACACTGGGTATAAACCCCAGATTTCAAGTTTTCTCCC 1454
Full_sequence   AGTTGTATATATTCCTGATCTCAACACTGGGTATAAACCCCAGATTTCAAGTTTTCTCCC 1560
                ************************************************************

V3              TGGGGGAAACCATCTCAGGAATGATGATGAAACAGCTTCCTCAATTCTGGAACCACCAGC 1648
V4              TGGGGGAAACCATCTCAGGAATGATGATGAAACAGCTTCCTCAATTCTGGAACCACCAGC 1530
V2              TGGGGGAAACCATCTCAGGAATGATGATGAAACAGCTTCCTCAATTCTGGAACCACCAGC 1601
V1              TGGGGGAAACCATCTCAGGAATGATGATGAAACAGCTTCCTCAATTCTGGAACCACCAGC 1652
V5              TGGGGGAAACCATCTCAGGAATGATGATGAAACAGCTTCCTCAATTCTGGAACCACCAGC 1514
Full_sequence   TGGGGGAAACCATCTCAGGAATGATGATGAAACAGCTTCCTCAATTCTGGAACCACCAGC 1620
                ************************************************************

V3              TGATTCCTTAAACTTGGGAAACAATGCCAGGTTTAAAAAATATCCTGATTTTGCTTTTTC 1708
V4              TGATTCCTTAAACTTGGGAAACAATGCCAGGTTTAAAAAATATCCTGATTTTGCTTTTTC 1590
V2              TGATTCCTTAAACTTGGGAAACAATGCCAGGTTTAAAAAATATCCTGATTTTGCTTTTTC 1661
V1              TGATTCCTTAAACTTGGGAAACAATGCCAGGTTTAAAAAATATCCTGATTTTGCTTTTTC 1712
V5              TGATTCCTTAAACTTGGGAAACAATGCCAGGTTTAAAAAATATCCTGATTTTGCTTTTTC 1574
Full_sequence    TGATTCCTTAAACTTGGGAAACAATGCCAGGTTTAAAAAATATCCTGATTTTGCTTTTTC 1680
                 ************************************************************

V3              TGTCTCAAGTACAAATTCACTAAGCAACACACTATTTCTTGAAGAATTAAGCCTCATTTT 1768
V4              TGTCTCAAGTACAAATTCACTAAGCAACACACTATTTCTTGAAGAATTAAGCCTCATTTT 1650
V2              TGTCTCAAGTACAAATTCACTAAGCAACACACTATTTCTTGAAGAATTAAGCCTCATTTT 1721
V1              TGTCTCAAGTACAAATTCACTAAGCAACACACTATTTCTTGAAGAATTAAGCCTCATTTT 1772
V5              TGTCTCAAGTACAAATTCACTAAGCAACACACTATTTCTTGAAGAATTAAGCCTCATTTT 1634
Full_sequence   TGTCTCAAGTACAAATTCACTAAGCAACACACTATTTCTTGAAGAATTAAGCCTCATTTT 1740
                ************************************************************

V3              AAATCAAGGAGAATGCAGTCCTCCGGACATGCAAAACTCAATAGAGGGGGAAACTGCCAT 1828
V4              AAATCAAGGAGAATGCAGTCCTCCGGACATGCAAAACTCAATAGAGGGGGAAACTGCCAT 1710
V2              AAATCAAGGAGAATGCAGTCCTCCGGACATGCAAAACTCAATAGAGGGGGAAACTGCCAT 1781
V1              AAATCAAGGAGAATGCAGTCCTCCGGACATGCAAAACTCAATAGAGGGGGAAACTGCCAT 1832
V5              AAATCAAGGAGAATGCAGTCCTCCGGACATGCAAAACTCAATAGAGGGGGAAACTGCCAT 1694
Full_sequence   AAATCAAGGAGAATGCAGTCCTCCGGACATGCAAAACTCAATAGAGGGGGAAACTGCCAT 1800
```

```
                    ************************************************************

V3                  GCTTCTGGAAGATGCATTACTGAATGAAACTATTCCAGAACAAACCCTGCTGCCTGATGA 1888
V4                  GCTTCTGGAAGATGCATTACTGAATGAAACTATTCCAGAACAAACCCTGCTGCCTGATGA 1770
V2                  GCTTCTGGAAGATGCATTACTGAATGAAACTATTCCAGAACAAACCCTGCTGCCTGATGA 1841
V1                  GCTTCTGGAAGATGCATTACTGAATGAAACTATTCCAGAACAAACCCTGCTGCCTGATGA 1892
V5                  GCTTCTGGAAGATGCATTACTGAATGAAACTATTCCAGAACAAACCCTGCTGCCTGATGA 1754
Full_sequence       GCTTCTGGAAGATGCATTACTGAATGAAACTATTCCAGAACAAACCCTGCTGCCTGATGA 1860
                    ************************************************************

V3                  ATTTGTCTCCTGTTTGGGGAGCATGAACAAAGAGTTGCCATCTATTAATTCTTATTTTCC 1948
V4                  ATTTGTCTCCTGTTTGGGGAGCATGAACAAAGAGTTGCCATCTATTAATTCTTATTTTCC 1830
V2                  ATTTGTCTCCTGTTTGGGGAGCATGAACAAAGAGTTGCCATCTATTAATTCTTATTTTCC 1901
V1                  ATTTGTCTCCTGTTTGGGGAGCATGAACAAAGAGTTGCCATCTATTAATTCTTATTTTCC 1952
V5                  ATTTGTCTCCTGTTTGGGGAGCATGAACAAAGAGTTGCCATCTATTAATTCTTATTTTCC 1814
Full_sequence       ATTTGTCTCCTGTTTGGGGAGCATGAACAAAGAGTTGCCATCTATTAATTCTTATTTTCC 1920
                    ************************************************************

V3                  ACAAAATATTTTGGAAAGCCACTTCAATCGATTTCACTCTTAGAAAACTAGATTTGTGTA 2008
V4                  ACAAAATATTTTGGAAAGCCACTTCAATCGATTTCACTCTTAGAAAACTAGATTTGTGTA 1890
V2                  ACAAAATATTTTGGAAAGCCACTTCAATCGATTTCACTCTTAGAAAACTAGATTTGTGTA 1961
V1                  ACAAAATATTTTGGAAAGCCACTTCAATCGATTTCACTCTTAGAAAACTAGATTTGTGTA 2012
V5                  ACAAAATATTTTGGAAAGCCACTTCAATCGATTTCACTCTTAGAAAACTAGATTTGTGTA 1874
Full_sequence       ACAAAATATTTTGGAAAGCCACTTCAATCGATTTCACTCTTAGAAAACTAGATTTGTGTA 1980
                    ************************************************************

V3                  GTCAGAACCTGCCTTGAAAA 2028
V4                  GTCAGAACCTGCCTTGAAAA 1910
V2                  GTCAGAACCTGCCTTGAAAA 1981
V1                  GTCAGAACCTGCCTTGAAAA 2032
V5                  GTCAGAACCTGCCTTGAAAA 1894
Full_sequence       GTCAGAACCTGCCTTGAAAA 2000
                    ********************
```

Figure 4.5. Nucleotide sequence alignment (ClustalW v2.1) showing *IL23R* full length and splice variant (V) sequences. Sections highlighted in yellow indicate variant regions; start codon is underlined.

4.3.2 Ovine IL23R translated protein sequence analysis

Nucleotide sequences of full length and alternatively spliced variant transcripts of ovine *IL23R* were translated into amino acid sequences and were then aligned using ClustalW v2.1 to identify changes to protein structure that occur as a result of transcriptional variation (figure 4.6). An alignment of human IL23R and ovine IL23R amino acid sequences identified conserved structural regions (Di Meglio *et al.*, 2011) of the IL23R protein and the protective R381Q variant found in humans, conserved regions are highlighted and described in 'KEY'. These alignments were used to form predicted protein structure maps and predicted biological functions based on human IL23R variant protein sequences (figure 4.7).

KEY:

Required for receptor activation: **'WSXWS' motif**

Transmembrane domain

R381Q variant site

```
V1              MNQVTIHWDVVIALYIFFSWCHGGITNINCSGHIWVEPATIFKMGMNISIYCQAAIRNCQ 60
V4              MNQVTIHWDVVIALYIFFSWCHGGITNIDCSGHIWVEPATIFKMGMNISIYCQAAIRNCQ 60
Full            MNQVTIHWDVVIALYIFFSWCHGGITNINCSGHIWVEPATIFKMGMNISIYCQAAIRNCQ 60
V5              MNQVTIHWDVVIALYIFFSWCHGGITNINCSGHIWVEPATIFKMGMNISIYCQAAIRNCQ 60
V3              MNQVTIHWDVVIALYIFFSWCHGGITNINCSGHIWVEPATIFKMGMNISIYCQAAIRNCQ 60
V2              MNQVTIHWDVVIALYIFFSWCHGGITNINCSGHIWVEPATIFKMGMNISIYCQAAIRNCQ 60
                 ***************************:******************************

V1              PRKLYFYKNGIKERFHITRINKTTARLWYNNFVEPQAFVYCTAECSRYFSETLICGKDIS 120
V4              PRKLYFYKNGIKERFHITRINKTTARLWYNNFVEPQAFVYCTAECSRYFPETLICGKDIS 120
Full            PRKLYFYKNGIKERFHITRINKTTARLWYNNFVEPQAFVYCTAECSRYFPETLICGKDIS 120
V5              PRKLYFYKNGIKERFHITRINKTTARLWYNNFVEPQAFVYCTAECSRYFPETLICGKDIS 120
V3              PRKLYFYKNGIKERFHITRINKTTARLWYNNFVEPQAFVYCTAECSRYFPETLICGKDIS 120
V2              PRKLYFYKNGIKERFHITRINKTTARLWYNNFVEPQAFVYCTAECSRYFPETLICGKDIS 120
                 ************************************************.**********

V1              SGYPPDVPDKVTCVIYEYSGNMTCTWNPGRPTYIDTKYVVYVKSLETEEEQEYLTSSYIN 180
V4              SGYPPDVPDKVTCVIYEYSGNMTCTWNPGRPTYIDTKYVVYVKSLETEEEQEYLTSSYIN 180
Full            SGYPPDVPDKVTCVIYEYSGNMTCTWNPGRPTYIDTKYVVYVKSLETEEEQEYLTSSYIN 180
V5              SGYPPDVPDKVTCVIYEYSGNMTCTWNPGRPTYIDTKYVVYVKSLETEEEQEYLTSSYIN 180
V3              SGYPPDVPDKVTCVIYEYSGNMTCTWNPGRPTYIDTKYVVYVKSLETEEEQEYLTSSYIN 180
V2              SGYPPDVPDKVTCVIYEYSGNMTCTWNPGRPTYIDTN-------LETEEEQEYLTSSYIN 173
                 ***********************************:       ****************

V1              ISTDSLQKGKKYLVWVQASNVLGTEKSKQLQIHLDDIVIPSASIISRAEDINTTVSKTVI 240
V4              ISTDSLQKGKKYLVWVQASNVLGTEKSKQLQIHLDDIVIPSASIISRAEDINTTVSKTVI 240
Full            ISTDSLQKGKKYLVWVQASNVLGTEKSKQLQIHLDDIVIPSASIISRAEDINTTVSKTVI 240
V5              ISTDSLQKGKKYLVWVQASNVLGTEKSKQLQIHLDDIVIPSASIISRAEDINTTVSKTVI 240
V3              ISTDSLQKGKKYLVWVQASNVLGTEKSKQLQIHLDDIVIPSASIISRAEDINTTVSKTVI 240
V2              ISTDSLQKGKKYLVWVQASNVLGTEKSKQLQIHLDDIVIPSASIISRAEDINTTVSKTVI 233
                 ************************************************************

V1              HWNSQTSIEKVSCEMRYKATTNQTWNVKEFDTNFTYEQQSEFYLQPNATYVFQVRCQETG 300
V4              HWNSQTSIEKVSCEMRYKATTNQTWNVKEFDTNFTYEQQSEFYLQPNATYVFQVRCQETG 300
Full            HWNSQTSIEKVSCEMRYKATTNQTWNVKEFDTNFTYEQQSEFYLQPNATYVFQVRCQETG 300
V5              HWNSQTSIEKVSCEMRYKATTNQTWNVKEFDTNFTYEQQSEFYLQPNATYVFQVRCQETG 300
V3              HWNSQTSIEKVSCEMRYKATTNQTWNVS------------------------------- 268
V2              HWNSQTSIEKVSCEMRYKATTNQTWNTLKNLIPILHMSNS-----QNST------CSQML 282
                 *************************. .

V1              KRYWQPWSSPFFHKTPEIVPQVTVKSFQHDTQNSGLLIASIFKKHLTSDNGKQDIGLLLG 360
V4              KRYWQPWSSPFFHKTPEID-----------------------------NGKQDIGLLLG 330
Full            KRYWQPWSSPFFHKTPEIVPQVTMKSFQHDTQNSGLLIASIFKKHLTSDNGKQDIGLLLG 360
V5              KRYWQPWSSPFFHKTPEIVPQVTMKSFQHDTQNSGLLIASIFKKHLTSELK--------- 351
V3              ------------------------------------------------------------
V2              HTYFK-------------------------------------------------------- 287


V1              MVFFAVMLSVLSLFGIFNKSLRTGIKRRILLLIPKWLYEDIPKLENSKVVKILQEKNEFM 420
V4              MVFFAVMLSVLSLFGIFNKSLRTGIKRRILLLIPKWLYEDIPKLENSKVVKILQEKNEFM 390
Full            MVFFAVMLSVLSLFGIFNKSLRTGIKRRILLLIPKWLYEDIPKLENSKVVKILQEKNEFM 420
V5              ---------------------------------------------EESYC----------- 356
V3              ------------------------------------------------------------
V2              ------------------------------------------------------------
```

```
V1              NNNSSEQVLYVDPVITEIEIILPEEKPIGYKKENNTGCLERKESLEKSLLTNTTVVYIPD 480
V4              NNNSSEQVLYVDPVITEIEIILPEEKPIGYKKENNTGCLERKESLEKSLLTNTTVVYIPD 450
Full            NNNSSEQVLYVDPVITEIEIILPEEKPIGYKKENNTGCLERKESLEKSLLTNTTVVYIPD 480
V5              ------------------------------------------------------------
V3              ------------------------------------------------------------
V2              ------------------------------------------------------------

V1              LNTGYKPQISSFLPGGNHLRNDDETASSILEPPADSLNLGNNARFKKYPDFAFSVSSTNS 540
V4              LNTGYKPQISSFLPGGNHLRNDDETASSILEPPADSLNLGNNARFKKYPDFAFSVSSTNS 510
Full            LNTGYKPQISSFLPGGNHLRNDDETASSILEPPADSLNLGNNARFKKYPDFAFSVSSTNS 540
V5              ------------------------------------------------------------
V3              ------------------------------------------------------------
V2              ------------------------------------------------------------

V1              LSNTLFLEELSLILNQGECSPPDMQNSIEGETAMLLEDALLNETIPEQTLLPDEFVSCLG 600
V4              LSNTLFLEELSLILNQGECSPPDMQNSIEGETAMLLEDALLNETIPEQTLLPDEFVSCLG 570
Full            LSNTLFLEELSLILNQGECSPPDMQNSIEGETAMLLEDALLNETIPEQTLLPDEFVSCLG 600
V5              ------------------------------------------------------------
V3              ------------------------------------------------------------
V2              ------------------------------------------------------------

V1              SMNKELPSINSYFPQNILESHFNRFHS 627
V4              SMNKELPSINSYFPQNILESHFNRFHS 597
Full            SMNKELPSINSYFPQNILESHFNRFHS 627
V5              ---------------------------
V3              ---------------------------
V2              ---------------------------
```

Figure 4.6 Predicted amino acid sequence alignment of IL23R full length and IL23Rsplicevariant sequences (ClustalW v2.1) with highlighted conserved sequence based on annotated human protein sequence.

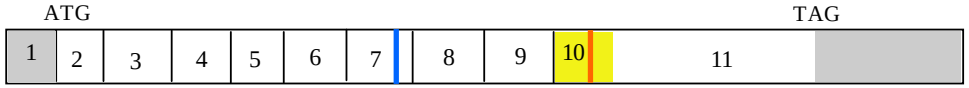
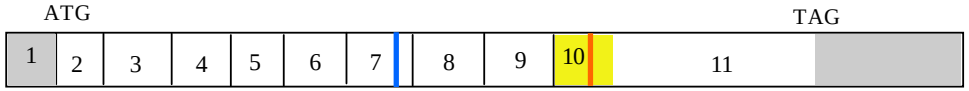
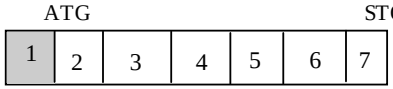
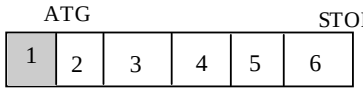
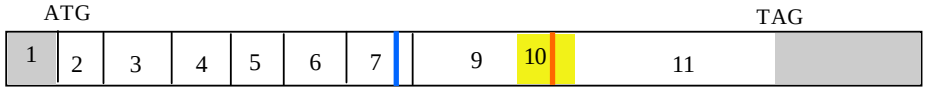
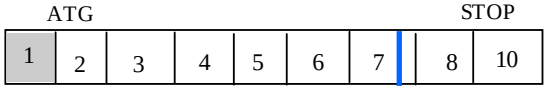
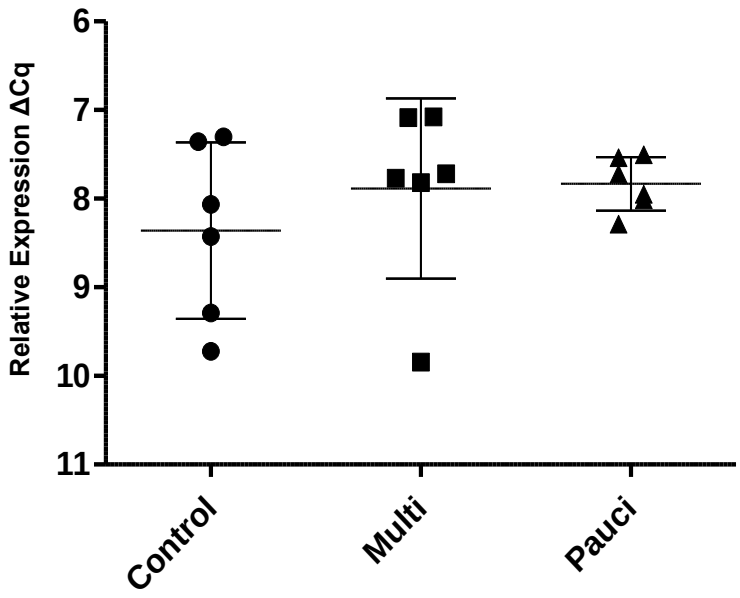
	Predicted protein expression	Amino acids	Predicted Biological Function
Full IL23R		627	Wild type IL23R
IL23Rv1		627	Same as wild type IL23R
IL23Rv2		287	Short peptide
IL23Rv3		268	Short peptide
IL23Rv4		597	Truncated extracellular domain
IL23Rv5		356	Soluble form

Figure 4.7 Predicted protein expression pattern from *IL23R* splice variants identified in this study. Each box represents the protein encoded by each exon. Grey boxes indicate untranslated regions. The 5' UTR is located in exon 1 and 5' end of exon 2. The 3' UTR is located at the 3' end of exon 11. The extracellular region spans exons 2-9 and contains the 'WSxWS' motif, in blue. The transmembrane region is encoded by exon 10, in yellow; this region contains the R318Q variant, in red. The 'ATG' indicates the initial start codon and 'STOP' indicates stop codons encoded by spice variants that differ from wild type *IL23R*. Annotations based on comparison of ovine sequence to annotated human sequence (Kan *et al.*, 2008).

4.3.3 Expression levels of *IL23R*and *IL23R* transcript variants

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for *IL23R* and individual *IL23R* transcript variants. Statistical analysis was carried out on genes where results were within a measureable range, with relative expression (ΔCq), fold change between groups and P-values shown in figure 4.8. Expression levels of total (primer in an area of gene that includes all *IL23R* transcripts) *IL23R*were measured; however transcript variants (*IL23Rv1 – v5*) were found to have expression levels >30 cycles and were considered above the level of expression for accurate measurement. Positive controls were carried out for each of the *IL23R* transcript variants as described in section 2.6.3. Melt curve and cycling analysis for *IL23Rv1*positive control is shown in figure 4.9. Positive control assays for *IL23Rv2 – v5*can be found in appendix A1-A4. Results show no significant difference of *IL23R* gene expression between multibacillary, paucibacillary and control groups.

A)



B)

	Fold Change	P-value
M vs P	0.98	0.99
M vs C	1.32	0.6
P vs C	1.35	0.54

Figure 4.8 *IL23R* RT-qPCR. A) Relative expression of total *IL23R*transcripts with B) fold change values and P-values of comparisons between multibacillary (M), paucibacillary (P) and control groups (C).

A)

B)

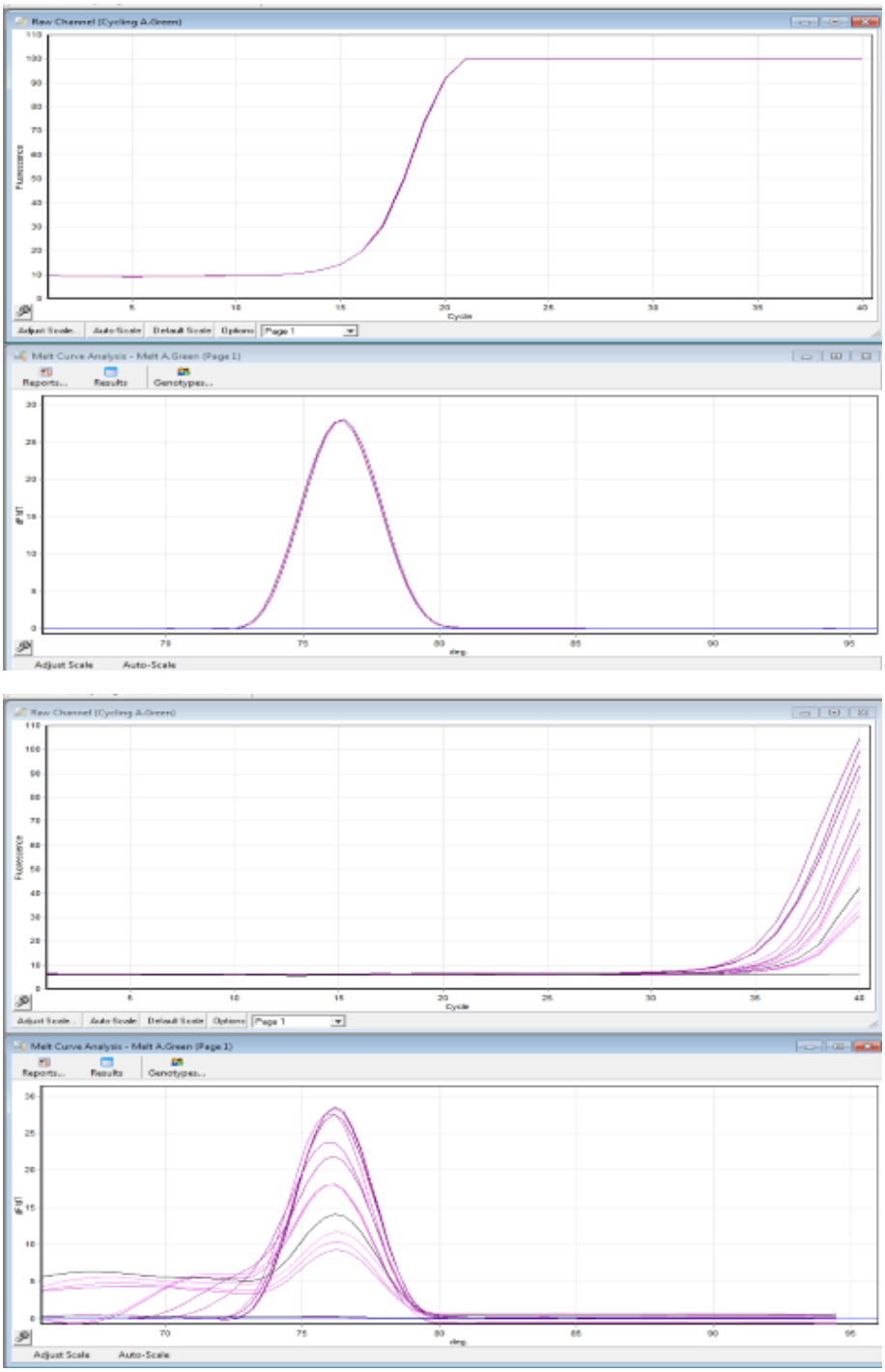


Figure 4.9 Positive control RT-qPCR assay for IL23Rv1.A) Cycling and melt curve analysis of pooled cDNA from paraTB infected and control animals spiked with *IL23Rv1* PCR product B) Cycling and melt curve analysis of dilution series (1:4-1:128) pooled cDNA from paraTB infected animals.

4.4 Ovine *IL12RB1* gene

4.4.1 Identifying *IL12RB1* transcript variants

Primers described in table 2.3 were designed to identify and amplify the ovine *IL12RB1* gene.

The ovine IL12RB1 gene is found on chromosome 5 and consists of 16 exons. Initially, when amplifying short overlapping sections of the ovine *IL12RB1*, transcript variants were identified in the extracellular region of the gene. Nested PCR was carried out on this region (exons 1-9) to amplify and confirm individual variants. Subsequently, full length transcripts were sequenced using nested PCR to obtain full length and variant transcripts including the 3' end of coding region. At least 6 clones, from separate RT-PCR reactions, were sequenced from each animal to give a representative sample of the transcripts present within individual sheep. Sequencing of the intracellular region of *IL12RB1* (exons 10-16) did not identify any variation. Transcript variants of the extracellular region were identified from several clones in both paucibacillary and multibacillary animals as summarised below; these are shown in figure 4.11. Full length ovine *IL12RB1* sequence aligned with predicted ovine *IL12RB1* sequence (OARv3.1) is shown in figure 4.10. Nucleotide and predicted amino acid sequence alignments of *IL12RB1* transcript variants can be found in appendix A5.

Cloning and sequencing identified 4 transcript variants of ovine *IL12RB1*. The full length*IL12RB1* gene is on the plus strand of chromosome 5 and is predicted to encode 730 amino acids (OARv3.1: NC_019462.1). Differences between full length *IL12RB1* (LN878970.1) and identified transcript variants are described below;

IL12RB1v1 (LN878971.1) contains a deletion spanning from the 3' end of exon 5 to exon 8 (Chr5: 4,857,176 – 4,859,356) resulting in a change of protein frame consequently leading to a premature stop codon at position 185 within the protein.

IL12RB1v2 (LN878972.1) contains a deletion spanning from the 3' end of exon 1 to the 5' end of exon 7 (Chr5: 4,853,183 – 4,858,549). This 651bp deletion is predicted to result in a 217 amino acid deletion at position 16-233 within the protein resulting

in a protein sequence of 515 amino acids. This region encodes part of the extracellular region of the protein.

IL12RB1v3 (LN878973.1) contains a deletion spanning from the 3’ end of exon 3 to the 5’ end of exon 7 (Chr5: 4,855,325 – 4,858,535). This 528bp deletion results in 176aa deletion at position 51-227 within the protein. This region encodes part of the extracellular region of the protein.

IL12RB1v4 (LN878974.1) contains a deletion spanning from the 3’ end of exon 4 to the 5’ end of exon 9 (Chr5: 4,855,845 – 4,860,357) resulting in a premature stop codon at position 163 within the predicted protein sequence. This region encodes part of the extracellular region of the protein.

```
OARV3.1           GGTGCCTAGGGCTTCACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGG        60
Ovine_IL12RB1     --------------TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGG        46
                                ********************************************

OARV3.1           ACCGGATGGGACAATGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGAC        120
Ovine_IL12RB1     ACCGGATGGGACAATGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGAC        106
                  ************************************************************

OARV3.1           AGGGCGCTGAAGCCTGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATG        180
Ovine_IL12RB1     AGGGCGCTGAAGCCTGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATG        166
                  ************************************************************

OARV3.1           CAGACTCAGGCTCAGCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCG        240
Ovine_IL12RB1     CAGACTCAGGCTCAGCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCG        226
                  ************************************************************

OARV3.1           CTGGTTATGAATGTTCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCA        300
Ovine_IL12RB1     CTGGTTATGAATGTTCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCA        286
                  ************************************************************

OARV3.1           GATGCTGCCTTCAGTCTGGCCGCTGCTGCTTCTTTGCCACAGGCTCAGCCACCAAGCTGC        360
Ovine_IL12RB1     GATGCTGCCTCCAGTCTGGCCGCTGCTGCTTCTTTGCCACAGGCTCAGCCACCAAGCTGC        346
                  ********** ***************************************************

OARV3.1           AGTTCTCCGACCAGGATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCC        420
Ovine_IL12RB1     AGTTCTCCGACCAGGATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCC        406
                  ************************************************************

OARV3.1           GGGCCGCCAACCGGACAGAGAAGTCCCCCAATGTTACTCTGAACCTCTACAGCTCGGTTA        480
Ovine_IL12RB1     GGGCCGCCAACCGGACAGAGAAGTCCCCCAATGTTACTCTGAACCTCTACAGCTCGGTTA        466
                  ************************************************************

OARV3.1           AATACGACCCTCCCCCAGGAAACATCAAGGTGTCCAAGTCAGCAGGGAAGCTGCGTATGG        540
Ovine_IL12RB1     AATACGACCCTCCCCCAGGAAACATCAAGGTGTCCAAGTCAGCAGGGAAGCTGCGTATGG        526
                  ************************************************************

OARV3.1           AGTGGGAGACCCCAGCCCGCCAGGATGGTGCCGAGGTACAGTTCCGGCACCGCACACCTG        600
Ovine_IL12RB1     AGTGGGAGACCCCAGCCCGCCAGGATGGTGCCGAGGTACAGTTCCGGCACCGCACACCTG        586
                  ************************************************************

OARV3.1           GCAGCCCGTGGAAGCTGGGTGACTGTGGACGTCAGGATGATGCTGGCTTCGAGTCATGCC        660
Ovine_IL12RB1     GCAGCCCGTGGAAGCTGGGTGACTGTGGACGTCAGGATGATGCTGGCTTCGAGTCATGCC        646
                  ************************************************************

OARV3.1           TCTGCCCCTTGGAGATGGACATGGCCCAGGAATTCCAGCTGCGGCGGCGTCTGGGGCCAG        720
Ovine_IL12RB1     TCTGCCCCTTGGAGATGGACATGGCCCAGGAATTCCAGCTGCGGCGGCGTCTGGGGCCAG        706
                  ************************************************************

OARV3.1           GGGTCCCCGGAGGTCCCTGGAGCAGCTGGAGCAGCCCTGTGTGCATCCCCCCTGAAACCC        780
Ovine_IL12RB1     GGGTCCCCGGAGGTCCCTGGAGCAGCTGGAGCAGCCCTGTGTGCATCCCCCCTGAAACCC        766
                  ************************************************************

OARV3.1           CTCCACAGGCTGAGGTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGG        840
Ovine_IL12RB1     CTCCACAGGCTGAGGTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGG        826
                  ************************************************************

OARV3.1           TGGCCTTACATGAGCAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACC        900
Ovine_IL12RB1     TGGCCTTACATGAGCAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACC        886
                  ************************************************************

OARV3.1           CTGGCATGGAGGTGACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCA        960
Ovine_IL12RB1     CTGGCATGGAGGTGACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCA        946
                  ************************************************************

OARV3.1           GGGCCAAGAAGACTCTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATC       1020
Ovine_IL12RB1     GGGCCAAGAAGACTCTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATC       1006
                  ************************************************************

OARV3.1           TGACCATCGTTTCCCAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTG       1080
Ovine_IL12RB1     TGACCATCGTTTCCCAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTG       1066
                  ************************************************************

OARV3.1           CCTGCATCCACTCAGAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCA       1140
Ovine_IL12RB1     CCTGCATCCACTCAGAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCA       1126
                  ************************************************************

OARV3.1           TGCATTGGCCAGCCCGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCC       1200
Ovine_IL12RB1     TGCATTGGCCAGCCCGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCC       1186
                  ************************************************************

OARV3.1           AGGAGGAGAACCTTGCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAA       1260
Ovine_IL12RB1     AGGAGGAGAACCTTGCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAA       1246
                  ************************************************************

OARV3.1           TGGCAACTCACAGCTGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACA       1320
Ovine_IL12RB1     TGGCAACTCACAGCTGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACA       1306
                  ************************************************************

OARV3.1           TCGCCATCTTTGCCTCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCA       1380
Ovine_IL12RB1     TCGCCATCTTTGCCTCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCA       1366
                  ************************************************************
```

```
OARV3.1           CCTACCATTTTGGAGGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGA       1440
Ovine_IL12RB1     CCTACCATTTTGGAGGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGA       1426
                  ************************************************************

OARV3.1           AGCTCAGCCAGGATTCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTG       1500
Ovine_IL12RB1     AGCTCAGCCAGGATTCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTG       1486
                  ************************************************************

OARV3.1           GCGTCCTGAAGGAGTACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGC       1560
Ovine_IL12RB1     GCGTCCTGAAGGAGTACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGC       1546
                  ************************************************************

OARV3.1           TGCACGTGAAGCCCACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTT       1620
Ovine_IL12RB1     TGCACGTGAAGCCCACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTT       1606
                  ************************************************************

OARV3.1           ACAAGGTGCAGGTTCGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGC       1680
Ovine_IL12RB1     ACAAGGTGCAGGTTCGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGC       1666
                  ************************************************************

OARV3.1           ATTTCACCATCGAAGTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTT       1740
Ovine_IL12RB1     ATTTCACCATCGAAGTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTT       1726
                  ************************************************************

OARV3.1           TGGGGAGTTTCGTGAGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGG       1800
Ovine_IL12RB1     TGGGGAGTTTCGTGAGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGG       1786
                  ************************************************************

OARV3.1           CTGTAAGGCACCTGTGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCT       1860
Ovine_IL12RB1     CTGTAAGGCACCTGTGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCT       1846
                  ************************************************************

OARV3.1           CTGGCAGCCAGGGGAAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGG       1920
Ovine_IL12RB1     CTGGCAGCCAGGGGAAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGG       1906
                  ************************************************************

OARV3.1           TGTCTCCACACGAGGCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACA       1980
Ovine_IL12RB1     TGTCTCCACACGAGGCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACA       1966
                  ************************************************************

OARV3.1           TGGGCACACTTGGGACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAG       2040
Ovine_IL12RB1     TGGGCACACTTGGGACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAG       2026
                  ************************************************************

OARV3.1           CCCCGGACACAGAGCTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTG       2100
Ovine_IL12RB1     CCCCGGACACAGAGCTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTG       2086
                  ************************************************************

OARV3.1           GGGCTCTGAGGCCTGGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTA       2160
Ovine_IL12RB1     GGGCTCTGAGGCCTGGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTA       2146
                  ************************************************************

OARV3.1           GACCCCCACTGCTGCTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAG       2220
Ovine_IL12RB1     GACCCCCACTGCTGCTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAG       2206
                  ************************************************************

OARV3.1           AAGCAGGAACAGCTGCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGG       2280
Ovine_IL12RB1     AAGCAGGAACAGCTGCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGG       2266
                  ************************************************************

OARV3.1           CTGTTGTTCACTGCACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCG       2340
Ovine_IL12RB1     CTGTTGTTCACTGCACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCG       2326
                  ************************************************************

OARV3.1           ATATGGCTGTCCTATTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCTAGTG       2400
Ovine_IL12RB1     ATATGGCTGTCCTATTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA----       2382
                  ***************************************************:

OARV3.1           GCTGGCCTGGGGCCACACGACACACGCAAG                                     2430
Ovine_IL12RB1     -----------------------------                                      2382
```

Figure 4.10 Nucleotide sequence alignment of full length ovine *IL12RB1* sequence and predicted *IL12RB1* ovine sequence (ClustalW v2.1) based on sheep genome OARv3.1. Start and stop codons are underlined.

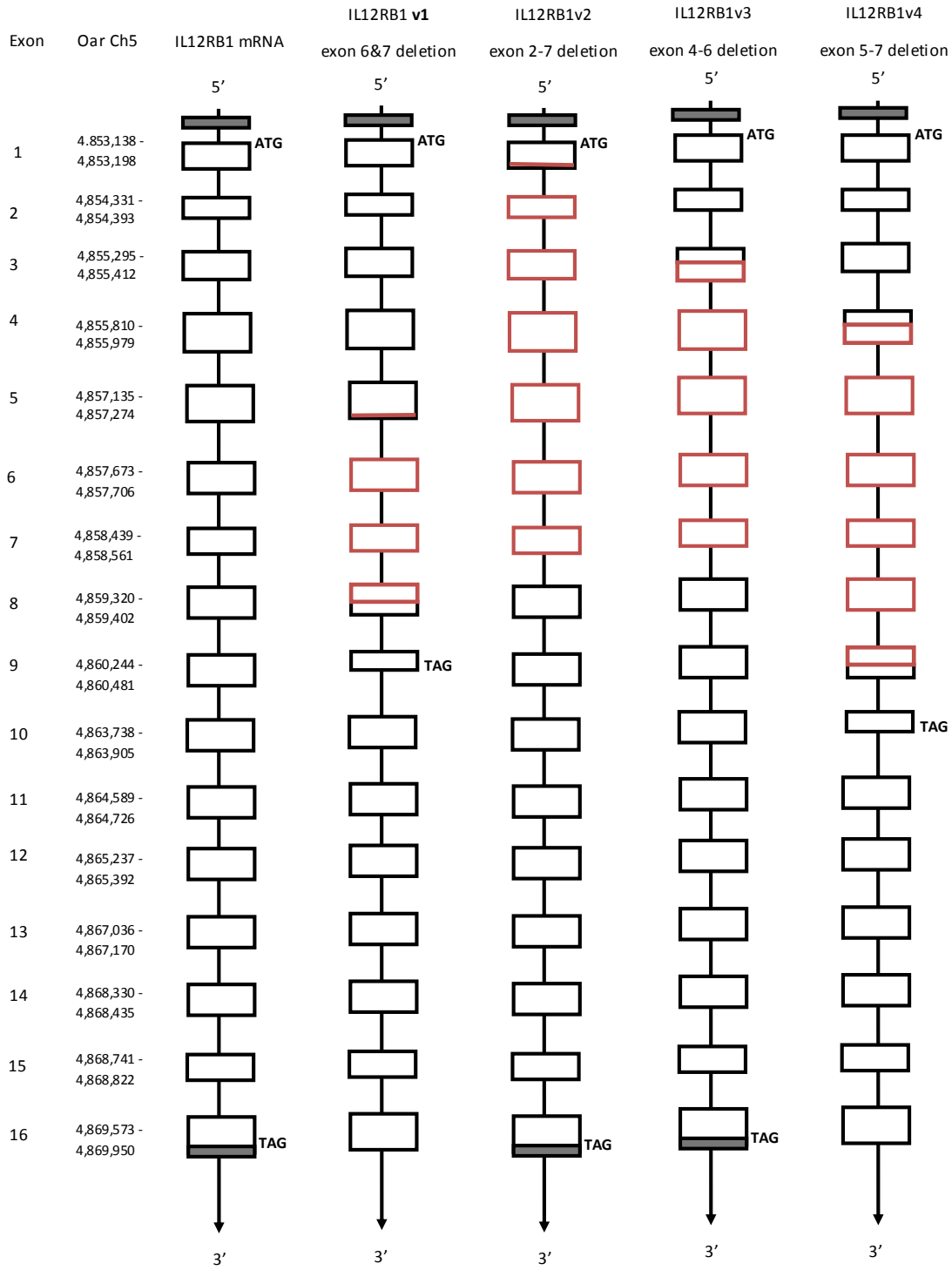


Figure 4.11 Schematic diagrams of the ovine *IL12RB1* gene and splice variants (*IL12RB1v1-4*). Boxes represent individual exons with chromosome location shown. Grey boxes indicate UTRs with deleted variant regions highlighted in red. The 'ATG' is initial start codon and stop codons are shown for each transcript.

4.4.2 Ovine IL12RB1 translated protein sequence analysis

Nucleotide sequences of full length and alternatively spliced variant transcripts of ovine *IL12RB1* were translated into amino acid sequences and aligned using ClustalW v2.1 to identify changes to protein structure that occur as a result of splice variation. Protein alignment of full length *IL12RB1* against sheep genome assembly OARv3.1 is shown in figure 4.12. Alignment of human IL12RB1 and ovine IL12RB1 amino acid sequences identified conserved structural regions of the IL12RB1 protein. Conserved regions are highlighted and described in 'KEY'. This alignment can be found in appendix A8. Alignments were used to form predicted protein structure maps and predict biological functions based on human IL12RB1 variant protein sequences figure 4.13.

KEY:

Required for receptor activation: **'WSXWS' motif**

Transmembrane domain

```
IL12RB1         MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
OARV3.1         MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
                ************************************************************

IL12RB1         YECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
OARV3.1         YECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
                ************************************************************

IL12RB1         ANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRTPGS 180
OARV3.1         ANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRTPGS 180
                ************************************************************

IL12RB1         PWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPETPP 240
OARV3.1         PWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPETPP 240
                ************************************************************

IL12RB1         QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 300
OARV3.1         QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 300
                ************************************************************

IL12RB1         KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 360
OARV3.1         KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 360
                ************************************************************

IL12RB1         WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 420
OARV3.1         WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 420
                ************************************************************

IL12RB1         IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 480
OARV3.1         IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 480
                ************************************************************

IL12RB1         LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 540
OARV3.1         LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 540
                ************************************************************

IL12RB1         TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 600
OARV3.1         TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 600
                ************************************************************

IL12RB1         SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 660
OARV3.1         SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 660
                ************************************************************

IL12RB1         DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 720
OARV3.1         DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 720
                ************************************************************

