Expression profiling reveals differences in immuno-inflammatory gene expression between the two disease forms of sheep paratuberculosis

Elsevier Editorial System(tm) for Veterinary Immunology and Immunopathology
Manuscript Draft

Manuscript Number: VETIMM-D-09-1576R1

Title: Expression profiling reveals differences in immuno-inflammatory gene expression between the two disease forms of sheep paratuberculosis

Article Type: Research Paper - Full length

Keywords: Mycobacterium avium subsp. paratuberculosis
MAP
Sheep
Chemokines
Genomics
Pathology.

Corresponding Author: Professor J. Hopkins,

Corresponding Author's Institution: University of Edinburgh, Royal (Dick) School of Vet. Studies

First Author: Jennifer A Smeed, BSc. PhD.

Order of Authors: Jennifer A Smeed, BSc. PhD. ; Craig A Watkins, BSc. PhD; Anton G Gossner, BSc. PhD; J. Hopkins

Revision Note

Professor John Hopkins
THE ROSLIN INSTITUTE
R(D)SVS
University of Edinburgh
Summerhall
Edinburgh, EH9 1QH
UK

Telephone +44 (0)131 650 6169
Fax +44 (0)131 650 6511
john.hopkins@ed.ac.uk
www.roslin.ed.ac.uk

The Editor
Veterinary Immunology & Immunopathology

26 Oct 2009

Ms. No. VETIMM-D-09-1576

Veterinary Immunology and Immunopathology

Many thanks to the referees for positive and informative comments. We have modified the manuscript in accordance with these comments.

Reviewer #1.

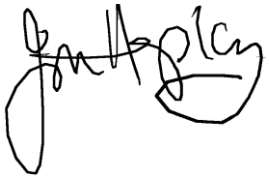
1. The title has now been changed to better reflect the content of the paper.
2. We have amended the Abstract to make unambiguous that we have measured transcripts and not protein.
3. Clinical signs clarified etc. Have explained and referenced IS900
4. We have included descriptions of results for chemokines/receptors/TCR molecules in the Results. I have reformatted Table 1 so that the genes are in strict alphabetic order; it is not possible to arrange them in comparison order as there would be some duplication.
5. The point of the RT-qPCR is validation of the array results – and this it does (section 3.2). The rationale for their selection is now explained in section 3.2.
6. IGFBP6 was represented in the array (see supplementary data Array Express) – but in these experiments it did not show differential expression of >1.5 and $p \leq 0.05$ in any comparison. As referenced in the manuscript we had previously identified it in a preliminary series of experiments (Roupaka MSc thesis 2004).
7. There was strong correlation between array and RT-qPCR (section 3.2) but difference in absolute values – this is a common finding and comments/references etc have now been added to the Discussion.
8. MMP9 and TYROBP were not discussed specifically because the purpose of the paper was not to cover all genes exhaustively, but to focus on areas of particular interest. If required I can add a paragraph to the discussion – but it will not add much.

Minor comments…. Asterisks explained in Table 1. SDHA/YWHAZ defined.

Reviewer #2.

1. See comment 1 and 2 (Reviewer #1).
2. Abbreviations now defined.
3. Further detailed choice of housekeeping genes. The two used were selected from six; including ribosomal protein L18, ATPase, GAPDH, β actin as well as SDHA and YWHAZ. I can list these in the text if you wish – but it is not very relevant to the paper.
4. MIQE guidelines followed for abbreviations.
5. Statistics further explained, end section 2.5.
6. References added and clarified. Rather that adding several references for well known facts I have included a reference for an informative review (Campbell *et al.* 2003).
7. Fig 1 legend better described.

Very many thanks,



Expression profiling reveals differences in immuno-inflammatory gene expression between the two disease forms of sheep paratuberculosis

J. A. Smeed[a,c], C. A. Watkins[b,c], A. G. Gossner[a] , J. Hopkins[a,*]

[a]*The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Summerhall, Edinburgh, EH9 1QH, UK.*

[b]*Moredun Research Institute, International Research Centre, Pentlands Science Park, Penicuik, Midlothian EH26 0PZ, UK.*

[c] Equal contribution

*Corresponding author at: The Roslin Institute & Royal (Dick) School of Veterinary Studies, University of Edinburgh, Summerhall, Edinburgh, EH9 1QH, UK. Tel. +44 131 650 6169; Fax. +44 131 650 6511.

E-mail address: john.hopkins@ed.ac.uk

Abstract

Paratuberculosis is a chronic enteropathy of ruminants caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP); infection of sheep results in two disease forms – paucibacillary (tuberculoid) and multibacillary (lepromatous) associated with the differential polarization of the immune response. In addition the majority of MAP-infected animals show no pathology and remain asymptomatic. Microarray and real-time RT-qPCR analyses were used to compare gene expression in ileum from sheep with the two disease forms and asymptomatic sheep, to further understand the molecular basis of the pathologies. Microarrays identified 36 genes with fold-change of >1.5 and $P \leq 0.05$ in at least one comparison; eight candidates were chosen for RT-qPCR validation. Sequence analysis of two candidates, CXCR4 and IGFBP6, identified three SNPs in each; five were found in all three forms of disease and showed no significant relationship to pathological type. The IGFBP6 G ^{3743} A SNP was not detected in asymptomatic sheep. The data show that the two forms of disease are associated with distinct molecular profiles highlighted by the differential expression of chemokine and chemokine receptor transcripts, the protein products of which might be implicated in the different cell infiltrates of the pathologies. The cells within the lesions also show evidence of abnormal activation; they express high levels of cytokine transcripts but have reduced expression levels of transcripts for T cell receptor associated molecules.

Keywords:

Mycobacterium avium subsp. *paratuberculosis*

MAP

Sheep

Chemokines

Genomics

Pathology

1. Introduction

Mycobacterium avium subspecies *paratuberculosis* (MAP) is a facultative intracellular bacterium that primarily infects macrophages and is the causative agent of paratuberculosis (Johne's disease), a common enteropathy of ruminants. In cattle an extended preclinical phase is dominated by a type 1 immune response that is followed by progression to fatal granulomatous disease, which is linked to a switch to a type 2 response and the production of non-protective IgG1 antibodies (Koets et al., 2002; Kurade et al., 2004). In sheep both type 1 and type 2 responses give rise to disease states (Clarke, 1997) and there seems to be no progression from the paucibacillary (type 1 or tuberculoid) to the multibacillary (type 2 or lepromatous) disease; both are fatal. As with homologous pathologies in tuberculosis and leprosy, the cytokines that dominate paucibacillary paratuberculosis are IFNγ and IL-12, while IL-10 dominates multibacillary disease (Smeed *et al.*, 2007). Only a minority of infected (IS900+) sheep progress to disease and the majority remain asymptomatic with no pathology (Koets et al., 2002; Smeed et al., 2007; de Silva et al., 2009).

We hypothesize that the polarization of the immune response is critical to the clinical outcome of paratuberculosis infection. Intestinal tissue damage that results from a type 1 response involves lymphocyte and eosinophil infiltration and is fundamentally different to that caused by a type 2 response, which leads to epithelioid macrophage infiltration and dissemination of infection (Smeed *et al.*, 2007). The related disease leprosy also has this dichotomy of immunology and pathology (Meisner *et al.*, 2001). Furthermore there is a clear relationship between host genetics and disease type (Fortin *et al.*, 2007); many genes have been implicated but one that has received much attention is solute carrier family 11 (SLC11A1, NRAMP-1 or Bcg), which is linked to both tuberculosis susceptibility (Vidal *et al.*, 1993) and discrimination of the two forms of human leprosy (Meisner *et al.*, 2001).

Resistance to MAP infection in mice is also partly linked to this gene, with animals carrying the Bcg^r allele (e.g. C3H) being refractory to infection (Veazey *et al.*, 1995).