IL12RB1         GTAASSYRED  730
OARv3.1         GTAASSYRED  730
                **********
```

Figure. 4.12 Predicted amino acid sequence alignment of ovine IL12RB1 full length and sequence from sheep genome assembly OARv3.1 (ClustalW v2.1) with highlighted conserved sequence based on annotated human and mouse protein sequence.

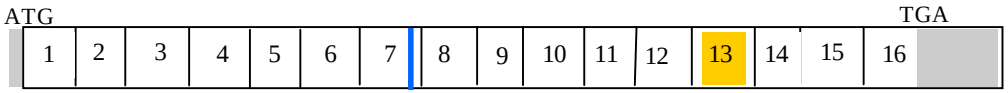
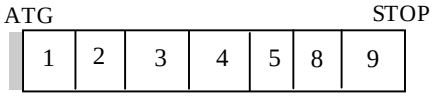
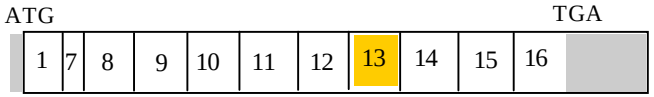
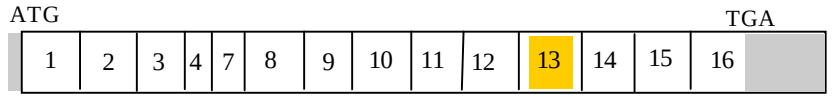
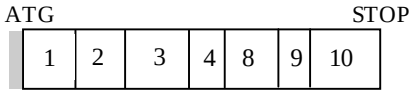
	Predicted protein expression	Amino acids	Predicted Biological Function
Full IL12RB1		730	Wild type IL12RB1
IL12RB1v1		185	Short peptide
IL12RB1v2		515	Truncated extracellular domain
IL12RB1v3		555	Truncated extracellular domain
IL12RB1v4		163	Short peptide

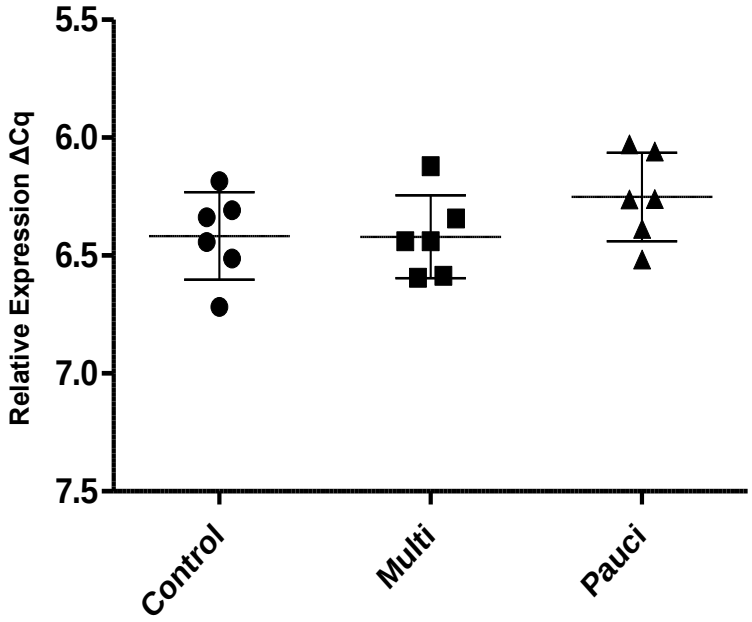
Figure 4.13 Predicted protein expression pattern from *IL12RB1* splice variants identified in this study. Each box represents the protein encoded by each exon. Grey boxes indicate UTRs. The 5' UTR is located within the 5' end of exon 1. The 3' UTR is located at the 3' end of exon 16. The extracellular region spans exons 1-13 and contains the 'WSxWS' motif, in blue. The transmembrane region is encoded by exon 14, in yellow. The 'ATG' indicates the initial start codon and 'STOP' indicates stop codons encoded by spice variants that differ from full length (wild-type) *IL12RB1*. Annotations based on comparison of ovine sequence to annotated human sequence.

4.4.3 Expression levels of *IL12RB1*and *IL12RB1* transcript variants

RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for total (primer in an area of gene that includes all *IL12RB1* transcripts) *IL12RB1*and individual*IL12RB1* transcript variants. Statistical analysis was carried out on genes where results were within a measureable range, with relative expression (ΔCq), fold change between groups and P-values shown. Expression levels of total *IL12RB1*, *IL12RB1v3* and *IL12RB1v4* were measured (figures 4.14, 4.15, 4.16); transcript variants *IL12RB1v1*and *v2*were detected only after >30 cycles and so expression levels were too low for accurate measurement. Positive controls were carried out for *IL12RB1v1* and *v2*as described in section 2.3.6. Melt curve and cycle analysis for each of these assays can be found in appendix A6 and A7.

The results show no significant difference in gene expression between groups for total *IL12RB1*and *IL12RB1v4*; however a 2.1 fold increase in *IL12RB1*v*3* expression was found in multibacillary sheep compared to paucibacillary, with a significant P-value of 0.04.

A)

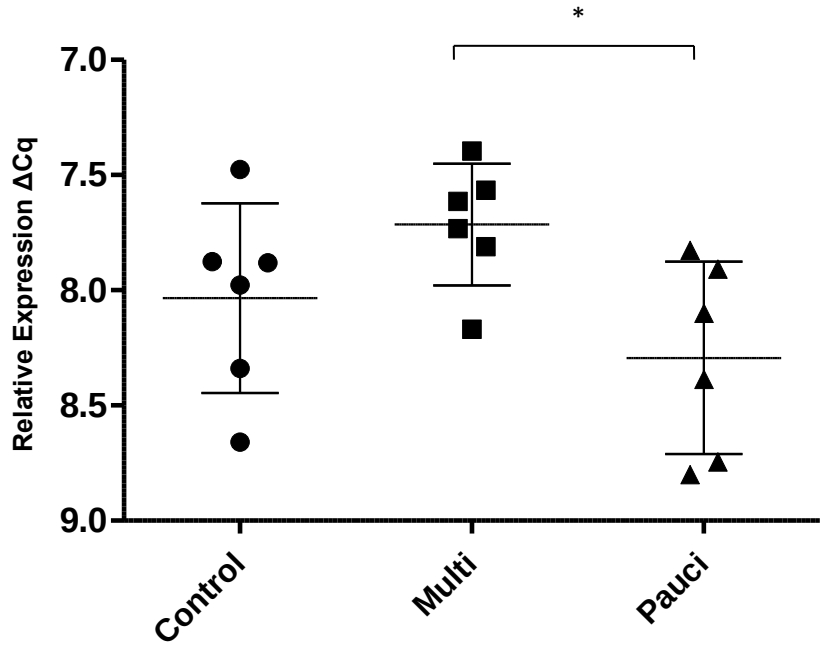


B)

	Fold Change	P-value
M vs P	0.62	0.28
M vs C	0.97	0.99
P vs C	1.57	0.28

Figure4.14*IL12RB1* RT-qPCR. A) Relative expression of total *IL12RB1* with B) fold change and P-values for comparisons between multibacillary(M), paucibacillary (P) and control (C) groups.

A)

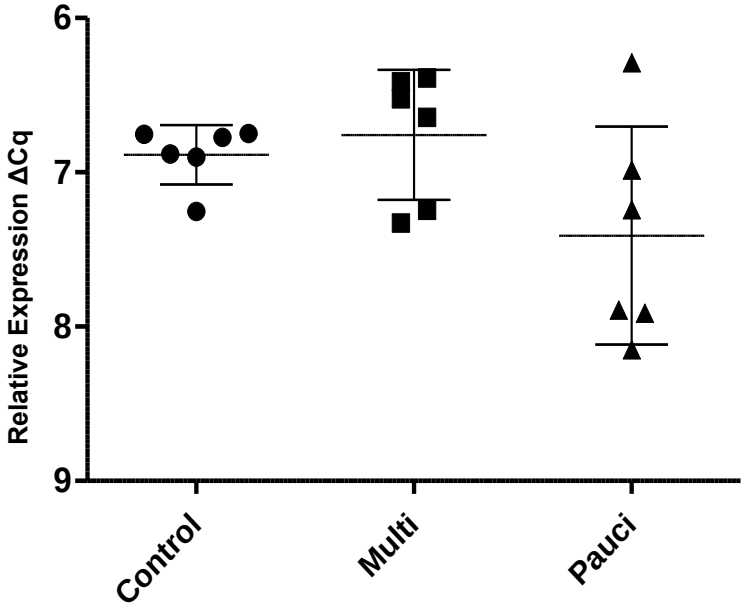


B)

	Fold Change	P-value
M vs P	2.1	0.04
M vs C	1.42	0.32
P vs C	-1.49	0.47

Figure 4.15*IL12RB1v3* RT-qPCR. A) Relative expression of *IL12RB1v3* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)



B)

	Fold Change	P-value
M vs P	1.87	0.08
M vs C	1.1	0.89
P vs C	-1.72	0.18

Figure 4.16 *IL12RB1v4* RT-qPCR. A) Relative expression of *IL12RB1v4* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

4.5 Conclusions

- Previous study has identified an allelic variant present in sheep IL23A (Gossner *et al.*, 2012b). This variant was not found in any of the sequenced clones from either multibacillary or paucibacillary animals in this study.
- Sequencing of the ovine *IL23R* has identified full length sequence and 5 spliced variants of this gene.
- Sequencing of the ovine *IL12RB1* has identified full length sequence and 4 spliced variants of this gene.
- All variants identified, excluding *IL23Rv1*, alter amino acid sequence and may have an effect of IL23R/IL23/IL12 ligand binding or signalling.
- No differential expression was found between paucibacillary and multibacillary groups for total *IL23R*, *IL12RB1*and *IL12RB1v4* gene expression.
- Expression of *IL12RB1v3* was significantly higher in multibacillary animals than in paucibacillary animals (P=0.04). This variant has a truncated extracellular region leading to a loss of the conserved WSxWS motif.

Chapter 5

IL25 genes in sheep paratuberculosis

5.1 Introduction

IL25 is a Th2 cytokine belonging to the IL17 cytokine family (also known as IL17E). The interleukin-17 (IL17) family consists of six cytokine members (IL17A-F); five function as pro-inflammatory mediators during Th17 immune responses (IL17A, B, C, D and F), and IL25 that is largely anti-inflammatory. IL25 was originally identified by sequence homology to other IL17 cytokine family members (Gu *et al.*, 2013). Production of IL17A and IL17F by Th17 cells has been shown to induce a Th1-like immune response through stimulation of IL6, IL8 and TNFα (Pappu *et al.*, 2011). In contrast, IL25 is markedly different as it plays a key role in the stimulation and maintenance of Th2 immune responses (Fort *et al.*, 2001). In human studies it has been shown that IL25 is present in several tissues including the lung, small intestine, colon and stomach. IL25 has been reported in many cell types including Th2 cells, macrophages, epithelial cells, eosinophils, basophils and mast cells and has been shown to link innate and adaptive immunity by enhancing Th2 cytokine production (Pappu *et al.*, 2011). Administration of IL25 induces expression of IL4, IL5 and IL13 in Th2 cells (Owyang *et al.*, 2006). Mice injected with exogenous IL-25 positively regulate Th2 responses in various organs and this has also been shown in transgenic mice over-expressing IL25. In these models, IL25 was shown to increase IL4 and IL13 expression and serum IgE levels in the lung and digestive tract (Fort *et al.*, 2001. Kim *et al.*, 2002. Hurst *et al.*, 2002). In other studies, IL25 deficient mice were unable to clear parasitic helminths including *Nippostrongylus* and *Trichuris*muris owing to failed Th2 responses (Fallon *et al.*, 2006. Owyang *et al.*, 2006).

The IL17 receptor family (IL17RA to IL17RE) are a unique group of proteins that share minimal structural and signal transduction properties with other receptors (Gaffen, 2009). IL17Rs are membrane-bound proteins that contain single transmembrane domains and possess a unique structural motif within the cytoplasmic tail termed the similar expression to fibroblast growth factor genes/IL17R (SEFIR) domain (Novatchkova *et al.*, 2003). NF-κB activator 1 (Act1) is also a member of the SEFIR protein family and acts as an adaptor protein in the signal transduction of several IL17Rs (Zhang *et al.*, 2014. Chang *et al.*, 2006. Swaidani *et al.*, 2009).

The IL25 receptor was initially identified as a receptor for IL17B; however recent studies have subsequently shown that this receptor exhibits higher affinity for IL25 than IL17B (Gu *et al.*, 2013). The IL25 cytokine receptor (IL25R) is composed of IL17RA and IL17RB and the biological effects of IL25 are mediated by this receptor complex. In humans IL17RB shares approximately 26% amino acid identity with IL17RA and IL25 is the high affinity ligand for IL17RB (Gaffen, 2009). In mice, cultured splenocytes from either IL17RA knockout or IL17RB knockout models are unable to produce IL5 and IL13 in response to IL25 stimulation (Rickel *et al.*, 2008). Furthermore, administration of antagonistic mAbs to either IL17RA or IL17RB completely blocked IL25 induced pulmonary inflammation; indicating that both receptor components are essential for signalling via interaction with the IL25 ligand (Rickel *et al.*, 2008). IL25 has been demonstrated to activate NFκB, however the molecular mechanisms of IL25R signalling pathways continue to be investigated; IL17RB contains a TNF receptor associated factor 6 (TRAF6) binding motif in its cytoplasmic tail whereas IL17RA does not have this motif (Gaffen, 2009). The TRAF family proteins serve as adaptor proteins in the assembly of receptor associated signalling complexes and link upstream receptors to downstream effector molecules, including activation of NF-κB. TRAF6 is recruited by Act1 during IL17R signalling. Maezawa *et al*. showed that IL25R-mediated NFκB activation, but not MAPK activation, was diminished in a TRAF6 deficient murine model; indicating that TRAF6 plays a critical role in IL25R mediated signal transduction and gene expression (Maezawa *et al.*, 2006).

Although IL17RB has been found to interact only with IL17RA, IL17RA has been shown to create a receptor complex also with IL17RC (figure 5.1). The IL17RA/C receptor complex is required for IL17A and IL17F induction and subsequent stimulation of Th1/Th17 responses (Gu *et al.*, 2013.Zepp *et al.*, 2011).

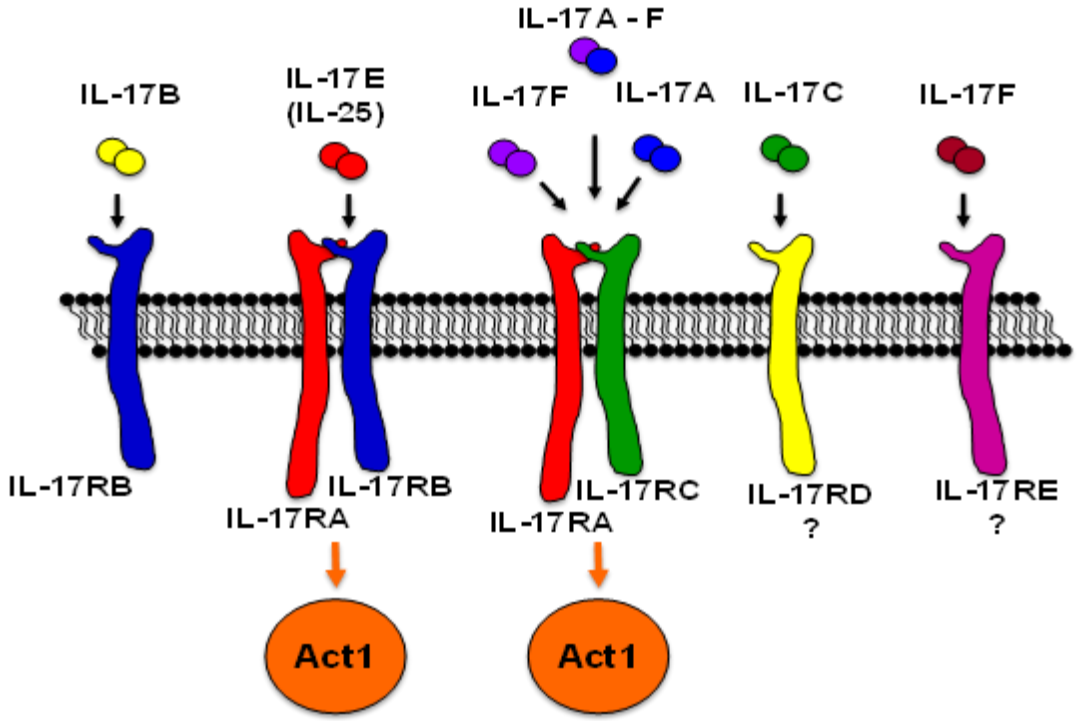


Figure 5.1 IL17 family cytokine and receptor interactions**.** Image adapted (Zepp *et al.*, 2011)

Responses induced by IL25 are characterised by elevated levels of type-2 cytokines, often leading to pathological changes in the gut and lungs. Elevated IgE and IgG1, epithelial cell hyperplasia and increased mucus secretion are typical pathological responses found in IL25 induced inflammation (Dong, 2008). During gut inflammation an anti-inflammatory role has been attributed to IL25 in type-1 models of disease (Caruso *et al.*, 2009). The mechanism by which IL25 acts to reduce/prevent inflammation of the gut has not yet been confirmed, although it has been shown that IL25 is associated with the induction of alternatively-activated macrophages with anti-inflammatory properties (Rizzo *et al.*, 2012). IL25 is able to prevent oxazolone-induced colitis, a murine model of Th2 driven colitis resembling UC; subsequently it has been suggested that IL25 may contribute to a broad negative effect on inflammatory properties of intestinal macrophages rather than specifically suppressing Th1/Th17 pathway components (Franze *et al.*, 2011). Treatment with exogenous IL25 was shown to increase expression of type 2 cytokines in mice

genetically susceptible to *Trichuris muris* helminth infection, whilst IL25 knockout mice developed severe intestinal inflammation associated with high IFNγ and IL17A expression (Owyang *et al.*, 2006). It is considered that IL25, including the IL25 receptors, act as a critical link between Th1/Th17 inflammatory cell-mediated and Th2 immune responses.

Few studies that have investigated *IL17RA* and *IL17RB* at the gene/transcript level have identified polymorphisms that may be associated with inflammation of the lung. Jung *et al.* analyzed the association of polymorphisms in the *IL17RB* gene using 954 asthmatic patients compared to 265 healthy control subjects. It was found that a rare allele of *IL17RB* (*5661G>A*) may have a protective role in the development of asthma (Jung *et al.*, 2009). Furthermore, 15 SNPs of the *IL17RA* gene were analyzed in 825 healthy control subjects, 143 subjects with Aspirin Exacerbated Respirator Disease (AERD) and 411 subjects with aspirin-tolerant asthma (ATA). Three of the SNPs (*1075A>G, 947A>G, 50C>T*) were shown potentially to decrease the risk of AERD via attenuation of *IL17RA gene* (Park *et al.*, 2013).

In this chapter sequencing of full gene transcripts and transcript variants of the IL25 cytokine receptors, *IL17RB* and *IL17RA*, is described. This chapter will also describe expression of the *IL25* cytokine and *IL25*receptor transcripts, including identified spliced variants, in multibacillary and paucibacillary paraTB infected and in uninfected control sheep.

5.2 Ovine *IL17RA* gene

5.2.1 Sequencing of ovine *IL17RA* gene

The ovine *IL17RA* sequence (LN878979.1) was obtained by amplifying gene sections by RT-PCR (described in Chapter 2) and compiling a consensus sequence of the gene. Initially primers were designed based on a predicted ovine *IL17RA* sequence (XM_012140663.1) template compiled from predicted bovine sequence (XM_010827960.1) and a BLAST alignment of known human (NM_014339.6) and rat *IL17RA*(NM_001107883.2) against the sheep genome OARv3.1, this template is depicted in figure 5.2. Partial *IL17RA* was sequenced; however attempts to amplify 5' and 3' coding region using primers designed on the predicted ovine template (exons 1 and 13) were unsuccessful. The partially amplified *IL17RA* sequence was used to design primers for 5' and 3' rapid amplification of cDNA ends (RACE) RT-PCR as described in Chapter 2. At least three clones were sequenced for both 5' and 3' RACE from different animals/RT-PCR reactions that gave a product of expected size and consensus sequences were compiled.

PCR products obtained from 5' RACE RT-PCR only ever generated sequence as far as predicted exon 2 in the 5' direction, despite several attempts. In order to amplify the *IL17RA* sequence further upstream of exon 2, a search of the ovine *IL17RA* gene was carried out on data collected from RNA-seq analysis (carried out by Dr. Anton Gossner). RNA-seq provides short (~100bp) sequenced reads of all RNA in a sample; these reads are subsequently mapped to the sheep genome OARv3.1. Results are shown as clusters of reads in exon regions using Integrative Genomics Viewer v2.3 (IGV - http://www.broadinstitute.org). A representative image of this is shown along with the *IL17RA* 5' region in figure 5.3. *IL17RA* 5' sequence data taken from IGV was used to design a primer within sequence upstream of the predicted ovine *IL17RA* start codon, that would amplify the start codon in predicted exon 1and the 5' UTR. *IL17RA* sequence obtained from this method gave amplified sequence 5' and 3' of the predicted ovine exon 1; however the predicted exon 1, containing the predicted start codon, was not found, this is shown in figure 5.4. When translated to protein sequence the ovine *IL17RA* sequence that did not contain the predicted exon

1 presented with a premature stop codon, thus no full protein sequence that matched the predicted ovine protein sequence was found (figure 5.5 and 5.6). No transcript variants were identified from the ovine *IL17RA* clones that were sequenced,

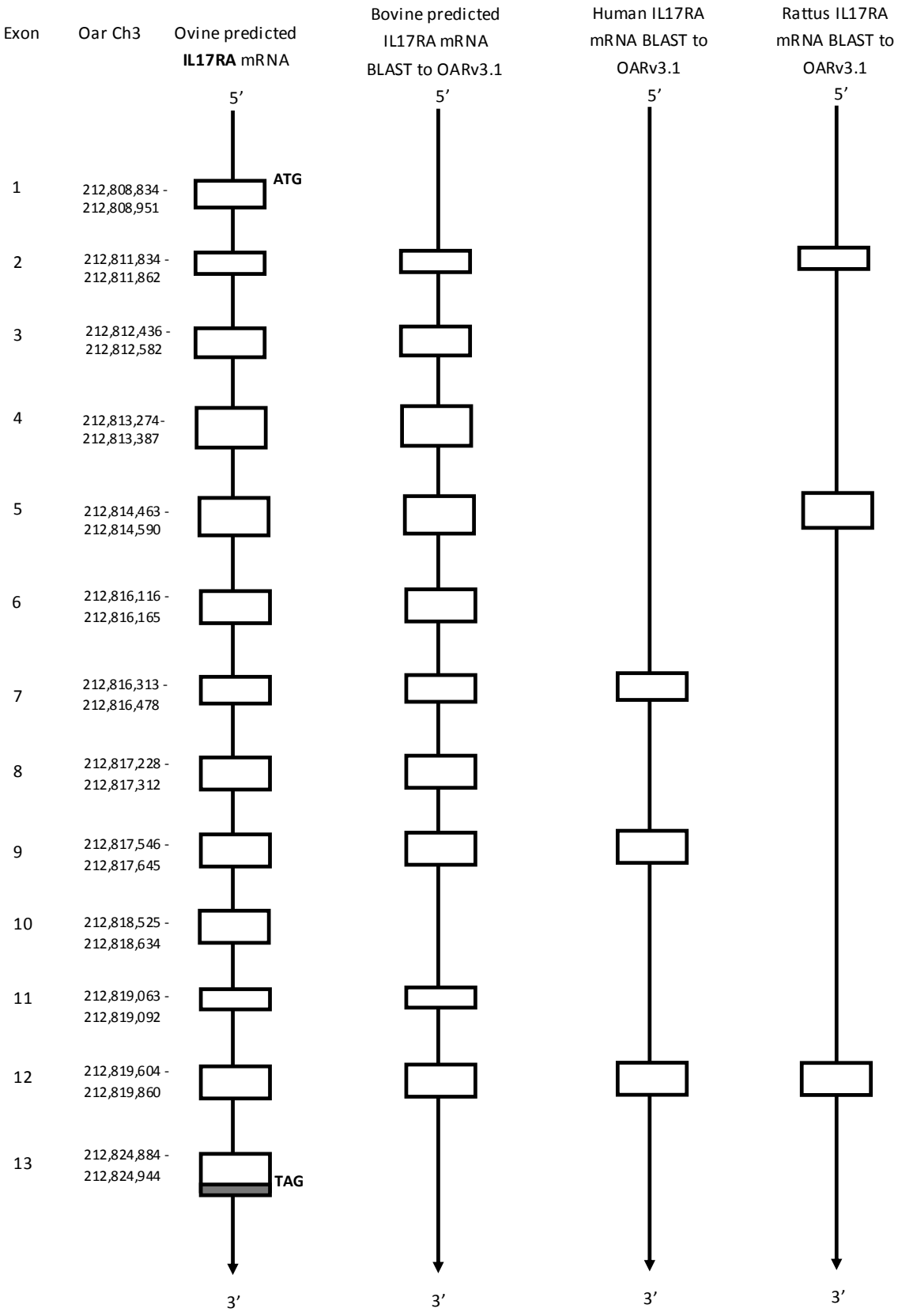
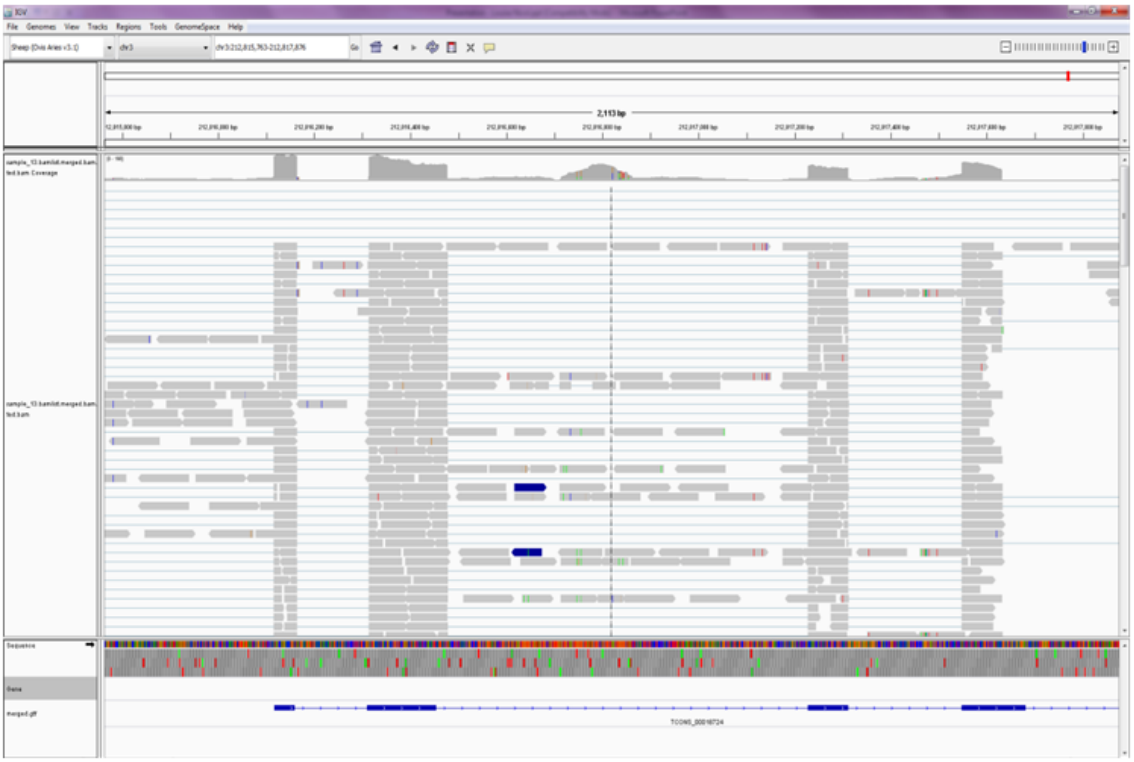


Figure 5.2 Gene map of predicted ovine *IL17RA* sequence template used to design primers for ovine *IL17RA* gene amplification. Template was compiled from predicted bovine sequence and sequence from a BLAST alignment of known human and rat *IL17RA* against the sheep genome OARv3.1. Chromosome locations of predicted ovine exons are shown.

A)



B)

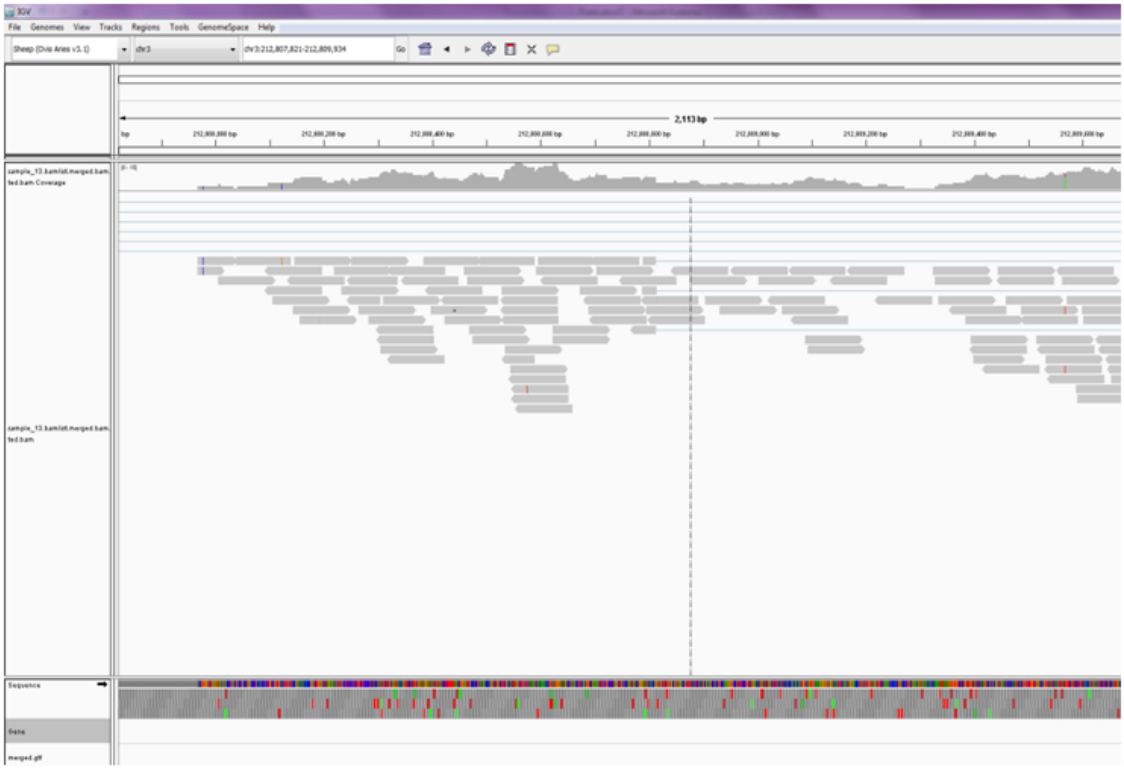


Figure 5.3 A) Representative image of RNA-seq data analysis using Integrative Genome Viewer software, showing short reads of nucleotide sequences mapped to chromosomal locations within sheep genome OARv3.1 and B) RNA-seq data at *IL17RA* 5' end with predicted ovine start codon region indicated by vertical dotted line. Primers were designed upstream and downstream of the predicted ovine *IL17RA* start codon.

5.2.2 *IL17RA* nucleotide sequence analysis

The ovine *IL17RA* sequence was aligned to the sheep genome OARv3.1. The 3' region of ovine *IL17RA* was unable to be mapped to a chromosome location of OARv3.1 as this area of the sheep genome has yet to be annotated. A schematic diagram of the sequence is shown in figure 5.4, with the predicted exon one highlighted. This exon was not found in any of the *IL17RA* clones sequenced. The nucleotide sequence alignment of ovine *IL17RA* mapped to chromosome 3 of OARv3.1 can be found in appendix section A9.

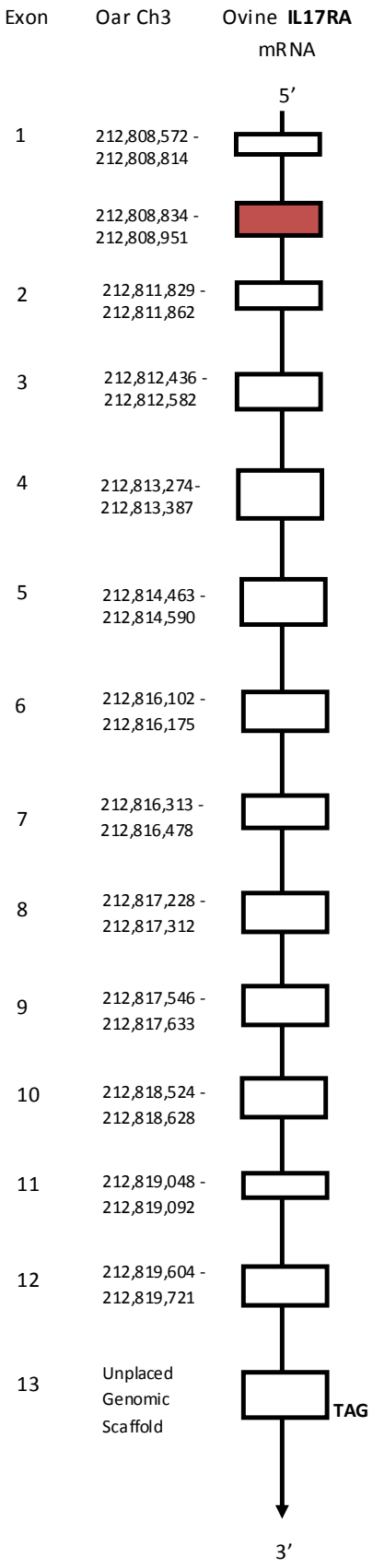


Figure 5.4 Gene map of amplified ovine *IL17RA*gene sequence. Obtained using primers designed 5' upstream of the predicted exon 1/start codon. The predicted exon 1 (red) was not found in any sequenced clones.

5.2.3 Ovine IL17RA translated protein sequence analysis

Consensus sequence of ovine *IL17RA* gene was translated into amino acid sequence and aligned to the coding region amino acid sequence of predicted ovine IL17RA, from this it could be seen that there is 77% sequence similarity between the ovine IL17RA protein sequence identified in this study and predicted IL17RA sequence (figure 5.5 and 5.6); however without the presence of the predicted start codon, the ovine IL17RA gene contains a premature stop codon. A second open reading frame is shown in sequence a) which may encode a truncated IL17RA protein. Furthermore the sequenced and predicted ovine amino acid sequences were aligned with the two known isoforms of human IL17RA (Isoform 1: NP_055154.3 and Isoform 2: NP_001276834.1); this shows the presence of a conserved transmembrane domain sequence in both the sequenced and predicted ovine IL17RA sequences figure 5.7, however both 5' and 3' ends of both ovine sequences differ to that of human IL17RA.

```
Predicted       ------------------MAGDQELERDPGLFVY---TCLLYTWPENRAPLPAGSPEK-- 37
Sequenced       MTDVGAVNQKSAWSRSLCQPGAMEFQGTEKLLLPGCPSCREGTWAALQESLEHSRRGKSL 60
                                  .*  *::    *::    :*   **.  : .*  .   *

Predicted       -RG--GLNCTVKNSTCLDDSWIHPRNLTPSSPKNVQTQLRFAHTQQGHLLPVVHIEWTLQ 94
Sequenced       SQGPKGLNCTVKNSTCLDDSWIHPRNLTPSSPKNVQTQLRFAHTQQGHLLPVVHIEWTLQ 120
                 :*  ************************************************

Predicted       TDASVLYLEGAELSILQLSTNERLCVRFEFLTTLRHHHKRWRFAFSHFVVEPREEYEVTV 154
Sequenced       TDASVLYLEGAELSILQLSTNERLCVRFEFLTTLRHHHKRWRFAFSHFVVEPREEYEVTV 180
                ************************************************************

Predicted       HHLPKPIPDGDPNHQSRNFLVPDCKDPRMKDTTPCVSSGSLWDPNITVETLEAHQLRLSF 214
Sequenced       HHLPKPIPDGDPNHQSRNFLVPDCKDPRMKDTTPCVSSGSLWDPNITVETLEAHQLRLSF 240
                ************************************************************

Predicted       TPWNESTSYQVLLHSFPPAENQSCFQHVVDMPVPAQEAAPQRCHITVTLRDSSWCCRHHV 274
Sequenced       TPWNESTSYQVLLHSFPPAENQSCFQHVVDMPVPAQEAAPQRCHITVTLLDSSWCCRHHV 300
                ********************************************** **********

Predicted       QIQPFFSSCLNDCLRHSVSVACPEVSHTPGGTRDHTPLWVSAFITGLSILLVGSVILLIL 334
Sequenced       QIQPFFSSCLNDCLRHSVSMACPEVSHTPDAAEDHTPLWVSAFITGLSILLVGSVILLIL 360
                ******************:*********..:.**************************

Predicted       CMTWRLPGKR--KHEDGTKDTEILPAATSLTPPPLKPRKVWIVYSADHPLYVDVVLKFAQ 392
Sequenced       CMTWRLPGFRQGKHEDGTKDTEILPAATSLTPPPLKPRKVWIVYSADHPLYVDVVLKFAH 420
                ******** *  *********************************************:

Predicted       FLLTVCGTEVALDLLEEQAISEVGVMTWVGRQKQEVADSNSKIVVLPRQPRKKRFLSMKT 452
Sequenced       ----------------GPSLGRPGGQAGVGASGPDAKD-QGFFTRMPCSPQRR--LS--- 458
                              ::.. *  : ** .  :. * :. :. :* .*:::  **

Predicted       ESGCWS 458
Sequenced       ------
```

Figure 5.5 Translated amino acid sequence alignment of ovine IL17RA and predicted ovine IL17RA.

A) **M**TDVGAVNQKSAWSRSLCQPGA**M**EFQGTEKLLLPGCPSCREGTWAALQESLEHSRRGK**S**L
SQGPKGLNCTVKNSTCLDDSWIHPRNLTPSSPKNVQTQLRFAHTQQGHLLPVVHIEWTLQT
DASVLYLEGAELSILQLSTNERLCVRFEFLTTLRHHHKRWRFAFSHFVVEPREEYEVTVHHLPK
PIPDGDPNHQSRNFLVPDCKDPR**M**KDTTPCVSSGSLWDPNITVETLEAHQLRLSFTPWNES
TSYQVLLHSFPPAENQSCFQHVVD**M**PVPAQEAAPQRCHITVTLLDSSWCCRHHVQIQPFFS
SCLNDCLRHSVS**M**ACPEVSHTPDAAEDHTPLWVSAFITGLSILLVGSVILLILC**M**TWRLPGFR
QGKHEDGTKDTEILPAATSLTPPPLKPRKVWIVYSADHPLYVDVVLKFAHGPSLGRPGGQAG
VGASGPDAKDQGFFTR**M**PCSPQRRL**S**

B) **M**AGDQELERDPGLFVYTCLLYTWPENRAPLPAGSPEKRGGLNCTVKNSTCLDDSWIHPRNL
TPSSPKNVQTQLRFAHTQQGHLLPVVHIEWTLQTDASVLYLEGAELSILQLSTNERLCVRFEFL
TTLRHHHKRWRFAFSHFVVEPREEYEVTVHHLPKPIPDGDPNHQSRNFLVPDCKDPR**M**KDT
TPCVSSGSLWDPNITVETLEAHQLRLSFTPWNESTSYQVLLHSFPPAENQSCFQHVVD**M**PV
PAQEAAPQRCHITVTLRDSSWCCRHHVQIQPFFSSCLNDCLRHSVSVACPEVSHTPGGTRD
HTPLWVSAFITGLSILLVGSVILLILC**M**TWRLPGKRKHEDGTKDTEILPAATSLTPPPLKPRKV
WIVYSADHPLYVDVVLKFAQFLLTVCGTEVALDLLEEQAISEVGV**M**TWVGRQKQEVADSNS
KIVVLPRQPRKKRFLS**M**KTESGCW**S**

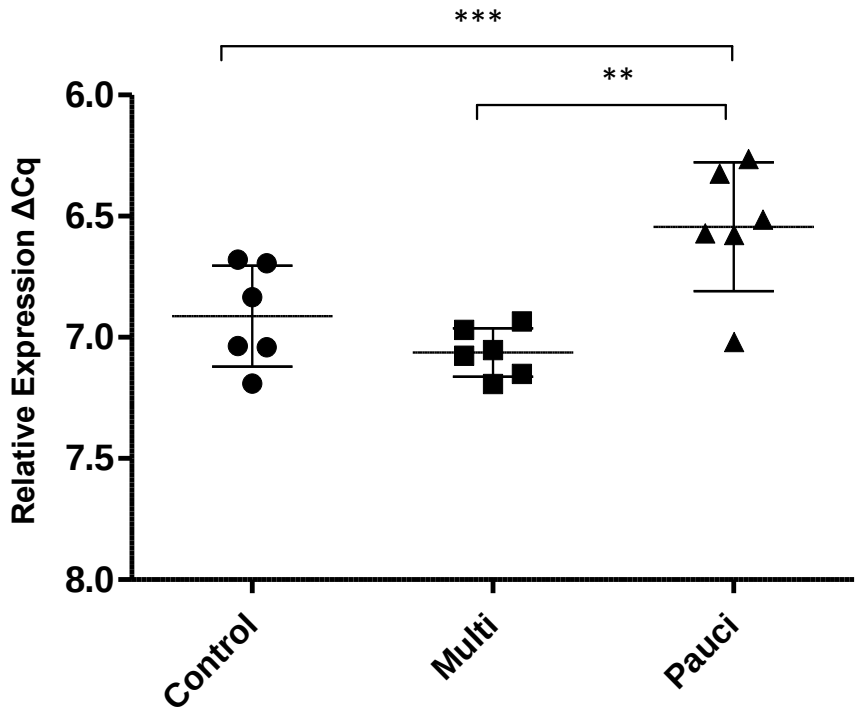
Figure 5.6 ExPASy TRANSLATE sequence output of A) sequenced ovine IL17RA protein and B) predicted ovine IL17RA protein. Open reading frames are highlighted in red with region of sequence similarity underlined.

Figure 5.7 Amino acid sequence alignment of translated ovine IL17RA, predicted ovine IL17RA and two human isoforms of IL17RA; Homo_v1 and Homo_v2. Human transmembrane domain is highlighted in yellow.

5.2.4 Expression levels of *IL17RA* transcripts

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for total *IL17RA* using available sequence. Statistical analysis was carried out on genes where results were within a measureable range, with relative expression (ΔCq), fold change between groups and P-values shown in figure 5.8. Expression of *IL17RA* was found to be significantly increased in the paucibacillary group when compared to multibacillary (2 fold, P=0.001) and control (2.2 fold, P=0.02) groups.

A)



B)

	Fold Change	P-value
M vs P	-2	0.001
M vs C	0.46	0.43
P vs C	2.2	0.02

Figure 5.8 *IL17RA* RT-qPCR. A) Relative expression of total *IL17RA* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

5.3Ovine *IL17RB* gene

5.3.1 Identifying *IL17RB* transcript variants

In order to identify whether there were any transcript variants present in the *IL17RB* gene, full length sequence for *IL17RB* was amplified by RT-PCR using primers described in table 2.3. At least six clones, from separate RT-PCR reactions, for each infected animal were sequenced within individual samples. Each clone was aligned against the sheep genome OARv3.1 and differences in sequence that were present in two or more clones from different animals were considered transcript variants. Transcript variants were identified from several clones in both paucibacillary and multibacillary animals, along with full length gene transcripts; summarised below. A gene map including transcript variants is shown in figure 5.9.

Full ovine *IL17RB* transcript sequence and variant transcripts were aligned using ClustalW v2.1, this is shown in figure 5.10.

Cloning and sequencing identified 3 transcript variants of ovine *IL17RB*. The *IL17RB* gene is encoded on the minus strand of chromosome 19 (OARv3.1: NC_019476.1) and is predicted to encode a protein sequence of 497 amino acids. Differences between full length *IL17RB* (LN878975.1) and identified transcript variants are described below;

IL17RB*v1* (LN878976.1) has a deletion of 96bp (bp 975-1071) resulting in a 32 amino acid deletion at position 311 within the predicted protein sequence. This results in a truncated intracellular region with a loss of the conserved TRAF6 binding motif.

IL17RB*v2* (LN878977.1) has a deletion of exon 4 (Chr19: 47,056,667 – 47,054,190) resulting in a change of frame and a premature stop codon at position 96 within the protein.

IL17RB*v3* (LN878978.1) has a deletion of exon 4 (Chr19: 47,056,667 – 47,054,190) and an insertion of 177bp (bp 651-828) at the 5’ end of exon 8 (Chr19: 47,050,145 – 47,049,971); resulting in a change of frame and a premature stop codon at position 96 within the protein.

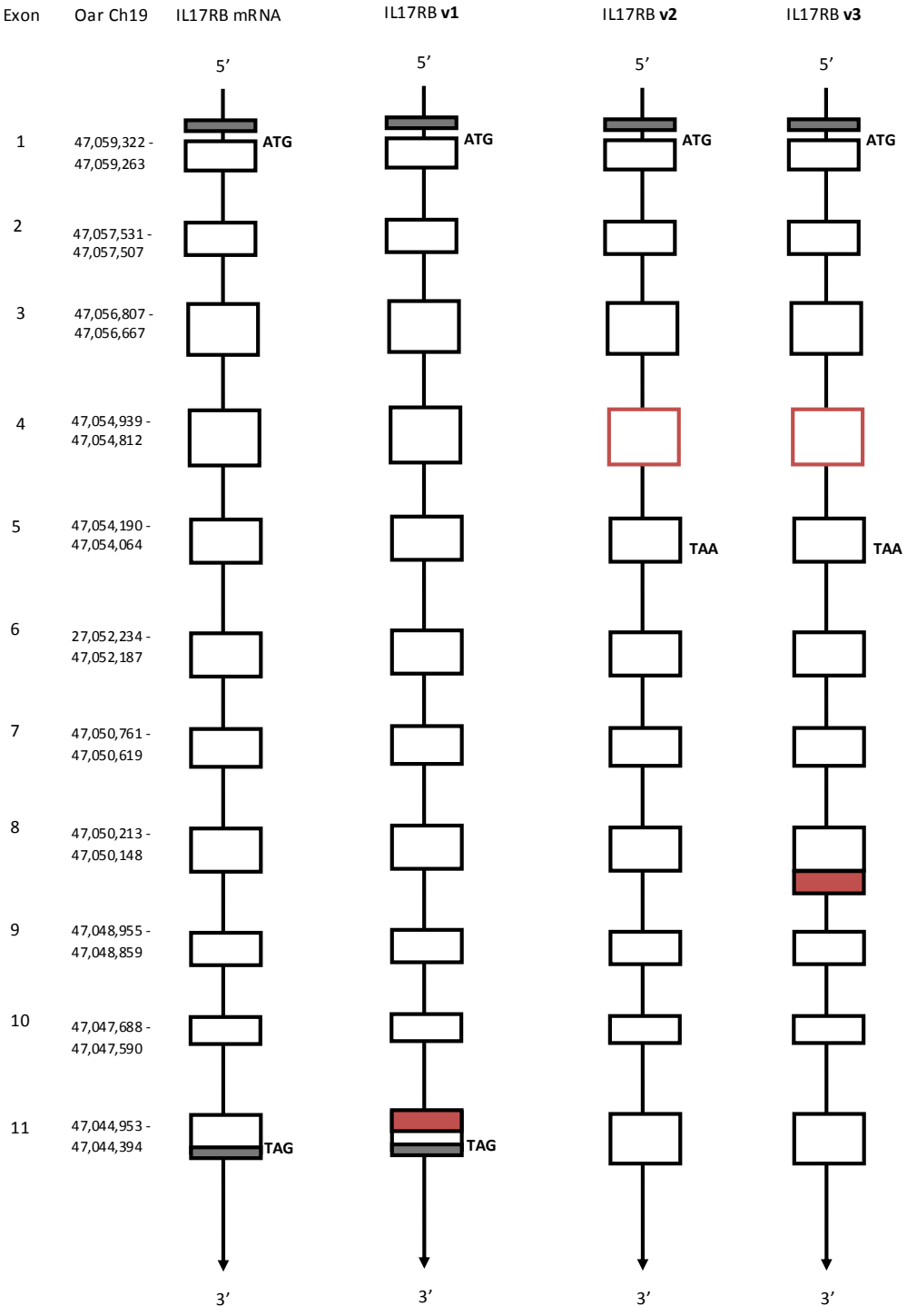


Figure 5.9 Schematic diagrams of the ovine *IL17RB* gene and splice variants (*IL17RB*v1-3) identified in this study. Boxes represent individual exons with chromosome location shown. Grey boxes indicate untranslated region (UTR) with variant regions highlighted in red. The 'ATG' is initial start codon and stop codons are shown for each transcript.