The principal cell populations associated with mycobacterial infection are cells of the macrophage lineage, a major function of which is the activation of T cells and initiation of the immune response. Antigen-presenting cell subsets play a crucial role in the polarization of responses through the differential expression of IL-10 and IL-12 (Pulendran *et al.*, 1999). They also show differential expression of pattern recognition receptors (PRRs) (Nalubamba *et al.*, 2008), and differential PRR activation can tailor the immune response (Kapsenberg, 2003). Both TLR2 and CARD15 (NOD2) have been shown to be up-regulated in clinical paratuberculosis (Nalubamba *et al.*, 2008) and involved in both recognition (Ferwerda *et al.*, 2007) and response (Weiss *et al.*, 2001) to MAP. Indeed, there is growing evidence for linkage of TLR2 and/or CARD15 mutations and susceptibility to human mycobacterial diseases (Bochud et al., 2003; Ben-Ali et al., 2004) and bovine MAP infection (Mucha *et al.*, 2009).

To begin to unravel the complex interactions between MAP and the immune system that gives rise to bovine paratuberculosis, Coussens and colleagues have used bovine leukocyte microarrays to examine transcript changes in peripheral blood mononuclear cells (PBMC) (Coussens *et al.*, 2002) and macrophages (Murphy *et al.*, 2006). In addition they have examined cattle ileal tissue (Aho *et al.*, 2003) in order to understand the mechanisms that lead to intestinal tissue damage and have identified a number of genes involved in the immunology and pathogenesis of bovine paratuberculosis (including insulin-like growth factor binding protein-6 - IGFBP6). Our own studies (Smeed et al., 2007; Nalubamba et al., 2008) have focussed on ovine paratuberculosis because of the defined pathologies in sheep and exploit an oligonucleotide microarray of ~600 immuno-inflammatory genes (Watkins *et al.*, 2008) to gain insight into the role of these genes in the development of the different pathologies. The defined nature of the disease phenotype in sheep facilitates the identification of genetic markers associated with pathological type, and might also inform studies in cattle.

Competitive hybridization of cDNA isolated from the ileum (the site of paratuberculosis lesions) between the three infected groups was performed to identify changes to transcript expression associated with each disease group and develop molecular signatures of each pathological form. Quantitative real-time RT-PCR (RT-qPCR) on selected candidate genes was used to validate the arrays and the data were analysed to identify and compare physiological pathways associated with the contrasting pathologies. Data sets gathered from such experiments provide the potential to follow pathways of immune reactivity as well as assessing disease states that can then be related to specific gene expression signatures. In addition, as the epidemiology of these mycobacterial diseases suggests a genetic susceptibility (Ben-Ali *et al.,* 2004) we investigated this link, by a preliminary single nucleotide polymorphism (SNP) analysis, between polymorphisms in two of the candidates and pathological form.

2. Materials and Methods

2.1. Experimental animals and tissues

MAP infected animals presented with clinical Johne's disease (diagnosed by a veterinary surgeon: prolonged weight loss, inappetance and depression, and occasional diarrhoea) were out bred, female sheep (Table S1) of a variety of breeds and ages. All sheep were euthanized and diagnosis was confirmed by histopathology, Ziehl Neelsen (ZN) staining and IS900 real-time quantitative PCR with a sensitivity of 5 genome equivalents (Eishi *et al.,* 2002); sheep of similar ages from the same flocks (Nalubamba *et al.*, 2008), with no signs of clinical Johne's disease and positive for IS900 DNA were considered asymptomatic (Smeed *et al*., 2007). IS900 is a specific marker for the precise identification of MAP (Green et al., 1989). Animal procedures were performed under a valid Animals (Scientific Procedures) Act 1986 Project Licence.

2.2. RNA isolation

Terminal ileum sections (~0.5 g) were placed in five volumes of RNA*later* (Qiagen, Crawley, UK), incubated overnight at 4ºC and stored at -80ºC. Tissues were homogenized in guanidine isothiocyanate lysis buffer (RNeasy, Qiagen) and RNA was isolated using the RNeasy Maxi kit, eluted in 0.8 ml of RNase-free water, precipitated with ethanol and resuspended in 0.3 ml of RNase-free water. Samples were DNase I digested (Qiagen) and genomic DNA contamination assessed by GAPDH PCR and a no reverse transcriptase control. RNA quality was assessed using a RNA 6000 Nano LabChip on the Agilent 2100 bioanalyzer and quantified using a NanoDrop ND-1000 spectrophotometer; all samples had an RNA integrity number >7; n = 6 for the microarray experiments.

2.3. Preparation of labelled cDNA and microarray hybridization

cDNA generation and indirect labelling was performed using the Fairplay III Microarray Labelling Kit (Stratagene, Cambridge, UK) incorporating Cy3 and Cy5 dyes (GE Healthcare, Amersham, UK); the DyeEx spin 2.0 kit (Qiagen) removed unincorporated dye. Pre-hybridization and hybridization was carried out using the Pronto!^{TM} Kit (Corning, Loughborough, UK). This analysis utilized the ruminant immuno-inflammatory gene universal array (Watkins *et al.*, 2008); 4,824 spots from 596 genes. A 22x22 mm lifterslip (VWR, VWR, Poole, UK) was then placed over the array and the slide pre-heated to 42ºC. Equal dye concentrations from two samples (30 pmols each) were mixed, dried and resuspended in 20 µl cDNA hybridization solution and introduced onto the array. The slides were incubated for 16 - 20 h at 42ºC with agitation every 7 s, using a SlideBooster (Implen, Stansted Mountfitchet, UK) and post-hybridization washes of the microarray slides carried out using the Pronto!TM Kit, using an AdvaWash AW400 slide washer (Implen). The slides were scanned using a GenePix Autoloader 4200AL, controlled by GenePix Pro v 6.1 (Molecular Devices, Sunnyvale, CA). In total, we performed 18 microarrays. All groups contained six animals and every sample was analysed on three arrays, with each animal being

compared with an animal in each of the other groups. The identity of the pairings between each group was chosen at random.

Protocols of the experimental procedures, methods of analysis and microarray data are available as supplementary information in the European Bioinformatics Institute's ArrayExpress database (www.ebi.ac.uk/arrayexpress). Login to www.ebi.ac.uk/aerep/login and query for the accession number of the experiment or array design E-TABM-487.

2.4. Microarray data analysis

Scanned images were aligned to the GenePix Array List (GAL file); background noise subtracted; spot intensity amplitudes generated and quality data extracted using BlueFuse v3.3 (BlueGnome, Cambridge, UK). The $\log_2$ ratios from the two channels for each spot were calculated by Bluefuse. Data generated by Bluefuse were pre-processed using Limma within BioConductor 2.2 (http://www.bioconductor.org). Normalization was carried out using a weighted print-tip Loess method. Spots were weighted by a confidence value provided by BlueFuse. Differentially expressed genes from the different group comparisons (Table 1) were classified based on biological process using GO annotation. GORetriever, GOanna and GoSlimViewer (http://www.agbase.msstate.edu) were used to annotate gene products based on sequence similarity.

2.5. Real time RT-qPCR validation

RT-qPCR was performed using a Corbett Rotor-Gene™ 3000 (Qiagen). First strand cDNA synthesis used 0.5 μg of total RNA and Oligo(dT) and Superscript™ III RT (Invitrogen, Paisley, UK), from all RNA samples used for microarrays with a further two samples from each of the groups (n = 8). Primers for candidate genes were designed (Table S2) and the amplicons cloned using pGEM-T Easy kit (Promega, Southampton, UK). PCR conditions were optimized using a standard curve generated from linearized plasmid. All serial dilutions were generated using a CAS-1200™ Precision Liquid Handling System