```
V2              TCGGCCCGAAGCCGATCCGGACCCCGCGCGGTGGCCCCGCGATGTTGCTAGTACTGCTGA 60
V3              TCGGCCCGAAGCCGATCCGGACCCCGCGCGGTGGCCCCGCGATGTTGCTAGTACTGCTGA 60
Full            TCGGCCCGAAGCCGATCCGGACCCCGCGCGGTGGCCCCGCGATGTTGCTAGTACTGCTGA 60
V1              TCGGCCCGAAGCCGATCCGGACCCCGCGCGGTGGCCCCGCGATGTTGCTAGTACTGCTGA 60
                 ************************************************************

V2              GCCTGGCCGCGCTGTGCTGGGGGGCCGTGCCTCCGGAGCCGACAATTCAGTGTGGCTCTG 120
V3              GCCTGGCCGCGCTGTGCTGGGGGGCCGTGCCTCCGGAGCCGACAATTCAGTGTGGCTCTG 120
Full            GCCTGGCCGCGCTGTGCTGGGGGGCCGTGCCTCCGGAGCCGACAATTCAGTGTGGCTCTG 120
V1              GCCTGGCCGCGCTGTGCTGGGGGGCCGTGCCTCCGGAGCCGACAATTCAGTGTGGCTCTG 120
                 ************************************************************

V2              AACCCGGACCGTCTCCAGAGTGGATGGTCCGACACACCCTGACCCCAGGAGACCTAAGGG 180
V3              AACCCGGACCGTCTCCAGAGTGGATGGTCCGACACACCCTGACCCCAGGAGACCTAAGGG 180
Full            AACCCGGACCGTCTCCAGAGTGGATGGTCCGACACACCCTGACCCCAGGAGACCTAAGGG 180
V1              AACCCGGACCGTCTCCAGAGTGGATGGTCCGACACACCCTGACCCCAGGAGACCTAAGGG 180
                 ************************************************************

V2              ACCTCCGAGTGGAACCTATTAAAAGCAGTGTTGACCTGGAGGACTCTTCAATTTTGATGA 240
V3              ACCTCCGAGTGGAACCTATTAAAAGCAGTGTTGACCTGGAGGACTCTTCAATTTTGATGA 240
Full            ACCTCCGAGTGGAACCTATTAAAAGCAGTGTTGACCTGGAGGACTCTTCAATTTTGATGA 240
V1              ACCTCCGAGTGGAACCTATTAAAAGCAGTGTTGACCTGGAGGACTCTTCAATTTTGATGA 240
                 ************************************************************

V2              ACATAAGCTGGATACTCCGGGCAGATG---------------------------------- 267
V3              ACATAAGCTGGATACTCCGGGCAGATG---------------------------------- 267
Full            ACATAAGCTGGATACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAAGATCT 300
V1              ACATAAGCTGGATACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAAGATCT 300
                 **************************

V2              ------------------------------------------------------------
V3              ------------------------------------------------------------
Full            GCGTGATGGGCAAAAGCCACTTCCAGTCCCGCAGCTGCATCAGGTGCAATTACACCCAGG 360
V1              GCGTGATGGGCAAAAGCCACTTCCAGTCCCGCAGCTGCATCAGGTGCAATTACACCCAGG 360


V2              ----------------------------------TGGACGTTTTCCTACATTGGTTTTC 292
V3              ----------------------------------TGGACGTTTTCCTACATTGGTTTTC 292
Full            CGTTCCGCACTCAGACCAGACACTCTGGTGGCAAATGGACGTTTTCCTACATTGGTTTTC 420
V1              CGTTCCGCACTCAGACCAGACACTCTGGTGGCAAATGGACGTTTTCCTACATTGGTTTTC 420
                                                  ************************

V2              CTGTCGAGCTGAACACAGTCTATTTCATTGGAGCCCATAACATCCCCAATGCAAATATGA 352
V3              CTGTCGAGCTGAACACAGTCTATTTCATTGGAGCCCATAACATCCCCAATGCAAATATGA 352
Full             CTGTCGAGCTGAACACAGTCTATTTCATTGGAGCCCATAACATCCCCAATGCAAATATGA 480
V1              CTGTCGAGCTGAACACAGTCTATTTCATTGGAGCCCATAACATCCCCAATGCAAATATGA 480
                 ************************************************************

V2              ATGAGGACGGCCCCTCCATGGCTGTGAACTTCACCTCCCCAGGGTGCCTGGACCATGTAA 412
V3              ATGAGGACGGCCCCTCCATGGCTGTGAACTTCACCTCCCCAGGGTGCCTGGACCACGTAA 412
Full            ATGAGGACGGCCCCTCCATGGCTGTGAACTTCACCTCCCCAGGGTGCCTGGACCATGTAA 540
V1              ATGAGGACGGCCCCTCCATGGCTGTGAACTTCACCTCCCCAGGGTGCCTGGACCATGTAA 540
                 ********************************************************** ****

V2              TGAAATATAAAAAAAAAATGCATCGAGGCAGGAAGCCTGTGGAAGCCAAACATCACCGCTT 472
V3              TGAAATAT-AAAAAAAATGCATCGAGGCAGGAAGCCTGTGGAAGCCAAACATCACCGCTT 471
Full            TGAAATATAAAAAAAAAATGCATCGAGGCAGGAAGCCTGTGGAAGCCAAACATCACCGCTT 600
V1              TGAAATATAAAAAAAAAATGCATCGAGGCAGGAAGCCTGTGGAAGCCAAACATCACCGCTT 600
                 ******** ***************************************************

V2              GCAAGAAGAGCGCGAACACGGTGGAAGTGAATTTTACCACCAGTCCCCTTGGAGACAGAT 532
V3              GCAAGAAGAGCGCGAACACGGTGGAAGTGAATTTTACCACCAGTCCCCTTGGAGACAGAT 531
Full            GCAAGAAGAGCGCGAACACGGTGGAAGTGAATTTTACCACCAGTCCCCTTGGAGACAGAT 660
V1              GCAAGAAGAGCGCGAACACGGTGGAAGTGAATTTTACCACCAGTCCCCTTGGAGACAGAT 660
                 ************************************************************

V2              ACATGGCGCTCATCCAAAATACCTCCGTGATTGGGACTTCCTATGTGTCAGAGAAGGAAC 592
V3              ACATGGCGCTCATCCAAAATACCTCCGTGATTGGGACTTCCTATGTGTCAGAGAAGGAAC 591
Full            ACATGGCGCTCATCCAAAATACCTCCGTGATTGGGACTTCCTATGTGTCAGAGAAGGAAC 720
V1              ACATGGCGCTCATCCAAAATACCTCCGTGATTGGGACTTCCTATGTGTCAGAGAAGGAAC 720
                 ************************************************************
```

```
V2              TAACTCGAACTTCCGTGGTGGTCCACGTGACTGGGGAAAGTGAAGGTGCTGTGGTCCAG- 651
V3              TAACTCGAACTTCTGTGGTGGTCCACGTGACTGGGGAAAGTGAAGGTGCTGTGGTCCAGG 651
Full            TAACTCGAACTTCCGTGGTGGTCCACGTGACTGGGGAAAGTGAAGGTGCTGTGGTCCAG- 779
V1              TAACTCGAACTTCCGTGGTGGTCCACGTGACTGGGGAAAGTGAAGGTGCTGTGGCCCAG- 779
                 ************* ****************************************** ****

V2              ------------------------------------------------------------
V3              TAAAGCTGAACAAGATGCTCTGACAGGCACAGAACGGGGCTCACTGAAGACCCACTGCAC 711
Full            ------------------------------------------------------------
V1              ------------------------------------------------------------


V2              ------------------------------------------------------------
V3              CGCCACTCGCTGCTTTTGAGCAGTCCATTCTGGATGGAGCAGATGCGTTCTCCCCTCGGC 771
Full            ------------------------------------------------------------
V1              ------------------------------------------------------------


V2              ----------------------------------------------------------CTGA 655
V3              TTGCTTCTGCCAGCAGCCTCTCTGGCATTACGTGATGCTGATGGGGAATCCATGAGCTGA 831
Full            ----------------------------------------------------------CTGA 783
V1              ----------------------------------------------------------CTGA 783
                                                                           ****

V2              CTCCGTATTTCCATACCTGTGGCAACGACTGCATCCGACAAAGAGGAACCGTGGTGCGGT 715
V3              CTCCGTATTTCCATACCTGTGGCAACGACTGCATCCGACAAAGAGGAACCGTGGTGCGGT 891
Full            CTCCGTATTTCCATACCTGTGGCAACGACTGCATCCGACAAAGAGGAACCGTGGTGCGGT 843
V1              CTCCGTATTTCCATACCTGTGGCAACGACTGCATCCGACAAAGAGGAACCGTGGTGCGGT 843
                ************************************************************

V2              GCCCACAAACAGGAGTCTCTCCTCAGGATCATGAGAGAAGTGTGCTGGGTGGCTGGCTGC 775
V3              GCCCACAAACAGGAGTCTCTCCTCAGGATCATGAGAGAAGCGTGCTGGGTGGCTGGCTGC 951
Full            GCCCACAAACAGGAGTCTCTCCTCAGGATCATGAGAGAAGTGTGCTGGGTGGCTGGCTGC 903
V1              GCCCACAAACAGGAGTCTCTCCTCAGGATCATGAGAGAAGTGTGCTGGGTGGCTGGCTGC 903
                ************************************ ******************

V2              CTCTCCTCCTGTCCGCTCTCCTTGTGGCCACATGGGTGCTGGTGGCTGGCATCTATCTGA 835
V3              CTCTCCTCCTGTCCGCTCTCCTTGTGGCCACATGGGTGCTGGTGGCTGGCATCTATCTGA 1011
Full            CTCTCCTCCTGTCCGCTCTCCTTGTGGCCACATGGGTGCTGGTGGCTGGCATCTATCTGA 963
V1              TCTCCTCCTGTCCGCTCTCCTTGTGGCCACATGGGTGCTGGTGGCTGGCATCTATCTGA  963
                ************************************************************

V2              TCTGGAGGCACGAAAGGATCAAGAAGACTTCCTTCTCAACTACCACACTACTGCCCTCTC 895
V3              TCTGGAGGCACGAAAGGATCAAGAAGACTTCCTTCTCAACTACCACACTACTGCCCTCTC 1071
Full            TCTGGAGGCACGAAAGGATCAAGAAGACTTCCTTCTCAACTACCACACTACTGCCCTCTC 1023
V1              TCTGGAGGCACG------------------------------------------------ 975
                ************

V2              TTAAGGTTCTTGTGGTTTACCCTTCTGAAATATGCTTCCATCACACAGTTTGTTACTTCA 955
V3              TTAAGGTTCTTGTGGTTTACCCTTCTGAAATATGCTTCCATCACACAGTTTGTTACTTCA 1131
Full            TTAAGGTTCTTGTGGTTTACCCTTCTGAAATATGCTTCCATCACACAGTTTGTTACTTCA 1083
V1              ------------------------------------------------TTTGTTACTTCA 987
                                                                ************

V2              CTGAATTTCTTCAAAACCGCTGCAGAAGTGAAGTTATCCTTGAAAAGTGGCGGAAAAAGA 1015
V3              CTGAATTTCTTCAAAACCGCTGCAGAAGTGAAGTTATCCTTGAAAAGTGGCAGAAAAAGA 1191
Full            CTGAATTTCTTCAAAACCGCTGCAGAAGTGAAGTTATCCTTGAAAAGTGGCAGAAAAAGA 1143
V1              CTGAATTTCTTCAAAACCGCTGCAGAAGTGAAGTTATTCTTGAAAAGTGGCAGAAAAAGA 1047
                ************************************ *************.********

V2              AAATAGCCGAGATGGGTCCCGTGCAGTGGCTCACCACTCAGAAGCAAGCAGCGGATAAGG 1075
V3              AAATAGCCGAGATGGGTCCCGTGCAGTGGCTCACCACTCAGAAGCAAGCAGCGGATAAGG 1251
Full            AAATAGCCGAGATGGGTCCCGTGCAGTGGCTCACCACTCAGAAGCAAGCAGCGGATAAGG 1203
V1              AAATAGCCGAGATGGGTCCCGTGCAGTGGCTCACCACTCAGAAGCAAGCAGCGGATAAGG 1107
                ************************************************************

V2              TCATTTTCCTTCTTTCCAATGATAACACCATGTGTGATGGTACCTGTGACAAGAAGGAGG 1135
V3              TCATTTTCCTTCTTTCCAATGATAACACCATGTGTGATGGTACCTGTGACAAGAAGGAGG 1311
Full            TCATTTTCCTTCTTTCCAATGATAACACCATGTGTGATGGTACCTGTGACAAGAAGGAGG 1263
V1              TCATTTTCCTTCTTTCCAATGATAACACCATGTGTGATGGTACCTGTGACAAGAAGGAGG 1167
                ************************************************************
```

```
V2              GCGGCCCCTGTGAGAACTCCCAAGACCTGTTCCACCTCGCCTTTAACCTCTTCTGCAGTG 1195
V3              GCGGCCCCTGTGAGAACTCCCAAGACCTGTTCCACCTCGCCTTTAACCTCTTCTGCAGTG 1371
Full            GCGGCCCCTGTGAGAACTCCCAAGACCTGTTCCACCTCGCCTTTAACCTCTTCTGCAGTG 1323
V1              GCGGCCCCTGTGAGAACTCCCAAGACCTGTTCCACCTCGCCTTTAACCTCTTCTGCAGTG 1227
                ************************************************************

V2              ATCTGAGAAGCCAGACTCATCTGCGCAAGTATGTGGTGGTCTACTTCAGAGAGGGTGACA 1255
V3              ATCTGAGAAGCCAGACTCATCTGCGCAAGTATGTGGTGGTCTACTTCAGAGAGGGTGACA 1431
Full            ATCTGAGAAGCCAGACTCATCTGCGCAAGTATGTGGTGGTCTACTTCAGAGAGGGTGACA 1383
V1              ATCTGAGAAGCCAGACTCATCTGCGCAAGTATGTGGTGGTCTACTTCAGAGAGGGTGACA 1287
                ************************************************************

V2              TCAAAGACAGCTACAGTGCGCTCAGTGTCTGCCCCACGTACCGCCTCACGAAGGACGCTA 1315
V3              TCAAAGACAGCTACAGTGCGCTCAGTGTCTGCCCCACGTACCGCCTCACGAAGGACGCTA 1491
Full            TCAAAGACAGCTACAGTGCGCTCAGTGTCTGCCCCACGTACCGCCTCACGAAGGACGCTA 1443
V1              TCAAAGACAGCTACAGTGCGCTCAGTGTCTGCCCCACGTACCGCCTCACGAAGGACGCTA 1347
                ************************************************************

V2              CAGATTTCTGTGCGGAGCTTCTCCATGCCAAGCAGCATGTGTCGGTGGGGAGAAGGTCAC 1375
V3              CAGATTTCTGTGCGGAGCTTCTCCATGCCAAGCAGCATGTGTCGGTGGGGAGAAGGTCAC 1551
Full            CAGATTTCTGTGCGGAGCTTCTCCATGCCAAGCAGCATGTGTCGGTGGGGAGAAGGTCAC 1503
V1              CAGATTTCTGTGCGGAGCTTCTCCATGCCAAGCAGCATGTGTCGGTGGGGAGAAGGTCAC 1407
                ************************************************************

V2              GCGCCCGTCACTACAGCTGCCTCTCCCTGTAGCCCAACCACGAAGCAGGAGACCCTAAAG 1435
V3              GCGCCCGTCACTACAGCTGCCTCTCCCTGTAGCCCAACCACGAAGCAGGAGACCCTAAAG 1611
Full            GCGCCCGTCACTACAGCTGCCTCTCCCTGTAGCCCAACCACGAAGCAGGAGACCCTAAAG 1563
V1              GCGCCCGTCACTACAGCTGCCTCTCCCTGTAGCCCAACCACGAAGCAGGAGACCCTAAAG 1467
                ************************************************************

V2              CCTTCTGACCCA 1447
V3              CCTTCTGACCCA 1623
Full            CCTTCTGACCCA 1575
V1              CCTTCTGACCCA 1479
                ************
```

Figure 5.10 Nucleotide sequence alignment of ovine *IL17RB* full length and splice variant sequences (ClustalW v2.1). Highlighted sections indicate variant regions and start codon is underlined.

5.3.2 IL17RB translated protein sequence analysis

Nucleotide sequences of full length and alternatively spliced variant transcripts of the *IL17RB* gene were translated into amino acid and then aligned to identify changes to protein structure that occur as a result of transcriptional variation figure 5.12. An alignment of human IL17RB and ovine IL17RB amino acid sequences identified conserved TRAF-6 structural domain (Zhang *et al.*, 2014) within the IL17RB protein which is highlighted in red. Alignment of protein sequences shows that IL17RBv2 and v3 contain a premature stop codon. Transcript variant 1 has a 32 amino acid deletion caused by the deletion of 96bp in exon 11. The region of this protein that is deleted contains the conserved TRAF-6 binding motif and partial SEFIR domain. Conserved regions of this protein were identified by comparison of sequences from human, mouse and cattle; the sequence alignment of ovine, human and murine IL17RB, showing conserved TRAF6 binding motif and SEFIR domain location can be found in appendix section A10.

Key: **TRAF6 binding motif**

Transmembrane domain

SEFIR domain

```
V2              MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
V3              MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
Full            MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
V1              MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
                ************************************************************

V2              DSSILMNISWILRADVDV------------------------FLHWFSCRAEHSLFHWS 95
V3              DSSILMNISWILRADVDV------------------------FLHWFSCRAEHSLFHWS 95
Full            DSSILMNISWILRADASIRLLKATKICVMGKSHFQSRSCIRCNYTQAFRTQTRHSGGKWT 120
V1              DSSILMNISWILRADASIRLLKATKICVMGKSHFQSRSCIRCNYTQAFRTQTRHSGGKWT 120
                ***************..:                         : : *  ::.**  :*:

V2              P---------------------------------------------------------- 96
V3              P---------------------------------------------------------- 96
Full            FSYIGFPVELNTVYFIGAHNIPNANMNEDGPSMAVNFTSPGCLDHVMKYKKKCIEAGSLW 180
V1              FSYIGFPVELNTVYFIGAHNIPNANMNEDGPSMAVNFTSPGCLDHVMKYKKKCIEAGSLW 180
                 *

V2              ----------------------------------------------------------
V3              ----------------------------------------------------------
Full            KPNITACKKSANTVEVNFTTSPLGDRYMALIQNTSVIGTSYVSEKELTRTSVVVHVTGES 240
V1              KPNITACKKSANTVEVNFTTSPLGDRYMALIQNTSVIGTSYVSEKELTRTSVVVHVTGES 240

V2              ----------------------------------------------------------
V3              ----------------------------------------------------------
Full            EGAVVQLTPYFHTCGNDCIRQRGTVVRCPQTGVSPQDHERSVLGGWLPLLLSALLVATWV 300
V1              EGAVAQLTPYFHTCGNDCIRQRGTVVRCPQTGVSPQDHERSVLGGWLPLLLSALLVATWV 300

V2              ----------------------------------------------------------
V3              ----------------------------------------------------------
Full            LVAGIYLIWRHERIKKTSFSTTTLLPSLKVLVVYPSEICFHHTVCYFTEFLQNRCRSEVI 360
V1              LVAGIYLIWRH------------------------------VCYFTEFLQNRCRSEVI 328

V2              ----------------------------------------------------------
V3              ----------------------------------------------------------
Full            LEKWQKKKIAEMGPVQWLTTQKQAADKVIFLLSNDNTMCDGTCDKKEGGPCENSQDLFHL 420
V1              LEKWQKKKIAEMGPVQWLTTQKQAADKVIFLLSNDNTMCDGTCDKKEGGPCENSQDLFHL 388

V2              ----------------------------------------------------------
V3              ----------------------------------------------------------
Full            AFNLFCSDLRSQTHLRKYVVVYFREGDIKDSYSALSVCPTYRLTKDATDFCAELLHAKQH 480
V1              AFNLFCSDLRSQTHLRKYVVVYFREGDIKDSYSALSVCPTYRLTKDATDFCAELLHAKQH 448