(Qiagen). Succinate dehydrogenase complex, subunit A (SDHA) and tyrosine 3/tryptophan 5 -monooxygenase activation protein, zeta (YWHAZ); were chosen from a group of six using GeNorm v3.4 (Vandesompele *et al.*, 2002) and NormFinder v 0.953 (Andersen et al., 2004), as the two most stably expressed genes to normalize the RT-qPCR assays. All reactions were 10 μl volumes containing 4 μl of cDNA at the optimum dilution (1:20), 5 μl 2x FastStart SYBR Green Master mix (Roche, Lewes, UK) and 0.5 μl of each primer at optimum concentration (Table S2). PCR conditions were: 94ºC for 5 min and 45 cycles of 94ºC for 20 s, 60ºC for 20 s, 72ºC for 20 s, followed by dissociation curve analysis to confirm a single gene product. Each run assayed one gene and contained all samples in duplicate, a cDNA standard curve and a non-template control. This was repeated twice using cDNA from two additional RT reactions, resulting in 6 data points for each gene. The average Cq value for each sample was calculated and converted into relative quantities, taking account of the gene specific efficiency of the reaction and multiple reference gene normalization, using qBase (Hellemans *et al.*, 2007). Statistical analyses were performed on data from individual animals using Kolmogorov-Smirnov to test for normality of distribution and two-tailed Student's t-test for each of the comparisons.

2.6. PCR and DNA sequencing

DNA was extracted from 250 mg ileum using the Qiagen DNEasy Blood and Tissue kit or Qbiogene Fast DNA kit and quantified using the Nanodrop ND-1000 spectrometer. Full length mRNA sequence was obtained for sheep CXCR4 and IGFBP6, and aligned to the bovine genome to identify the coding regions in the genomic sequence. Primers (Table S3) were designed to amplify the whole bovine coding region and the intron / exon boundaries where possible (Fig. S1). Three PCR reactions were performed with each set of primers on ten animals in each infected group (Table S1). The PCR products from each sample were mixed, purified, ligated and cloned into the PGEM-T easy. Three independent sequences were produced for each PCR product and consensus sequences produced using VectorNTI AlignX

(Invitrogen). Differences between the genomic sequences were considered to be SNPs when the same change was found consistently in two or more of the consensus sequences. The distribution of alleles and genotypes was compared between the pathological groups using the χ^2-test.

3. Results

3.1. Microarray analysis

Of the 596 genes analysed, 36 fulfilled the criteria of fold-change of ≥1.5 and P≤0.05 in at least one comparison (Table 1 and supplementary information). The fold-changes shown are means of all probes for each gene. Intrinsic variation within replicate spots on single arrays (technical) and between arrays (biological) was determined. Technical variation was determined by comparing normalized $\log_2$ ratios of the spots printed in triplicate within each array, for six replicate array experiments of each comparison. The median standard deviation for the 3 comparisons varied from 0.24 (for multi- verses paucibacillary (M v P) to 0.26 (for multi- verses asymptomatic (M v A) with a variance of between 0.06 and 0.07 across the three comparisons. The biological variation was analysed by comparing the $\log_2$ ratios (CH1:CH2) of identical spots between the 6 replicate arrays in each comparison. The median standard deviation for each comparison ranged from 0.37 to 0.42, with a variance between 0.14 and 0.18 across all comparisons, demonstrating that within chip variability (technical variation) was between third and a half of the between chip variation (biological variation).

ITGAL was the only gene significantly affected in all three comparisons (Table 1, Fig. 1) and two other genes (SELL and TFRC) were up-regulated but P≥0.05. Five genes were differentially expressed between the two disease groups and the asymptomatic group (Table 1). Fifteen genes were significantly different in the paucibacillary vs asymptomatic comparison including a 1.52 fold repression of CCL5 and a 2.25 fold increase in CXCL10. Twenty three were significantly different in the multibacillary vs asymptomatic comparison including a 1.64 fold increase in CCL2; 1.54 and 1.82 fold increases in CXCR3 and CXCR4 and 1.59 and 1.67 decreases in CD3E and CD3F. The pauci- vs multibacillary comparison

identified fifteen affected genes, but only four (C9, CLIC5, ITGAM and MMP7) were unique to that comparison, as the other comparative data was not significant (<1.5 and p>0.05). Within the limitations of this focussed microarray, it is clear that the differentially expressed genes associated with each of the paratuberculosis pathologies were connected with a distinct range of biological processes (Fig. 2). Both clinical diseases showed changes to genes linked to responses to biotic stimuli, cell differentiation and signal transduction but only in the multibacillary form were there changes to genes associated with cell communication.