V2              ------------------
V3              ------------------
Full            VSVGRRSRARHYSCLSL 497
V1              VSVGRRSRARHYSCLSL 465
```

Figure 5.11 predicted amino acid sequence alignment of ovine IL17RB full length and IL17RB splice variant sequences (ClustalW v2.1) with highlighted conserved sequence based on human, mouse and cattle protein sequences.

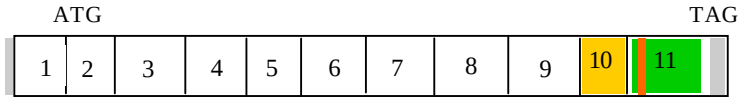
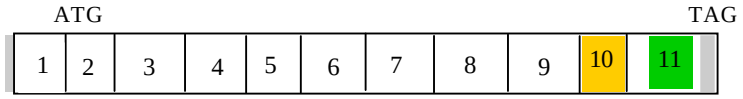
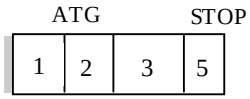
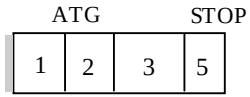
	Predicted protein expression	Amino acids	Predicted Biological Function
Full IL17RB		497	Wild type IL17RB
IL17RBv1		465	Truncated intracellular region. Loss of TRAF6 binding motif
IL17RBv2		96	Short peptide
IL17RBv3		96	Short peptide

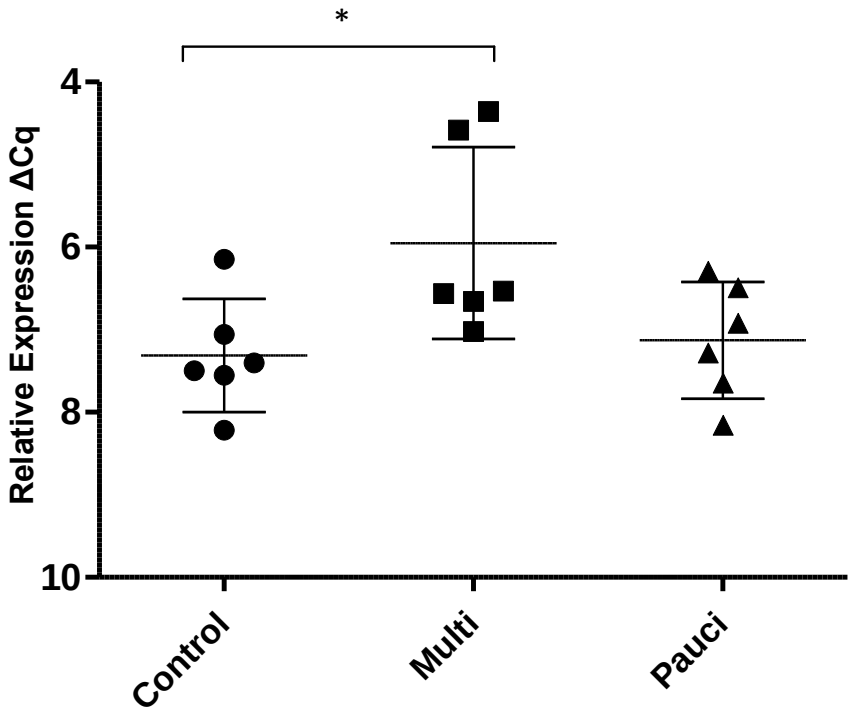
Figure 5.12 Predicted protein expression pattern *IL17RB* splice variants identified in this study. Each box represents the protein encoded by each exon. Grey boxes indicate UTR. The 5' UTR is located 5' upstream of exon 1. The 3' UTR is located at the 3' end of exon 11. The extracellular region spans exons 1-9.The transmembrane region is encoded by exon 10, in yellow. The intracellular region contains the conserved TRAF6 binding motif indicated in red and the SEFIR domain indicated in green. The 'ATG' indicates the initial start codon and 'STOP' indicates stop codons encoded by spice variants that differ from wild type *IL17RB*. Annotations based on comparison of ovine sequence to annotated human and murine sequence.

5.3.3 Expression levels of *IL17RB* gene and *IL17RB* transcript variants

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for the *IL17RB* gene and individual*IL17RB* transcript variants. Statistical analysis was carried out on genes where results were within a measureable range, with relative expression (ΔCq), fold change between groups and P-values shown. Expression levels of total *IL17RB, IL17RBv2* and *IL17RBv3* were measured (figures 5.13 – 5.15); however *IL17RBv1* expression levels were >30 cycles and was considered too lowly expressed for accurate measurement. A positive control assay was carried out for *IL17RBv1* as described in section 2.6.3. Melt curve and cycling analysis for this assay can be found in appendix A11.

Expression levels of total *IL17RB* gene were significantly increased in the multibacillary group when compared to control (2.49 fold, P=0.04) but no significant difference was found when compared with the paucibacillary group (2.08 fold, but P = 0.08). No differential expression was found when comparing *IL17RBv2*or *IL17RBv3*between each of the groups.

A)

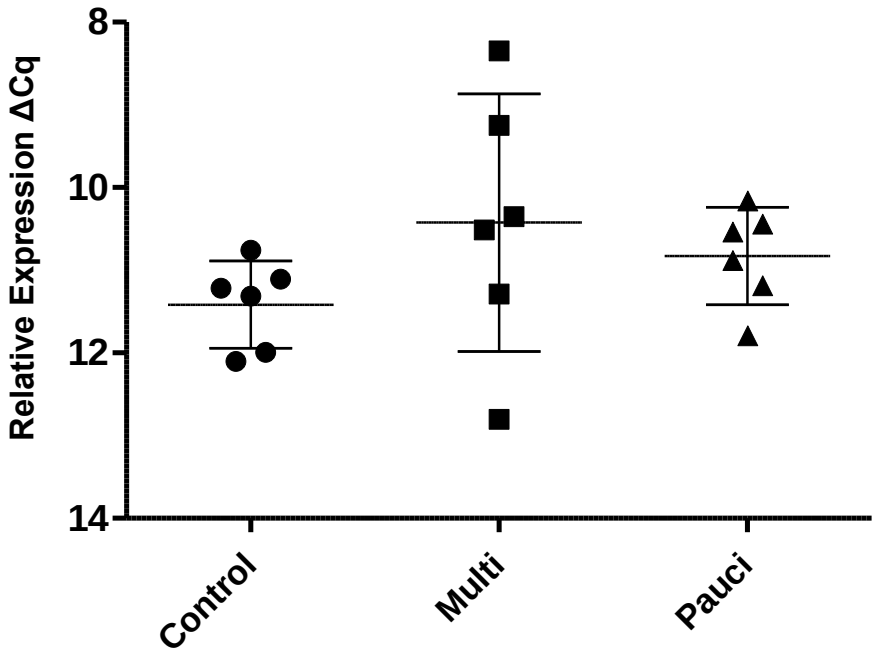


B)

	Fold Change	P-value
M vs P	2.08	0.08
M vs C	2.49	0.04
P vs C	1.19	0.93

Figure 5.13 *IL17RB* RT-qPCR. A) Relative expression of full length *IL17RB* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)

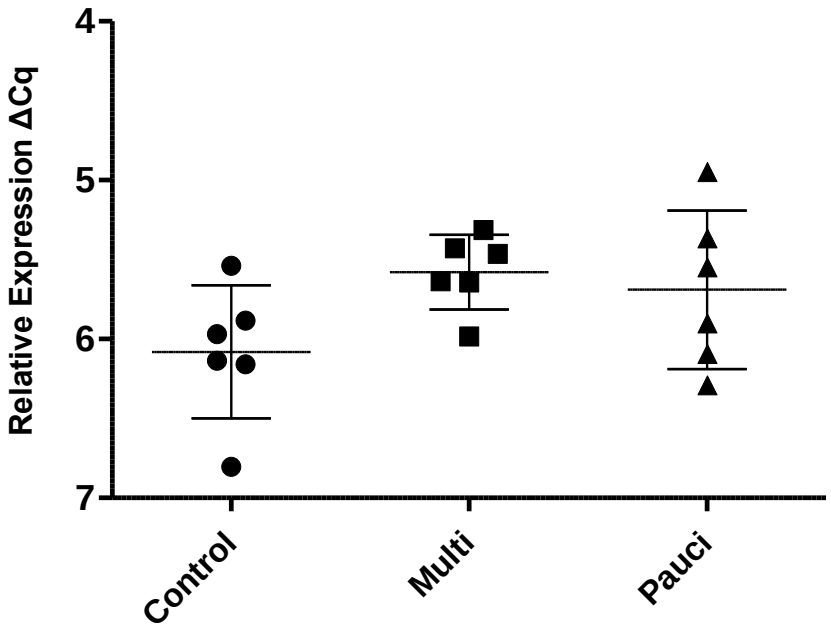


B)

	Fold Change	P-value
M vs P	1.2	0.77
M vs C	1.7	0.24
P vs C	1.4	0.58

Figure 5.14 *IL17RBv2* RT-qPCR. A) Relative expression of *IL17RB*v2 gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)



B)

	Fold Change	P-value
M vs P	1.09	0.88
M vs C	1.6	0.12
P vs C	1.52	0.24

Figure 5.15 *IL17RB*v3 RT-qPCR. A) Relative expression of *IL17RB*v3 gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

5.4 Expression levels of *IL25* transcripts

Primers designed by Gossner *et al.* (Gossner *et al.*, 2011) were used to quantify expression levels of the *IL25* gene. *IL25* gene had expression levels >30 cycles and was considered too lowly expressed for accurate measurement. A positive control assay was carried out for *IL25* as described in section 2.6.3. Melt curve and cycling analysis for this assay can be found in appendix A12.

5.5 Conclusion

- Sequencing of the ovine *IL17RA and IL17RB* genes has identified alternatively spliced variants in the *IL17RB* gene and these variants alter the protein structure of this gene. No transcript variants were found in the ovine *IL17RA* gene
- Increased expression of total *IL17RB* was found in multibacillary animals when compared to control group; however no differential expression was found for the identified *IL17RB*splice variants between each infected group.
- Differential gene expression was identified for *IL17RA* with increased levels of expression found in paucibacillary animals when compared to multibacillary and control.
- *IL25* expression was too low for measurement in the samples used in this study.

Chapter 6

Transcription factor genes in sheep paratuberculosis

6.1 Introduction

Cells control the rate at which genes are transcribed by utilising regulatory DNA regions in the vicinity of where gene transcription is initiated. This regulation occurs in all cells and requires two components i) short stretches of defined DNA sequence (motifs) within the enhancer region and ii) gene regulatory proteins (transcription factors) that specifically recognise and bind to these DNA sequences to initiate transcription . Transcription factor activity is largely influenced by the levels at which the proteins are expressed; T-helper (Th) cell polarisation is controlled by differential expression of master transcription factors as well as cytokines and their receptors (Zhu *et al.*, 2010). Master transcription factors stabilize Th cell phenotypes and regulate the production of cytokines by Th cells and this is critical for the maintenance of appropriate immune responses to specific antigens (Zhu and Paul, 2010). The differentiation and regulation of Th1, Th2 and Th17cell subsets are associated with the master transcription factors; T-bet, GATA3 and RORα/RORγt, respectively.

6.1.1Th1 cell transcriptional regulation

T-bet is an immune cell specific transcription factor of the T-box family, encoded by the *TBX21* gene. T-box family proteins contain conserved sequences of amino acids that bind a specific DNA binding domain known as the T-box domain; all members known to date have a conserved region that bind to the DNA sequence TCACACCT (Wilson and Conlon, 2002). T-bet is the master transcription factor responsible for inducing and stabilizing the Th1 cell lineage, primarily by initiating transcription of IFNγ (Szabo *et al.*, 2002).

T-bet is not expressed in naive CD4+ T cells; its expression is induced in these cells in response to TCR and IL12R-STAT4/IFNγ-STAT1 signalling. Initially T-bet expression is induced independently of IL12R via TCR/IFNγ-STAT1 signalling. IFNγ promotes Th1 development by inducing IL12 expression in macrophages and IL12R expression on T-helper cells that have been activated by antigen; subsequent T-bet induction by IL12R-STAT4 signalling stabilizes the Th1 cell phenotype

(Luckheeram *et al.*, 2012). T-bet directly activates *IFNG* gene expression by binding to the regulatory elements at the *IFNG* locus and recruiting RNA polymerase for gene transcription (Sekimata *et al.*, 2009). IFNγ production by Th1 cells acts in a positive feedback loop to further increase T-bet expression by these cells. Furthermore, ectopic expression of T-bet in differentiated Th17 or Th2 cells results in expression of IFNγ from these cells and their transformation into 'Th1-like cells' (Afkarian *et al.*, 2002). Mice deficient in IL12, IL12RB1 or STAT4 have significantly reduced Th1 cell responses (Gately *et al.*, 1998).

IFNγ produced by Th1 cells is required for efficient control of intracellular pathogens. T-bet deficient CD4+ T cells produce only small amounts of IFNγ under Th1 polarizing conditions and are unable to induce protective Th1 responses during infection with *Mycobacterium tuberculosis, Leishmania major* and *Salmonella enterica* subsp. *enteric*a (Ravindran *et al.*, 2005.Szabo *et al.*, 2002). However, given that over-activation of Th1 cell responses can lead to inflammation associated tissue damage, studies investigating Th1/cell mediated inflammation found that T-bet deficient mice can incur greater resistance to the development of several autoimmune inflammatory disorders including IBD (Neurath *et al.*, 2002), arthritis (Wang *et al.*, 2006) and experimental autoimmune encephalomyelitis (EAE) (Nath *et al.*, 2006).

6.1.2 Th2 cell transcriptional regulation

The GATA3 transcription factor is part of the GATA family of transcription factors and is encoded by the *GATA3* gene (Merika and Orkin, 1993). The GATA family of transcription factors consist of conserved proteins that contain one or two C2-C2-type zinc finger binding motifs that are able to recognise specific DNA sequences required for initiation of transcription of downstream genes (Drevet *et al.*, 1994). Many DNA binding motifs contain zinc atoms as structural components. Zinc plays a key role in maintaining the tertiary structure of several proteins which bind DNA. To do this complexes are formed between zinc ions and cysteine and histidine residues which stabilizes the polypeptide conformation; they often present as protrusions, hence the name zinc 'finger' (Alberts *et al.*, 2002). There are several types of zinc finger motif, those that are recognised by GATA3 consist of four cysteine side chains

(C2C2) (Alberts *et al.*, 2002). These protrusions allow the GATA3 protein to bind the target DNA sequence and stabilize the protein-DNA interaction in order for GATA3 to initiate transcription of downstream Th2 cell associated cytokines (Yagi *et al.*, 2011). GATA3 binds to multiple sites in the *IL4, IL5* and *IL13* locus and initiates transcription of these genes (Zheng and Flavell, 1997.Zhang *et al.*, 1997.Kishikawa *et al.*, 2001). GATA3 is up-regulated by IL4-STAT6 mediated signalling and induces expression of IL4, IL5 and IL13. GATA3 is regarded as the master transcription factor for Th2 cell differentiation as ectopic expression of GATA3 has been shown to cause a shift from Th1 to Th2 cell cytokine production in developing Th1 cells (Ouyang *et al.*, 1998). Furthermore, GATA3 deficient mice fail to produce Th2 cytokines, IL5 and IL13 (Zhu *et al.*, 2004).

GATA3 is associated with the intestinal inflammatory disorder, ulcerative colitis. Ohtani *et al.* found that mucosal expression of GATA3 and STAT4 were significantly increased in the acute phase of paediatric UC when compared with normal, control mucosa (Ohtani *et al.*, 2010). Furthermore, variants and mutations of the *GATA3* gene have been associated with human inflammatory pathologies, including asthma and antibody mediated allergy. Combinations of *GATA3* and *IL13* polymorphisms and their interactions confer increased risk of rhinitis and allergic sensitization at a young age (10 years) (Huebner *et al.*, 2008). A genotyping study found three haplotypes of *GATA3* that were associated with asthma and atopy related phenotypes when compared with control subjects (Pykalainen *et al.*, 2005).

6.1.3 Th17 cell transcriptional regulation

The transcriptional mechanisms by which Th17 cell activity is regulated continues to be investigated, however studies have shown that the transcription factor orphan nuclear receptor-RORγt is associated with expression of Th17 cytokines (Ivanov *et al.*, 2006). RORγ is encoded by the *RORC* gene of which in humans there are two isoforms generated by alternative splicing; isoform 1 (encodes RORγ) is a key regulator of cellular differentiation and metabolism, expressed in many tissues and is mostly associated with metabolism processes (Jetten, 2009.Lee *et al.*, 1995), whereas

isoform 2, also known as *RORC2* (encodes RORγt) is immune specific and plays a key role in the differentiation of uncommitted CD4+ T cells into Th17 cells (Ivanov *et al.*, 2006). The two isoforms differ in the 5' end of the coding region (figure 6.1). RORγt is known to induce Th17 activities synergistically with RORα, encoded by the *RORA* gene. RORγt expression is regulated by STAT3 (Laurence et al 2007) and over expression of RORγt promotes greater Th17 differentiation and subsequent expression of Th17 cytokines including IL23 and IL21 (Ivanov *et al.*, 2006). RORγt deficiency was found to reduce Th17 cell differentiation but did not abolish Th17 cell activity; suggesting that additional factors were involved in Th17 differentiation. This is when further investigation found that Th17 cells highly express orphan nuclear receptor, RORα (Yang *et al.*, 2008). Both RORγt and RORα contain zinc finger binding domains that function to initiate transcription of Th17 associated genes (Yang *et al.*, 2008). These transcription factors promote Th17 cell differentiation and IL17 and IL17F expression in a TGFβ-IL6-STAT3 dependant manner (Hwang, 2010). Furthermore, RORα and RORγt co-expression synergistically drive greater Th17 cell differentiation. Double deficiency of RORα and RORγt results in entirely impaired Th17 cell differentiation *in vitro* and so synergistic expression of RORα and RORγt is considered essential for efficient generation of Th17 cells and subsequent Th17 activities (Yang *et al.*, 2008).

Variants of the *RORA* and *RORC* genes have been associated with disease in humans. It is well reported that Th17 associated molecules including IL17F, IL21, IL23, IL23R and RORγt are upregulated in the inflamed mucosal tissues of IBD patients (Fina *et al.*, 2008.Kamada *et al.*, 2008.Liu *et al.*, 2011). As previously described RORγt directs transcriptional activation of IL17. Rauen *et al.* reported the presence of an isoform of RORγt, lacking exons 5-8 within the encoding gene; expression of this isoform results in repressed expression of *IL17* and *IL21* gene expression in human T cells, reducing further Th17 cytokine expression and it is thought this isoform may act as a negative regulator of RORγt-mediated gene expression (Rauen *et al.*, 2012). In humans the *RORA* gene encodes four isoforms (RORA1-4); a GWAS study identified that the isoform RORA1 may be implicated in susceptibility to asthma (Moffatt *et al.*, 2007). Furthermore, a microarray analysis found that SNPs in exon 1 of *RORA* were significantly associated with susceptibility

to human neovascular age-related macular degeneration (AMD); AMD is of unknown aetiology and results in a loss of vision in older adults caused by damage to the retina (Schaumberg *et al.*, 2010).

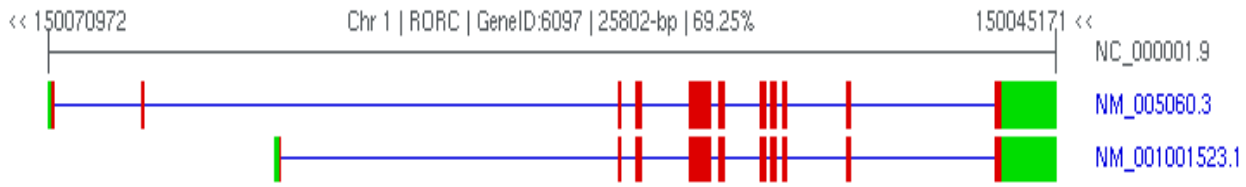


Figure 6.1 Gene maps of human *RORC* isoforms with differing 5' regions (NM_005060.3) and (NM_001001523.1) found on human chromosome 1 (NC_000001.9). Red boxes indicate exons with UTR shown in green boxes (Image captured from NCBI database (www.ncbi.nlm.nih.gov)).

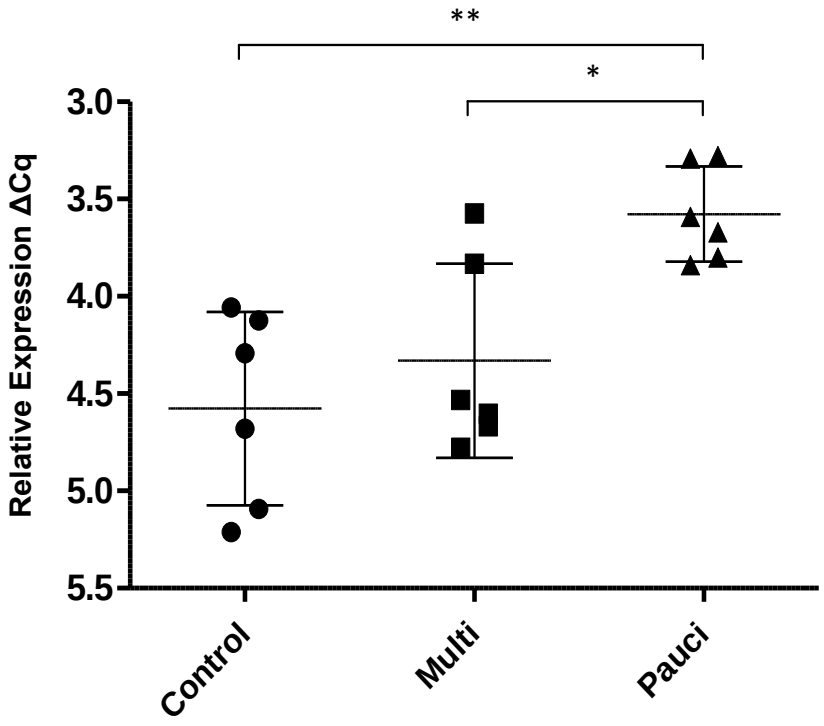
A recent study by Wilkie and colleagues investigated expression levels, and presence of transcript variants, of the *TBX21, GATA3, RORA* and *RORC2* genes in sheep infected with *Teladorsagia circumcinta* (Wilkie *et al.*, 2016). This chapter describes the expression levels of the master transcription factor genes; *TBX21, GATA3, RORA* and *RORC2*, in the two pathological forms of sheep paraTB Transcript variants reported by Wilkie *et al*. are described and the expression levels of individual transcript variants are also reported. The aim of this section was to determine whether the expression of *TBX21, GATA3, RORA* and *RORC2* genes, and transcript variants thereof, are associated with the different pathological outcomes of paraTB.

6.2 TBX21

Sequencing of the *TBX21* gene did not identify any transcript variants; therefore the expression of total *TBX21* (full length transcript) was measured (Wilkie *et al.*, 2016).

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for total *TBX21* gene. Relative expression (ΔCq), fold change between groups and P-values are shown in figure 6.2. Results show a significant 1.88 fold (P=0.02) increase in gene expression in paucibacillary infected sheep when compared with multibacillary sheep and a 2.6 fold increase (P=0.003) when compared to uninfected controls.

A)



B)

	Fold Change	P-value
M vs P	-1.88	0.02
M vs C	1.38	0.59
P vs C	2.6	0.003

Figure 6.2 *TBX21* RT-qPCR. A) Relative expression of total *TBX21* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

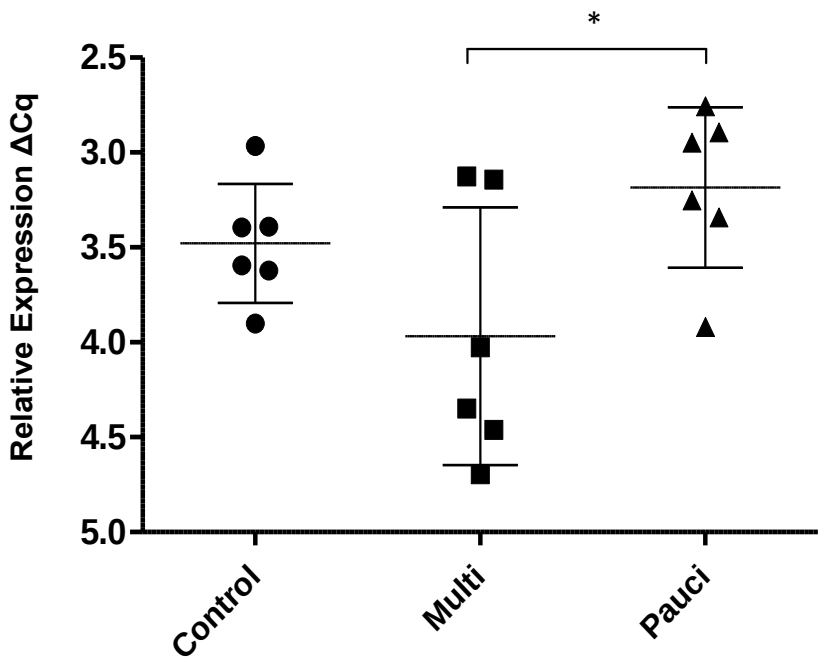
6.3 GATA3

Sequencing of the *GATA3* gene identified a transcript variant of this gene; *GATA3v1* differs to that of the full length gene (*GATA3)* by the deletion of a single codon at the 5' end of exon 3, this represents the deletion of a single amino acid within the coding region (Wilkie *et al.,* 2016)

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for total*GATA3* gene and *GATA3v1*. Relative expression (ΔCq), fold change between groups and P-values shown in figures 6.3 and 6.4.

A 2.08 fold increase of *GATA3* gene expression was found when comparing paucibacillary to multibacillary sheep group, this finding was significant with a P-value of 0.04. No significant difference in gene expression was found between multibacillary, paucibacillary and control groups for *GATA3v1*.

A)

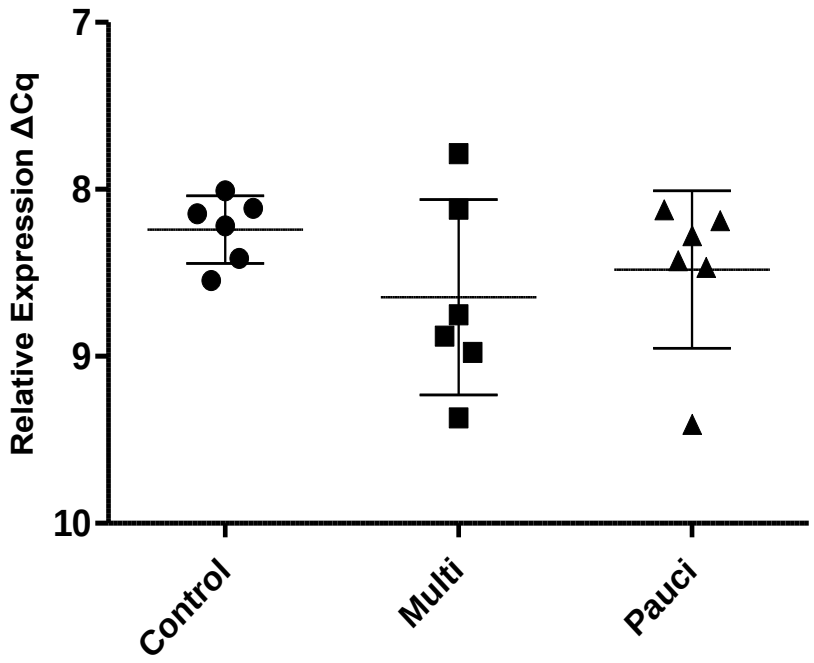


B)

	Fold Change	P-value
M vs P	-2.08	0.04
M vs C	-1.6	0.24
P vs C	1.2	0.57

Figure 6.3 *GATA3* RT-qPCR. A) Relative expression of total *GATA3* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)



B)

	Fold Change	P-value
M vs P	-1.2	0.79
M vs C	-1.5	0.29
P vs C	-1.25	0.64

Figure 6.4*GATA3v1* RT-qPCR. A) Relative expression of total *GATA3v1*gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

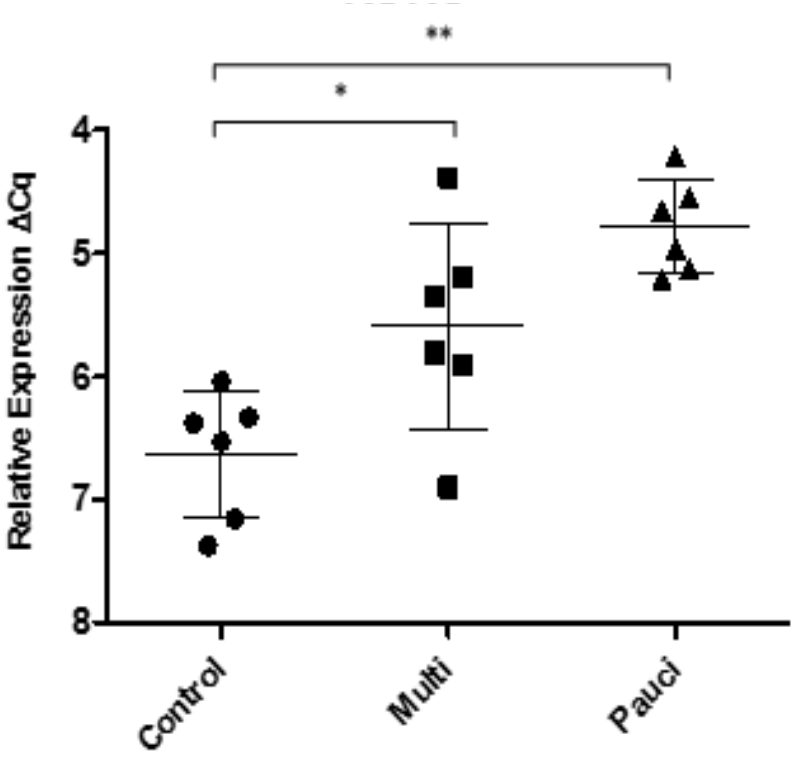
6.4 RORC2

Sequencing of the *RORC2* gene identified a transcript variant of this gene; *RORC2v1* differs to that of the full length gene (*RORC2*) by a 36 bp deletion at the 3' end of exon 7, resulting in 12 amino acid deletion within the coding region (Wilkie *et al.*, 2016).

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for total *RORC2* gene and *RORC2v1*. Relative expression (ΔCq), fold change between groups and P-values shown in figures 6.5 and 6.6.

Significant differences were found between multibacillary and control (2.4 fold, P=0.02) and paucibacillary and control (3.48 fold, P=0.0003) groups for total *RORC2* gene. Similarly for *RORCv1*, a 3.09 fold change (P=0.01) was found when comparing multibacillary to control and a fold change of 3.49 (P=0.0003) when comparing paucibacillary to control. No significant differences were seen in the M vs P comparisons; however total *RORC2* was 1.45 fold greater in M than P group with a P-value of only 0.09.

A)

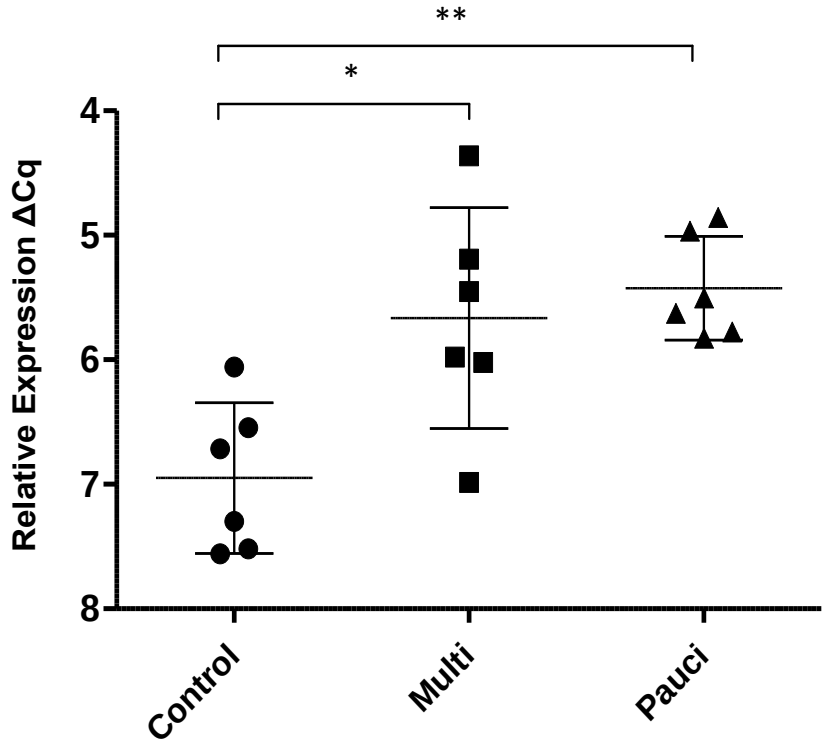


B)

	Fold Change	P-value
M vs P	-1.45	0.09
M vs C	2.4	0.02
P vs C	3.48	0.0003

Figure 6.5 *RORC2* RT-qPCR. A) Relative expression of total *RORC2* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)



B)

	Fold Change	P-value
M vs P	-1.12	0.81
M vs C	3.09	0.01
P vs C	3.49	0.003

Figure 6.6 *RORC2v1* RT-qPCR. A) Relative expression of total *RORC2v1*gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

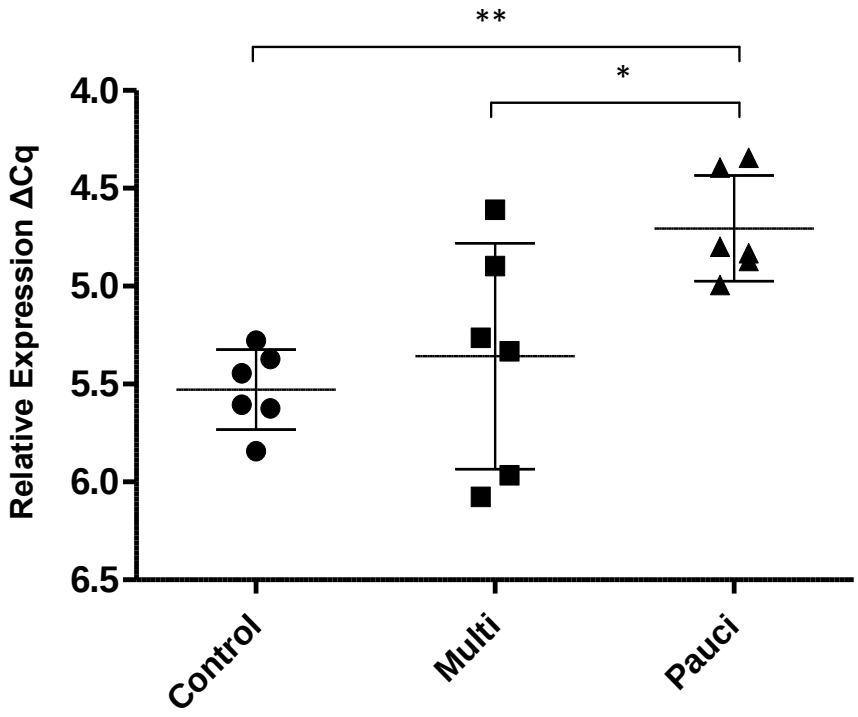
6.5 RORA

Sequencing of the *RORA* gene identified five transcript variants of this gene; all five variants are identical at the 3' end of the gene from the 5' end of exon 6 onwards, however all variants have differential usage of exons 1-6 and therefore have differing 5' UTR lengths. *RORAv2 and RORAv4* have identical translation start sites and subsequently identical protein sequences; whereas *RORAv1, v2, v3* and *v5* have unique start sites (Wilkie *et al.*, 2016).

Real time RT-qPCR analysis was performed in multibacillary, paucibacillary and uninfected control animals for all five *RORA* variants, (*RORAv1, v2, v3, v4 and v5*). Expression levels of *RORAv1, RORAv4* and *RORAv5* were measured figures 6.7 – 6.9; transcript variants *RORAv2* and *RORAv3* were found to have expression levels >30 cycles and were considered below the expression level for accurate measurement. Positive controls were carried out for *RORAv2* and *v3*. Melt curve and cycling analysis for each of these assays can be found in appendix A13 and A14.

A significant fold change was found for *RORAv1* when comparing multibacillary to paucibacillary animals (-1.9, P=0.02) and paucibacillary to control animals (2.49, P=0.005). *RORAv4* was found to have increased expression in both multibacillary (P=0.0008) and paucibacillary (P=0.0004) groups when compared to control with a fold change of 2.78 and 2.89 respectively; however no significant difference was seen for *RORAv5* expression between the three groups.

A)

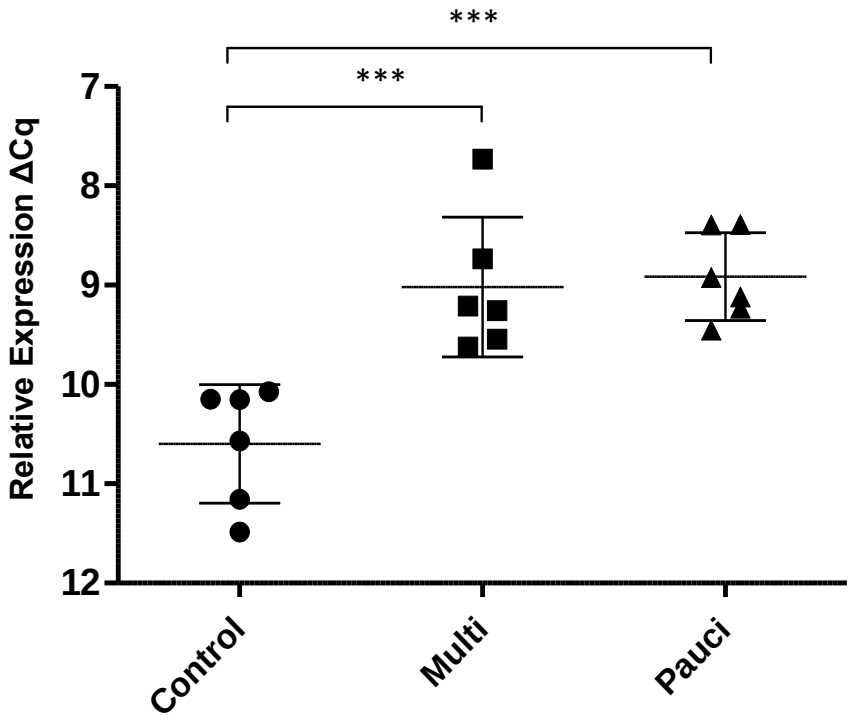


B)

	Fold Change	P-value
M vs P	-1.9	0.02
M vs C	1.2	0.73
P vs C	2.49	0.005

Figure 6.7 *RORAv1* RT-qPCR. A) Relative expression of *RORAv1* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)

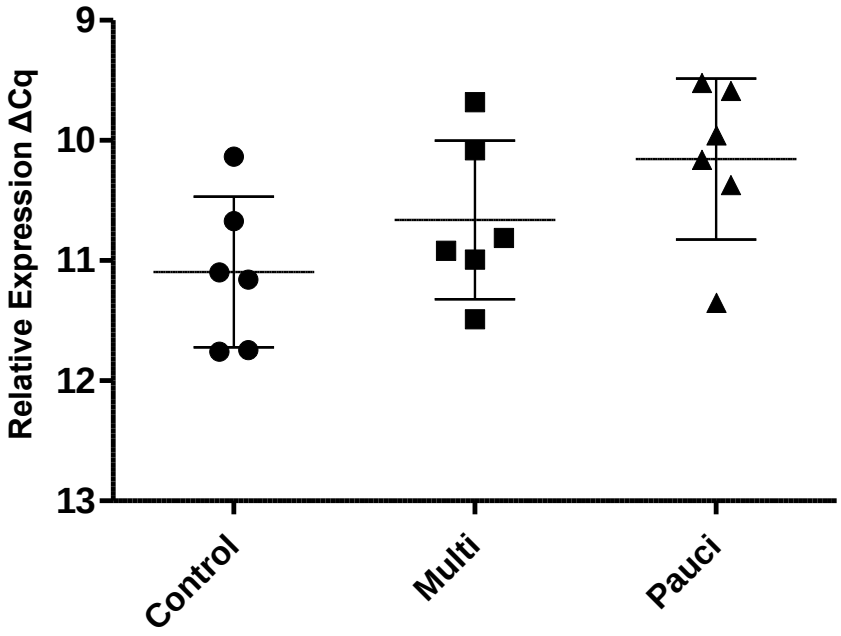


B)

	Fold Change	P-value
M vs P	-1.04	0.95
M vs C	2.78	0.0008
P vs C	2.89	0.0004

Figure 6.8 *RORAv4* RT-qPCR. A) Relative expression of *RORAv4* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

A)



B)

	Fold Change	P-value
M vs P	-1.47	0.39
M vs C	1.8	0.5
P vs C	2.67	0.061

Figure 6.9 *RORAv5* RT-qPCR. A) Relative expression of *RORAv5* gene with B) fold change and P-values for comparisons between multibacillary (M), paucibacillary (P) and control (C) groups.

6.6 Conclusions

- Sequencing of the *GATA3*, *RORC2* and *RORA* transcription factor genes (Wilkie *et al.* 2016) has identified transcript variants, these variants were investigated in both multibacillary and paucibacillary paraTB infected sheep.
- Increased expression of the *TBX21* gene was found when comparing both paucibacillary and multibacillary infected groups to control. No differential expression was identified between the infected groups.
- Increased expression of the *GATA3* gene was found when comparing paucibacillary to multibacillary infected group, No significant difference in gene expression was found between multibacillary, paucibacillary and control groups for *GATA3v1*.
- Increased expression of the *RORC2* gene was found when comparing both paucibacillary and multibacillary infected groups to control. Similarly for *RORCv1*, increased gene expression was identified when comparing both paucibacillary and multibacillary infected groups to control. No significant differences were seen in the paucibacillary vs multibacillary comparisons; however total *RORC2* was 1.45 fold greater in multibacillary than paucibacillary, but not significant.
- Differential expression was identified for *RORAv1* when comparing multibacillary to paucibacillary animals (-1.9, P=0.02) and paucibacillary to control animals (2.49, P=0.005). *RORAv4* was found to have increased expression in both multibacillary and paucibacillary groups when compared to control; however no significant difference was seen for *RORAv5* expression between the three groups.

Chapter 7

Discussion

This project aimed to contribute to the knowledge of sheep immune responses to MAP infection by investigating genes that are known to influence differential T cell activation in response to pathogen. We know that differential T cell activation is associated with the pathological outcomes of human diseases including IBD, TB and leprosy. The two main forms of IBD; Crohn's disease and ulcerative colitis are linked to Th1/Th17 and Th2 responses respectively. Similar immune responses are seen inhuman TB and leprosy with differential pathologies relating to differential Th cell polarization. Leprosy in humans presents as a spectrum between two distinct forms of pathology; tuberculoid (predominantly Th1 response)and lepromatous(predominantly Th2 response) lesions (Modlin, 1994). In human leprosy the host immune response determines the clinical outcome of infection, with a Th1 immune response conferring greater host protection by preventing bacterial replication (Ottenhoff *et al.*, 1997). When comparing these human diseases with paraTB infection in sheep we find that all have multiple clinical manifestations with striking similarities in pathology and Th cell polarization (Ottenhoff *et al.*, 1997. Coussens, 2004. Greenstein, 2003); the pathological forms of paraTB in sheep are defined as paucibacillary (tuberculoid) and multibacillary (lepromatous) and these forms are associated with inflammatory Th1/Th17 and Th2 immune responses respectively. In some species there is a natural progression from paucibacillary to multibacillary form of disease; this is particularly seen in paraTB in cattle (Coussens, 2004). However it is not known what causes the progression from one disease state to another (Coussens, 2001) and this progression is not often seen in sheep (Smeed *et al.*, 2008). The mycobacterial causative agents of TB, leprosy and paraTB induce immune responses via similar mechanisms (Woo and Czuprynski, 2008.Greenstein, 2003); and so host immune responses to these pathogen may be similar between species. When investigating the pathogenesis of these diseases, the underlying questions are the same in both ruminants and humans; what causes the T cell differentiation that leads to different pathological forms of disease? Why do certain infected individuals present with lepromatous lesions whereas others present with tuberculoid lesions? Possible explanations of these questions lie in the genetics of the host. Studies in human disease have associated genes that influence differential T cell activation with differential pathology, and this has provided us with a basis for the

hypothesis that gene pathways associated with differential Th cell activation in the human diseases may also be associated with Th cell activities and subsequent pathology in sheep paraTB.

The polarization of Th cell lineages is influenced by many proteins including; IL23/IL25 cytokines, IL23/IL25 cytokine receptors, and the transcription factors that are associated with T cell differentiation (Zhu *et al.*, 2010). We know that changes in mRNA expression levels and/or transcript variation of the genes that encode these proteins alter Th cell activity in other species and in some cases can alter the outcome of disease; this is discussed in more detail throughout this chapter. Significant knowledge gaps still remain in regard to immune response to paraTB in sheep and the genes that are involved in the immune responses that drive pathology of this disease. Previously, no work has investigated IL23 or IL25 receptor genes and T cell transcription factors in relation to sheep paraTB pathology and these genes were selected for this study to determine whether or not they are associated with differential T cell activation and consequently pathological outcome of sheep paraTB infection, in the same way that they are associated with altered Th cell activation and pathology of disease in other species (Tesmer *et al.*, 2008. Kidd, 2003).

The hypothesis tested was that differential gene expression and/or genetic variants of IL23 and IL25 cytokines, their receptors and the master transcription factors associated with Th1, Th2 and Th17 cell development, contribute to the differential polarization of the Th cell immune responses seen in the two pathological disease forms of MAP infection in sheep. This hypothesis was tested by sequencing the IL23 and IL25 cytokine receptors; IL23R/IL12RB1 and IL17RA/IL17RB to identify alternatively spliced transcript variants; and then by comparing the expression levels of each transcript and variant in multibacillary, paucibacillary and control (uninfected) groups.

There are very few studies that have investigated the role of alternative splicing in relation to paraTB (Casey *et al.*, 2015.Ruiz-Larranaga *et al.*, 2010). Alternative splicing (AS) involves selectively including or removing exon sequences leading to a variety of transcripts from pre-mRNA. Splicing events occur more frequently in genes involved in signalling and transcriptional regulation and immune processes,

including receptors and transcription factors (Martinez and Lynch, 2013). It is thought that these events, and expression of subsequent variants, are valuable in the regulation of immune processes where responses can often occur differently at different time points and where cells require the ability to adapt to environmental changes. It is therefore likely that the key events during immune response to antigen are influenced by AS products of immune stimulating genes (Modrek and Lee, 2002). An aberrant regulation and expression of AS mRNA transcripts is associated with disease in other species and it has been hypothesised that AS is likely to be an associated cause or consequence of disease (Faustino and Cooper, 2003). This information has provided a further reason to investigate the role of AS in the genes associated with immune activation in paraTB. Furthermore, there is an increasing literature on the association of these transcripts in T cell development and responses upon antigen stimulation (Atamas, 1997. Levine, 2008. Talavera *et al.*, 2009). For some genes, the evidence of changes in immune responses linked to spliced transcripts is based solely on analysis of mRNA transcripts, without confirmation that the encoded protein isoforms are expressed and only predicted functions of the protein isoforms can be described; consequently there is debate on whether identified splicing events are of functional relevance (Lynch, 2015.Martinez and Lynch, 2013). However in many other genes there is evidence that multiple protein isoforms are produced from a single gene and the function of these isoforms can have opposing effects on the outcome of immune responses (Chen *et al.*, 2012.Lynch, 2004). There is also increasing evidence of conserved AS transcripts between species and although some transcripts are degraded during mRNA processing, the majority are thought to contribute to a diverse functional protein pool and act as key influencers in the differentiation of immune responses (Kelemen *et al.*, 2013.Chen *et al.*, 2012). Here we have investigated the presence and expression of AS transcripts of cytokine receptor and transcription factor genes in sheep and sought to determine whether they contribute to differential Th cell activity and pathological outcomes of paraTB in sheep. The data collected in this study is based on mRNA transcripts. There is no evidence of whether these transcripts are translated into protein isoforms in sheep, however sequence analysis comparisons between sheep transcripts and those of other species has allowed us to identify similarities in gene structure and predict the role

that these transcripts may play in relation to paraTB in sheep based on what is known in other species.

The present study sought to better define the immune response in the different pathologies seen in paraTB infection in sheep; however there were limitations to consider. The time of infection in diseased animals in the present study is unknown as they all have naturally-acquired infections; although it is presumed that that they were infected as neonates. The tissue samples from infected animals were collected at time of euthanasia when chronic infection was already established, therefore it is possible that observed differences from other studies may be as a result of differences in time since infection. Furthermore some immune responses associated with differential pathology of sheep used in this study may have occurred at an earlier stage of infection. Variation within groups can also be explained by the possibility that these animals have been exposed to MAP for varying lengths of time; therefore animals within groups may be a different stages/severity of disease. In a study of paraTB in cattle, Coussens *et al*. identified differential expression of immune genes in late sub-clinically infected animals compared to controls (Coussens *et al.*, 2004), whereas Robinson *et al*. found no differential expression between early sub-clinically infected deer in comparison to controls; although this may be attributed to species difference it was suggested that the time since initial infection was a major contributing factor to the level immune gene expression (Robinson *et al.*, 2011). In the present study, this limitation was unavoidable due to the selection of animals with naturally-acquired infection. As previously mentioned, many studies of immune response to paraTB are carried out in cattle (Coussens *et al.*, 2002. Coussens *et al.*, 2004. Lee *et al.*, 2001) with fewer studies done in sheep however the pathological forms of disease are better defined in sheep with clearer criteria available to define multibacillary and paucibacillary pathologies (Clarke, 1997), similar to those used in human lepromatous and tuberculoid disease, and so makes sheep a good model when investigating immune responses to paraTB pathologies.

7.1 Sheep selection & definition of pathological forms of paratuberculosis

The initial stage of this study involved careful selection of sheep with defined pathological forms of paraTB. All infected animals were naturally-infected but variability between animals was reduced as much as possible by selecting only females of similar age and breed. Sheep with no/low gastrointestinal nematode infection (as assessed by faecal egg count) were selected where possible, this was important to ensure that the data collected from gene expression assays showed a result from immune response to MAP infection and not a parasite infection. It was critical to correctly categorise each animal into three defined study groups; multibacillary, paucibacillary or uninfected control, by both histological and molecular analysis, so that gene expression results could be assessed in relation to defined paraTB pathologies and characteristics. All of the infected sheep had signs of inflammation in the terminal ileum and loss of villus structure but neither of these signs were observed in control sheep. Molecular methods of paraTB testing are regularly used as *IS*900 is specific to MAP DNA (Collins, 1996) and the PCR techniques used to detect this insertion sequence are considered the 'gold standard' in providing a quick, highly specific marker for precise identification of MAP in samples (Green *et al.*, 1989. Moss *et al.*, 1991). All sheep were tested by two independent *IS*900 PCR assays. By using two PCR assays with different primer sets for the same insertion sequence we can be confident in our confirmation of infected animals that tested positive for *IS*900 or control animals that were *IS*900 negative; given that these animals did not contain MAP DNA in their tissues they were classified as uninfected. Faecal culture can also be used to confirm the presence of MAP in infected animals and although this method is highly specific, it is only moderately sensitive (30-50%) and is considered very slow since the mycobacteria can take up to 16 weeks to grow in culture (Clark *et al.*, 2008) and so in the interest of time this method was not used in the current study. Gross pathology observations of the terminal ileum along with *IS*900 PCR analysis provided confirmation of whether animals were infected or uninfected; following which it was critical to

correctly define the infected animals as having multibacillary or paucibacillary lesions. Lepromatous and tuberculoid forms of human mycobacterial disease present with very distinct differences in histopathology and fundamentally different pathologies; this is also the case in sheep paraTB. Early definition studies of human leprosy have described the histology of infected tissues; with both forms of disease containing granulomatous lesions (Bhat and Prakash, 2012). In humans, the lepromatous form is characterised by a predominant macrophage infiltration with most cells infected with numerous mycobacteria; whereas the tuberculoid form is characterised by lymphocyte infiltration with low numbers of bacteria, multi-nucleated giant cells can also be characteristic of tuberculoid disease (Ridley, 1974. Modlin *et al.*, 1988); however very few giant cells were noted in tuberculoid tissues in the current study. The defining characteristics of these pathologies are mimicked in multibacillary and paucibacillary paraTB in sheep (Clarke and Little, 1996). The tuberculoid form of human leprosy is considered protective and is sometimes 'self-curing' (Rea and Modlin, 1991); which is not the case in end-stage tuberculoid disease in sheep where animals are unable to control intestinal inflammation and eventually die or are euthanized. The defining characteristics described in human disease were used as the basis of disease characterisation for the animals used in the present study. Infected sheep that showed high number of macrophages mostly infected with bacteria in the ileum tissues were defined as multibacillary infection status. Whereas animals with lymphocytic infiltration of the ileum tissues with low or no bacteria present were defined as paucibacillary infection status; assessed by ZN staining and histology. The histological and molecular criteria used to distinguish between multibacillary and paucibacillary forms of disease in the current study have been used in other studies defining the pathological forms of sheep paraTB (Clarke and Little, 1996. Gillan *et al.*, 2010. Smeed *et al.*, 2010) and form a robust foundation of paraTB infection status for subsequent gene expression analysis.

7.2 IL23 axis genes in sheep paratuberculosis

Sequencing of the IL23receptor genes identified five splice variants of *IL23R* and four splice variants of *IL12RB1*. This is not an unusual finding as multiple transcripts of these genes have been identified in other species. This work has provided the first investigation of ovine *IL23R* and *IL12RB1* genes and their splice variants, and the first study of these genes in relation to the two forms of paraTB pathology in sheep.

Translated amino acid sequences of the *IL23R* full length nucleotide sequence and variant nucleotide sequences in sheep were annotated based on known human *IL23R* sequence(Kan *et al.*, 2008). This was done by creating sequence alignments of known human sequences and the sheep transcripts to identify regions of variation that could provide information on the functional consequences of the spliced variants found in sheep. Kan *et al.* described 24 isoforms of human IL23R and discussed the predicted biological function of these variants. When analysing protein sequence comparisons, similar patterns were noted between the human isoforms and sheep IL23R. *IL23Rv1* has an extra 29bp sequence at the start of exon 2 which may be a consequence of alternate splice sites being used, however this variant uses the same start codon as full length *IL23R* and so the coding sequence is the same and so this transcript is thought to possess the same functional abilities as the wild-type (full length) transcript. The insertion 5' of the start codon could possibly have an effect on regulatory functions within the 5' UTR. There are no similar insertions sequences reported in this region of the IL23R protein in other species. Four of the sheep splice variants (*IL23Rv2*-v5) encode truncated protein sequences when compared with wild-type *IL23R* (Figure 4.7). The sheep *IL23Rv2* and *IL23Rv3* only have partial extracellular domains and a loss of transmembrane/intracellular regions. *IL23Rv2* has a 21bp deletion in the 3' end of exon 4; in human transcripts, exon 4 deletion has been identified and encodes a similar truncated protein sequence to that seen in sheep *IL23Rv2;* this transcript in humans is thought to be a target for nonsense mediated decay (NMD), a regulatory degradation of malformed mRNA transcripts, and so will not be translated into protein (Kan *et al.*, 2008). The sheep *IL23Rv3* encodes a truncated protein that only contains exons 1-6 due to a premature stop codon in the extracellular region; a similar isoform consisting of exons 1-7 is found in humans

when a premature stop codon is caused by deletion of exon 6; again this isoform is predicted to be a target of NMD. Given the similarities between these sheep and human isoforms it is predicted that these variants in sheep lead to non-processed transcripts and that these will be of little biological relevance in the immune response to paraTB. *IL23Rv4* encodes a truncated extracellular region of the *IL23R* caused by the deletion of exon 8. This is an in-frame deletion and this isoform still contains the transmembrane domain and intracellular regions; the same isoform with deletion of exon 8 is expressed in humans and is predicted to have an effect on the conformation of the extracellular ligand binding region and/or changes in downstream signalling compared to wild-type IL23R, since the intracellular signalling components are still intact (Kan *et al.*, 2008). *IL23Rv5* does not contain a complete exon 10 sequence (encodes transmembrane domain) and lacks the intracellular region. It is predicted that this transcript encodes a soluble isoform of the IL23R. Evidence shows that soluble receptors can serve as either antagonists or agonists during disease and normal homeostasis (Jones 2005). Comparison of human and sheep *IL23R* shows that exon 10 in sheep corresponds with exon 9 in humans; exon 9 of *IL23R* in humans encodes the transmembrane domain of the receptor protein and deletion of this exon creates a naturally occurring soluble isoform of the IL23R (Δ9). This isoform has been shown to inhibit STAT3 phosphorylation and Th17 cell development and the expression of IL17A and IL17F is abolished in the presence of the Δ9 isoform; however expression levels of RORγt are unaffected (Yu and Gallagher, 2010). This isoform is strongly associated with protection against CD in humans by reducing inflammatory Th17 responses (Gallagher *et al.*, 2012). Zhang and colleagues also identified six alternatively spliced *IL23R* variants in humans and three of these have deleted sequence that encodes the transmembrane domain. These isoforms are expressed in human lung carcinoma tissues and although their precise function is unknown it is though that they may regulate IL23 immune activities (Zhang *et al.*, 2006). Although the IL23R (Δ9) isoform provides protection in human inflammatory disease of the gut, it seems that this variant does not provide the same protection in sheep paraTB as expression levels of *IL23Rv5* were too low to accurately quantify in all study groups of the present study and so we are unable to conclude that it plays an important role in this disease. This was also the case for

IL23Rv4 where expression levels were too low to quantify and so is unlikely to be biologically relevant in sheep paraTB. However these variants may play a role in other diseases. Furthermore it could be suggested that these variants are expressed as a means to regulate Th17 activities during infection and that their low expression confers a loss of Th17 regulation in the animals of this study, subsequently contributing to the severity of inflammation in the gut tissues. The R381Q variant was identified in the predicted protein sequence translated from *IL23Rv1, IL23Rv4* and full length *IL23R* nucleotide sequences. This variant has been shown to reduce Th17 inflammatory responses in humans (Di Meglio *et al.,* 2011); and confers strong protection against CD in human populations (Duerr *et al.,* 2006). This protective variant may play a similar role in sheep however it was not possible to investigate correlation between the presence of this variant and paraTB pathology as the number of samples used in this study are far too few to identify a relationship between SNP/single amino acid change and pathology; in future studies it would be beneficial to analyse the relationship of this variant with paraTB pathology in a larger sample population. The sequence encoding the 'WSXWS' motif, required for receptor activation (Dagil *et al.,* 2012) was identified in all *IL23R* transcripts in this study, however the effect that the spliced regions within the sheep variant transcripts have on the function of this motif is unknown. Investigation into the expression of the proteins encoded by these transcripts is required to determine their functional significance in paraTB infection and sheep immune responses in general.

Human studies have shown that the IL23/IL17 axis plays an important role in the intestinal mucosal immunity, with IL23 being the key cytokine that promotes IL17A production by Th17 cells (Iwakura and Ishigame, 2006. Gaffen *et al.,* 2014). IL23 axis genes were investigated in IBD patients and found that *IL23R*mRNA expression was up-regulated in $CD4^+$ cells isolated from the lamina propria of UC and CD patients; furthermore *IL23A* mRNA expression correlated with increased *IL17A* expression in UC, and increased *IFNG* expression in CD (Kobayashi *et al.*, 2008). The current study investigated expression levels of *IL23R* and found no significant differential expression of total (this assay detects full length and variants) *IL23R* between the three study groups. Expression levels of *IL23A* were also investigated in the three groups; however levels were too low for accurate measurement and so it is

thought that this cytokine may not play an important role in Th17 immune responses at the time these tissues were collected. This was an unusual finding given that *IL17A* mRNA levels are increased in paraTB infection in other species (Robinson *et al.*, 2011) and expression of this cytokine is driven by IL23R-mediated signalling. Interestingly Robinson *et al.* found no differential expression of *IL23R or IL23A* in paraTB infected red deer, even though *IFNG* and *IL17A* mRNA levels were elevated in animals with minimal (largely paucibacillary) or severe (multibacillary) paraTB infection compared to controls. They suggested that Th17 activities may not play an important role in protective immunity to paraTB in deer (Robinson *et al.*, 2011) and this may also be true in sheep paraTB. Furthermore, in another study*IL23A* expression was up-regulated in the ileal mucosa of paucibacillary infected sheep compared to multibacillary infection and control but no differential expression between the three groups was identified in mesenteric lymph node tissues (Gossner *et al.*, 2012b). This could reflect the MAP-induced inflammation at the site of infection and may explain results of low *IL23A* gene expression found in ileocecal lymph node tissues in the current study (figure 4.3) if IL23A production is occurring in the ileum tissues during paraTB infection. The *IL23R* gene is one of the strongest associated genes with CD as determined by GWAS (Roberts *et al.*, 2007); many of the disease associated variants of *IL23R* have also been described in UC, ankylosing spondylitis and psoriasis patients (Duerr *et al.*, 2006). Here we have shown that expression levels of each of the individual *IL23R* splice variants were too low to quantify (≥30 cycles) in the RT-qPCR assays developed. This is a limitation of RT-qPCR when investigating lowly expressed transcripts. Positive control assays using template sample known to contain the variant being measured were used to confirm that the expression levels were an accurate reflection of the mRNA transcripts within the sample and not a technical assay error. It would be possible in future study to increase the volume of initial RNA and subsequently increase the presence of these variants in cDNA templates; however it is questionable whether these variants are of significant biological relevance in the immune response to paraTB infection if expressed at such low levels. IL23 protection against mycobacterial infection was demonstrated in a mouse model; whereby double deficient IL12-IL23-knockout mice were more susceptible to infection than IL12-deficiency alone. This is due to the

presence of the protective IL12p40 subunit and highlights the importance of this subunit in IL23-mediated Th17 responses to mycobacterial infection (Cooper *et al.*, 2002). In contrast the results of the current study show very low expression of *IL23A*, *IL23R* or *IL23R* transcript variants in sheep paraTB. The low expressions of the *IL23R* variants suggests a lack of regulation of Th17 responses if the function of these variants were to modulate IL23 activities as seen in similar IL23R variants in humans. If this were the case it would be expected that *IL23A* levels would be up-regulated and driving Th17 inflammation but expression of this gene was also low in all study groups; perhaps *IL23A* is expressed at greater levels in the gut tissues than in the lymph node tissues as described by Gossner et al. and is driving inflammation at the site of infection (Gossner *et al.*, 2012a). Furthermore, it may be possible that the IL23/IL23R activities have occurred earlier in infection during initial activation of adaptive responses and so are not seen at the late stage of disease that the animals in this study are at.

IL12RB1 is critical for IL12 and IL23- dependent immune responses in humans: CD4+ T cells of individuals with non-functioning *IL12RB1* alleles display defective Th1 and Th17 polarization (Altare *et al.*, 1998). As a result of these combined immunological defects, these individuals are most susceptible to infection with mycobacteria, Salmonella and Candida species (de Beaucoudrey *et al.*, 2010). Furthermore, IL12R- deficient mice are unable to build cell mediated immune responses to sub-lethal doses of Listeria, another intracellular bacterial pathogen (Brombacher *et al.*, 1999). Mutations within the *IL12RB1* gene have been shown to impair the development of IL23-induced Th17 activities and alter host defence against infection (de Beaucoudrey *et al.*, 2010). The majority of deleterious mutations identified in humans with *IL12RB1* deficiency result in premature truncation of the protein, in most cases by encoding premature stop codons prior to the transmembrane domain (van de Vosse *et al.*, 2003) and here we show that similar transcripts are found in sheep.

In this study sequencing of the sheep *IL12RB1* gene identified four splice variants (*IL12RB1v1-4)* that are predicted to encode four distinct proteins (figure 4.12); all of which are derived from variation in mRNA sequence that encodes the extracellular

region of the protein (exons 1-9). The sheep *IL12RB1v1-v4* translated protein sequences lack the 'WSXWS' motif required for receptor activation and ligand binding (Dagil *et al.*, 2012) encoded within exon 7, suggesting that these receptor isoforms are unable to initiate downstream signalling in response to IL12 ligand in sheep but may function to regulate IL12-induced activities (Levine, 2008). *IL12RB1v1* and *v4*both encode two short peptide sequences with truncated extracellular region, no transmembrane domain and no intracellular region; truncated IL12RB1 isoforms lacking transmembrane and/or intracellular regions (*IL12RB1Δtm*) have been identified in humans and mice (van de Vosse *et al.*, 2003.van de Vosse *et al.*, 2013). Studies from both human and mice have shown that expression of the *IL12RB1Δtm* variant and the IL12RB1Δtm isoform is induced in dendritic cells upon stimulation with mycobacteria or mycobacteria-derived products; and in mice promotes *M. tuberculosis*-specific T cell activation, however the mechanisms by which this occurs are unknown (Robinson *et al.*, 2010). Furthermore, the absence of alternatively spliced secreted *IL12RB1Δtm*in mice compromised their ability to control *M. tuberculosis* infection in extra pulmonary organs (Ray *et al.*, 2015). These studies have shown that this spliced *IL12RB1* variant lacking a transmembrane domain is required to initiate effective immune responses to mycobacterial infection in other species. However this may not be the case in the current study as the sheep transcripts differ to those of mouse and human variants in that they have truncated extracellular sequence as well as no transmembrane or intracellular region. The isoforms in mice and humans contain the full extracellular region and so are still able to bind ligand. It is still possible that the sheep isoforms bind IL12/IL23 ligands and modulate their activity during infection, however functional analysis is required in future to determine whether this is possible and also whether these transcripts are even expressed at the protein level or are degraded. The *IL12RB1v1* transcript was too lowly expressed for accurate measurement and so unlikely to be involved in immune responses in the infected animals. *IL12RB1v4* showed no significant differential expression between any of the study groups, although multibacillary sheep had an almost significant 1.87 fold increase in expression compared to paucibacillary ($p=0.08$). A possible explanation for this is that the truncated protein may act as a soluble receptor and modulate

IL12/IL23 activities in these animals by binding to the ligands and preventing downstream signalling. This could contribute to reduced Th1/Th17 responses seen in multibacillary animals. *IL12RB1v2* and *v3* both encode proteins with truncated extracellular regions. Fieschi *et al*. identified a novel form of complete IL12RB1 deficiency associated with expression of a non-functional cell surface expressed IL12RB1. This receptor protein lacks part of the extracellular region and is unable to bind IL12 and IL23, reducing downstream activities of STAT4 and IFNγ (Fieschi *et al.*, 2004); the clinical implications of this isoform are yet to be determined, however several studies have identified IL12RB1 deficiency in patients susceptible to mycobacteria and/or Salmonella infections (van de Vosse *et al.*, 2013). The presence of these variants in sheep would suggest that they play a similar role to that seen in humans and are expressed on the cell surface, bind IL12/IL23 ligands but do not initiate downstream signalling; subsequently reducing Th1/Th17 cell activities. Expression levels of *IL12RB1v2* were too low to quantify; however *IL12RB1v3* was significantly up-regulated in the ICLN of multibacillary compared to paucibacillary animals with a fold change of 2.1 ($p=0.04$), no difference in expression was noted when comparing multibacillary to control or paucibacillary to control. The increased expression levels of this transcript in multibacillary paraTB diseased sheep suggests a lack of IL12RB1 signalling, which contributes to reduced Th1/Th17 activity. Consequently this would contribute to the loss of control of bacterial replication, which is characteristic in multibacillary paraTB infection in sheep.

An investigation of *IL12R* expression in leprosy found that the *IL12RB2* was highly expressed in tuberculoid lesions with far lower expression in lepromatous lesions, however *IL12RB1* expression was similar in both tuberculoid and lepromatous lesions; demonstrating that the inability of lepromatous patients to up-regulate *IL12RB2* correlates with their inability to induce effective Th1 response to *M. leprae* (Kim *et al.*, 2001). Gene expression analysis of total *IL12RB1* in the current study found no differential expression of this gene between paucibacillary, multibacillary and control sheep groups. As is seen in humans, the ability of sheep to mount Th1 responses to MAP infection may in part rely on the presence of the IL12R subunit, IL12RB2. Future studies to investigate expression of *IL12RB2* in relation to sheep paraTB pathology may provide further insight into the role of IL12R signalling in

Th1 cell differentiation and inflammatory responses in tuberculoid and lepromatous forms of paraTB.

The function of many *IL12RB1* splice variants identified in humans is still unknown; however changes in abundance of each isoform are likely to have an effect on the fine tuning of IL12 and IL23 responses to infection and this may also be the case in sheep paraTB. Robinson et al. found that although the alternatively spliced *IL12RB1Δtm*did not contain a transmembrane domain it was still membrane associated and was able to enhance DC migration to local lymph nodes, and as a result increase T cell activation in response to *M. tuberculosis* infection (Robinson *et al.*, 2010). These studies that have been carried out in human and mouse describe *IL1RB1* variants expressed in specific cell types. The current study was carried out using RNA from sheep lymph node tissues containing mixed cell populations, including macrophages and DCs, and so it is unclear of which specific cell types are expressing each of the alternatively spliced transcripts. This information could provide further insight on the function of these variants and could clarify further their role in activation of Th cells in response to paraTB infection.

7.3 IL25 and IL25 receptor in sheep paratuberculosis

IL25 has been shown to induce and maintain Th2 cell mediated inflammation in human and animal studies via IL17RA/IL17RB signalling. Differential expression of the IL25 cytokine and its receptors are associated with chronic inflammatory disorders including asthma, IBD and helminth infection; it has been demonstrated that IL25 deficiency can provide protection from Th2-mediated lung and intestinal inflammatory diseases (Kleinschek *et al.*, 2007. Ballantyne *et al.*, 2007. Wang *et al.*, 2014). Furthermore, there are two main signalling pathways mediated by the shared cytokine receptor subunit IL17RA; IL17A/F and IL25, meaning that this receptor plays a key role in the balance between IL17A/F induced Th1/Th17 cell mediated inflammation and IL25-mediated Th2 inflammation (Gaffen, 2009). IL25 activities are regulated by differential expression of the IL25 receptor subunits, IL17RA and IL17RB (Rickel *et al.*, 2008).

This work has provided the first sequencing investigation of ovine *IL17RA* and *IL17RB* genes, and splice variants of these genes, in relation to the two forms of paraTB pathology in sheep. Previous study has described a skew towards Th2 responses in multibacillary paraTB disease in sheep (Smeed *et al.*, 2007); therefore it is thought that these genes may play a role on the regulation of IL25 induced Th2 activities during paraTB infection. Sequencing of the IL25 receptor genes, *IL17RA* and *IL17RB*, did not identify any variants in the *IL17RA* gene, however three splice variants of *IL17RB* were identified (figure 5.9).Variants of both *IL17RA* and *IL17RB* genes have been identified in other species (Jung *et al.*, 2009.Sohda *et al.*, 2013).

Characterization of the *IL17RA* gene was unable to identify the full length sequence of this gene according to predicted sequence described in the NCBI nucleotide sequence database. Attempts to sequence the 5' end of the gene did not identify a start codon; the start codon of this gene in sheep may be much further 5' than initially anticipated, although attempts to sequence this region using 5'RACE in the current study were unsuccessful (figure 5.4). Furthermore when comparing the sequence obtained in this project to that of RNA-seq study carried out by colleagues (Gossner *et al.* 2016 – in review), there was no indication of a start codon present in

the predicted sequence region from the next generation sequence data. This may be a consequence of the predicted sequence that is available in the NCBI database being inaccurate. Comparison of the translated sheep IL17RA amino acid sequence with the two known human IL17RA isoforms identified the presence of a conserved transmembrane domain (figure 5.7); however sequence at the full length of the sequenced sheep *IL17RA* gene did not match to that of human *IL17RA*. This may be an indication that there is transcript variation within the 5' and 3'ends of the sheep *IL17RA* or it is possibly due to sequence difference between species.

In humans, a nonsense mutation within the *IL17RA*gene results in a premature stop codon and loss of IL17RA protein expression; IL17RA deficiency abolishes the cellular responses to IL17A and IL17F and is associated with susceptibility to chronic mucocutaneous candidiasis disease and *Staphylococcus aureus* infection (Puel *et al.*, 2011). Also the expression of an alternatively spliced soluble form of IL17RA, lacking the transmembrane region, has been reported in various human tissues at the mRNA and protein levels and although it is hypothesised that this isoform regulates IL17 and IL25 activities, the biological functions of this protein are yet to be confirmed (Sohda *et al.*, 2013). Although transcript variants have been identified in humans and mice, in this study there were no variants identified within the sequence obtained for the sheep *IL17RA* gene however further work is required to confirm the full gene sequence of ovine *IL17RA* and so there may be variants present that have not been identified in this study.

Expression levels of *IL17RA* using primers designed within the obtained sequence showed significantly higher expression in paucibacillary animals than both multibacillary and control (M vs P fold change of -2, p=0.001 and P vs C fold change of 2.2, p= 0.02). No differential expression was noted when comparing multibacillary to control group. This result suggests that there is an increase of IL17RA associated activities in the paucibacillary animals compared to multibacillary and control groups. It has been shown that IL17RA expression is required for both IL25 and IL17A/F signalling, however this receptor exhibits greater binding affinity for IL17A over IL25 and IL17F (Ely *et al.*, 2009); therefore higher expression of *IL17RA* in the paucibacillary sheep when compared to multibacillary may be relevant to the

development of Th17 associated inflammation in paucibacillary infection and is consistent with increased levels of IL17A seen in MAP induced lesions in other species (Robinson et al 2011). These results are similar to what is seen in humans where the role of IL17RA-mediated activities has been described in human inflammatory diseases of the gut; Yen et al. found that IL17RA was required for IL17A mediated inflammation in colitis and suggested that blockingIL17RA signalling would be more efficient at ameliorating intestinal inflammation in IBD patients than blocking IL17A alone by simultaneously reducing the activities of IL17A and IL17F (Yen *et al.*, 2006).

Sequencing of the sheep *IL17RB* gene identified three splice variants (*IL17RBv1-v3)* which are predicted to encode three distinct proteins (figure 5.12). The *IL17RB* gene was sequenced in unrelated human DNA samples and 18 genetic polymorphisms were identified; four deletions/insertions and 14 SNPs. Of these only one, an intronic polymorphism, was significantly associated with the development of asthma and conferred protection by regulating IL17RB mediated IL25 activity at the transcriptional level (Jung *et al.*, 2009). This study described a variant within human *IL17RB* that confers protection against Th2 mediated inflammation; and it was thought that variants within the sheep *IL17RB* gene may play a similar role in inflammatory Th2 responses during paraTB infection. Sheep *IL17RBv2* and *v3* gene sequences both contain a deletion of exon 4 and *IL17RBv3* has an insertion of 177bp further downstream at the 5' end of exon 8. Both of these transcripts are predicted to encode identical short extracellular proteins. A change of frame caused by exon 4 deletion leads to a premature stop codon resulting in a truncated extracellular region in both of these variants. Since these variants are predicted to encode proteins that do not contain a transmembrane domain or intracellular region; it is presumed that these variants would be unable to induce IL17RB-mediated cellular signalling but may potentially bind and modulate IL25 activity. This is similar to what is seen in experimental models of IL25 induced inflammation where a soluble IL17RB-Fc protein is used to block/reduce IL25 activities (Rickel et al., 2008). Here I have shown that there is no differential expression present between study groups for *IL17RBv2* or *IL17RBv3.* This result suggests that these transcripts do not play a role in immune response to paraTB infection but may play a role in modulating IL25

activities in other diseases. When comparing sheep IL17RB to that of other species, the TRAF-6 binding motif and SEFIR domain were identified in the sheep IL17RB sequence (appendix section A10). *IL17RBv1* encodes a truncated protein which lacks the TRAF6 binding motif and has only partial sequence of the SEFIR domain within the intracellular region. Maezawa and colleagues reported that TRAF6 plays a critical role in IL25 receptor-mediated gene expression of *IL6, TGFβ* and *CSF3*. They also found that TRAF6 signalling is required for IL25 receptor-mediated NFκB activation and subsequent NFκB-mediated gene expression in mice (Maezawa *et al.*, 2006). It is predicted that *IL17RBv1* in sheep may encode a protein that is unable to bind TRAF6 and induce the TRAF6-associated activities during paraTB infection. Furthermore the conserved SEFIR domain is required as a 'binding surface' for Act1 recruitment during receptor-ligand signalling. Given that this variant only encodes a partial SEFIR domain (as well as loss of the TRAF-6 binding motif), it can be deduced that this receptor isoform may act to regulate IL17RB-mediated activities by binding ligand without carrying out downstream signalling activities. In the case of paraTB infection, this receptor might act to modulate IL25-induced Th2 inflammation during infection. However when investigating gene expression levels, *IL17RBv1* was too lowly expressed for accurate measurement and so is not likely to be of biological relevance in the immune response to paraTB infection but may play a role in other infections in sheep.

Full length *IL17RB* expression levels were significantly higher in the multibacillary group compared to control with a 2.49 fold change ($p=0.04$). Expression was also higher (2.08 fold) in the multibacillary group compared with the paucibacillary but $p=0.08$. Given that *IL17RB* expression is required for IL25 mediated activity(Rickel *et al.*, 2008) these results indicate that there is greater IL25 activity in multibacillary animals than paucibacillary and control sheep. This is consistent with the observed increase of Th2 responses in multibacillary animals (Smeed *et al.*, 2007). In order to confirm whether there is a true increase of *IL17RB* expression in multibacillary compared to paucibacillary groups (not currently significant) this assay should be repeated in a greater sample population in future study. When looking at the results from this assay (figure 5.13) there was a wide spread of values within the multibacillary group; this may be due to differences in time since infection of these

animals, the severity of disease or the different rates of progression of disease in these sheep. In future studies, functional analysis of these splice variants is required to determine the role of each variant during IL25-mediated activities and to determine whether *IL17RBv1* could have an effect on *IL6, TGFβ* and *CSF3* expression in paraTB infected sheep, as seen in other species (Maezawa *et al.*, 2006).

Transgenic expression of IL25 in mice has been shown to increase eosinophilia, serum IgG1 and IgE and gene expression of *IL4, IL5, IL10* and *IL13* in several tissues. In contrast, *IL25* knockout mice have impaired Th2 responses to infections, particularly helminth infections at mucosal surfaces; resulting in susceptibility to infection and chronic inflammation (Zhao *et al.*, 2010. Fallon *et al.*, 2006). In induced pulmonary inflammation models, blocking IL25 activity using a soluble IL17RB-Fc protein or antagonistic antibodies to IL25 results in reduced Th2-mediated inflammation (Camelo *et al.*, 2012. Rickel *et al.*, 2008). IL25 mediates suppression of CD4+ T cell differentiation into Th1 and Th17 cells, whilst promoting Th2 responses. Levels of IL25 were found to be significantly reduced in the sera and mucosal tissues of patients with active IBD in comparison to control subjects and so these individuals were unable to provide protection by suppressing Th1/Th17 inflammation; Su et al have postulated that the anti-inflammatory role of IL25 in IBD is associated with stimulation of IL10 and subsequent inhibition of Th1/Th17 responses rather than stimulation of Th2 responses and have suggested that administration of IL25 supplements in IBD patients exhibiting low IL25 expression may provide a therapy by reducing Th1/Th17 induced inflammation (Su *et al.*, 2013). Investigation of *IL25* in the current study found expression levels were too low to quantify in all groups; a lack of IL25-mediated control of Th1/Th17 responses could be associated with the amplification of intestinal inflammation in paraTB infected sheep, as is described in IBD patients (Su *et al.*, 2013). Another possibility is that there is little expression of IL25 in the lymph node tissues as this cytokine is produced by epithelial tissues during pathogen infection (Peterson and Artis, 2014) and so future studies could look at the expression of this cytokine in the ileum tissues of paraTB infected sheep to provide further insight on the role of this cytokine in this disease.

7.4 Transcription factors in sheep paratuberculosis

The differentiation of Th subsets is finely tuned by the cytokine milieu in which the T cells are activated. A further contributing factor is the expression of T cell polarizing transcription factors. We know that these transcription factors determine the expression of specific genes by Th cells and their expression is required for the maintenance of immune responses by each subset. Sequencing of the sheep master regulator transcription factors that are associated with T cell differentiation by Wilkie *et al.* identified splice variants of these genes (Wilkie *et al.*, 2016) as described in Chapter 6 and the expression levels of these and full length transcripts was investigated in this study to identify whether they are associated with differential T cell polarization and pathological outcomes of paraTB infection in sheep.

7.4.1 *TBX21*

T-bet is required to mount efficient Th1 responses by inducing expression of IFNγ; it has been shown that T-bet deficient mice exhibit profound defects in developing Th1 immunity (Szabo *et al.*, 2002). Also, deficient mice infected with *Mycobacterium avium* complex were highly susceptible to infection and developed severe pulmonary inflammation compared to T-bet-overexpressing and control mice that were able to control infection (Matsuyama *et al.*, 2014). These studies highlight the role of T-bet in controlling mycobacterial infection and developing Th1 responses. Given that Th1 responses are required to control bacterial replication during paraTB infection and that this is typically seen in paucibacillary infection, it was thought that expression and or variants of this gene may be associated with the paucibacillary form of paraTB in sheep. This is similar to what is seen in humans where increased *TBX21*levels are observed in patients with inflammatory diseases of the gut; predominantly inflammation associated with increased IFNγ production (Salvati *et al.*, 2002. Neurath *et al.*, 2002). Only one transcript of this gene has been identified in mice, cattle and humans. SNPs have been identified in human *TBX21*and are associated with mucosal inflammatory diseases; one such SNP was suggested to promote resistance to TB when comparing 466 unrelated pulmonary TB patients to 300 healthy controls (Sanchez *et al.*, 2013) whilst others have identified SNPs associated with airway

hyper- responsiveness in asthma (Raby *et al.,* 2006). Many of these SNPs were identified by GWAS and do not provide information on the function of the variants. Although there are variants present in the *TBX21* gene in other species, no transcript variants were identified in sheep. Expression levels of the full length *TBX21* sequence were measured in the paraTB groups and showed significantly greater expression in paucibacillary animals when compared with both multibacillary and control groups (M vs P fold change of -1.88, p=0.02 and P vs C fold change of 2.6, p= 0.003). No significant difference was noted between expression in multibacillary and control groups. T-bet is known to initiate *IFNG* gene expression and higher expression of *IFNG* has been reported in paucibacillary paraTB infection compared to multibacillary and control sheep (Smeed *et al* 2007). Therefore the increased expression of *TBX21* in paucibacillary animals found in this study is consistent with increased *IFNG* expression observed in other studies (Smeed *et al* 2007, Coussens *et al.,* 2008). My data suggest that the high levels of IFNG and consequent Th1 activation in paucibacillary animals are likely to be as a result of T-bet expression. This result is different to that seen in red deer infected with paraTB where no differential expression of T-bet was noted; deer were experimentally challenged and samples collected from clinically affected animals at 20-25 weeks post challenge (Robinson *et al.*, 2011). These variables, as well as species differences, may account for differences in mRNA expression observed. Similarities were identified between the paucibacillary form of sheep paraTB and CD in humans, in terms of pathology and immune responses. High levels of *IFNG* is associated with increased inflammation within the gut tissues of both diseases (Smeed *et al.*, 2008. Fais *et al.*, 1994). *TBX21* expression in human Th1 cells contributes to intestinal inflammation in CD patients and higher levels of T-bet protein and IFNγ expression were detected in the lamina propria CD4+ T cells of CD patients, but not of UC patients or healthy controls (Matsuyama *et al.*, 2014). These results correspond to the results that are seen in sheep paraTB infection and highlight the role of *TBX21* expression in inducing Th1 mediated inflammation during paraTB infection.

7.4.2 GATA3

GATA3 is required for the differentiation of naive $CD4^+$ T cells to Th2 cells and induces production of IL4, IL5 and IL13 (Ho *et al.*, 2009). The GATA3 transcription factor is required for Th2 responses but has also been shown to represses the development of Th1 responses (Yagi *et al.*, 2011). The expression of *GATA3* is down-regulated during Th1 cell development and up-regulated during Th2 cell development (Yagi *et al.*, 2011). The modulation of Th1/Th2 responses by *GATA3*expression is associated with differential T cell polarization in inflammatory diseases including IBD, asthma and allergy in humans (Ohtani *et al.*, 2010. Barnes, 2008).

Two sequence transcripts were identified in the sheep *GATA3* gene, one full length and one variant sequence. Differences in sequence bet ween the two sheep *GATA3* transcripts are also found in two known human *GATA3* transcripts; however no phenotype consequences have been reported for the human variant transcript. The single deleted amino acid (aa260) does not lie within either of the two conserved zinc finger binding motifs or the DNA-binding motif. It is known that mutation of amino acids within the DNA-binding motif abolishes Th2-cytokine production and DNA binding of GATA3 in other species (Shinnakasu *et al.*, 2006). It is difficult to predict the biological role of the sheep *GATA3* variant without carrying out functional analysis, a possibility for future investigation; however this variant is close to the NH2-zinc finger binding domain and could possibly alter protein structure. Human studies have identified SNPs of the *GATA3* gene which are associated with autoimmune diseases, but no such variants were identified in sheep. Combinations of SNPs identified in the 5' promoter region and 3' UTR of *GATA3* and SNPs identified in *IL13*are associated with rhinitis and atopy, both of which display Th2-mediated inflammation (Huebner *et al.*, 2008). Similarly, another study identified association of SNPs in the *GATA3* gene with inflammatory phenotypes in asthmatic patients compared to control (Pykalainen *et al.*, 2005). These studies highlight the ways in which the *GATA3* gene is associated with Th2-associated immune diseases and demonstrate the effects that variants within this gene can have on the outcome of

disease in humans, in the current study the *GATA3* variant sequence was not found to be associated with the pathological forms of paraTB.

Expression of total *GATA3* was significantly lower only in multibacillary vs paucibacillary comparison, (fold change -2.08, p=0.04). No significant change in expression was noted in either of the diseased groups in comparison to controls; expression levels seem to be lower in multibacillary compared to control (fold change -1.6), although not significantly different (figure 6.3). This result differs to human IBD when considering differential T cell activation. Although an overlap in the spectrum of IBD is noted, in simplistic terms CD seems to exhibit a Th1 profile (similar to paucibacillary paraTB), whereas UC exhibits a Th2 profile (similar to multibacillary paraTB); increased expression of *GATA3* and *IL4* is associated with the pathogenesis of UC compared to CD patients and healthy controls (Ohtani *et al.*, 2010). In the current study the expression of *GATA3* was not associated with a Th2 profile seen with multibacillary paraTB lesions in sheep (Smeed *et al.*, 2007). The reduced *GATA3* expression in multibacillary form of the disease may indicate a loss of Th2 responses in these animals. Similar expression patterns were found in red deer, with significantly higher expression of *IL4* and *GATA3* in minimally diseased (paucibacillary) and control groups compared to severely diseased (multibacillary) group. It is suggested that Th2 responses act to control immunopathology in paraTB infected animals, and loss of these responses leads to the severity of multibacillary clinical disease state (Robinson *et al.*, 2011).

7.4.3 RORA & RORC2

Th17 cell differentiation is regulated by the TFs RORγt (encoded by *RORC2*) and RORα (encoded by *RORA*). Although RORγt is the master regulator of transcription in Th17 cells, both TFs are required for the induction of Th17 cytokine gene expression; blocking either component reduces the expression of *IL17A, IL17F* and *IL23* (Skepner *et al.,* 2014. Yang *et al.,* 2008). Human and mouse studies have highlighted the importance of Th17 cells in the pathogenesis of a range of immune-mediated (Yang *et al.,* 2014. Tesmer *et al.,* 2008).

There are two transcripts of the human *RORC* gene; *RORC1* appears to be expressed broadly and is associated with embryo development and liver function, whereas *RORC2* is expressed exclusively in cells of the immune system and is the functional RORC isoform in Th17 development. Both of these transcripts share sequence similarity from exon 3 through to the last exon; however *RORC2* is shorter than *RORC1* 5' of exon 3 (Lee *et al.,* 1995. Ivanov *et al.,* 2006). Two *RORC2* transcripts were identified in sheep (*RORC2* and *RORC2v1*) and analysis of these sequences found homology between both of the sheep transcripts and two known human transcript variants (Wilkie *et al.,* 2016). *RORC2v1* has a 12 amino acid deletion in the predicted ligand binding domain; the functional activity of this variant is unknown, however it can be predicted that this variant has impaired binding to target ligand which may result in lower expression of downstream Th17 associated genes. A developmental link between Th17 cells and Treg cells, both requiring TGFβ rich environment for differentiation, has been described(Hatton and Weaver, 2009). Antigen-activated naive Tcells exposed to TGFβ are induced to co-express Foxp3and RORC in the presence of IL6, with the expression of RORC being inhibited when Treg differentiation develops as Foxp3expression increases (Hatton and Weaver, 2009). It is possible that differential expression levels of *RORC2* and *RORC2v1*in paraTB infected sheep could play a role modulation of Th17/Treg activity and inflammation in the gut of multibacillary and paucibacillary infection.

Expression level analysis found that total *RORC2* and *RORC2v1* expression levels were significantly increased in both paucibacillary and multibacillary groups in

comparison to the control group. These data imply that RORγt-mediated Th17 activity is increased in response to paraTB infection, regardless of pathology status. RORγt expression is associated with a range of inflammatory diseases in humans; psoriasis is a chronic inflammatory disease of the skin characterised by hyperplasia, monocyte and T lymphocyte infiltration (Manel *et al.*, 2008. Tesmer *et al.*, 2008). This disease is associated with a Th1/Th17 autoimmune response with increased levels of *IFNG* and *IL17A (*Manel *et al.*, 2008*)*; up-regulation of these cytokines has been shown in paraTB infection (Smeed *et al.*, 2007. Robinson *et al.*, 2011). Increased expression of *RORC2* is correlated with increased production of Th17 cytokines and enhanced autoimmune inflammation in psoriatic skin compared to healthy controls (Tesmer *et al.*, 2008). Furthermore, *RORC2* expression was increased in asthmatic children compared to healthy controls (Hamzaoui *et al.*, 2011). Although asthma is typically associated with Th2-inflammatory response, it was suggested that Th17 cells could be associated with sustaining inflammatory processes in asthmatic patients; in this respect up-regulation of *RORC2* in both multibacillary and paucibacillary paraTB infected sheep may be associated with the sustained Th17 driven chronic inflammation in these animals.

In humans, four isoforms of RORA have been identified (a-d) which differ in their N-terminal regions and exhibit distinct DNA-binding properties (Aschrafi *et al.*, 2006). Five transcripts of the sheep *RORA* gene (*RORAv1-v5*) were identified with no 'full length' gene assigned (Wilkie *et al*., 2016). *RORAv1* encodes a protein sequence containing DNA- and ligand-binding domains and a B domain upstream of the DNA-binding domain; this isoform is identical to the human RORA isoform d. *RORAv2* and *v4* derived proteins are identical and contain the conserved ligand-binding domain but do not have a complete DNA-binding domain. The sheep RORAv3andv5 proteins both contain DNA- and ligand-binding domains and identical B domains; these sections are of the same sequence as human isoform a, however the sheep variants differ in their A domain with RORAv3 possessing an A domain of 13 amino acids and RORAv5, an A domain of 44 amino acids; this A domain variation has not been identified in any of the human isoforms. Human RORA isoform a has been associated with brain function in autism spectrum disorder (ASD) (Sarachana and Hu, 2013). In mice, RORA isoforms display the '*staggerer*' phenotype with specific cerebellar

abnormalities showing that RORA plays a key role in the development of the cerebellum (Steinmayr *et al.*, 1998). Furthermore, RORA is suggested to be involved in lipid metabolism and regulation of the cell cycle (Fitzsimmons *et al.*, 2012. Sato *et al.*, 2004). The association of RORA in these human and mice phenotypes does not explain the up-regulation of this gene in sheep paraTB, however it is thought that the up-regulation of this gene along with *RORC2* is key to Th17 development in these animals.

When investigating *RORA* expression levels, *RORAv1* was significantly expressed in paucibacillary sheep compared to multibacillary and control groups. *RORAv4* had significantly higher expression in paucibacillary and multibacillary groups compared to control. The A/B domains of human isoform d, identical to sheep RORAv1, is known to mediate cytoplasmic localization and is a weak transcription factor with a narrow range of targets (Giguere *et al.*, 1994). Furthermore, this isoform may be responsible for residual Th17 cell differentiation in the absence of RORγt (Yang *et al.*, 2008) and so is likely involved in the induction of Th17 cytokine responses during MAP infection. The functional relevance of *RORAv4* derived protein is unknown; however since this transcription factor only contains the conserved ligand-binding domain, it may function to regulate ligand-binding activity (i.e. binding STAT3). No differential expression between the three groups was noted for *RORAv5* and expression levels of *RORAv3* and *RORA2* were too low for quantification, suggesting that these variants may not be biologically relevant in paraTB infection.

It is known that both RORA and RORC are required for Th17 cytokine production and so increased expression of these genes in both paucibacillary and multibacillary paraTB groups may signify a role for Th17 activity in the immune response to MAP. Robinson et al. suggested that Th17 cells do not play a protective role in paraTB infection in deer (Robinson *et al.*, 2011) and so it may be that Th17 responses are having an opposite effect; studies have identified that Th17 responses confer long-lasting, chronic inflammation and that IL17A sustains rather than induces inflammation (Shi *et al.*, 2009.Maione *et al.*, 2009). This could suggest that Th17 cell activity is responsible for prolonged, chronic pathology seen in both multibacillary and paucibacillary paraTB sheep.

7.5 Concluding remarks& future study

In conclusion, this study found that transcript variants of the IL23 and IL25 cytokine receptors were identified in sheep using sequencing methods. Alongside full length transcripts there are five variants of *IL23R*, four variants of *IL12RB1* and three variants of *IL17RB*. Expression analysis of the cytokine receptor genes, transcription factor genes and variants identified *IL12RB1v3, IL17RA, IL17RB, TBX21* and *GATA3* to be significantly associated with different pathologies of sheep paraTB. When considering the results of this study it is suggested that paucibacillary disease is linked to increased Th1/Th17 cell activity and multibacillary animals show a lack of strong Th2 or Th1/Th17 cell activities; this may be a contributing factor to the loss of control of bacterial replication in multibacillary animals. *IL17RB, RORC2, RORC2v1, RORA and RORA* showed increased expression in infected animals compared to uninfected control animals with no differentiation between paucibacillary and multibacillary disease forms.

Alternative splicing is a key regulator of gene expression and protein function; however the functional significance of splice variants identified in this study is unclear. It is estimated that one third of alternatively spliced mRNA transcripts containing premature stop codons are likely to be targets of NMD (Lewis *et al.*, 2003). As mentioned throughout the discussion, in future studies, further analysis of the identified transcript variants to determine whether these mRNA transcripts are present at the protein level would be useful when trying to relate mRNA levels to pathology, particularly in transcripts where differential expression was found to be associated with outcome of paraTB pathology in sheep. However this is a challenge in itself as there are very few methods available for isoform detection and those that are available are difficult to optimise (Stastna and Van Eyk, 2012).

Due to time limitations, this study was unable to investigate individual SNPs in the cytokine and cytokine receptor genes; future study could investigate whether any SNPs present in the gene sequences are associated with pathological forms of paraTB. This work would require a large number of sheep samples which are difficult to obtain when investigating naturally acquired paraTB infection;

controlling variables in such a study would also be challenging when working with outbred animals.

Other considerations for future study are; do these transcripts play a role in the ileum tissues of these animals? It is possible that some of the immune processes are occurring at the site of MAP infection rather than in the lymph node and so it would be useful to investigate this further using similar methods carried out in the current study. Are these variants present in other ruminants? It may be worth investigating these genes and their variants in other ruminant species, particularly cattle, to determine whether the role they play in immune response to paraTB is different to that of sheep. Do they play a role in other diseases? Although many of these transcripts have not been found to be associated with the different pathological forms of paraTB, it may be that they are associated with susceptibility or resistance to other diseases and infections in sheep.

In summary, this study has identified transcript variation in genes that are known to influence T cell responses in other species and has provided gene expression profiles of these gene and their variants. The key findings of this project are summarised in table 7.1. This work contributes to the overall knowledge of sheep immune responses to MAP infection, and has provided a foundation upon which further investigation of these gene pathways in relation to pathological outcomes of disease can be derived.

Table 7.1 Summary of key findings.

Gene /variant	Predicted protein isoform	Gene expression
IL23R	WT	NDE between groups
IL12RB1	WT	NDE between groups
IL12RB1v3	Truncated extracellular domain	M v P: ↑ M v C: NDE P v C: NDE
IL12RB1v4	Short peptide	NDE between groups
IL17RA	WT	M v P: ↓ M v C: NDE P v C: ↑
IL17RB	WT	M v P: NDE M v C: ↑ P v C: NDE
IL17RBv2	Short peptide	NDE between groups
IL17RBv3	Short peptide	NDE between groups
TBX21	WT	M v P: ↓ M v C: NDE P v C: ↑
GATA3	WT	M v P: ↓ M v C: NDE P v C: NDE
GATA3v1	Single amino acid deletion within coding region	NDE between groups
RORC2	WT	M v P: ↓ M v C: NDE P v C: ↑
RORC2v1	12 amino acid deletion within coding region	M v P: ↓ M v C: NDE P v C: ↑
RORAv1	Altered 5' region	M v P: ↓ M v C: NDE P v C: ↑
RORAv4	Altered 5' region	M v P: ↓ M v C: NDE P v C: ↑
RORAv5	Altered 5' region	NDE

Summary of results from both wild type (WT) and variant (V) genes investigated in the current project. Predicted protein isoforms and expression patterns of each gene in multibacillary (M), paucibacillary (P) and control (C) sheep groups are described. NDE: no differential expression.

Bibliography

ABBOT, K., TAYLOR, M. & STUBBINGS, L. 2009. Sustainable worm control strategies for sheep. Packington, UK: Context Publications.

ABEL, L. & CASANOVA, J. L. 2000. Genetic predisposition to clinical tuberculosis: bridging the gap between simple and complex inheritance. *Am J Hum Genet,* 67, 274-7.

AFKARIAN, M., SEDY, J. R., YANG, J., JACOBSON, N. G., CEREB, N., YANG, S. Y., MURPHY, T. L. & MURPHY, K. M. 2002. T-bet is a STAT1-induced regulator of IL-12R expression in naive CD4+ T cells. *Nat Immunol,* 3, 549-57.

ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2002. DNA-Binding Motifs in Gene Regulatory Proteins.

ALTARE, F., DURANDY, A., LAMMAS, D., EMILE, J. F., LAMHAMEDI, S., LE DEIST, F., DRYSDALE, P., JOUANGUY, E., DOFFINGER, R., BERNAUDIN, F., JEPPSSON, O., GOLLOB, J. A., MEINL, E., SEGAL, A. W., FISCHER, A., KUMARARATNE, D. & CASANOVA, J. L. 1998. Impairment of mycobacterial immunity in human interleukin-12 receptor deficiency. *Science,* 280, 1432-5.

ALZUHERRI, H., LITTLE, D. & CLARKE, C. 1997. Altered intestinal macrophage phenotype in ovine paratuberculosis. *Research in veterinary science,* 63, 139-143.

ANDERSSON, A., KOKKOLA, R., WEFER, J., ERLANDSSON-HARRIS, H. & HARRIS, R. A. 2004. Differential macrophage expression of IL-12 and IL-23 upon innate immune activation defines rat autoimmune susceptibility. *J Leukoc Biol,* 76, 1118-24.

ARSENAULT, R. J., MAATTANEN, P., DAIGLE, J., POTTER, A., GRIEBEL, P. & NAPPER, S. 2014. From mouth to macrophage: mechanisms of innate immune subversion by Mycobacterium avium subsp. paratuberculosis. *Vet Res,* 45, 54.

ASCHRAFI, A., MEINDL, N., FIRLA, B., BRANDES, R. P. & STEINHILBER, D. 2006. Intracellular localization of RORalpha is isoform and cell line-dependent. *Biochim Biophys Acta,* 1763, 805-14.

ATAMAS, S. P. 1997. Alternative splice variants of cytokines: Making a list. *Life Sciences,* 61, 1105-1112.

ATAMAS, S. P., CHOI, J., YUROVSKY, V. V. & WHITE, B. 1996. An alternative splice variant of human IL-4, IL-4 delta 2, inhibits IL-4-stimulated T cell proliferation. *J Immunol,* 156, 435-41.

AUTSCHBACH, F., EISOLD, S., HINZ, U., ZINSER, S., LINNEBACHER, M., GIESE, T., LOFFLER, T., BUCHLER, M. W. & SCHMIDT, J. 2005. High prevalence of Mycobacterium avium subspecies paratuberculosis IS900 DNA in gut tissues from individuals with Crohn's disease. *Gut,* 54, 944-9.

AYELE, W. Y., MACHACKOVA, M. & PAVLIK, I. 2001. The transmission and impact of paratuberculosis infection in domestic and wild ruminants. *Veterinarni Medicina,* 46, 205-224.

BALLANTYNE, S. J., BARLOW, J. L., JOLIN, H. E., NATH, P., WILLIAMS, A. S., CHUNG, K. F., STURTON, G., WONG, S. H. & MCKENZIE, A. N. 2007. Blocking IL-25 prevents airway hyperresponsiveness in allergic asthma. *Journal of Allergy and Clinical Immunology,* 120, 1324-1331.

BANNANTINE, J. P. & BERMUDEZ, L. E. 2013. No holes barred: invasion of the intestinal mucosa by Mycobacterium avium subsp. paratuberculosis. *Infect Immun,* 81, 3960-5.

BARLOW, J. L., PEEL, S., FOX, J., PANOVA, V., HARDMAN, C. S., CAMELO, A., BUCKS, C., WU, X., KANE, C. M., NEILL, D. R., FLYNN, R. J., SAYERS, I., HALL, I. P. & MCKENZIE, A. N. J. 2013. IL-33 is more potent than IL-25 in provoking IL-13–producing nuocytes (type 2 innate lymphoid cells) and airway contraction. *Journal of Allergy and Clinical Immunology,* 132, 933-941.

BARNES, P. J. 2008. Role of GATA-3 in allergic diseases. *Curr Mol Med,* 8, 330-4.

BASTIDA, F. & JUSTE, R. 2011. Paratuberculosis control: a review with a focus on vaccination. *Journal of Immune Based Therapies and Vaccines,* 9, 1-17.

BAUERFEIND, R., BENAZZI, S., WEISS, R., SCHLIESSER, T., WILLEMS, H. & BALJER, G. 1996. Molecular characterization of Mycobacterium paratuberculosis isolates from sheep, goats, and cattle by hybridization with a DNA probe to insertion element IS900. *J Clin Microbiol,* 34, 1617-21.

BEGARA-MCGORUM, I., WILDBLOOD, L. A., CLARKE, C. J., CONNOR, K. M., STEVENSON, K., MCINNES, C. J., SHARP, J. M. & JONES, D. G. 1998. Early immunopathological events in experimental ovine paratuberculosis. *Vet Immunol Immunopathol,* 63, 265-87.

BEHR, M. A. & COLLINS, D. M. 2010. *Paratuberculosis: organism, disease, control*, CABI.

BELLAMY, R. 1998. Genetic susceptibility to tuberculosis in human populations. *Thorax,* 53, 588-593.

BERGIN, A. M., BALDER, B., KISHORE, S., SWARD, K., HAHN-ZORIC, M., LOWHAGEN, O., HANSON, L. A. & PADYUKOV, L. 2006. Common variations in the IL4R gene affect splicing and influence natural expression of the soluble isoform. *Hum Mutat,* 27, 990-8.

BERMUDEZ, L. E., PETROFSKY, M., SOMMER, S. & BARLETTA, R. G. 2010. Peyer's Patch-Deficient Mice Demonstrate That Mycobacterium avium subsp. paratuberculosis Translocates across the Mucosal Barrier via both M Cells and Enterocytes but Has Inefficient Dissemination. *Infection and Immunity,* 78, 3570-3577.

BHAT, R. M. & PRAKASH, C. 2012. Leprosy: an overview of pathophysiology. *Interdisciplinary perspectives on infectious diseases,* 2012.

BOUMA, G. & STROBER, W. 2003. The immunological and genetic basis of inflammatory bowel disease. *Nat Rev Immunol,* 3, 521-33.

BROMBACHER, F., DORFMULLER, A., MAGRAM, J., DAI, W. J., KOHLER, G., WUNDERLIN, A., PALMER-LEHMANN, K., GATELY, M. K. & ALBER, G. 1999. IL-12 is dispensable for innate and adaptive immunity against low doses of Listeria monocytogenes. *Int Immunol,* 11, 325-32.

BULL, T. J., HERMON-TAYLOR, J., PAVLIK, I., EL-ZAATARI, F. & TIZARD, M. 2000. Characterization of IS900 loci in Mycobacterium avium subsp. paratuberculosis and development of multiplex PCR typing. *Microbiology,* 146 (Pt 9), 2185-97.

BULL, T. J., MCMINN, E. J., SIDI-BOUMEDINE, K., SKULL, A., DURKIN, D., NEILD, P., RHODES, G., PICKUP, R. & HERMON-TAYLOR, J. 2003. Detection and verification of Mycobacterium avium subsp. paratuberculosis in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. *J Clin Microbiol,* 41, 2915-23.

BURRELLS, C., CLARKE, C. J., COLSTON, A., KAY, J. M., PORTER, J., LITTLE, D. & SHARP, J. M. 1998a. A study of immunological responses of sheep clinically-affected with paratuberculosis (Johne's disease). The relationship of blood, mesenteric lymph node and intestinal lymphocyte responses to gross and microscopic pathology. *Vet Immunol Immunopathol,* 66, 343-58.

BURRELLS, C., CLARKE, C. J., COLSTON, A., KAY, J. M., PORTER, J., LITTLE, D. & SHARP, J. M. 1998b. A study of immunological responses of sheep clinically-affected with paratuberculosis (Johne's disease): The relationship of blood, mesenteric lymph node and intestinal lymphocyte responses to gross and microscopic pathology. *Veterinary Immunology and Immunopathology,* 66, 343-358.

CAMELO, A., BARLOW, J. L., DRYNAN, L. F., NEILL, D. R., BALLANTYNE, S. J., WONG, S. H., PANNELL, R., GAO, W., WRIGLEY, K., SPRENKLE, J. & MCKENZIE, A. N. 2012. Blocking IL-25 signalling protects against gut inflammation in a type-2 model of colitis by suppressing nuocyte and NKT derived IL-13. *J Gastroenterol,* 47, 1198-211.

CARUSO, R., SARRA, M., STOLFI, C., RIZZO, A., FINA, D., FANTINI, M. C., PALLONE, F., MACDONALD, T. T. & MONTELEONE, G. 2009. Interleukin-25 inhibits interleukin-12 production and Th1 cell-driven inflammation in the gut. *Gastroenterology,* 136, 2270-9.

CASEY, M. E., MEADE, K. G., NALPAS, N. C., TARAKTSOGLOU, M., BROWNE, J. A., KILLICK, K. E., PARK, S. D., GORMLEY, E., HOKAMP, K., MAGEE, D. A. & MACHUGH, D. E. 2015. Analysis of the Bovine Monocyte-Derived Macrophage Response to Mycobacterium avium Subspecies Paratuberculosis Infection Using RNA-seq. *Front Immunol,* 6, 23.

CHACON, O., BERMUDEZ, L. E. & BARLETTA, R. G. 2004. Johne's disease, inflammatory bowel disease, and Mycobacterium paratuberculosis. *Annu Rev Microbiol,* 58, 329-63.

CHAFFEY, N. 2003. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. Molecular biology of the cell. 4th edn. *Annals of Botany,* 91, 401.

CHANG, S. H., PARK, H. & DONG, C. 2006. Act1 adaptor protein is an immediate and essential signaling component of interleukin-17 receptor. *Journal of Biological Chemistry,* 281, 35603-35607.

CHEN, L., TOVAR-CORONA, J. M. & URRUTIA, A. O. 2012. Alternative splicing: a potential source of functional innovation in the eukaryotic genome. *International journal of evolutionary biology,* 2012.

CHEN, M. & MANLEY, J. L. 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol,* 10, 741-54.

CHEN, X., ZHOU, B., LI, M., DENG, Q., WU, X., LE, X., WU, C., LARMONIER, N., ZHANG, W., ZHANG, H., WANG, H. & KATSANIS, E. 2007. CD4+CD25+FoxP3+ regulatory T cells suppress Mycobacterium tuberculosis immunity in patients with active disease. *Clinical Immunology,* 123, 50-59.

CHIODINI, R., VAN KRUININGEN, H. & MERKAL, R. 1984. Ruminant paratuberculosis (Johne's disease): the current status and future prospects. *Cornell Vet,* 74, 218 - 262.

CLARK, D., KOZICZKOWSKI, J., RADCLIFF, R., CARLSON, R. & ELLINGSON, J. 2008. Detection of Mycobacterium avium subspecies paratuberculosis: comparing fecal culture versus serum enzyme-linked immunosorbent assay and direct fecal polymerase chain reaction. *Journal of dairy science,* 91, 2620-2627.

CLARK, D. L., JR., ANDERSON, J. L., KOZICZKOWSKI, J. J. & ELLINGSON, J. L. 2006. Detection of Mycobacterium avium subspecies paratuberculosis genetic components in retail cheese curds purchased in Wisconsin and Minnesota by PCR. *Mol Cell Probes,* 20, 197-202.

CLARK, R. G., GRIFFIN, J. F. & MACKINTOSH, C. G. 2010. Johne's disease caused by Mycobacterium avium subsp. paratuberculosis infection in red deer (Cervus elaphus): an histopathological grading system, and comparison of paucibacillary and multibacillary disease. *N Z Vet J,* 58, 90-7.

CLARKE, C. J. 1997. The pathology and pathogenesis of paratuberculosis in ruminants and other species. *J Comp Pathol,* 116, 217-61.

CLARKE, C. J. & LITTLE, D. 1996. The pathology of ovine paratuberculosis: gross and histological changes in the intestine and other tissues. *J Comp Pathol,* 114, 419-37.

CLEMENS, D. L. 1996. Characterization of the Mycobacterium tuberculosis phagosome. *Trends Microbiol,* 4, 113-8.

COLLINS, M. 1996. Diagnosis of paratuberculosis. *Vet Clin North Am: Food Anim Practice,* 12, 357 - 371.

COLLINS, M., WELLS, S. & PETRINI, K. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. *Clin Diag Lab Immun,* 12, 685 - 692.

COLLINS, M. T. 1994. Clinical approach to control of bovine paratuberculosis. *JOURNAL-AMERICAN VETERINARY MEDICAL ASSOCIATION,* 204, 208-208.

COOPER, A. M., DALTON, D. K., STEWART, T. A., GRIFFIN, J. P., RUSSELL, D. G. & ORME, I. M. 1993. Disseminated tuberculosis in interferon gamma gene-disrupted mice. *J Exp Med,* 178, 2243-7.

COOPER, A. M., KIPNIS, A., TURNER, J., MAGRAM, J., FERRANTE, J. & ORME, I. M. 2002. Mice lacking bioactive IL-12 can generate protective, antigen-specific cellular responses to mycobacterial infection only if the IL-12 p40 subunit is present. *J Immunol,* 168, 1322-7.

CORPA, J. M., GARRIDO, J., GARCIA MARIN, J. F. & PEREZ, V. 2000. Classification of lesions observed in natural cases of paratuberculosis in goats. *J Comp Pathol,* 122, 255-65.

COUSSENS, P., LAMONT, E. A., KABARA, E. & SREEVATSAN, S. 2010. 11 Host–Pathogen Interactions and Intracellular Survival of Mycobacterium avium subsp. paratuberculosis. *Paratuberculosis: organism, disease, control*, 109.

COUSSENS, P. M. 2001. Mycobacterium paratuberculosis and the bovine immune system. *Anim Health Res Rev,* 2, 141-61.

COUSSENS, P. M. 2004. Model for immune responses to Mycobacterium avium subspecies paratuberculosis in cattle. *Infect Immun,* 72, 3089-96.

COUSSENS, P. M., COLVIN, C. J., WIERSMA, K., ABOUZIED, A. & SIPKOVSKY, S. 2002. Gene expression profiling of peripheral blood mononuclear cells from cattle infected with Mycobacterium paratuberculosis. *Infect Immun,* 70, 5494-502.

COUSSENS, P. M., SIPKOVSKY, S., MURPHY, B., ROUSSEY, J. & COLVIN, C. J. 2012. Regulatory T cells in cattle and their potential role in bovine paratuberculosis. *Comparative Immunology, Microbiology and Infectious Diseases,* 35, 233-239.

COUSSENS, P. M., VERMAN, N., COUSSENS, M. A., ELFTMAN, M. D. & MCNULTY, A. M. 2004. Cytokine gene expression in peripheral blood mononuclear cells and tissues of cattle infected with Mycobacterium avium subsp. paratuberculosis: evidence for an inherent proinflammatory gene expression pattern. *Infect Immun,* 72, 1409-22.

CRANE, I. J. & FORRESTER, J. V. 2005. Th1 and Th2 lymphocytes in autoimmune disease. *Crit Rev Immunol,* 25, 75-102.

CUMMINGS, J. R., AHMAD, T., GEREMIA, A., BECKLY, J., COONEY, R., HANCOCK, L., PATHAN, S., GUO, C., CARDON, L. R. & JEWELL, D. P. 2007. Contribution of the novel inflammatory bowel disease gene IL23R to disease susceptibility and phenotype. *Inflamm Bowel Dis,* 13, 1063-8.

DAGIL, R., KNUDSEN, MAIKEN J., OLSEN, JOHAN G., O'SHEA, C., FRANZMANN, M., GOFFIN, V., TEILUM, K., BREINHOLT, J. & KRAGELUND, BIRTHE B. 2012. The WSXWS Motif in Cytokine Receptors Is a Molecular Switch Involved in Receptor Activation: Insight from Structures of the Prolactin Receptor. *Structure,* 20, 270-282.

DAMSKER, J. M., HANSEN, A. M. & CASPI, R. R. 2010. Th1 and Th17 cells: adversaries and collaborators. *Ann N Y Acad Sci,* 1183, 211-21.

DAVIS, W. C. & MADSEN-BOUTERSE, S. A. 2012a. Crohn's disease and Mycobacterium avium subsp. paratuberculosis: The need for a study is long overdue. *Veterinary Immunology and Immunopathology,* 145, 1-6.

DAVIS, W. C. & MADSEN-BOUTERSE, S. A. 2012b. Crohn's disease and Mycobacterium avium subsp. paratuberculosis: the need for a study is long overdue. *Vet Immunol Immunopathol,* 145, 1-6.

DE BEAUCOUDREY, L., SAMARINA, A., BUSTAMANTE, J., COBAT, A., BOISSON-DUPUIS, S., FEINBERG, J., AL-MUHSEN, S., JANNIERE, L., ROSE, Y., DE SUREMAIN, M., KONG, X. F., FILIPE-SANTOS, O., CHAPGIER, A., PICARD, C., FISCHER, A., DOGU, F., IKINCIOGULLARI, A., TANIR, G., AL-HAJJAR, S., AL-JUMAAH, S., FRAYHA, H. H., ALSUM, Z., AL-AJAJI, S., ALANGARI, A., AL-GHONAIUM, A., ADIMI, P., MANSOURI, D., BEN-MUSTAPHA, I., YANCOSKI, J., GARTY, B. Z., RODRIGUEZ-GALLEGO, C., CARAGOL, I., KUTUKCULER, N., KUMARARATNE, D. S., PATEL, S., DOFFINGER, R., EXLEY, A., JEPPSSON, O., REICHENBACH, J., NADAL, D., BOYKO, Y., PIETRUCHA, B., ANDERSON, S., LEVIN, M., SCHANDENE, L., SCHEPERS, K., EFIRA, A.,

MASCART, F., MATSUOKA, M., SAKAI, T., SIEGRIST, C. A., FRECEROVA, K., BLUETTERS-SAWATZKI, R., BERNHOFT, J., FREIHORST, J., BAUMANN, U., RICHTER, D., HAERYNCK, F., DE BAETS, F., NOVELLI, V., LAMMAS, D., VERMYLEN, C., TUERLINCKX, D., NIEUWHOF, C., PAC, M., HAAS, W. H., MULLER-FLECKENSTEIN, I., FLECKENSTEIN, B., LEVY, J., RAJ, R., COHEN, A. C., LEWIS, D. B., HOLLAND, S. M., YANG, K. D., WANG, X., WANG, X., JIANG, L., YANG, X., ZHU, C., XIE, Y., LEE, P. P., CHAN, K. W., CHEN, T. X., CASTRO, G., NATERA, I., CODOCEO, A., KING, A., BEZRODNIK, L., DI GIOVANI, D., GAILLARD, M. I., DE MORAES-VASCONCELOS, D., GRUMACH, A. S., DA SILVA DUARTE, A. J., ALDANA, R., ESPINOSA-ROSALES, F. J., BEJAOUI, M., BOUSFIHA, A. A., BAGHDADI, J. E., OZBEK, N., AKSU, G*., et al.* 2010. Revisiting human IL-12Rbeta1 deficiency: a survey of 141 patients from 30 countries. *Medicine (Baltimore),* 89, 381-402.

DE JONG, R., ALTARE, F., HAAGEN, I. A., ELFERINK, D. G., BOER, T., VAN BREDA VRIESMAN, P. J., KABEL, P. J., DRAAISMA, J. M., VAN DISSEL, J. T., KROON, F. P., CASANOVA, J. L. & OTTENHOFF, T. H. 1998. Severe mycobacterial and Salmonella infections in interleukin-12 receptor-deficient patients. *Science,* 280, 1435-8.