3.2. RT-qPCR validation of microarray results

Of the 36 genes found to be differentially expressed in the microarray experiment, seven candidates were chosen for RT-qPCR validation of the array; CD63, CXCR4, IGF2R, ITGB2, MMP9, TLR2, and TYROBP; on the basis that they represent different biological processes (Fig. 2) and were differentially expressed between pauci- and multibacillary samples. In addition we identified IGFBP6 as it had been highlighted as being differentially expressed in bovine paratuberculosis (Coussens *et al.*, 2004) and in a preliminary array analysis (Roupaka, 2004). Table 2 shows the results expressed as relative fold-changes in the same comparisons as performed in the microarray. The RT-qPCR analyses for seven original candidate genes validated the results from the microarray experiment with a correlation coefficient (r) 0.63 – 0.97, mean of 0.87. Comparison (two-tailed Students t test) of expression levels of all eight candidates between Blackface x Blue du Maine vs other breeds in the paucibacillary group and Blackface x Blue du Maine vs other breeds in the asymptomatic group showed that breed had no significant impact ($P > 0.8$) on the distribution of the data. Similar comparisons of the 4-5 year old vs the 2-3 year old Blackface sheep in the multibacillary group also showed that age also had no effect on the data ($P > 0.9$).

3.2. Single nucleotide polymorphism (SNP) analysis

Analysis of the consensus CXCR4 sequences revealed the presence of three synonymous SNPs; in exon 2 a C to T change at position 2080 and a G to A change at 2275

and in exon 3 a G to A change at 3361; the last was 3' of the stop codon. Analysis of IGFBP6 sequence also revealed three SNPs – a non-synonymous C to A change at position 12 in exon 1, which changes a histidine to a glutamine; a synonymous G to A change at 2730 in intron 3 and a G to A change at 3743 in exon 4, 3' of the stop codon at position 3612 (Fig. S1). When these six SNPs were analysed in relation to infection there was no significant relationship between any SNP and disease form (Table 3). Each of the three CXCR4 SNPs and two IGFBP6 SNPs were found in all infected groups. However, IGFBP6 SNP3 (G 3822 A) was only found in clinically diseased animals, and was absent from the asymptomatic group. The animals tested were of a variety of breeds and crosses (Table S1) but there was no discernable, significant relationship between breed and any genotype (χ^2, $P \geq 0.79$).

4. Discussion

This is the first microarray analysis of ileal gene expression in sheep with clinical paratuberculosis and directly compares gene expression at the site of infection in each of the pathologies, and has identified novel pathways involved in MAP pathogenesis. Previous work on sheep paratuberculosis has shown that the pauci- and multibacillary pathologies are associated type 1 and type 2 T cell responses respectively, with the differential activation of T cells (Burrells *et al.*, 1998) and the expression of quite distinct panels of both innate receptors (Nalubamba *et al.*, 2008) and cytokines (Smeed *et al.*, 2007). Analysis of the data for the eight candidate genes indicates that it is highly unlikely that these changes are associated with the breed or age (Table SI). Paratuberculosis is characterized by extensive inflammation and cellular infiltration of the terminal ileum; paucibacillary disease is associated with lymphocytes and eosinophils, and multibacillary lesions with epithelioid macrophages (Smeed *et al.*, 2007). These variations in the cellular composition are reflected in the differences in gene expression profiles detected by microarray, with 36 genes found to be differentially expressed (>1.5 fold and $P \leq 0.05$) in at least one comparison (Table 1). The RT-

qPCR data justify our decision to use 1.5 fold (and $p \leq 0.05$) as the cut-off point for our candidate gene selection as they confirmed the array data in relation to the direction of fold change. In addition, fold changes of just under 1.5 (but $p \leq 0.05$) were shown to be significantly different by RT-qPCR, e.g. MMP9. RT-qPCR also confirmed the array data, both quantitatively and qualitatively, when microarray showed no change (>–1.5 to <1.5), e.g. CD63 in the pauci- vs asymptomatic comparison. The correlation coefficient (r) between array and RT-qPCR varied between candidate genes (0.63 – 0.97, mean r = 0.87) and was similar to that described in other systems (Dallas et al., 2005), but was much lower for genes with a fold change of <1.5. The magnitude of fold change with RT-qPCR was generally much higher than with microarray, which is probably due to the different sensitivities of the two assays (Morey et al., 2006); nevertheless RT-qPCR is the preferred validation method for microarrays (Dallas et al., 2005).