DEFRA 2010. Guidance on control of Johne's Disease in Dairy Herds.

http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/documents/johnesguidance.pdf.

DENIS, M., KEEN, D. L., PARLANE, N. A., STORSET, A. K. & BUDDLE, B. M. 2007. Bovine natural killer cells restrict the replication of Mycobacterium bovis in bovine macrophages and enhance IL-12 release by infected macrophages. *Tuberculosis (Edinb),* 87, 53-62.

DENNIS, M. M., REDDACLIFF, L. A. & WHITTINGTON, R. J. 2011. Longitudinal Study of Clinicopathological Features of Johne's Disease in Sheep Naturally Exposed to Mycobacterium avium Subspecies Paratuberculosis. *Veterinary Pathology Online,* 48, 565-575.

DEO, S. S., MISTRY, K. J., KAKADE, A. M. & NIPHADKAR, P. V. 2010. Role played by Th2 type cytokines in IgE mediated allergy and asthma. *Lung India: Official Organ of Indian Chest Society,* 27, 66.

DI CESARE, A., DI MEGLIO, P. & NESTLE, F. O. 2009. The IL-23/Th17 axis in the immunopathogenesis of psoriasis. *J Invest Dermatol,* 129, 1339-50.

DI MEGLIO, P., DI CESARE, A., LAGGNER, U., CHU, C. C., NAPOLITANO, L., VILLANOVA, F., TOSI, I., CAPON, F., TREMBATH, R. C., PERIS, K. & NESTLE, F. O. 2011. The IL23R R381Q gene variant protects against immune-mediated diseases by impairing IL-23-induced Th17 effector response in humans. *PLoS One,* 6, e17160.

DONG, C. 2008. Regulation and pro-inflammatory function of interleukin-17 family cytokines. *Immunol Rev,* 226, 80-6.

DRACHMAN, J. G. & KAUSHANSKY, K. 1995. Structure and function of the cytokine receptor superfamily. *Curr Opin Hematol,* 2, 22-8.

DREVET, J. R., SKEIKY, Y. & IATROU, K. 1994. GATA-type zinc finger motif-containing sequences and chorion gene transcription factors of the silkworm Bombyx mori. *Journal of Biological Chemistry,* 269, 10660-10667.

DUERR, R. H., TAYLOR, K. D., BRANT, S. R., RIOUX, J. D., SILVERBERG, M. S., DALY, M. J., STEINHART, A. H., ABRAHAM, C., REGUEIRO, M., GRIFFITHS, A., DASSOPOULOS, T., BITTON, A., YANG, H., TARGAN, S., DATTA, L. W., KISTNER, E. O., SCHUMM, L. P., LEE, A. T., GREGERSEN, P. K., BARMADA, M. M., ROTTER, J. I., NICOLAE, D. L. & CHO, J. H. 2006. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science,* 314, 1461-3.

EDA, S., BANNANTINE, J. P., WATERS, W. R., MORI, Y., WHITLOCK, R. H., SCOTT, M. C. & SPEER, C. A. 2006. A highly sensitive and subspecies-specific surface antigen enzyme- linked immunosorbent assay for diagnosis of Johne's disease. *Clin Vaccine Immunol,* 13, 837-44.

EISHI, Y., SUGA, M., ISHIGE, I., KOBAYASHI, D., YAMADA, T., TAKEMURA, T., TAKIZAWA, T., KOIKE, M., KUDOH, S., COSTABEL, U., GUZMAN, J., RIZZATO, G., GAMBACORTA, M., DU BOIS, R., NICHOLSON, A. G., SHARMA, O. P. & ANDO, M. 2002. Quantitative analysis of mycobacterial and propionibacterial DNA in lymph nodes of Japanese and European patients with sarcoidosis. *J Clin Microbiol,* 40, 198-204.

ELLINGSON, J. L., ANDERSON, J. L., KOZICZKOWSKI, J. J., RADCLIFF, R. P., SLOAN, S. J., ALLEN, S. E. & SULLIVAN, N. M. 2005. Detection of viable Mycobacterium avium subsp. paratuberculosis in retail pasteurized whole milk by two culture methods and PCR. *J Food Prot,* 68, 966-72.

ELY, L. K., FISCHER, S. & GARCIA, K. C. 2009. Structural basis of receptor sharing by interleukin 17 cytokines. *Nature immunology,* 10, 1245-1251.

EPPLESTON, J. & WINDSOR, P. A. 2007. Lesions attributed to vaccination of sheep with Gudair for the control of ovine paratuberculosis: post farm economic impacts at slaughter. *Aust Vet J,* 85, 129-33.

ERNST, J. D. 1998. Macrophage receptors for Mycobacterium tuberculosis. *Infection and immunity,* 66, 1277-1281.

FAIS, S., CAPOBIANCHI, M. R., SILVESTRI, M., MERCURI, F., PALLONE, F. & DIANZANI, F. 1994. Interferon expression in Crohn's disease patients: increased interferon-gamma and -alpha mRNA in the intestinal lamina propria mononuclear cells. *J Interferon Res,* 14, 235-8.

FALLON, P. G., BALLANTYNE, S. J., MANGAN, N. E., BARLOW, J. L., DASVARMA, A., HEWETT, D. R., MCILGORM, A., JOLIN, H. E. & MCKENZIE, A. N. 2006. Identification of an interleukin (IL)-25-dependent cell population that provides IL-4, IL-5, and IL-13 at the onset of helminth expulsion. *J Exp Med,* 203, 1105-16.

FAUSTINO, N. A. & COOPER, T. A. 2003. Pre-mRNA splicing and human disease. *Genes Dev,* 17, 419-37.

FAWCETT, A. R., GODDARD, P. J., MCKELVEY, W. A., BUXTON, D., REID, H. W., GREIG, A. & MACDONALD, A. J. 1995. Johne's disease in a herd of farmed red deer. *Vet Rec,* 136, 165-9.

FERNÁNDEZ, B., GILARDONI, L. R., JOLLY, A., COLAVECCHIA, S. B., PAOLICCHI, F. A. & MUNDO, S. L. 2012. Detection of Bovine IgG Isotypes in a PPA-ELISA for Johne's Disease Diagnosis in Infected Herds. *Veterinary medicine international,* 2012.

FERWERDA, G., KULLBERG, B. J., DE JONG, D. J., GIRARDIN, S. E., LANGENBERG, D. M., VAN CREVEL, R., OTTENHOFF, T. H., VAN DER MEER, J. W. & NETEA, M. G. 2007. Mycobacterium paratuberculosis is recognized by Toll-like receptors and NOD2. *J Leukoc Biol,* 82, 1011-8.

FIESCHI, C., BOSTICARDO, M., DE BEAUCOUDREY, L., BOISSON-DUPUIS, S., FEINBERG, J., SANTOS, O. F., BUSTAMANTE, J., LEVY, J., CANDOTTI, F. & CASANOVA, J. L. 2004. A novel form of complete IL-12/IL-23 receptor beta1 deficiency with cell surface-expressed nonfunctional receptors. *Blood,* 104, 2095-101.

FILIPE-SANTOS, O., BUSTAMANTE, J., CHAPGIER, A., VOGT, G., DE BEAUCOUDREY, L., FEINBERG, J., JOUANGUY, E., BOISSON-DUPUIS, S., FIESCHI, C., PICARD, C. & CASANOVA, J.-L. 2006. Inborn errors of IL-12/23- and IFN-γ-mediated immunity: molecular, cellular, and clinical features. *Seminars in Immunology,* 18, 347-361.

FINA, D., SARRA, M., FANTINI, M. C., RIZZO, A., CARUSO, R., CAPRIOLI, F., STOLFI, C., CARDOLINI, I., DOTTORI, M., BOIRIVANT, M., PALLONE, F., MACDONALD, T. T. & MONTELEONE, G. 2008. Regulation of gut inflammation and th17 cell response by interleukin-21. *Gastroenterology,* 134, 1038-48.

FITNESS, J., TOSH, K. & HILL, A. V. S. 2002. Genetics of susceptibility to leprosy. *Genes Immun,* 3, 441-453.

FITZSIMMONS, R. L., LAU, P. & MUSCAT, G. E. 2012. Retinoid-related orphan receptor alpha and the regulation of lipid homeostasis. *J Steroid Biochem Mol Biol,* 130, 159-68.

FLYNN, J. L., CHAN, J., TRIEBOLD, K. J., DALTON, D. K., STEWART, T. A. & BLOOM, B. R. 1993. An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection. *J Exp Med,* 178, 2249-54.

FORT, M. M., CHEUNG, J., YEN, D., LI, J., ZURAWSKI, S. M., LO, S., MENON, S., CLIFFORD, T., HUNTE, B., LESLEY, R., MUCHAMUEL, T., HURST, S. D., ZURAWSKI, G., LEACH, M. W., GORMAN, D. M. & RENNICK, D. M. 2001. IL-25 induces IL-4, IL-5, and IL-13 and Th2-associated pathologies in vivo. *Immunity,* 15, 985-95.

FORTIN, A., ABEL, L., CASANOVA, J. L. & GROS, P. 2007. Host genetics of mycobacterial diseases in mice and men: forward genetic studies of BCG-osis and tuberculosis. *Annu Rev Genomics Hum Genet,* 8, 163-92.

FRANZE, E., RIZZO, A., CARUSO, R., PALLONE, F. & MONTELEONE, G. 2011. Interleukin-25 negatively controls pathogenic responses in the gut. *Inflamm Allergy Drug Targets,* 10, 187-91.

GAFFEN, S. L. 2009. Structure and signalling in the IL-17 receptor family. *Nat Rev Immunol,* 9, 556-567.

GAFFEN, S. L., JAIN, R., GARG, A. V. & CUA, D. J. 2014. The IL-23-IL-17 immune axis: from mechanisms to therapeutic testing. *Nature Reviews Immunology,* 14, 585-600.

GALLAGHER, G., YU, R., KAN, S. & MANCINI, G. 2012. SPLICE VARIANTS OF HUMAN IL-23 RECEPTOR (IL-23R) mRNA AND USE OF A DELTA9 ISOFORM IN PREDICTING INFLAMMATORY BOWEL DISEASES. Google Patents.

GATELY, M. K., RENZETTI, L. M., MAGRAM, J., STERN, A. S., ADORINI, L., GUBLER, U. & PRESKY, D. H. 1998. The interleukin-12/interleukin-12-receptor system: role in normal and pathologic immune responses. *Annu Rev Immunol,* 16, 495-521.

GAZOULI, M., ATSAVES, V., MANTZARIS, G., ECONOMOU, M., NASIOULAS, G., EVANGELOU, K., ARCHIMANDRITIS, A. J. & ANAGNOU, N. P. 2008. Role of functional polymorphisms of NRAMP1 gene for the development of Crohn's disease. *Inflamm Bowel Dis,* 14, 1323-30.

GERAGHTY, T., GRAHAM, D. A., MULLOWNEY, P. & MORE, S. J. 2014. A review of bovine Johne's disease control activities in 6 endemically infected countries. *Prev Vet Med,* 116, 1-11.

GIGUERE, V., TINI, M., FLOCK, G., ONG, E., EVANS, R. M. & OTULAKOWSKI, G. 1994. Isoform-specific amino-terminal domains dictate DNA-binding properties of ROR alpha, a novel family of orphan hormone nuclear receptors. *Genes Dev,* 8, 538-53.

GILLAN, S., O'BRIEN, R., HUGHES, A. D. & GRIFFIN, J. F. 2010. Identification of immune parameters to differentiate disease states among sheep infected with Mycobacterium avium subsp. paratuberculosis. *Clin Vaccine Immunol,* 17, 108-17.

GOSSNER, A., PEERS, A., VENTURINA, V. & HOPKINS, J. 2011. Expressed gene sequences of two variants of sheep interleukin-25. *Vet Immunol Immunopathol,* 139, 319-23.

GOSSNER, A. G., VENTURINA, V. M., PEERS, A., WATKINS, C. A. & HOPKINS, J. 2012a. Expression of sheep interleukin 23 (IL23A, alpha subunit p19) in two distinct gastrointestinal diseases. *Vet Immunol Immunopathol,* 150, 118-22.

GOSSNER, A. G., VENTURINA, V. M., PEERS, A., WATKINS, C. A. & HOPKINS, J. 2012b. Expression of sheep interleukin 23 (IL23A, alpha subunit p19) in two distinct gastrointestinal diseases. *Veterinary Immunology and Immunopathology,* 150, 118-122.

GRANT, I. 2005. Zoonotic potential of Mycobacterium avium ssp. paratuberculosis: the current position. *Journal of applied microbiology,* 98, 1282-1293.

GREEN, E. P., TIZARD, M. L., MOSS, M. T., THOMPSON, J., WINTERBOURNE, D. J., MCFADDEN, J. J. & HERMON-TAYLOR, J. 1989. Sequence and characteristics of IS900, an insertion element identified in a human Crohn's disease isolate of Mycobacterium paratuberculosis. *Nucleic Acids Res,* 17, 9063-73.

GREENSTEIN, R. J. 2003. Is Crohn's disease caused by a mycobacterium? Comparisons with leprosy, tuberculosis, and Johne's disease. *Lancet Infect Dis,* 3, 507-14.

GREIG, A., STEVENSON, K., HENDERSON, D., PEREZ, V., HUGHES, V., PAVLIK, I., HINES, M. E., 2ND, MCKENDRICK, I. & SHARP, J. M. 1999. Epidemiological study of paratuberculosis in wild rabbits in Scotland. *J Clin Microbiol,* 37, 1746-51.

GU, C., WU, L. & LI, X. 2013. IL-17 family: cytokines, receptors and signaling. *Cytokine,* 64, 477-85.

HAMZAOUI, A., MAALMI, H., BERRAÏES, A., ABID, H., AMMAR, J. & HAMZAOUI, K. 2011. Transcriptional characteristics of CD4(+) T cells in young asthmatic children: RORC and FOXP3 axis. *Journal of Inflammation Research,* 4, 139-146.

HATTON, R. D. & WEAVER, C. T. 2009. Duality in the Th17-Treg developmental decision. *F1000 Biology Reports,* 1, 5.

HILBORN, E. D., COVERT, T. C., YAKRUS, M. A., HARRIS, S. I., DONNELLY, S. F., RICE, E. W., TONEY, S., BAILEY, S. A. & STELMA, G. N., JR. 2006. Persistence of nontuberculous mycobacteria in a drinking water system after addition of filtration treatment. *Appl Environ Microbiol,* 72, 5864-9.

HO, I. C., TAI, T.-S. & PAI, S.-Y. 2009. GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. *Nat Rev Immunol,* 9, 125-135.

HOSTETTER, J., STEADHAM, E., HAYNES, J., BAILEY, T. & CHEVILLE, N. 2003. Phagosomal maturation and intracellular survival of Mycobacterium avium subspecies paratuberculosis in J774 cells. *Comp Immunol Microbiol Infect Dis,* 26, 269-83.

HUEBNER, M., KIM, D.-Y., EWART, S., KARMAUS, W., SADEGHNEJAD, A. & ARSHAD, S. H. 2008. Patterns of GATA3 and IL13 gene polymorphisms associated with childhood rhinitis and atopy in a birth cohort. *Journal of Allergy and Clinical Immunology,* 121, 408-414.

HUGOT, J.-P., CHAMAILLARD, M., ZOUALI, H., LESAGE, S., CEZARD, J.-P., BELAICHE, J., ALMER, S., TYSK, C., O'MORAIN, C. A., GASSULL, M., BINDER, V., FINKEL, Y., CORTOT, A., MODIGLIANI, R., LAURENT-PUIG, P., GOWER-ROUSSEAU, C., MACRY, J., COLOMBEL, J.-F., SAHBATOU, M. & THOMAS, G. 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature,* 411, 599-603.

HUNTER, C. A. 2005. New IL-12-family members: IL-23 and IL-27, cytokines with divergent functions. *Nature Reviews Immunology,* 5, 521-531.

HURST, S. D., MUCHAMUEL, T., GORMAN, D. M., GILBERT, J. M., CLIFFORD, T., KWAN, S., MENON, S., SEYMOUR, B., JACKSON, C., KUNG, T. T., BRIELAND, J. K., ZURAWSKI, S. M., CHAPMAN, R. W., ZURAWSKI, G. & COFFMAN, R. L. 2002. New IL-17 family members promote Th1 or Th2 responses in the lung: in vivo function of the novel cytokine IL-25. *J Immunol,* 169, 443-53.

HWANG, E. S. 2010. Transcriptional regulation of T helper 17 cell differentiation. *Yonsei medical journal,* 51, 484-491.

IVANOV, II, MCKENZIE, B. S., ZHOU, L., TADOKORO, C. E., LEPELLEY, A., LAFAILLE, J. J., CUA, D. J. & LITTMAN, D. R. 2006. The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell,* 126, 1121-33.

IWAKURA, Y. & ISHIGAME, H. 2006. The IL-23/IL-17 axis in inflammation. *J Clin Invest,* 116, 1218-22.

IWASAKI, A. & MEDZHITOV, R. 2010. Regulation of Adaptive Immunity by the Innate Immune System. *Science,* 327, 291-295.

JETTEN, A. M. 2009. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. *Nucl Recept Signal,* 7, e003.

JOSEFOWICZ, S. Z., LU, L. F. & RUDENSKY, A. Y. 2012. Regulatory T cells: mechanisms of differentiation and function. *Annu Rev Immunol,* 30, 531-64.

JUNG, J. S., PARK, B. L., CHEONG, H. S., BAE, J. S., KIM, J. H., CHANG, H. S., RHIM, T., PARK, J. S., JANG, A. S., LEE, Y. M., KIM, K. U., UH, S. T., NA, J. O., KIM, Y. H., PARK, C. S. & SHIN, H. D. 2009. Association of IL-17RB gene polymorphism with asthma. *Chest,* 135, 1173-80.

KAEVSKA, M., LVONCIK, S., LAMKA, J., PAVLIK, I. & SLANA, I. 2014. Spread of Mycobacterium avium subsp. paratuberculosis through soil and grass on a mouflon (Ovis aries) pasture. *Curr Microbiol,* 69, 495-500.

KAMADA, N., HISAMATSU, T., OKAMOTO, S., CHINEN, H., KOBAYASHI, T., SATO, T., SAKURABA, A., KITAZUME, M. T., SUGITA, A., KOGANEI, K., AKAGAWA, K. S. & HIBI, T. 2008. Unique CD14 intestinal macrophages contribute to the pathogenesis of Crohn disease via IL-23/IFN-gamma axis. *J Clin Invest,* 118, 2269-80.

KAN, S. H., MANCINI, G. & GALLAGHER, G. 2008. Identification and characterization of multiple splice forms of the human interleukin-23 receptor [alpha] chain in mitogen-activated leukocytes. *Genes Immun,* 9, 631-639.

KANG, T. J. & CHAE, G. T. 2001. Detection of Toll-like receptor 2 (TLR2) mutation in the lepromatous leprosy patients. *FEMS Immunol Med Microbiol,* 31, 53-8.

KELEMEN, O., CONVERTINI, P., ZHANG, Z., WEN, Y., SHEN, M., FALALEEVA, M. & STAMM, S. 2013. Function of alternative splicing. *Gene,* 514, 1-30.

KEREN, H., LEV-MAOR, G. & AST, G. 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet,* 11, 345-55.

KIDD, P. 2003. Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease. *Altern Med Rev,* 8, 223-46.

KIM, J., UYEMURA, K., VAN DYKE, M. K., LEGASPI, A. J., REA, T. H., SHUAI, K. & MODLIN, R. L. 2001. A role for IL-12 receptor expression and signal transduction in host defense in leprosy. *J Immunol,* 167, 779-86.

KIM, M. R., MANOUKIAN, R., YEH, R., SILBIGER, S. M., DANILENKO, D. M., SCULLY, S., SUN, J., DEROSE, M. L., STOLINA, M., CHANG, D., VAN, G. Y., CLARKIN, K., NGUYEN, H. Q., YU, Y. B., JING, S., SENALDI, G., ELLIOTT, G. & MEDLOCK, E. S. 2002. Transgenic overexpression of human IL-17E results in eosinophilia, B-lymphocyte hyperplasia, and altered antibody production. *Blood,* 100, 2330-40.

KIMURA, A. & KISHIMOTO, T. 2010. IL-6: regulator of Treg/Th17 balance. *Eur J Immunol,* 40, 1830-5.

KISHIKAWA, H., SUN, J., CHOI, A., MIAW, S. C. & HO, I. C. 2001. The cell type-specific expression of the murine IL-13 gene is regulated by GATA-3. *J Immunol,* 167, 4414-20.

KLEINSCHEK, M. A., OWYANG, A. M., JOYCE-SHAIKH, B., LANGRISH, C. L., CHEN, Y., GORMAN, D. M., BLUMENSCHEIN, W. M., MCCLANAHAN, T., BROMBACHER, F., HURST, S. D., KASTELEIN, R. A. & CUA, D. J. 2007. IL-25 regulates Th17 function in autoimmune inflammation. *The Journal of Experimental Medicine,* 204, 161-170.

KOBAYASHI, T., OKAMOTO, S., HISAMATSU, T., KAMADA, N., CHINEN, H., SAITO, R., KITAZUME, M. T., NAKAZAWA, A., SUGITA, A. & KOGANEI, K. 2008. IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. *Gut,* 57, 1682-1689.

KOO, H. C., PARK, Y. H., HAMILTON, M. J., BARRINGTON, G. M., DAVIES, C. J., KIM, J. B., DAHL, J. L., WATERS, W. R. & DAVIS, W. C. 2004. Analysis of the immune response to Mycobacterium avium subsp. paratuberculosis in experimentally infected calves. *Infect Immun,* 72, 6870-83.

KORN, T., BETTELLI, E., OUKKA, M. & KUCHROO, V. K. 2009. IL-17 and Th17 Cells. *Annu Rev Immunol,* 27, 485-517.

KRUSE, S., FORSTER, J., KUEHR, J. & DEICHMANN, K. A. 1999. Characterization of the membrane-bound and a soluble form of human IL-4 receptor alpha produced by alternative splicing. *Int Immunol,* 11, 1965-70.

KURADE, N., TRIPATHI, B., RAJUKUMAR, K. & PARIHAR, N. 2004. Sequential development of histologic lesions and their relationship with bacterial isolation, fecal shedding, and immune responses during progressive stages of experimental infection of lambs with Mycobacterium avium subsp. paratuberculosis. *Veterinary Pathology Online,* 41, 378-387.

KUSUHARA, K., YAMAMOTO, K., OKADA, K., MIZUNO, Y. & HARA, T. 2007. Association of IL12RB1 polymorphisms with susceptibility to and severity of tuberculosis in Japanese: a gene-based association analysis of 21 candidate genes. *Int J Immunogenet,* 34, 35-44.

LANGRISH, C. L., CHEN, Y., BLUMENSCHEIN, W. M., MATTSON, J., BASHAM, B., SEDGWICK, J. D., MCCLANAHAN, T., KASTELEIN, R. A. & CUA, D. J. 2005. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J Exp Med,* 201, 233-40.

LANKFORD, C. S. & FRUCHT, D. M. 2003. A unique role for IL-23 in promoting cellular immunity. *J Leukoc Biol,* 73, 49-56.

LARSEN, A. B., STALHEIM, O. H., HUGHES, D. E., APPELL, L. H., RICHARDS, W. D. & HIMES, E. M. 1981. Mycobacterium paratuberculosis in the semen and genital organs of a semen-donor bull. *J Am Vet Med Assoc,* 179, 169-71.

LEE, C.-H., COPELAND, N. G., GILBERT, D. J., JENKINS, N. A. & WEI, L.-N. 1995. Genomic structure, promoter identification, and chromosomal mapping of a mouse nuclear orphan receptor expressed in embryos and adult testes. *Genomics,* 30, 46-52.

LEE, H., STABEL, J. R. & KEHRLI, M. E., JR. 2001. Cytokine gene expression in ileal tissues of cattle infected with Mycobacterium paratuberculosis. *Vet Immunol Immunopathol,* 82, 73-85.

LENNARD-JONES, J. 1989. Classification of inflammatory bowel disease. *Scandinavian Journal of Gastroenterology,* 24, 2-6.

LEVINE, S. J. 2008. Molecular mechanisms of soluble cytokine receptor generation. *J Biol Chem,* 283, 14177-81.

LEWIS, B. P., GREEN, R. E. & BRENNER, S. E. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences of the United States of America,* 100, 189-192.

LITTLE, D., ALZUHERRI, H. M. & CLARKE, C. J. 1996. Phenotypic characterisation of intestinal lymphocytes in ovine paratuberculosis by immunohistochemistry. *Vet Immunol Immunopathol,* 55, 175-87.

LIU, Z., YADAV, P. K., XU, X., SU, J., CHEN, C., TANG, M., LIN, H., YU, J., QIAN, J., YANG, P. C. & WANG, X. 2011. The increased expression of IL-23 in inflammatory bowel disease promotes intraepithelial and lamina propria lymphocyte inflammatory responses and cytotoxicity. *J Leukoc Biol,* 89, 597-606.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods,* 25, 402-408.

LOVELL, R., LEVI, M. & FRANCIS, J. 1944. Studies on the survival of Johne's bacilli. *Journal of Comparative Pathology and Therapeutics,* 54, 120-129.

LUCKHEERAM, R. V., ZHOU, R., VERMA, A. D. & XIA, B. 2012. CD4+T Cells: Differentiation and Functions. *Clinical and Developmental Immunology,* 2012, 12.

LUND, S., WALFORD, H. H. & DOHERTY, T. A. 2013. Type 2 Innate Lymphoid Cells in Allergic Disease. *Curr Immunol Rev,* 9, 214-221.

LYNCH, K. W. 2004. Consequences of regulated pre-mRNA splicing in the immune system. *Nature Reviews Immunology,* 4, 931-940.

LYNCH, K. W. 2015. Thoughts on NGS, alternative splicing and what we still need to know. *Rna,* 21, 683-4.

MACATONIA, S. E., HOSKEN, N. A., LITTON, M., VIEIRA, P., HSIEH, C. S., CULPEPPER, J. A., WYSOCKA, M., TRINCHIERI, G., MURPHY, K. M. & O'GARRA, A. 1995. Dendritic cells produce IL-12 and direct the development of Th1 cells from naive CD4+ T cells. *J Immunol,* 154, 5071-9.

MACHACKOVA, M., SVASTOVA, P., LAMKA, J., PARMOVA, I., LISKA, V., SMOLIK, J., FISCHER, O. A. & PAVLIK, I. 2004. Paratuberculosis in farmed and free-living wild ruminants in the Czech Republic (1999–2001). *Veterinary Microbiology,* 101, 225-234.

MAEZAWA, Y., NAKAJIMA, H., SUZUKI, K., TAMACHI, T., IKEDA, K., INOUE, J., SAITO, Y. & IWAMOTO, I. 2006. Involvement of TNF receptor-associated factor 6 in IL-25 receptor signaling. *J Immunol,* 176, 1013-8.

MAIONE, F., PASCHALIDIS, N., MASCOLO, N., DUFTON, N., PERRETTI, M. & D'ACQUISTO, F. 2009. Interleukin 17 sustains rather than induces inflammation. *Biochemical Pharmacology,* 77, 878-887.

MANEL, N., UNUTMAZ, D. & LITTMAN, D. R. 2008. The differentiation of human TH-17 cells requires transforming growth factor-[beta] and induction of the nuclear receptor ROR[gamma]t. *Nat Immunol,* 9, 641-649.

MANNING, E. J. & COLLINS, M. T. 2001. Mycobacterium avium subsp. paratuberculosis: pathogen, pathogenesis and diagnosis. *Rev Sci Tech,* 20, 133-50.

MANNON, P. J., FUSS, I. J., MAYER, L., ELSON, C. O., SANDBORN, W. J., PRESENT, D., DOLIN, B., GOODMAN, N., GRODEN, C., HORNUNG, R. L., QUEZADO, M., YANG, Z., NEURATH, M. F., SALFELD, J., VELDMAN, G. M., SCHWERTSCHLAG, U. & STROBER, W. 2004. Anti-interleukin-12 antibody for active Crohn's disease. *N Engl J Med,* 351, 2069-79.

MARTINEZ, N. M. & LYNCH, K. W. 2013. Control of alternative splicing in immune responses: many regulators, many predictions, much still to learn. *Immunol Rev,* 253, 216-36.

MATSUYAMA, M., ISHII, Y., YAGETA, Y., OHTSUKA, S., ANO, S., MATSUNO, Y., MORISHIMA, Y., YOH, K., TAKAHASHI, S., OGAWA, K., HOGABOAM, C. M. & HIZAWA, N. 2014. Role of Th1/Th17 balance regulated by T-bet in a mouse model of Mycobacterium avium complex disease. *J Immunol,* 192, 1707-17.

MCKENZIE, B. S., KASTELEIN, R. A. & CUA, D. J. 2006. Understanding the IL-23–IL-17 immune pathway. *Trends in immunology,* 27, 17-23.

MERIKA, M. & ORKIN, S. H. 1993. DNA-binding specificity of GATA family transcription factors. *Mol Cell Biol,* 13, 3999-4010.

MILLAR, D., FORD, J., SANDERSON, J., WITHEY, S., TIZARD, M., DORAN, T. & HERMON-TAYLOR, J. 1996. IS900 PCR to detect Mycobacterium paratuberculosis in retail supplies of whole pasteurized cows' milk in England and Wales. *Appl Environ Microbiol,* 62, 3446-52.

MILLER, H. E. & ROBINSON, R. T. 2012. Early control of Mycobacterium tuberculosis infection requires il12rb1 expression by rag1-dependent lineages. *Infect Immun,* 80, 3828-41.

MODLIN, R. L. 1994. Th1-Th2 Paradigm: Insights from Leprosy. *J Investig Dermatol,* 102, 828-832.

MODLIN, R. L., MELANCON-KAPLAN, J., YOUNG, S., PIRMEZ, C., KINO, H., CONVIT, J., REA, T. H. & BLOOM, B. R. 1988. Learning from lesions: patterns of tissue inflammation in leprosy. *Proceedings of the National Academy of Sciences,* 85, 1213-1217.

MODREK, B. & LEE, C. 2002. A genomic view of alternative splicing. *Nat Genet,* 30, 13-9.

MOFFATT, M. F., KABESCH, M., LIANG, L., DIXON, A. L., STRACHAN, D., HEATH, S., DEPNER, M., VON BERG, A., BUFE, A., RIETSCHEL, E., HEINZMANN, A., SIMMA, B., FRISCHER, T.,

WILLIS-OWEN, S. A., WONG, K. C., ILLIG, T., VOGELBERG, C., WEILAND, S. K., VON MUTIUS, E., ABECASIS, G. R., FARRALL, M., GUT, I. G., LATHROP, G. M. & COOKSON, W. O. 2007. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature,* 448, 470-3.

MOLLER, M. & HOAL, E. G. 2010. Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. *Tuberculosis (Edinb),* 90, 71-83.

MORRIS, C. A., HICKEY, S. M. & HENDERSON, H. V. 2006. The effect of Johne's disease on production traits in Romney, Merino and Merino x Romney-cross ewes. *N Z Vet J,* 54, 204-9.

MOSS, M., GREEN, E., TIZARD, M., MALIK, Z. & HERMON-TAYLOR, J. 1991. Specific detection of Mycobacterium paratuberculosis by DNA hybridisation with a fragment of the insertion element IS900. *Gut,* 32, 395-398.

NASER, S. A., GHOBRIAL, G., ROMERO, C. & VALENTINE, J. F. 2004. Culture of Mycobacterium avium subspecies paratuberculosis from the blood of patients with Crohn's disease. *Lancet,* 364, 1039-44.

NATH, N., PRASAD, R., GIRI, S., SINGH, A. K. & SINGH, I. 2006. T-bet is essential for the progression of experimental autoimmune encephalomyelitis. *Immunology,* 118, 384-91.

NAVARRO, J. A., RAMIS, G., SEVA, J., PALLARES, F. J. & SANCHEZ, J. 1998. Changes in lymphocyte subsets in the intestine and mesenteric lymph nodes in caprine paratuberculosis. *J Comp Pathol,* 118, 109-21.

NAZARETH, N., MAGRO, F., APPELBERG, R., SILVA, J., GRACIO, D., COELHO, R., CABRAL, J. M., ABREU, C., MACEDO, G., BULL, T. J. & SARMENTO, A. 2015. Increased viability but decreased culturability of Mycobacterium avium subsp. paratuberculosis in macrophages from inflammatory bowel disease patients under Infliximab treatment. *Med Microbiol Immunol,* 204, 647-56.

NEILL, S. D., SKUCE, R. A. & POLLOCK, J. M. 2005. Tuberculosis--new light from an old window. *J Appl Microbiol,* 98, 1261-9.

NEURATH, M. F. 2014. Cytokines in inflammatory bowel disease. *Nat Rev Immunol,* 14, 329-342.

NEURATH, M. F., WEIGMANN, B., FINOTTO, S., GLICKMAN, J., NIEUWENHUIS, E., IIJIMA, H., MIZOGUCHI, A., MIZOGUCHI, E., MUDTER, J., GALLE, P. R., BHAN, A., AUTSCHBACH, F., SULLIVAN, B. M., SZABO, S. J., GLIMCHER, L. H. & BLUMBERG, R. S. 2002. The transcription factor T-bet regulates mucosal T cell activation in experimental colitis and Crohn's disease. *J Exp Med,* 195, 1129-43.

NICOL, L., GOSSNER, A., WATKINS, C., CHIANINI, F., DALZIEL, R. & HOPKINS, J. 2016. Variations in IL-23 and IL-25 receptor gene structure, sequence and expression associated with the two disease forms of sheep paratuberculosis. *Veterinary Research,* 47, 1-9.

NIELSEN, S., THAMSBORG, S. & HOUE, H. 2000. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. *Prev Vet Med,* 44, 1 - 7.

NIELSEN, S. S. & TOFT, N. 2009. A review of prevalences of paratuberculosis in farmed animals in Europe. *Prev Vet Med,* 88, 1-14.

NOBEN-TRAUTH, N., HU-LI, J. & PAUL, W. E. 2002. IL-4 secreted from individual naive CD4+ T cells acts in an autocrine manner to induce Th2 differentiation. *Eur J Immunol,* 32, 1428-33.

NORTH, R. J. & JUNG, Y. J. 2004. Immunity to tuberculosis. *Annu Rev Immunol,* 22, 599-623.

NOVATCHKOVA, M., LEIBBRANDT, A., WERZOWA, J., NEUBUSER, A. & EISENHABER, F. 2003. The STIR-domain superfamily in signal transduction, development and immunity. *Trends Biochem Sci,* 28, 226-9.

NUGENT, G., WHITFORD, E. J., HUNNAM, J. C., WILSON, P. R., CROSS, M. & DE LISLE, G. W. 2011. Mycobacterium avium subsp. paratuberculosis infection in wildlife on three deer farms with a history of Johne's disease. *N Z Vet J,* 59, 293-8.

O'GARRA, A., REDFORD, P. S., MCNAB, F. W., BLOOM, C. I., WILKINSON, R. J. & BERRY, M. P. 2013. The immune response in tuberculosis. *Annu Rev Immunol,* 31, 475-527.

OGURA, Y., BONEN, D. K., INOHARA, N., NICOLAE, D. L., CHEN, F. F., RAMOS, R., BRITTON, H., MORAN, T., KARALIUSKAS, R., DUERR, R. H., ACHKAR, J. P., BRANT, S. R., BAYLESS, T. M., KIRSCHNER, B. S., HANAUER, S. B., NUNEZ, G. & CHO, J. H. 2001. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature,* 411, 603-6.

OHTANI, K., OHTSUKA, Y., IKUSE, T., BABA, Y., YAMAKAWA, Y., AOYAGI, Y., FUJII, T., KUDO, T., NAGATA, S. & SHIMIZU, T. 2010. Increased mucosal expression of GATA-3 and STAT-4 in pediatric ulcerative colitis. *Pediatr Int,* 52, 584-9.

ORME, I. M., GRIFFIN, J. P., ROBERTS, A. D. & ERNST, D. N. 1993. Evidence for a defective accumulation of protective T cells in old mice infected with Mycobacterium tuberculosis. *Cell Immunol,* 147, 222-9.

OTT, S., WELLS, S. & WAGNER, B. 1999. Herd level economic losses associated with Johne's disease in US dairy operations. *Prev Vet Med,* 40, 179 - 192.

OTTENHOFF, T. H., SPIERINGS, E., NIBBERING, P. H. & DE JONG, R. 1997. Modulation of protective and pathological immunity in mycobacterial infections. *Int Arch Allergy Immunol,* 113, 400-8.

OUYANG, W., RANGANATH, S. H., WEINDEL, K., BHATTACHARYA, D., MURPHY, T. L., SHA, W. C. & MURPHY, K. M. 1998. Inhibition of Th1 development mediated by GATA-3 through an IL-4-independent mechanism. *Immunity,* 9, 745-55.

OWYANG, A. M., ZAPH, C., WILSON, E. H., GUILD, K. J., MCCLANAHAN, T., MILLER, H. R., CUA, D. J., GOLDSCHMIDT, M., HUNTER, C. A., KASTELEIN, R. A. & ARTIS, D. 2006. Interleukin 25 regulates type 2 cytokine-dependent immunity and limits chronic inflammation in the gastrointestinal tract. *J Exp Med,* 203, 843-9.

PAOLICCHI, F., CIRONE, K., MORSELLA, C. & GIOFFRE, A. 2012. First isolation of Mycobacterium avium subsp Paratuberculosis from commercial pasteurized milk in Argentina. *Braz J Microbiol,* 43, 1034-7.

PAPPU, R., RAMIREZ-CARROZZI, V. & SAMBANDAM, A. 2011. The interleukin-17 cytokine family: critical players in host defence and inflammatory diseases. *Immunology,* 134, 8-16.

PARK, J. S., PARK, B. L., KIM, M. O., HEO, J. S., JUNG, J. S., BAE, D. J., UH, S. T., KIM, M. K., CHOI, I. S., CHO, S. H., HONG, C. S., LEE, J. Y., CHOI, B. W., SHIN, H. D. & PARK, C. S. 2013. Association of single nucleotide polymorphisms on Interleukin 17 receptor A (IL17RA) gene with aspirin hypersensitivity in asthmatics. *Hum Immunol,* 74, 598-606.

PAVLIK, I., BARTL, J., DVORSKA, L., SVASTOVA, P., DU MAINE, R., MACHACKOVA, M., YAYO AYELE, W. & HORVATHOVA, A. 2000. Epidemiology of paratuberculosis in wild ruminants

studied by restriction fragment length polymorphism in the Czech Republic during the period 1995-1998. *Vet Microbiol,* 77, 231-51.

PEREZ, V., TELLECHEA, J., CORPA, J. M., GUTIERREZ, M. & GARCIA MARIN, J. F. 1999. Relation between pathologic findings and cellular immune responses in sheep with naturally acquired paratuberculosis. *Am J Vet Res,* 60, 123-7.

PETERSON, L. W. & ARTIS, D. 2014. Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nature Reviews Immunology,* 14, 141-153.

PHILPOTT, D. J., SORBARA, M. T., ROBERTSON, S. J., CROITORU, K. & GIRARDIN, S. E. 2014. NOD proteins: regulators of inflammation in health and disease. *Nat Rev Immunol,* 14, 9-23.

PIERCE, E. S. 2009. Possible transmission of Mycobacterium avium subspecies paratuberculosis through potable water: lessons from an urban cluster of Crohn's disease. *Gut Pathog,* 1, 17.

PINEDO, P. J., BUERGELT, C. D., DONOVAN, G. A., MELENDEZ, P., MOREL, L., WU, R., LANGAEE, T. Y. & RAE, D. O. 2009. Candidate gene polymorphisms (BoIFNG, TLR4, SLC11A1) as risk factors for paratuberculosis infection in cattle. *Prev Vet Med,* 91, 189-96.

POHL, M., BORTFELDT, R. H., GRUTZMANN, K. & SCHUSTER, S. 2013. Alternative splicing of mutually exclusive exons--a review. *Biosystems,* 114, 31-8.

PUEL, A., CYPOWYJ, S., BUSTAMANTE, J., WRIGHT, J. F., LIU, L., LIM, H. K., MIGAUD, M., ISRAEL, L., CHRABIEH, M., AUDRY, M., GUMBLETON, M., TOULON, A., BODEMER, C., EL-BAGHDADI, J., WHITTERS, M., PARADIS, T., BROOKS, J., COLLINS, M., WOLFMAN, N. M., AL-MUHSEN, S., GALICCHIO, M., ABEL, L., PICARD, C. & CASANOVA, J. L. 2011. Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity. *Science,* 332, 65-8.

PUYANG, X., LEE, K., PAWLICHUK, C. & KUNIMOTO, D. Y. 1999. IS1626, a new IS900-related Mycobacterium avium insertion sequence. *Microbiology,* 145 (Pt 11), 3163-8.

PYKALAINEN, M., KINOS, R., VALKONEN, S., RYDMAN, P., KILPELAINEN, M., LAITINEN, L. A., KARJALAINEN, J., NIEMINEN, M., HURME, M., KERE, J., LAITINEN, T. & LAHESMAA, R. 2005. Association analysis of common variants of STAT6, GATA3, and STAT4 to asthma and high serum IgE phenotypes. *J Allergy Clin Immunol,* 115, 80-7.

QU, H. Q., FISHER-HOCH, S. P. & MCCORMICK, J. B. 2011. Molecular immunity to mycobacteria: knowledge from the mutation and phenotype spectrum analysis of Mendelian susceptibility to mycobacterial diseases. *Int J Infect Dis,* 15, e305-13.

RABY, B. A., HWANG, E.-S., STEEN, K. V., TANTISIRA, K., PENG, S., LITONJUA, A., LAZARUS, R., GIALLOURAKIS, C., RIOUX, J. D., SPARROW, D., SILVERMAN, E. K., GLIMCHER, L. H. & WEISS, S. T. 2006. T-Bet Polymorphisms Are Associated with Asthma and Airway Hyperresponsiveness. *American Journal of Respiratory and Critical Care Medicine,* 173, 64-70.

RAUEN, T., JUANG, Y. T., HEDRICH, C. M., KIS-TOTH, K. & TSOKOS, G. C. 2012. A novel isoform of the orphan receptor RORgammat suppresses IL-17 production in human T cells. *Genes Immun,* 13, 346-50.

RAVINDRAN, R., FOLEY, J., STOKLASEK, T., GLIMCHER, L. H. & MCSORLEY, S. J. 2005. Expression of T-bet by CD4 T cells is essential for resistance to Salmonella infection. *J Immunol,* 175, 4603-10.

RAY, A. A., FOUNTAIN, J. J., MILLER, H. E., COOPER, A. M. & ROBINSON, R. T. 2015. IL12Rβ1ΔTM Is a Secreted Product of il12rb1 That Promotes Control of Extrapulmonary Tuberculosis. *Infection and Immunity,* 83, 560-571.

REA, T. H. & MODLIN, R. L. 1991. Immunopathology of leprosy skin lesions. *Semin Dermatol,* 10, 188-93.

REDDACLIFF, L., EPPLESTON, J., WINDSOR, P., WHITTINGTON, R. & JONES, S. 2006. Efficacy of a killed vaccine for the control of paratuberculosis in Australian sheep flocks. *Veterinary Microbiology,* 115, 77-90.

RICKEL, E. A., SIEGEL, L. A., YOON, B. R., ROTTMAN, J. B., KUGLER, D. G., SWART, D. A., ANDERS, P. M., TOCKER, J. E., COMEAU, M. R. & BUDELSKY, A. L. 2008. Identification of functional roles for both IL-17RB and IL-17RA in mediating IL-25-induced activities. *J Immunol,* 181, 4299-310.

RIDLEY, D. S. 1974. Histological classification and the immunological spectrum of leprosy. *Bulletin of the World Health Organization,* 51, 451-465.

RIZZO, A., MONTELEONE, I., FINA, D., STOLFI, C., CARUSO, R., FANTINI, M. C., FRANZE, E., SCHWENDENER, R., PALLONE, F. & MONTELEONE, G. 2012. Inhibition of colitis by IL-25 associates with induction of alternatively activated macrophages. *Inflamm Bowel Dis,* 18, 449-59.

ROBERTS, R. L., GEARRY, R. B., HOLLIS-MOFFATT, J. E., MILLER, A. L., REID, J., ABKEVICH, V., TIMMS, K. M., GUTIN, A., LANCHBURY, J. S., MERRIMAN, T. R., BARCLAY, M. L. & KENNEDY, M. A. 2007. IL23R R381Q and ATG16L1 T300A are strongly associated with Crohn's disease in a study of New Zealand Caucasians with inflammatory bowel disease. *Am J Gastroenterol,* 102, 2754-61.

ROBINSON, M. W., O'BRIEN, R., MACKINTOSH, C. G., CLARK, R. G. & GRIFFIN, J. F. 2011. Immunoregulatory cytokines are associated with protection from immunopathology following Mycobacterium avium subspecies paratuberculosis infection in red deer. *Infect Immun,* 79, 2089-97.

ROBINSON, R. T., KHADER, S. A., MARTINO, C. A., FOUNTAIN, J. J., TEIXEIRA-COELHO, M., PEARL, J. E., SMILEY, S. T., WINSLOW, G. M., WOODLAND, D. L., WALTER, M. J., CONEJO-GARCIA, J. R., GUBLER, U. & COOPER, A. M. 2010. Mycobacterium tuberculosis infection induces il12rb1 splicing to generate a novel IL-12Rβ1 isoform that enhances DC migration. *The Journal of Experimental Medicine,* 207, 591-605.

ROMAGNANI, S. 1996. Th1 and Th2 in human diseases. *Clin Immunol Immunopathol,* 80, 225-35.

ROMAGNANI, S. 1999. Th1/Th2 cells. *Inflamm Bowel Dis,* 5, 285-94.

ROMAGNANI, S. 2006. Regulation of the T cell response. *Clin Exp Allergy,* 36, 1357-66.

ROUSSEY, J. A., STEIBEL, J. P. & COUSSENS, P. M. 2014. Regulatory T cell activity and signs of T cell unresponsiveness in bovine paratuberculosis. *Frontiers in Veterinary Science,* 1, 20.

RUIZ-LARRANAGA, O., GARRIDO, J. M., IRIONDO, M., MANZANO, C., MOLINA, E., KOETS, A. P., RUTTEN, V. P., JUSTE, R. A. & ESTONBA, A. 2010. Genetic association between bovine NOD2 polymorphisms and infection by Mycobacterium avium subsp. paratuberculosis in Holstein-Friesian cattle. *Anim Genet,* 41, 652-5.

SAKABE, N. J. & DE SOUZA, S. J. 2007. Sequence features responsible for intron retention in human. *BMC genomics,* 8, 59.

SALVATI, V. M., MACDONALD, T. T., BAJAJ-ELLIOTT, M., BORRELLI, M., STAIANO, A., AURICCHIO, S., TRONCONE, R. & MONTELEONE, G. 2002. Interleukin 18 and associated markers of T helper cell type 1 activity in coeliac disease. *Gut,* 50, 186-90.

SANCHEZ, D., LEFEBVRE, C., GARCIA, L. F. & BARRERA, L. F. 2013. Variants in the IFNgamma transcription factor genes TBET, STAT1, STAT4, and HLX and the risk of pulmonary tuberculosis in a Colombian population: a case-control study. *Biomedica,* 33, 259-67.

SARACHANA, T. & HU, V. W. 2013. Genome-wide identification of transcriptional targets of RORA reveals direct regulation of multiple genes associated with autism spectrum disorder. *Mol Autism,* 4, 14.

SARTOR, R. B. 2005. Does Mycobacterium avium subspecies paratuberculosis cause Crohn's disease? *Gut,* 54, 896-898.

SARTOR, R. B. 2006. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nature clinical practice Gastroenterology & hepatology,* 3, 390-407.

SATO, T. K., PANDA, S., MIRAGLIA, L. J., REYES, T. M., RUDIC, R. D., MCNAMARA, P., NAIK, K. A., FITZGERALD, G. A., KAY, S. A. & HOGENESCH, J. B. 2004. A Functional Genomics Strategy Reveals Rora as a Component of the Mammalian Circadian Clock. *Neuron,* 43, 527-537.

SAUNDERS, B. M. & COOPER, A. M. 2000. Restraining mycobacteria: role of granulomas in mycobacterial infections. *Immunol Cell Biol,* 78, 334-41.

SCHAUMBERG, D. A., CHASMAN, D., MORRISON, M. A., ADAMS, S. M., GUO, Q., HUNTER, D. J., HANKINSON, S. E. & DEANGELIS, M. M. 2010. Prospective study of common variants in the retinoic acid receptor-related orphan receptor alpha gene and risk of neovascular age-related macular degeneration. *Arch Ophthalmol,* 128, 1462-71.

SCHURR, E. 2011. The contribution of host genetics to tuberculosis pathogenesis. *Kekkaku,* 86, 17-28.

SCHWARTZ, D., SHAFRAN, I., ROMERO, C., PIROMALLI, C., BIGGERSTAFF, J., NASER, N., CHAMBERLIN, W. & NASER, S. A. 2000. Use of short-term culture for identification of

Mycobacterium avium subsp. paratuberculosis in tissue from Crohn's disease patients. *Clin Microbiol Infect,* 6, 303-7.

SECOTT, T. E., LIN, T. L. & WU, C. C. 2004. Mycobacterium avium subsp. paratuberculosis fibronectin attachment protein facilitates M-cell targeting and invasion through a fibronectin bridge with host integrins. *Infect Immun,* 72, 3724-32.

SEITZ, S. E., HEIDER, L. E., HEUSTON, W. D., BECH-NIELSEN, S., RINGS, D. M. & SPANGLER, L. 1989. Bovine fetal infection with Mycobacterium paratuberculosis. *J Am Vet Med Assoc,* 194, 1423-6.

SEKIMATA, M., PEREZ-MELGOSA, M., MILLER, S. A., WEINMANN, A. S., SABO, P. J., SANDSTROM, R., DORSCHNER, M. O., STAMATOYANNOPOULOS, J. A. & WILSON, C. B. 2009. CCCTC-binding factor and the transcription factor T-bet orchestrate T helper 1 cell-specific structure and function at the interferon-gamma locus. *Immunity,* 31, 551-64.

SHI, G., RAMASWAMY, M., VISTICA, B. P., COX, C. A., TAN, C., WAWROUSEK, E. F., SIEGEL, R. M. & GERY, I. 2009. Unlike Th1, Th17 cells mediate sustained autoimmune inflammation and are highly resistant to restimulation-induced cell death. *J Immunol,* 183, 7547-56.

SHINNAKASU, R., YAMASHITA, M., SHINODA, K., ENDO, Y., HOSOKAWA, H., HASEGAWA, A., IKEMIZU, S. & NAKAYAMA, T. 2006. Critical YxKxHxxxRP motif in the C-terminal region of GATA3 for its DNA binding and function. *J Immunol,* 177, 5801-10.

SIDOTI, F., BANCHE, G., ASTEGIANO, S., ALLIZOND, V., CUFFINI, A. M. & BERGALLO, M. 2011. Validation and standardization of IS900 and F57 real-time quantitative PCR assays for the specific detection and quantification of Mycobacterium avium subsp. paratuberculosis. *Can J Microbiol,* 57, 347-54.

SMEED, J. A., WATKINS, C. A., GOSSNER, A. G. & HOPKINS, J. 2010. Expression profiling reveals differences in immuno-inflammatory gene expression between the two disease forms of sheep paratuberculosis. *Veterinary Immunology and Immunopathology,* 135, 218-225.

SMEED, J. A., WATKINS, C. A., RHIND, S., HOPKINS, J., SMEED, J. A., WATKINS, C. A., RHIND, S. & HOPKINS, J. 2008. Differential cytokine gene expression profiles in the three pathological forms of sheep paratuberculosis. *Immunology,* 125, 105-105.

SMEED, J. A., WATKINS, C. A., RHIND, S. M. & HOPKINS, J. 2007. Differential cytokine gene expression profiles in the three pathological forms of sheep paratuberculosis. *BMC Vet Res,* 3, 18.

SOHDA, M., MISUMI, Y., TASHIRO, K., YAMAZAKI, M., SAKU, T. & ODA, K. 2013. Identification of a soluble isoform of human IL-17RA generated by alternative splicing. *Cytokine,* 64, 642-5.

SOUZA, C., DAVIS, W. C., ECKSTEIN, T. M., SREEVATSAN, S. & WEISS, D. J. 2013. Mannosylated Lipoarabinomannans from <italic>Mycobacterium Avium</italic> Subsp. Paratuberculosis Alters the Inflammatory Response by Bovine Macrophages and Suppresses Killing of <italic>Mycobacterium</italic> Avium Subsp. <italic>Avium</italic> Organisms. *PLoS ONE,* 8, e75924.

SOUZA, C. D., EVANSON, O. A., SREEVATSAN, S. & WEISS, D. J. 2007. Cell membrane receptors on bovine mononuclear phagocytes involved in phagocytosis of Mycobacterium avium subsp paratuberculosis. *Am J Vet Res,* 68, 975-80.

SOUZA, C. D., EVANSON, O. A. & WEISS, D. J. 2006. Mitogen activated protein kinase p38 pathway is an important component of the anti-inflammatory response in Mycobacterium avium subsp. paratuberculosis-infected bovine monocytes. *Microb Pathog,* 41, 59-66.

STASTNA, M. & VAN EYK, J. E. 2012. Analysis of protein isoforms: can we do it better? *Proteomics,* 12, 2937-48.

STEINMAYR, M., ANDRÉ, E., CONQUET, F., RONDI-REIG, L., DELHAYE-BOUCHAUD, N., AUCLAIR, N., DANIEL, H., CRÉPEL, F., MARIANI, J., SOTELO, C. & BECKER-ANDRÉ, M. 1998. staggerer phenotype in retinoid-related orphan receptor α-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America,* 95, 3960-3965.

STEWART, W. C., POLLOCK, K. G., BROWNING, L. M., YOUNG, D., SMITH-PALMER, A. & REILLY, W. J. 2005. Survey of zoonoses recorded in Scotland between 1993 and 2002. *Vet Rec,* 157, 697-702.

STRAIN, S. A., MCNAIR, J., MCDOWELL, S. W. & BRANCH, B. 2011. Bovine tuberculosis: A review of diagnostic tests for M. bovis infection in cattle. *Agri-Food and Biosciences Institute*.

STRAIN, S. A. J., J. MCNAIR, AND S. W. J. MCDOWELL,. 2011. *Bovine Tuberculosis: A Review of Diagnostic Tests for M. bovis Infection*

in Cattle. [Online].

SU, J., CHEN, T., JI, X. Y., LIU, C., YADAV, P. K., WU, R., YANG, P. & LIU, Z. 2013. IL-25 downregulates Th1/Th17 immune response in an IL-10-dependent manner in inflammatory bowel disease. *Inflamm Bowel Dis,* 19, 720-8.

SWAIDANI, S., BULEK, K., KANG, Z., LIU, C., LU, Y., YIN, W., ARONICA, M. & LI, X. 2009. The critical role of epithelial-derived Act1 in IL-17- and IL-25-mediated pulmonary inflammation. *J Immunol,* 182, 1631-40.

SWEENEY, R. W., JONES, D. E., HABECKER, P. & SCOTT, P. 1998. Interferon-gamma and interleukin 4 gene expression in cows infected with Mycobacterium paratuberculosis. *Am J Vet Res,* 59, 842-7.

SZABO, S. J., SULLIVAN, B. M., STEMMANN, C., SATOSKAR, A. R., SLECKMAN, B. P. & GLIMCHER, L. H. 2002. Distinct effects of T-bet in TH1 lineage commitment and IFN-gamma production in CD4 and CD8 T cells. *Science,* 295, 338-42.

TALAVERA, D., OROZCO, M. & DE LA CRUZ, X. 2009. Alternative splicing of transcription factors' genes: beyond the increase of proteome diversity. *Comparative and functional genomics,* 2009.

TANAKA, S., SATO, M., ONITSUKA, T., KAMATA, H. & YOKOMIZO, Y. 2005. Inflammatory cytokine gene expression in different types of granulomatous lesions during asymptomatic stages of bovine paratuberculosis. *Vet Pathol,* 42, 579-88.

TAYLOR, R. H., FALKINHAM, J. O., 3RD, NORTON, C. D. & LECHEVALLIER, M. W. 2000. Chlorine, chloramine, chlorine dioxide, and ozone susceptibility of Mycobacterium avium. *Appl Environ Microbiol,* 66, 1702-5.

TESMER, L. A., LUNDY, S. K., SARKAR, S. & FOX, D. A. 2008. Th17 cells in human disease. *Immunol Rev,* 223, 87-113.

TESSEMA, M., KOETS, A., RUTTEN, V. & GRUYS, E. 2001. Bacteriology: Review paratuberculosis: How does mycobacterium avium subsp. Paratuberculosis resist intracellular degradation? *Veterinary quarterly,* 23, 153-162.

THOREL, M. F., KRICHEVSKY, M. & LEVY-FREBAULT, V. V. 1990. Numerical taxonomy of mycobactin-dependent mycobacteria, emended description of Mycobacterium avium, and description of Mycobacterium avium subsp. avium subsp. nov., Mycobacterium avium subsp. paratuberculosis subsp. nov., and Mycobacterium avium subsp. silvaticum subsp. nov. *Int J Syst Bacteriol,* 40, 254-60.

TOOKER, B. C. & COUSSENS, P. M. 2004. Phagocytosis of M. paratuberculosis fails to activate expression of NADH dehydrogenase and nucleolin-related protein in bovine macrophages. *Immunol Lett,* 93, 137-42.

VAN DE VOSSE, E., HAVERKAMP, M. H., RAMIREZ-ALEJO, N., MARTINEZ-GALLO, M., BLANCAS-GALICIA, L., METIN, A., GARTY, B. Z., SUN-TAN, C., BROIDES, A., DE PAUS, R. A., KESKIN, O., CAGDAS, D., TEZCAN, I., LOPEZ-RUZAFA, E., AROSTEGUI, J. I., LEVY, J., ESPINOSA-ROSALES, F. J., SANAL, O., SANTOS-ARGUMEDO, L., CASANOVA, J. L., BOISSON-DUPUIS, S., VAN DISSEL, J. T. & BUSTAMANTE, J. 2013. IL-12Rbeta1 deficiency: mutation update and description of the IL12RB1 variation database. *Hum Mutat,* 34, 1329-39.

VAN DE VOSSE, E., LICHTENAUER-KALIGIS, E. G., VAN DISSEL, J. T. & OTTENHOFF, T. H. 2003. Genetic variations in the interleukin-12/interleukin-23 receptor (beta1) chain, and implications for IL-12 and IL-23 receptor structure and function. *Immunogenetics,* 54, 817-29.

VAN DER POUW KRAAN, T. C., VAN DER ZEE, J. S., BOEIJE, L. C., DE GROOT, E. R., STAPEL, S. O. & AARDEN, L. A. 1998. The role of IL-13 in IgE synthesis by allergic asthma patients. *Clin Exp Immunol,* 111, 129-35.

VELAZ-FAIRCLOTH, M., COBB, A. J., HORSTMAN, A. L., HENRY, S. C. & FROTHINGHAM, R. 1999. Protection against Mycobacterium aviumby DNA Vaccines Expressing Mycobacterial Antigens as Fusion Proteins with Green Fluorescent Protein. *Infection and immunity,* 67, 4243-4250.

WANG, A.-J., SMITH, A., LI, Y., URBAN, J. F., RAMALINGAM, T. R., WYNN, T. A., LU, N., SHEA-DONOHUE, T., YANG, Z. & ZHAO, A. 2014. Genetic deletion of IL-25 (IL-17E) confers resistance to dextran sulfate sodium-induced colitis in mice. *Cell & Bioscience,* 4, 72.

WANG, C. T. 1999. Diagnosing and treating asymptomatic tuberculosis infection. *Canadian Family Physician,* 45, 2397.

WANG, J., FATHMAN, J. W., LUGO-VILLARINO, G., SCIMONE, L., VON ANDRIAN, U., DORFMAN, D. M. & GLIMCHER, L. H. 2006. Transcription factor T-bet regulates inflammatory arthritis through its function in dendritic cells. *J Clin Invest,* 116, 414-21.

WANG, Y. A. N., LIU, J., HUANG, B. O., XU, Y.-M., LI, J., HUANG, L.-F., LIN, J. I. N., ZHANG, J., MIN, Q.-H., YANG, W.-M. & WANG, X.-Z. 2015. Mechanism of alternative splicing and its regulation. *Biomedical Reports,* 3, 152-158.

WANG, Y. H., ANGKASEKWINAI, P., LU, N., VOO, K. S., ARIMA, K., HANABUCHI, S., HIPPE, A., CORRIGAN, C. J., DONG, C., HOMEY, B., YAO, Z., YING, S., HUSTON, D. P. & LIU, Y. J. 2007. IL-25 augments type 2 immune responses by enhancing the expansion and functions of TSLP-DC-activated Th2 memory cells. *J Exp Med,* 204, 1837-47.

WATERS, W. R., MILLER, J. M., PALMER, M. V., STABEL, J. R., JONES, D. E., KOISTINEN, K. A., STEADHAM, E. M., HAMILTON, M. J., DAVIS, W. C. & BANNANTINE, J. P. 2003. Early induction of humoral and cellular immune responses during experimental Mycobacterium avium subsp. paratuberculosis infection of calves. *Infect Immun,* 71, 5130-8.

WEBER, M. F. 2006. Risk management of paratuberculosis in dairy herds. *Irish veterinary journal,* 59, 555.

WHITLOCK, R. H. & BUERGELT, C. 1996. Preclinical and clinical manifestations of paratuberculosis (including pathology). *Vet Clin North Am Food Anim Pract,* 12, 345-56.

WHITTINGTON, R. J., MARSH, I. B. & REDDACLIFF, L. A. 2005. Survival of Mycobacterium avium subsp. paratuberculosis in dam water and sediment. *Appl Environ Microbiol,* 71, 5304-8.

WILKIE, H., GOSSNER, A., BISHOP, S. & HOPKINS, J. 2016. Variations in T cell transcription factor sequence and expression associated with resistance to sheep nematode *Teladorsagia circumcinta. PLOS ONE* 10, 1371.

WILSON, V. & CONLON, F. L. 2002. The T-box family. *Genome Biol,* 3, Reviews3008.

WOLTERINK, R. G. J. K., KLEINJAN, A., VAN NIMWEGEN, M., BERGEN, I., DE BRUIJN, M., LEVANI, Y. & HENDRIKS, R. W. 2012. Pulmonary innate lymphoid cells are major producers of IL-5 and IL-13 in murine models of allergic asthma. *European Journal of Immunology,* 42, 1106-1116.

WOO, S. R. & CZUPRYNSKI, C. J. 2008. Tactics of Mycobacterium avium subsp. paratuberculosis for intracellular survival in mononuclear phagocytes. *J Vet Sci,* 9, 1-8.

WU, S., GEBNER, R., TAUBE, T., VON STACKELBERG, A., HENZE, G. & SEEGER, K. 2005. Expression of interleukin-10 splicing variants is a positive prognostic feature in relapsed childhood acute lymphoblastic leukemia. *Journal of clinical oncology,* 23, 3038-3042.

WURTZ, O., BAJENOFF, M. & GUERDER, S. 2004. IL-4-mediated inhibition of IFN-gamma production by CD4+ T cells proceeds by several developmentally regulated mechanisms. *Int Immunol,* 16, 501-8.

YAGI, R., ZHU, J. & PAUL, W. E. 2011. An updated view on transcription factor GATA3-mediated regulation of Th1 and Th2 cell differentiation. *Int Immunol,* 23, 415-20.

YANG, J., SUNDRUD, M. S., SKEPNER, J. & YAMAGATA, T. 2014. Targeting Th17 cells in autoimmune diseases. *Trends in pharmacological sciences,* 35, 493-500.

YANG, X. O., PAPPU, B. P., NURIEVA, R., AKIMZHANOV, A., KANG, H. S., CHUNG, Y., MA, L., SHAH, B., PANOPOULOS, A. D., SCHLUNS, K. S., WATOWICH, S. S., TIAN, Q., JETTEN, A. M. & DONG, C. 2008. T helper 17 lineage differentiation is programmed by orphan nuclear receptors ROR alpha and ROR gamma. *Immunity,* 28, 29-39.

YEN, D., CHEUNG, J., SCHEERENS, H., POULET, F., MCCLANAHAN, T., MCKENZIE, B., KLEINSCHEK, M. A., OWYANG, A., MATTSON, J., BLUMENSCHEIN, W., MURPHY, E., SATHE, M., CUA, D. J., KASTELEIN, R. A. & RENNICK, D. 2006. IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. *J Clin Invest,* 116, 1310-6.

YOSHIMURA, A., SUZUKI, M., SAKAGUCHI, R., HANADA, T. & YASUKAWA, H. 2012. SOCS, inflammation and autoimmunity. *Frontiers in Immunology,* 3.

YU, R. Y. & GALLAGHER, G. 2010. A naturally occurring, soluble antagonist of human IL-23 inhibits the development and in vitro function of human Th17 cells. *J Immunol,* 185, 7302-8.

ZARE, Y., SHOOK, G. E., COLLINS, M. T. & KIRKPATRICK, B. W. 2014. Genome-Wide Association Analysis and Genomic Prediction of <italic>Mycobacterium avium</italic> Subspecies <italic>paratuberculosis</italic> Infection in US Jersey Cattle. *PLoS ONE,* 9, e88380.

ZEPP, J., WU, L. & LI, X. 2011. IL-17 receptor signaling and T helper 17-mediated autoimmune demyelinating disease. *Trends in Immunology,* 32, 232-239.

ZHANG, B., LIU, C., QIAN, W., HAN, Y., LI, X. & DENG, J. 2014. Structure of the unique SEFIR domain from human interleukin 17 receptor A reveals a composite ligand-binding site containing a conserved alpha-helix for Act1 binding and IL-17 signaling. *Acta Crystallogr D Biol Crystallogr,* 70, 1476-83.

ZHANG, D. H., COHN, L., RAY, P., BOTTOMLY, K. & RAY, A. 1997. Transcription factor GATA-3 is differentially expressed in murine Th1 and Th2 cells and controls Th2-specific expression of the interleukin-5 gene. *J Biol Chem,* 272, 21597-603.

ZHANG, X. Y., ZHANG, H. J., ZHANG, Y., FU, Y. J., HE, J., ZHU, L. P., WANG, S. H. & LIU, L. 2006. Identification and expression analysis of alternatively spliced isoforms of human interleukin-23 receptor gene in normal lymphoid cells and selected tumor cells. *Immunogenetics,* 57, 934-43.

ZHAO, A., URBAN, J. F., JR., SUN, R., STILTZ, J., MORIMOTO, M., NOTARI, L., MADDEN, K. B., YANG, Z., GRINCHUK, V., RAMALINGAM, T. R., WYNN, T. A. & SHEA-DONOHUE, T. 2010. Critical role of IL-25 in nematode infection-induced alterations in intestinal function. *J Immunol,* 185, 6921-9.

ZHENG, W. & FLAVELL, R. A. 1997. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell,* 89, 587-96.

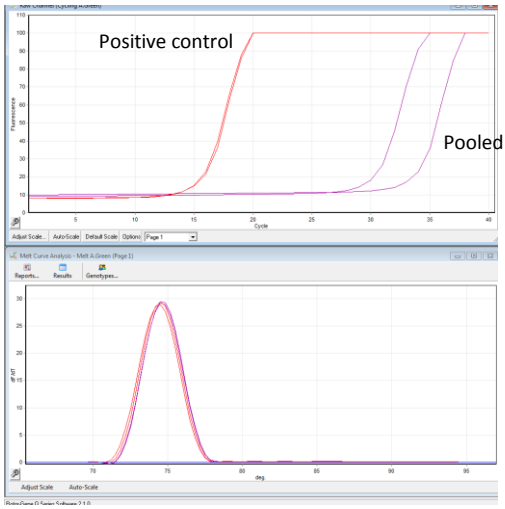
ZHU, J., MIN, B., HU-LI, J., WATSON, C. J., GRINBERG, A., WANG, Q., KILLEEN, N., URBAN, J. F., JR., GUO, L. & PAUL, W. E. 2004. Conditional deletion of Gata3 shows its essential function in T(H)1-T(H)2 responses. *Nat Immunol,* 5, 1157-65.

ZHU, J. & PAUL, W. E. 2010. Peripheral CD4+ T-cell differentiation regulated by networks of cytokines and transcription factors. *Immunological Reviews,* 238, 247-262.

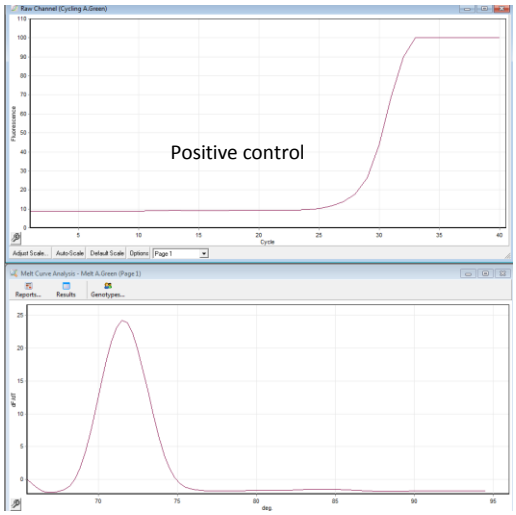
ZHU, J., YAMANE, H. & PAUL, W. E. 2010. Differentiation of effector CD4 T cell populations. *Annual review of immunology,* 28, 445.

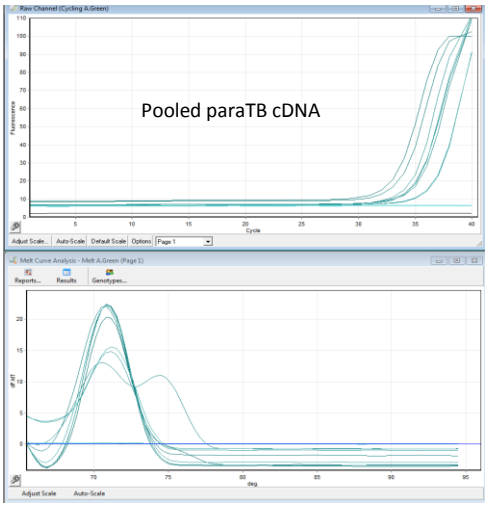
Appendix

A1. *IL23RV2* RT-qPCR positive control assay

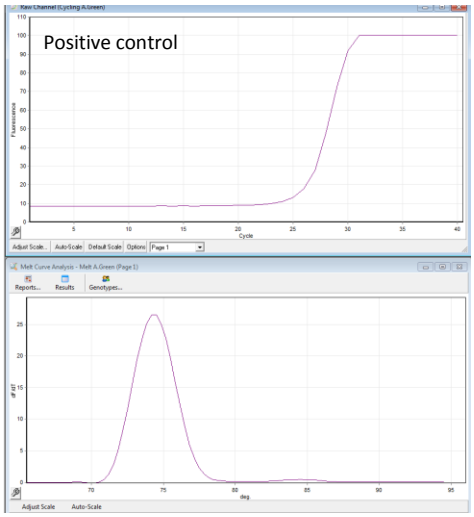


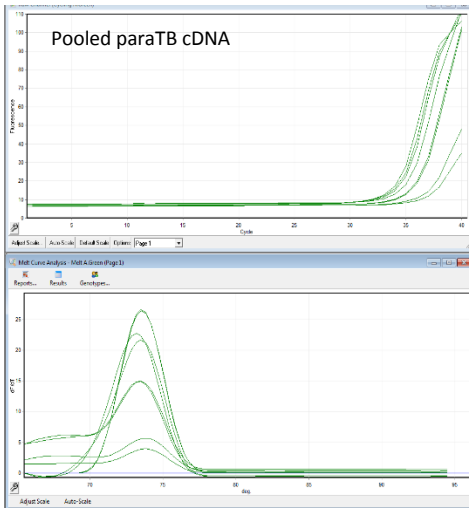
A2. *IL23RV3* RT-qPCR positive control assay



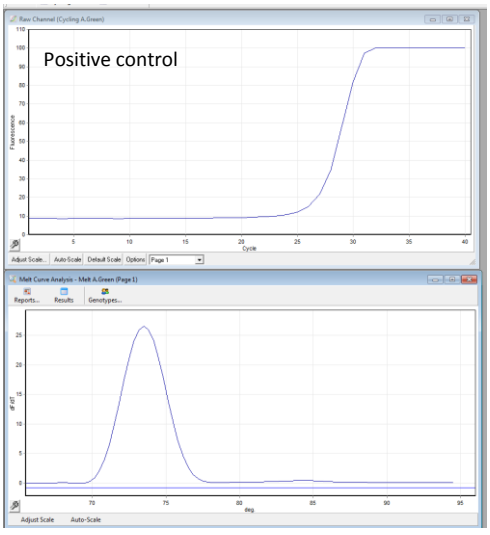


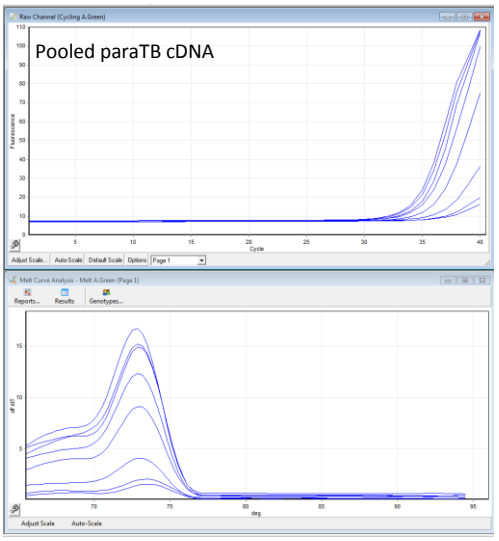
A3. *IL23R V4* RT-qPCR positive control assay





A4. *IL23R V5* RT-qPCR positive control assay





A5. IL12RB1 variant nucleotide and protein sequences.

IL12RB1v1 nucleotide

```
full            TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
TV1             TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
                ************************************************************

full            TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
TV1             TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
                ************************************************************

full            TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
TV1             TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
                ************************************************************

full            GCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCGCTGGTTATGAATGT 240
TV1             GCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCGCTGGTTATGAATGT 240
                ************************************************************

full            TCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCAGATGCTGCCTCCAG 300
TV1             TCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCAGCCACTTCCTCAGATGCTGCCTCCAG 300
                ********************************* ************************

full            TCTGGCCGCTGCTGCTTCTTTGCCACAGGCTCAGCCACCAAGCTGCAGTTCTCCGACCAG 360
TV1             TCTGGCCGCTGCTGCTACTTTGCCACAGGCTCAGCCACCAAGCTGCAGTTCTCCGACCAG 360
                ****************:*******************************************

full            GATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCCGGGCCGCCAACCGG 420
TV1             GATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCCGGGCCGCCAACCGG 420
                ************************************************************

full            ACAGAGAAGTCCCCCAATGTTACTCTGAACCTCTACAGCTCGGTTAAATACGACCCTCCC 480
TV1             ACAGAGAAGTCCCCCAATGTTACCCTGAACCTCTACAGCTCGGTTAAATACGACCCTCCC 480
                ********************** ***********************************

full            CCAGGAAACATCAAGGTGTCCAAGTCAGCAGGGAAGCTGCGTATGGAGTGGGAGACCCCA 540
TV1             CCAGGAAACATCAAGATGTCCAAG---------------------------------- 504
                ***************.********

full            GCCCGCCAGGATGGTGCCGAGGTACAGTTCCGGCACCGCACACCTGGCAGCCCGTGGAAG 600
TV1             ------------------------------------------------------------


full            CTGGGTGACTGTGGACGTCAGGATGATGCTGGCTTCGAGTCATGCCTCTGCCCCTTGGAG 660
TV1             ------------------------------------------------------------


full            ATGGACATGGCCCAGGAATTCCAGCTGCGGCGGCGTCTGGGGCCAGGGGTCCCCGGAGGT 720
TV1             ------------------------------------------------------------


full            CCCTGGAGCAGCTGGAGCAGCCCTGTGTGCATCCCCCCTGAAACCCCTCCACAGGCTGAG 780
TV1             ------------------------------------------------------------


full            GTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTACATGAG 840
TV1             ----------------AGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTGTATGAG 548
                                ***********************************. *****

full            CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 900
TV1             CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 608
                ************************************************************

full            ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 960
TV1             ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 668
                ************************************************************

full            CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 1020
TV1             CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 728
```

```
                ************************************************************

full            CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 1080
TV1             CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 788
                ************************************************************

full            GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 1140
TV1             GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 848
                ************************************************************

full            CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 1200
TV1             CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 908
                ************************************************************

full            GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 1260
TV1             GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 968
                ************************************************************

full            TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 1320
TV1             TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 1028
                ************************************************************

full            TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 1380
TV1             TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 1088
                ************************************************************

full            GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 1440
TV1             GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 1148
                ************************************************************

full            TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 1500
TV1             TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 1208
                ************************************************************

full            TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 1560
TV1             TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 1268
                ************************************************************

full            ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 1620
TV1             ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 1328
                ************************************************************

full            CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1680
TV1             CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1388
                ************************************************************

full            GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1740
TV1             GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1448
                ************************************************************

full            AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1800
TV1             AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1508
                ************************************************************

full            TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1860
TV1             TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1568
                ************************************************************

full            AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1920
TV1             AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1628
                ************************************************************

full            GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1980
TV1             GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1688
                ************************************************************

full            ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 2040
TV1             ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 1748
                ************************************************************

full            CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 2100
TV1             CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 1808
                ************************************************************
```

```
full            GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 2160
TV1             GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 1868
                ************************************************************

full            CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 2220
TV1             CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 1928
                ************************************************************

full            GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 2280
TV1             GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 1988
                ************************************************************

full            ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 2340
TV1             ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 2048
                ************************************************************

full            TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 2382
TV1             TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 2090
                *******************************************
```

IL12RB1v1 protein

```
full            MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
TV1             MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
                ************************************************************

full            YECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
TV1             YECSWEYEGPAAGVSHFLRCCLQSGRCCYFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
                ************** ************:****************************

full            ANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRTPGS 180
TV1             ANRTEKSPNVTLNLYSSVKYDPPPGNIKMSKSS---------------------------- 153
                ****************************:***:

full            PWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPETPP 240
TV1             ---FAQMG---------------------------------GGRWPCMSS---------- 167
                   :.: *                                 ** *.. **

full            QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 300
TV1             -----------CPN------------------------------------------------ 170
                          **:

full            KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 360
TV1             ----------------------WSFQKAALGLTLAWR------------------------ 185
                                     * :: .** . :**

full            WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 420
TV1             ------------------------------------------------------------


full            IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 480
TV1             ------------------------------------------------------------


full            LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 540
TV1             ------------------------------------------------------------


full            TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 600
TV1             ------------------------------------------------------------


full            SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 660
TV1             ------------------------------------------------------------


full            DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 720
TV1             ------------------------------------------------------------


full            GTAASSYRED 730
TV1             ----------
```

***IL12RB1v2* nucleotide**

```
full            TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
TV2             ------------------------TGGGCTACTGCTCTTCTCAAGGACCGGATGGGACAA 36
                                        ************  ********************

full            TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
TV2             TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGT----------------------- 73
                ***********************************

full            TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
TV2             ------------------------------------------------------------

full            GCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCGCTGGTTATGAATGT 240
TV2             ------------------------------------------------------------

full            TCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCAGATGCTGCCTCCAG 300
TV2             ------------------------------------------------------------

full            TCTGGCCGCTGCTGCTTCTTTGCCACAGGCTCAGCCACCAAGCTGCAGTTCTCCGACCAG 360
TV2             ------------------------------------------------------------

full            GATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCCGGGCCGCCAACCGG 420
TV2             ------------------------------------------------------------

full            ACAGAGAAGTCCCCCAATGTTACTCTGAACCTCTACAGCTCGGTTAAATACGACCCTCCC 480
TV2             ------------------------------------------------------------

full            CCAGGAAACATCAAGGTGTCCAAGTCAGCAGGGAAGCTGCGTATGGAGTGGGAGACCCCA 540
TV2             ------------------------------------------------------------

full            GCCCGCCAGGATGGTGCCGAGGTACAGTTCCGGCACCGCACACCTGGCAGCCCGTGGAAG 600
TV2             ------------------------------------------------------------

full            CTGGGTGACTGTGGACGTCAGGATGATGCTGGCTTCGAGTCATGCCTCTGCCCCTTGGAG 660
TV2             ------------------------------------------------------------

full            ATGGACATGGCCCAGGAATTCCAGCTGCGGCGGCGTCTGGGGCCAGGGGTCCCCGGAGGT 720
TV2             ------------------------------------------------------------

full            CCCTGGAGCAGCTGGAGCAGCCCTGTGTGCATCCCCCCTGAAACCCCTCCACAGGCTGAG 780
TV2             ----------------------------GCATCCCCCCTGAAACCCCTCCACAGGCTGAG 105
                                            ********************************

full            GTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTACATGAG 840
TV2             GTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTACATGAG 165
                ************************************************************

full            CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 900
TV2             CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 225
                ************************************************************

full            ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 960
TV2             ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 285
                ************************************************************

full            CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 1020
TV2             CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 345
                ************************************************************

full            CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 1080
TV2             CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 405
```

```
                ************************************************************

full            GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 1140
TV2             GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 465
                ************************************************************

full            CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 1200
TV2             CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 525
                ************************************************************

full            GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 1260
TV2             GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 585
                ************************************************************

full            TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 1320
TV2             TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 645
                ************************************************************

full            TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 1380
TV2             TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 705
                ************************************************************

full            GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 1440
TV2             GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 765
                ************************************************************

full            TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 1500
TV2             TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 825
                ************************************************************

full            TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 1560
TV2             TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 885
                ************************************************************

full            ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 1620
TV2             ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 945
                ************************************************************

full            CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1680
TV2             CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1005
                ************************************************************

full            GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1740
TV2             GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1065
                ************************************************************

full            AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1800
TV2             AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1125
                ************************************************************

full            TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1860
TV2             TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1185
                ************************************************************

full            AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1920
TV2             AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1245
                ************************************************************

full            GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1980
TV2             GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1305
                ************************************************************

full            ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 2040
TV2             ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 1365
                ************************************************************

full            CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 2100
TV2             CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 1425
                ************************************************************

full            GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 2160
TV2             GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 1485
                ************************************************************
```

```
full            CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 2220
TV2             CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 1545
                ************************************************************

full            GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 2280
TV2             GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 1605
                ************************************************************

full            ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 2340
TV2             ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 1665
                ************************************************************

full            TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 2382
TV2             TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 1707
                *******************************************
```

IL12RB1v2 protein

```
full            MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
TV2             MGQWGFRLVAFLLLLC---------------------------------------------- 16
                ***************

full            YECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
TV2             ------------------------------------------------------------


full            ANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRTPGS 180
TV2             ------------------------------------------------------------


full            PWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPETPP 240
TV2             ----------------------------------------------------IPPETPP 23
                                                                     *******

full            QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 300
TV2             QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 83
                ************************************************************

full            KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 360
TV2             KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 143
                ************************************************************

full            WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 420
TV2             WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 203
                ************************************************************

full            IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 480
TV2             IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 263
                ************************************************************

full            LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 540
TV2             LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 323
                ************************************************************

full            TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 600
TV2             TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 383
                ************************************************************

full            SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 660
TV2             SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 443
                ************************************************************

full            DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 720
TV2             DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 503
                ************************************************************

full            GTAASSYRED 730
TV2             GTAASSYRED 515
                **********
```

***IL12RB1v3* nucleotide**

```
full            TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
TV3             TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
                ************************************************************

full            TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
TV3             TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
                ************************************************************

full            TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
TV3             TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
                ************************************************************

full            GCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCGCTGGTTATGAATGT 240
TV3             GCTTCAGGCCCCCGGGCCCTGAGCT---------------------------------- 205
                ************************

full            TCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCAGATGCTGCCTCCAG 300
TV3             ------------------------------------------------------------

full            TCTGGCCGCTGCTGCTTCTTTGCCACAGGCTCAGCCACCAAGCTGCAGTTCTCCGACCAG 360
TV3             ------------------------------------------------------------

full            GATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCCGGGCCGCCAACCGG 420
TV3             ------------------------------------------------------------

full            ACAGAGAAGTCCCCCAATGTTACTCTGAACCTCTACAGCTCGGTTAAATACGACCCTCCC 480
TV3             ------------------------------------------------------------

full            CCAGGAAACATCAAGGTGTCCAAGTCAGCAGGGAAGCTGCGTATGGAGTGGGAGACCCCA 540
TV3             ------------------------------------------------------------

full            GCCCGCCAGGATGGTGCCGAGGTACAGTTCCGGCACCGCACACCTGGCAGCCCGTGGAAG 600
TV3             ------------------------------------------------------------

full            CTGGGTGACTGTGGACGTCAGGATGATGCTGGCTTCGAGTCATGCCTCTGCCCCTTGGAG 660
TV3             ------------------------------------------------------------

full            ATGGACATGGCCCAGGAATTCCAGCTGCGGCGGCGTCTGGGGCCAGGGGTCCCCGGAGGT 720
TV3             ------------------------------------------------------------

full            CCCTGGAGCAGCTGGAGCAGCCCTGTGTGCATCCCCCCTGAAACCCCTCCACAGGCTGAG 780
TV3             ------------GGAGCAGCCCTGTGTGCATCCCCCCTGAAACCCCTCCACAGGCTGAG 252
                            ******************************************

full            GTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTACATGAG 840
TV3             GTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTACATGAG 312
                ************************************************************

full            CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 900
TV3             CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 372
                ************************************************************

full            ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 960
TV3             ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 432
                ************************************************************

full            CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 1020
TV3             CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 492
```

```
                ************************************************************

full            CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 1080
TV3             CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 552
                ************************************************************

full            GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 1140
TV3             GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 612
                ************************************************************

full            CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 1200
TV3             CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 672
                ************************************************************

full            GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 1260
TV3             GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 732
                ************************************************************

full            TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 1320
TV3             TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 792
                ************************************************************

full            TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 1380
TV3             TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 852
                ************************************************************

full            GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 1440
TV3             GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 912
                ************************************************************

full            TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 1500
TV3             TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 972
                ************************************************************

full            TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 1560
TV3             TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 1032
                ************************************************************

full            ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 1620
TV3             ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 1092
                ************************************************************

full            CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1680
TV3             CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1152
                ************************************************************

full            GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1740
TV3             GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1212
                ************************************************************

full            AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1800
TV3             AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1272
                ************************************************************

full            TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1860
TV3             TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1332
                ************************************************************

full            AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1920
TV3             AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1392
                ************************************************************

full            GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1980
TV3             GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1452
                ************************************************************

full            ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 2040
TV3             ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 1512
                ************************************************************

full            CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 2100
TV3             CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 1572
                ************************************************************
```

```
full            GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 2160
TV3             GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 1632
                ************************************************************

full            CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 2220
TV3             CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 1692
                ************************************************************

full            GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 2280
TV3             GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 1752
                ************************************************************

full            ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 2340
TV3             ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 1812
                ************************************************************

full            TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 2382
TV3             TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 1854
                *******************************************
```

IL12RB1v3 protein

```
full            MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
TV3             MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALS--------- 51
                *************************************************

full            YECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
TV3             ------------------------------------------------------------


full            ANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRTPGS 180
TV3             ------------------------------------------------------------


full            PWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPETPP 240
TV3             ----------------------------------------------WSSPVCIPPETPP 64
*************

full            QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 300
TV3             QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 124
                ************************************************************

full            KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 360
TV3             KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 184
                ************************************************************

full            WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 420
TV3             WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 244
                ************************************************************

full            IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 480
TV3             IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 304
                ************************************************************

full            LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 540
TV3             LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 364
                ************************************************************

full            TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 600
TV3             TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 424
                ************************************************************

full            SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 660
```

```
  TV3             SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 484
                  ************************************************************

  full            DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 720
  TV3             DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 544
                  ************************************************************

  full            GTAASSYRED 730
  TV3             GTAASSYRED 554
            **********
```

IL12RB1v4 nucleotide

```
full            TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
  TV4             TACCGGCAGCACAGAAGTTCATCTGGGGCTACTGCTCTCTTCAAGGACCGGATGGGACAA 60
                  ************************************************************

  full            TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
  TV4             TGGGGGTTCAGGCTGGTTGCCTTCCTTCTCCTCTTGTGCCACAGACAGGGCGCTGAAGCC 120
                  ************************************************************

  full            TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
  TV4             TGCGGTACCGTCGGATGCTGTTTTCAGAACCCACCATACCCAGATGCAGACTCAGGCTCA 180
                  ************************************************************

  full            GCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCGCTGGTTATGAATGT 240
  TV4             GCTTCAGGCCCCCGGGCCCTGAGCTGCTACCGGCTGTCCAGCAGCGCTGGTTATGAATGT 240
                  ************************************************************

  full            TCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCAGATGCTGCCTCCAG 300
  TV4             TCCTGGGAGTATGAGGGCCCTGCAGCTGGAGTCATCCACTTCCTCAGATGCTGCCTCCAG 300
                  ************************************************************

  full            TCTGGCCGCTGCTGCTTCTTTGCCACAGGCTCAGCCACCAAGCTGCAGTTCTCCGACCAG 360
  TV4             TCTGGCTGCTGCTGCTTCTTTGCCACAG-------------------------------- 328
                  ****** ********************

  full            GATGGCATATCCGTGCTCCACAATGTCACTCTCTGGGTGGAATCCCGGGCCGCCAACCGG 420
  TV4             ------------------------------------------------------------

  full            ACAGAGAAGTCCCCCAATGTTACTCTGAACCTCTACAGCTCGGTTAAATACGACCCTCCC 480
  TV4             ------------------------------------------------------------

  full            CCAGGAAACATCAAGGTGTCCAAGTCAGCAGGGAAGCTGCGTATGGAGTGGGAGACCCCA 540
  TV4             ------------------------------------------------------------

  full            GCCCGCCAGGATGGTGCCGAGGTACAGTTCCGGCACCGCACACCTGGCAGCCCGTGGAAG 600
  TV4             ------------------------------------------------------------

  full            CTGGGTGACTGTGGACGTCAGGATGATGCTGGCTTCGAGTCATGCCTCTGCCCCTTGGAG 660
  TV4             ------------------------------------------------------------

  full            ATGGACATGGCCCAGGAATTCCAGCTGCGGCGGCGTCTGGGGCCAGGGGTCCCCGGAGGT 720
  TV4             ------------------------------------------------------------

  full            CCCTGGAGCAGCTGGAGCAGCCCTGTGTGCATCCCCCCTGAAACCCCTCCACAGGCTGAG 780
  TV4             ------------------------------------------------------------

  full            GTGAGGTTCTCAGCAGAGCAGCTTTGCCCAGATGGGAGGAGGCAGGTGGCCTTACATGAG 840
  TV4             ------------------------------------------------------------

  full            CAGCTGCCCCAACTGGAGCTTCCAGAAGGCTGCCTTGGGCCTGACCCTGGCATGGAGGTG 900
```

```
TV4             ------------------------------------------------------------

full            ACCTACAAGATCCATCTCCACATGCTGTCCTGCCCATGTAAGACCAGGGCCAAGAAGACT 960
TV4             ------------------------------------------------------------GCT 331
                                                                             .**

full            CTGCGCCTGAAGAGAAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 1020
TV4             CTGCGCCTGAAGAGCAAGCTCGTCCTCTCGAGTGCTGCCTACGATCTGACCATCGTTTCC 391
                **************.**********************************************

full            CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 1080
TV4             CAGAATCGCTTTGGCCTCGGCCCCAATCAGACATGGCGCATTCCTGCCTGCATCCACTCA 451
                ************************************************************

full            GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 1140
TV4             GAACCAGGGCTTCTGAATATCAGTGCGGGAGCCAACGGGACCACCATGCATTGGCCAGCC 511
                ************************************************************

full            CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 1200
TV4             CGGGCCCAGGGAATGAGGTACTGCATTGAGTGGCAGCTCCAGGGCCAGGAGGAGAACCTT 571
                ************************************************************

full            GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 1260
TV4             GCCGCCTGCATCGTGACGGCACCCCAGGACGCAGACGCTGCTGGAATGGCAACTCACAGC 631
                ************************************************************

full            TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 1320
TV4             TGGAGCCAAGAATCTGGAGCACTGGCACAGAAGGCATGTTACCACATCGCCATCTTTGCC 691
                ************************************************************

full            TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 1380
TV4             TCTGCACACCCGGAGAAGCTAACCTCCTGGTCTACAGTCTTGTCCACCTACCATTTTGGA 751
                ************************************************************

full            GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 1440
TV4             GGCAATGCCTCAGCGGCCGGAAGCCCACAGCACGTGTCGGTGAGGAAGCTCAGCCAGGAT 811
                ************************************************************

full            TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 1500
TV4             TCGGTGTCTGTGGACTGGACACCATCTCTGCTGAGCACCTGCCCTGGCGTCCTGAAGGAG 871
                ************************************************************

full            TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 1560
TV4             TACGTTGTGTACTTCCAGGAGGAGGACAGCAACCAAGCGTCTGAGCTGCACGTGAAGCCC 931
                ************************************************************

full            ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 1620
TV4             ACAGAGACCCAGGTCACCCTCCAAGGCCTGCGGGCTGGCACAGCTTACAAGGTGCAGGTT 991
                ************************************************************

full            CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1680
TV4             CGCGCAGACACAGCCAAGTGGAGGGGCGCCTGGAGCCAGCCCCTGCATTTCACCATCGAA 1051
                ************************************************************

full            GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1740
TV4             GTCCAGGTTTCTGAGTTGTCCGACTTGTCCATCTTCCTCGCATCTTTGGGGAGTTTCGTG 1111
                ************************************************************

full            AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1800
TV4             AGCATCCTTCTCCTGGGAATCTTTGGGTACCTCAGCTTGAACAGGGCTGTAAGGCACCTG 1171
                ************************************************************

full            TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1860
TV4             TGCCCACCCTTGCCCACACCTGGTGCCAGCACTGCCATCAAGTTCTCTGGCAGCCAGGGG 1231
                ************************************************************

full            AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1920
TV4             AAGCAGGTTTGGCAGTGGACCAGCCCAGCAGACTTCCCAGAGGAGGTGTCTCCACACGAG 1291
                ************************************************************

full            GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1980
TV4             GCCCTGATGGTGAATATATCCTGGGAAAAAGGCGAGGGAGCTGACATGGGCACACTTGGG 1351
                ************************************************************
```

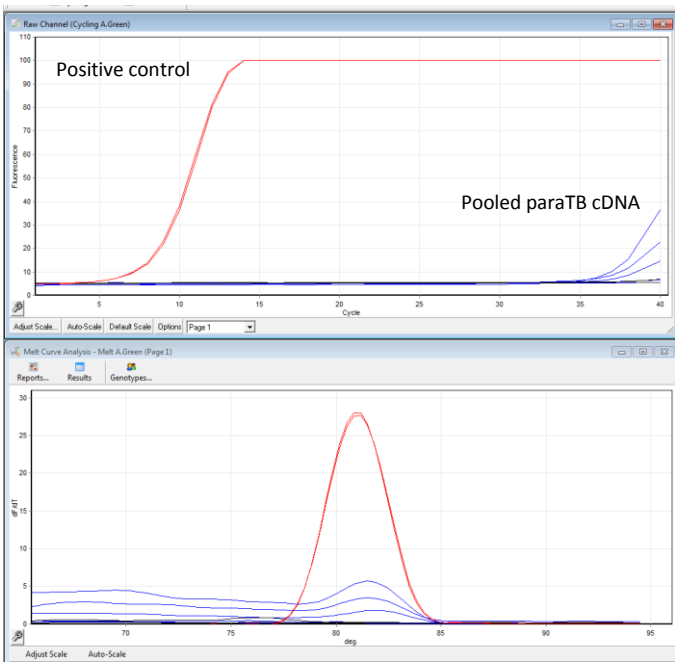
```
full            ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 2040
TV4             ACTCTCAAGGAGAAGATGGAGCTGCCTCTGCGTGCCCCTAAGCCAGCCCCGGACACAGAG 1411
                ************************************************************

full            CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 2100
TV4             CTGCCCTTGAAGGACAGGAAGCAGATGCAAGGATGCCCTGAGGCTGGGGCTCTGAGGCCT 1471
                ************************************************************

full            GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 2160
TV4             GGCTGGCAGGACGGTCTGGTGGAGGACAGCCTTGCCCAGGTGGCTAGACCCCCACTGCTG 1531
                ************************************************************

full            CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 2220
TV4             CTCCTGGGGGGCCTGAGGCAGGCCCCCAGGTTTGGTTCCCAGGGAGAAGCAGGAACAGCT 1591
                ************************************************************

full            GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 2280
TV4             GCCTCCTCTTACAGAGAAGACTGACAACTGCCTCCAGGACGGCTGGCTGTTGTTCACTGC 1651
                ************************************************************

full            ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 2340
TV4             ACACTTGGCCCGTTTTTTCCAGACACTCACGTGATGCCCTTGAGCGATATGGCTGTCCTA 1711
                ************************************************************

full            TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 2382
TV4             TTGACCTATTTCAGTGAGCCTCAGTGAGGTTCAGAGAGGGCA 1753
                ******************************************
```

IL12RB1v4 protein

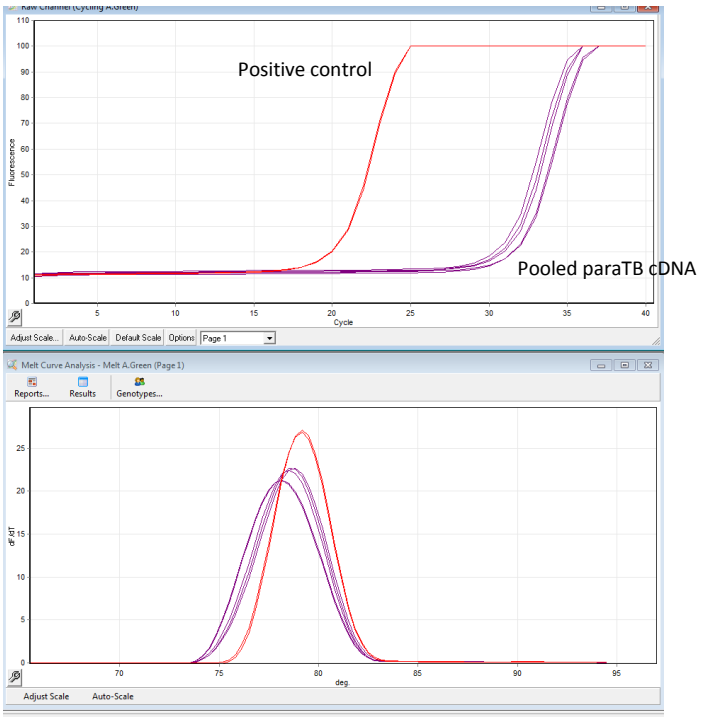
```
Full            MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
TV4             MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSSSAG 60
                ************************************************************

Full            YECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVESRA 120
TV4             YECSWEYEGPAAGVIHFLRCCLQSGCCCFFATGSAP------------------------ 96
                ************************ *********.

Full            ANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRTPGS 180
TV4             ----------------------------------------EEQARPLECCLRSDHRFPES 116
                                                         *  **   . ::  ** * *

Full            PWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPETPP 240
TV4             LWP-----RPQSDMAHSCLHP----------LRTR------------------------- 136
                 *      * :.   .*** *          ** *

Full            QAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCKTRA 300
TV4             ------ASEYQCGSQRDHHALAS-------------PGPGNEVLHS-------------- 163
                      ::*  * . * : ** .             *.** ** :.

Full            KKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGTTMH 360
TV4             ------------------------------------------------------------


Full            WPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACYHIA 420
TV4             ------------------------------------------------------------


Full            IFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTCPGV 480
TV4             ------------------------------------------------------------


Full            LKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQPLHF 540
TV4             ------------------------------------------------------------


Full            TIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIKFSG 600
TV4             ------------------------------------------------------------
```

```
Full            SQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPKPAP 660
TV4             ------------------------------------------------------------

Full            DTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQGEA 720
TV4             ------------------------------------------------------------

Full            GTAASSYRED 730
TV4             ----------
```

A6. *IL12RB1 v1* RT-QPCR positive control assay



A7. *IL12RB1 v2* RT-QPCR positive control assay



A8. IL12RB1 WSxWS motif and transmembrane analysis

Human IL12RB1: AAI37407.1

WSxWS motif

Transmembrane

```
human      MEPLVTWVVP---LLFLFLLSRQGAACRTSECCFQDPPYPDADSGSASGPRDLRCYRISS
sheep      ---MGQWGFRLVAFLLLLCHRQGAEACGTVGCCFQNPPYPDADSGSASGPRALSCYRLSS
              :  * .    :*:*:   : . ** *  ****:*************** * ***:**

human      -DRYECSWQYEGPTAGVSHFLRCCLSSGRCCYFAAGSATRLQFSDQAGVSVLYTVTLWVE
sheep      SAGYECSWEYEGPAAGVIHFLRCCLQSGRCCFFATGSATKLQFSDQDGISVLHNVTLWVE
              *****:****:*** *******.*****:**:****:****** *:***:.******

human      SWARNQTEKSPEVTLQLYNSVKYEPPLGDIKVSKLAGQLRMEWETPDNQVGAEVQFRHRT
sheep      SRAANRTEKSPNVTLNLYSSVKYDPPPGNIKVSKSAGKLRMEWETPARQDGAEVQFRHRT
           * * *:*****:***:**.****:** *:***** **:******** .* **********

human      PSSPWKLGDCGPQDD-DTESCLCPLEMNVAQEFQLRRRQ-LGSQGSSWSKWSSPVCVPPE
sheep      PGSPWKLGDCGRQDDAGFESCLCPLEMDMAQEFQLRRRLGPGVPGGPWSSWSSPVCIPPE
           *.********* ***   *********::*********   *  *. **.******:***

human      NPPQPQVRFSVEQLGQDGRRRLTLKEQPTQLELPEGCQGLAPGTEVTYRLQLHMLSCPCK
sheep      TPPQAEVRFSAEQLCPDGRRQVALHEQLPQLELPEGCLGPDPGMEVTYKIHLHMLSCPCK
           .*** :****.***  ****:::*:**  ******** *  ** ****:::**********

human      AKATRTLHLGKMPYLSGAAYNVAVISSNQFGPGLNQTWHIPADTHTEPVALNISVGTNGT
sheep      TRAKKTLRLKRKLVLSSAAYDLTIVSQNRFGLGPNQTWRIPACIHSEPGLLNISAGANGT
           ::*.:**:* :   **.***:::::*.*:** * ****:***  *:**  ****.*:***

human      TMYWPARAQSMTYCIEWQPVGQDGGLATCSLTAPQDPDPAGMATYSWSRESGAMGQEKCY
sheep      TMHWPARAQGMRYCIEWQLQGQEENLAACIVTAPQDADAAGMATHSWSQESGALAQKACY
           **:******.* ******  **:  **:* :***** * *****:***:****:.*: **

human      YITIFASAHPEKLTLWSTVLSTYHFGGNASAAGTPHHVSVKNHSLDSVSVDWAPSLLSTC
sheep      HIAIFASAHPEKLTSWSTVLSTYHFGGNASAAGSPQHVSVRKLSQDSVSVDWTPSLLSTC
           :*:********** ******************:*:****:: * *******:*******

human      PGVLKEYVVRCRDEDSKQVSEHPVQPTETQVTLSGLRAGVAYTVQVRADTAWLRGVWSQP
sheep      PGVLKEYVVYFQEEDSNQASELHVKPTETQVTLQGLRAGTAYKVQVRADTAKWRGAWSQP
           *********  ::***:*.**  *:********.*****.**.********  **.****

human      QRFSIEVQVS---DWLIFFASLGSFLSILLVGVLGYLGLNRAARHLCPPLPTPCASSAIE
sheep      LHFTIEVQVSELSDLSIFLASLGSFVSILLLGIFGYLSLNRAVRHLCPPLPTPGASTAIK
            :*:******   *  **:******:****:*::***.****.********** **:**:

human      FPG--GKETWQWINPVDFQEEASLQEALVVEMSWDKGERTEP------LEKTELPEGAPE
sheep      FSGSQGKQVWQWTSPADFPEEVSPHEALMVNISWEKGEGADMGTLGTLKEKMELPLRAPK
           * *  **:.*** .*.** **.* :***:*::**:*** ::        ** ***  **:

human      LALDTELSLEDGDRCKAKM-----------------------------------------
sheep      PAPDTELPLKDRKQMQGCPEAGALRPGWQDGLVEDSLAQVARPPLLLLGGLRQAPRFGSQ
            * **** *:* .: :.

human      --------------
sheep      GEAGTAASSYREDS
```

A9. *IL17RA* gene sequence aligned to OARv3.1

OARV3.1 Chromosome 3 IL17RA

```
Query  2          CTGTCTGC-AAATACAAGGTCCTGAAACAACAAAGACACACTCGCTGATGACGGATGTTG  60
                  ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212808572  CTGTCTGCAAAATACAAGGTCCTGAAACAACAAAGACACACTCGCTGATGACGGATGTTG  212808631

Query  61         GGGCAGTAAATCAGAAGTCAGCCTGGAGCAGGAGCCTCTGTCAGCCAGGAGCGATGGAGT  120
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212808632  GGGCAGTAAATCAGAAGTCAGCCTGGAGCAGGAGCCTCTGTCAGCCAGGAGCGATGGAGT  212808691

Query  121        TCCAGGGAACAGAGAAGCTCCTGCTGCCCGGGTGCCCTTCCTGCAGAGAGGGCACTTGGG  180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212808692  TCCAGGGAACAGAGAAGCTCCTGCTGCCCGGGTGCCCTTCCTGCAGAGAGGGCACTTGGG  212808751

Query  181        CAGCCCTCCAAGAGAGTCTAGAGCATTCCAGGAGGGGGAAATGACTGAGCCAAGGCCCGA  240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212808752  CAGCCCTCCAAGAGAGTCTAGAGCATTCCAGGAGGGGGAAATGACTGAGCCAAGGCCCGA  212808811

Query  241        AGG  243
                  |||
Sbjct  212808812  AGG  212808814
```

>>>>>> EXON 1??<<<<<<<

```
Query  237        CCGAAGGGGCTGAACTGCACAGTCAAGAACAGTA  270
                  ||| ||||||||||||||||||||||||||||||
Sbjct  212811829  CCGCAGGGGCTGAACTGCACAGTCAAGAACAGTA  212811862

Query  268        GTACCTGCCTGGACGACAGCTGGATCCACCCTCGGAACCTCACCCCCTCATCCCCCAAAA  327
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212812436  GTACCTGCCTGGACGACAGCTGGATCCACCCTCGGAACCTCACCCCCTCATCCCCCAAAA  212812495

Query  328        ACGTGCAGACCCAGCTGCGCTTCGCCCACACACAGCAAGGACACTTGCTTCCCGTGGTTC  387
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212812496  ACGTGCAGACCCAGCTGCGCTTCGCCCACACACAGCAAGGACACTTGCTTCCCGTGGTTC  212812555

Query  388        ACATCGAGTGGACGCTGCAGACAGATG  414
                  ||||||||||||||||||||||||||
Sbjct  212812556  ACATCGAGTGGACGCTGCAGACAGATG  212812582
```

```
Query  414        GCCAGCGTCCTCTACCTGGAGGGGGCGGAGCTGTCCATCCTGCAGCTGAGCACCAATGAG  473
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212813274  GCCAGCGTCCTCTACCTGGAGGGGGCGGAGCTGTCCATCCTGCAGCTGAGCACCAATGAG  212813333

Query  474        CGTCTGTGTGTCAGGTTTGAGTTTCTGACCACACTGAGGCATCATCACAAGCGG  527
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212813334  CGTCTGTGTGTCAGGTTTGAGTTTCTGACCACACTGAGGCATCATCACAAGCGG  212813387
```

```
Query  527        GTGGCGATTTGCCTTCAGCCACTTTGTGGTAGAACCCAGAGAGGAGTACGAGGTGACCGT  586
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212814463  GTGGCGATTTGCCTTCAGCCACTTTGTGGTAGAACCCAGAGAGGAGTACGAGGTGACCGT  212814522

Query  587        CCACCACCTGCCTAAGCCCATCCCTGACGGGGACCCAAACCACCAATCCAGAAACTTCCT  646
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212814523  CCACCACCTGCCTAAGCCCATCCCTGACGGGGACCCAAACCACCAATCCAGAAACTTCCT  212814582

Query  647        GGTGCCCG  654
                  ||||||||
Sbjct  212814583  GGTGCCCG  212814590
```

```
Query  641        CTTCCTGGTGCCC-GACTGCAAGGACCCCAGGATGAAGGACACCACGCCATGCGTGAGCT  699
                  ||||| ||| | | ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212816102  CTTCCCGGTCCACAGACTGCAAGGACCCCAGGATGAAGGACACCACGCCATGCGTGAGCT  212816161

Query  700        CAGGCAGCCTGTGGG  714
                  |||| | || |||||
```

```
Sbjct  212816162  CAGGTAACC-GTGGG  212816175


Query  701        AGGCAGCCTGTGGGACCCCAACATCACCGTGGAGACCCTTGAGGCCCACCAGCTGCGGCT  760
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212816313  AGGCAGCCTGTGGGACCCCAACATCACCGTGGAGACCCTTGAGGCCCACCAGCTGCGGCT  212816372

Query  761        GAGCTTCACCCCATGGAATGAGTCCACCAGTTACCAGGTCCTGCTGCACAGCTTCCCGCC  820
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212816373  GAGCTTCACCCCATGGAATGAGTCCACCAGTTACCAGGTCCTGCTGCACAGCTTCCCGCC  212816432

Query  821        TGCAGAGAACCAGAGCTGCTTCCAACATGTTGTCGACATGCCCGTG  866
                  |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212816433  TGCAGAGAACCAGAGCTGCTTCCAACATGTTGTCGACATGCCCGTG  212816478


Query  866        GCCCGCACAGGAAGCCGCCCCGCAGCGCTGCCACATCACGGTCACCCTGCTGGACTCCAG  925
                  ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  212817228  GCCCGCACAGGAAGCCGCCCCGCAGCGCTGCCACATCACGGTCACCCTGCGGGACTCCAG  212817287

Query  926        CTGGTGCTGCCGCCACCACGTGCAG  950
                  |||||||||||||||||||||||||
Sbjct  212817288  CTGGTGCTGCCGCCACCACGTGCAG  212817312

Query  948        CAGATCCAGCCCTTCTTCAGCAGCTGCCTCAACGACTGCCTCCGACACTCGGTGTCCATG  1007
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  212817546  CAGATCCAGCCCTTCTTCAGCAGCTGCCTCAACGACTGCCTCCGACACTCGGTGTCCGTG  212817605

Query  1008       GCCTGCCCGGAGGTCTCACACACCCCAG  1035
                  ||||||||||||||||||||||||||||
Sbjct  212817606  GCCTGCCCGGAGGTCTCACACACCCCAG  212817633


Query  1046       AGACCACACGCCTCTGTGGGTGTCTGCGTTCATCACAGGCCTCTCCATCCTACTGGTGGG  1105
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212818524  AGACCACACGCCTCTGTGGGTGTCTGCGTTCATCACAGGCCTCTCCATCCTACTGGTGGG  212818583

Query  1106       CTCCGTCATCCTGCTGATCCTCTGCATGACCTGGAGGCTACCAGG  1150
                  |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212818584  CTCCGTCATCCTGCTGATCCTCTGCATGACCTGGAGGCTACCAGG  212818628


Query  1147       CAGGGTTCCGTCAAGGAAAACATGAAGATGGCACCAAAGACACAG  1191
                  |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212819048  CAGGGTTCCGTCAAGGAAAACATGAAGATGGCACCAAAGACACAG  212819092


Query  1189CAGAGATCCTGCCCGCTGCCACCAGCCTGAcccccccGCCCCTGAAGCCCAGGAAGGTCT  1248
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212819604  CAGAGATCCTGCCCGCTGCCACCAGCCTGACCCCCCCGCCCCTGAAGCCCAGGAAGGTCT  212819663

Query  1249       GGATCGTCTACTCCGCCGACCACCCCCTCTACGTGGACGTGGTCCTCAAGTTCGCCCA  1306
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  212819664  GGATCGTCTACTCCGCCGACCACCCCCTCTACGTGGACGTGGTCCTCAAGTTCGCCCA  212819721
```

Ovisaries breed Texel unplaced genomic scaffold, Oar_v3.1 scaffold005641
Sequence ID: ref|NW_004083240.1| **Length: 4151 Number of Matches: 1**
Related Information

```
Query  1302GCCCATGGGCCCAGCTTGGGGAGGcctgggggccaggctggggtTGGGGCATCAGGCCCT  1361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3571  GCCCATGGGCCCAGCTTGGGGAGGCCTGGGGGCCAGGCTGGGGTTGGGGCATCAGGCCCT  3512

Query  1362  GATGCTAAAGACCAGGGCTTCTTTACCAGGATGCCCTGTTCCCCGCAGAGGAGGCTGTAG  1421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3511  GATGCTAAAGACCAGGGCTTCTTTACCAGGATGCCCTGTTCCCCGCAGAGGAGGCTGTAG  3452

Query  1422  ATCAGACACACTGGTCATTTCCAGAGAAAAACTGGATAGAAGTGAATGAGCCCGGATGGC  1481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  3451  ATCAGACACACTGGTCATTTCCAGAGAAAAACTGGATAGAAGTGAATGAGCCCGGATGGC  3392

Query  1482  CTCAGCAGGGCCTGAACCCGCATGTGGCCTCGGTCAGTGGGAGAGCTTTGGCCAAAGGAT  1541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  3391  CTCAGCAGGGCCTGAACCCGCATGTGGCCTCGGTCAGTGGGAGAGCTTTGGCCAA-GGAT  3333

Query  1542  TCTTCCAGGGTTGCGGGCTGAGCACTGGGCTATAGCATCCAGGCTCAAGAAGAACAAGGA  1601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3332  TCTTCCAGGGTTGCGGGCTGAGCACTGGGCTATAGCATCCAGGCTCAAGAAGAACAAGGA  3273

Query  1602  GCCCCCCGCCCTGTGCATACTGGGGTCACTGTAGGTTCTCTTTGCCTGAACCTGAGTCCT  1661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3272  GCCCCCCGCCCTGTGCATACTGGGGTCACTGTAGGTTCTCTTTGCCTGAACCTGAGTCCT  3213

Query  1662  C-ACTGGAGGTTCAAGCACCTGCAAGGAG-AGGGGAT-CTGAAATTCA-G-AG-AATCAG  1715
              | |||||||||||||||||||||||||| ||||||| |||||||||| | || ||||||
Sbjct  3212  CCACTGGAGGTTCAAGCACCTGCAAGGAGGAGGGGATTCTGAAATTCAAGGAGAAATCAG  3153

Query  1716  ATACCGCGTGTAACATGTGTGTGGTGTTGTGAA-GCGAGCCTCCACTGATAAA-G-ATGG  1772
              |||||||||||| |||||||||| ||| ||||| |||||| || ||||||||| | |||
Sbjct  3152  ATACCGCGTGTA-CATGTGTGTG-TGT-GTGAAAGCGAGC-TC-ACTGATAAAAGAATGA  3098

Query  1773  ATAAGTTTT-AA-ctctga  1789
              |||| |||| || ||||||
Sbjct 3097  ATAAATTTTTAAACTCTGA  307
```

A10. *IL17RB* TRAF-6 binding domain and SEFIR region analysis

TRAF-6 binding domain

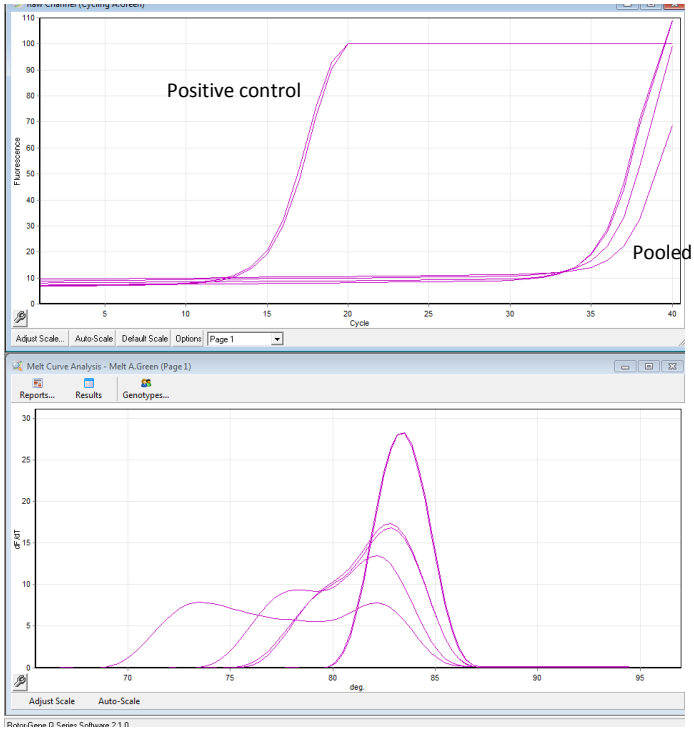
```
Ovine_V2        MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
Ovine_V3        MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
Ovine           MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
Ovine_V1        MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60
Homo            MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATG 60
Mus             MLLVLLLILAASCRSALPREPTIQCGSETGPSPEWMVQHTLTPGDLRDLQVELVKTSVAAE 60
                * **** *** * .*:* ***:*****.*******::* * *******:** :.:**

Ovine_V2        DSSILMNISWILRADVDV------------------------FLHWFSCRAEHSLFHWS 95
Ovine_V3        DSSILMNISWILRADVDV------------------------FLHWFSCRAEHSLFHWS 95
Ovine           DSSILMNISWILRADASIRLLKATKICVMGKSHFQSRSCIRCNYTQAFRTQTRHSGGKWT 120
Ovine_V1        DSSILMNISWILRADASIRLLKATKICVMGKSHFQSRSCIRCNYTQAFRTQTRHSGGKWT 120
Homo            DYSILMNVSWVLRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWT 120
Mus             EFSILMNISWILRADASIRLLKATKICVSGKNNMNSYSCVRCNYTEAFQSQTRPSGGKWT 120
                : *****:**:****..:                       : . *  ::. *  :*:

Ovine_V2        PS---------------------------------------------------------- 97
Ovine_V3        PS---------------------------------------------------------- 97
Ovine           FSYIGFPVELNTVYFIGAHNIPNANMNEDGPSMAVNFTSPGCLDHVMKYKKKCIEAGSLW 180
Ovine_V1        FSYIGFPVELNTVYFIGAHNIPNANMNEDGPSMAVNFTSPGCLDHVMKYKKKCIEAGSLW 180
Homo            FSYIGFPVELNTVYFIGAHNIPNANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLW 180
Mus             FSYVGFPVELSTLYLISAHNIPNANMNEDSPSLSVNFTSPGCLNHVMKYKKQCTEAGSLW 180
                 *

Ovine_V2        ------------------------------------------------------------
Ovine_V3        ------------------------------------------------------------
Ovine           KPNITACKKSANTVEVNFTTSPLGDRYMALIQNTSVIGTSYVSEK---ELTRTSVVVHVT 237
Ovine_V1        KPNITACKKSANTVEVNFTTSPLGDRYMALIQNTSVIGTSYVSEK---ELTRTSVVVHVT 237
Homo            DPNITACKKNEETVEVNFTTTPLGNRYMALIQHSTIIGFSQVFEPHQKKQTRASVVIPVT 240
Mus             DPDITACKKNEKMVEVNFTTNPLGNRYTILIQRDTTLGFSRVLEN---KLMRTSVAIPVT 237


Ovine_V2        ------------------------------------------------------------
Ovine_V3        ------------------------------------------------------------
Ovine           GESEGAVVQLTPYFHTCGNDCIRQRGTVVRCPQT-GVSPQDHERSVLGGWLPLLLSALLV 296
Ovine_V1        GESEGAVAQLTPYFHTCGNDCIRQRGTVVRCPQT-GVSPQDHERSVLGGWLPLLLSALLV 296
Homo            GDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNNKSKPGGWLPLLLLSLLV 300
Mus             EESEGAVVQLTPYLHTCGNDCIRREGTVVLCSETSAPIPPDDNRRMLGGWLPLFLV-LLV 296


Ovine_V2        ------------------------------------------------------------
Ovine_V3        ------------------------------------------------------------
Ovine           ATWVLVAGIYLIWRHERIKKTSFST-TTLLPSLKVLVVYPSEICFHHTVCYFTEFLQNRC 355
Ovine_V1        ATWVLVAGIYLIWRH---------------------------------VCYFTEFLQNRC 323
Homo            ATWVLVAGIYLMWRHERIKKTSFST-TTLLPPIKVLVVYPSEICFHHTICYFTEFLQNHC 359
Mus             AVWVLAAGIYLTWRQGRSTKTSFPISTMLLPLIKVLVVYPSEICFHHTVCRFTDFLQNYC 356


Ovine_V2        ------------------------------------------------------------
Ovine_V3        ------------------------------------------------------------
Ovine           RSEVILEKWQKKKIAEMGPVQWLTTQKQAADKVIFLLSNDN-TMCDGTCDKKEGGPCENS 414
Ovine_V1        RSEVILEKWQKKKIAEMGPVQWLTTQKQAADKVIFLLSNDN-TMCDGTCDKKEGGPCENS 382
Homo            RSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENS 419
Mus             RSEVILEKWQKKKIAEMGPVQWLTTQKQAADKVVFLLPSDVPTLCDSACGHNEGSARENS 416


Ovine_V2        ------------------------------------------------------------
Ovine_V3        ------------------------------------------------------------
Ovine           QDLFHLAFNLFCSDLRSQTHLRKYVVVYFREGDIKDSYSALSVCPTYRLTKDATDFCAEL 474
Ovine_V1        QDLFHLAFNLFCSDLRSQTHLRKYVVVYFREGDIKDSYSALSVCPTYRLTKDATDFCAEL 442
Homo            QDLFPLAFNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAEL 479
Mus             QDLFPLAFNLFCSDFSSQTHLHKYLVVYLGGADLKGDYNALSVCPQYHLMKDATAFHTEL 476


Ovine_V2        ------------------------
Ovine_V3        ------------------------
Ovine           LHAKQHVSVGRRSRARHYSCLSL 497
```

```
Ovine_V1        LHAKQHVSVGRRSRARHYSCLSL 465
Homo            LHVKQQVSAGKRSQACHDGCCSL 502
Mus             LKATQSMSVKKRSQACHDSCSPL 499
```

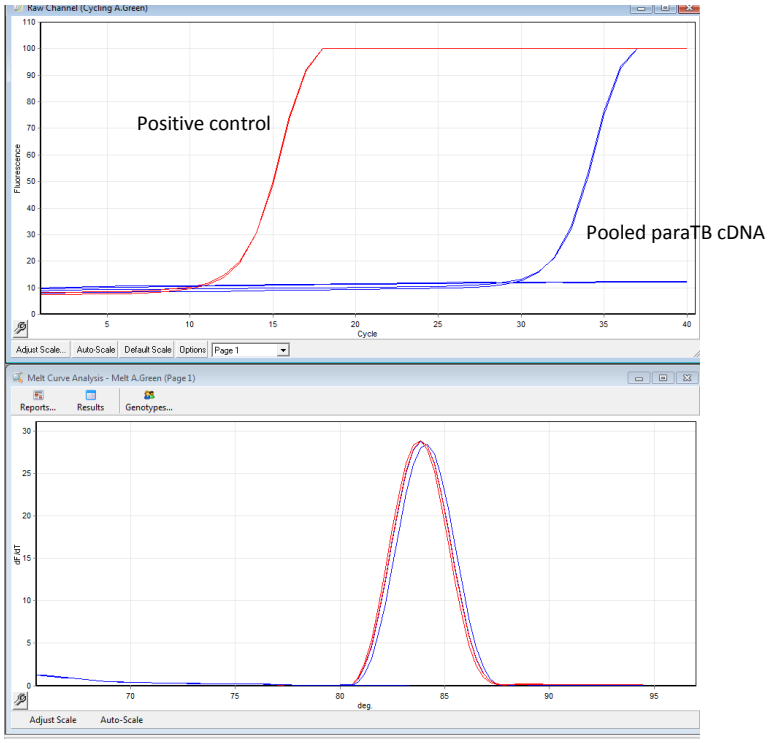
Alignment of ovine protein sequence (IL17RB) with Bos Taurus SEFIR domain: GenBank: AAI33637.1

```
bos             ------------------------------------------------------------
ovine           MLLVLLSLAALCWGAVPPEPTIQCGSEPGPSPEWMVRHTLTPGDLRDLRVEPIKSSVDLE 60

bos             ------------------------------------------------------------
ovine           DSSILMNISWILRADASIRLLKATKICVMGKSHFQSRSCIRCNYTQAFRTQTRHSGGKWT 120

bos             ------------------------------------------------------------
ovine           FSYIGFPVELNTVYFIGAHNIPNANMNEDGPSMAVNFTSPGCLDHVMKYKKKCIEAGSLW 180

bos             ------------------------------------------------------------
ovine           KPNITACKKSANTVEVNFTTSPLGDRYMALIQNTSVIGTSYVSEKELTRTSVVVHVTGES 240

bos             ------------------------------------------------------------
ovine           EGAVVQLTPYFHTCGNDCIRQRGTVVRCPQTGVSPQDHERSVLGGWLPLLLSALLVATWV 300

bos             ----------------------------IKVLVVYPSEICFHHTVCYFTEFLQNRCRSEVI 33
ovine           LVAGIYLIWRHERIKKTSFSTTTLLPSLKVLVVYPSEICFHHTVCYFTEFLQNRCRSEVI 360
                                            :*******************************

bos             LEKWQKKKIAEMGPVQWLTTQKEAADKVIFLLSNGNTTCDGNCDEKEGGPCESSRDLFHL 93
ovine           LEKWQKKKIAEMGPVQWLTTQKQAADKVIFLLSNDNTMCDGTCDKKEGGPCENSQDLFHL 420
                ***********************:************.** ***.**:*******.*:*****

bos             AFNLFCSDLRSQAHLHKYVVVYFREGDIADSYRALSVCPTYRLTKDATGFCAEL------ 147
ovine           AFNLFCSDLRSQTHLRKYVVVYFREGDIKDSYSALSVCPTYRLTKDATDFCAELLHAKQH 480
                ************:**:************ *** ***************.*****

bos             -----------------
ovine           VSVGRRSRARHYSCLSL 497
```

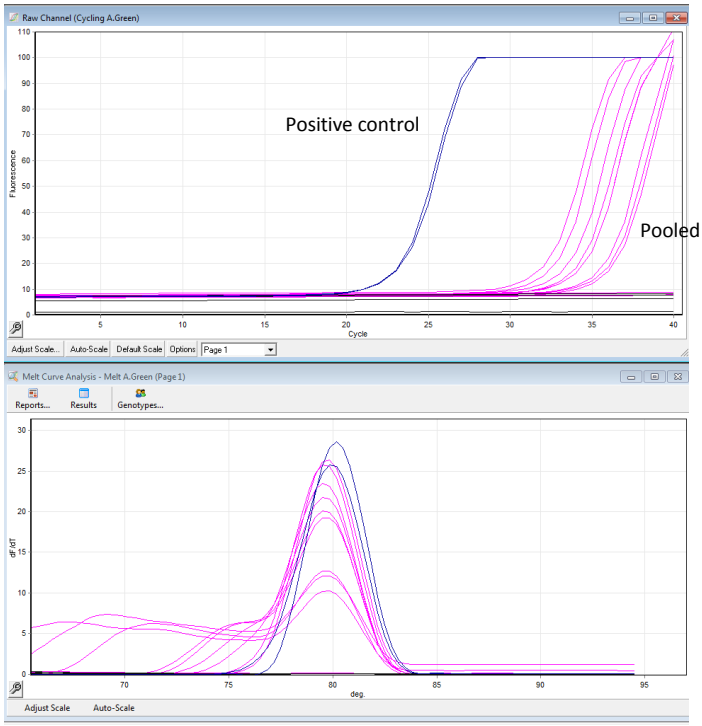
A11. *IL17RB v1* RT-QPCR positive control assay



A12. *IL25* RT-QPCR positive control assay



A13. *RORAv2* RT-QPCR positive control assay



A14. *RORAv3* RT-QPCR positive control assay