Within the limitations of the focussed microarray, it is clear that the differentially expressed genes associated with each of the pathologies were connected with a distinct range of biological processes (Fig. 2) probably associated with the different cell infiltrates in the two disease forms. The multibacillary samples displayed a consistent repression of the several genes encoding elements of the T cell receptor complex - CD3E, CD3G (and CD3Z with –1.47 fold change and $P \leq 0.05$). This might be related to the fact that the principal cells in multibacillary ileum were infected epithelioid macrophages and not T cells. It is possible that this reduction in antigen-receptor expression is a result of infection-associated aberrant activation and may represent a specific bacterial mechanism for the modification of the host immune response towards the development of disease as is seen in tuberculosis (Smith, 2003). Related to this were changes to transcripts of the different immunoglobulin molecules. IGHA1 and IGHM were raised in both disease forms and IGJ was increased in paucibacillary lesions. In view of the fact that B cell receptor associated genes were not raised in infected tissues, these increases in immunoglobulin heavy chain transcripts may be a result of greater numbers of plasma cells rather than mature B cells. This pattern of immunoglobulin heavy

chain expression does not fit exactly into the Th1/Th2 paradigm, which would predict high levels of IGHA only in the multibacillary (lepromatous) lesions (Mosmann and Coffman, 1989).

Representatives of the β1 and β2 integrins (ITGA4 and ITGAL) were up-regulated in the paucibacillary forms. However, their behaviour was less consistent in the multibacillary form where the β1 integrin ITGA4 is repressed and the β2 integrins ITGAL, ITGAM and ITGB2 were increased. The four transmembrane domain proteins (e.g. CD63) interact with integrins and modulate their function; and the high levels of CD63 in multi- relative to paucibacillary lesions may be indicative of different cell populations associated with those lesions. The major cellular ligand for the β2 integrins is ICAM1. This, and the lymphocyte adhesion molecule, SELL (CD62L or L selectin), were high in multibacillary diseased tissues, which reflects major cellular infiltration and also inflammation, as both are up-regulated by inflammatory cytokines (Dustin and Springer, 1991).

The stimulus for leukocyte infiltration into lesions is probably different for each disease form. CCL2 was raised in multibacillary lesions (and 1.65 fold but $P \geq 0.05$ in the pauci- vs asymptomatic comparison) and was probably produced by a variety of cells on stimulation with pro-inflammatory cytokines and IFNγ. It is chemotactic for lymphocytes and monocytes but not eosinophils and its role in inflammation is as an activator of neutrophils (Campbell et al., 2003). CCL5 is associated with T cell and dendritic cell recruitment, but it was repressed in the paucibacillary group. CCL5 is a product of activated T cells, but only at the late stage of activation, and is down-regulated by TGFβ and IFNγ both of which were increased in paucibacillary tissues (Smeed *et al.*, 2007). This is additional evidence for the aberrant activation of T cells in the paratuberculosis lesions. The chemokine CXCL10 is associated with chemoattraction of type 1 T cells (Campbell *et al.*, 2003) and was significantly raised in paucibacillary lesions and might explain the T cell infiltration. However its receptor, CXCR3 was up-regulated in multibacillary tissues, probably due to the effect of high levels of TGFβ. The closely related chemokine receptor, CXCR4 is also

responsive to TGFβ and was significantly increased in multibacillary tissues. This is the receptor for CXCL12 (SDF-1) and one of its major functions is as a chemoattractant for Th2 and regulatory T cells (Campbell et al., 2003).

IGF2R, the receptor for IGF2 and mannose-6-phosphate, was increased in the multibacillary disease form and plays a role in the activation of TGFβ (Melnick *et al.*, 1998), a gene previously shown to be up-regulated in these samples (Smeed *et al.*, 2007). IGFBP6 was differentially expressed in sheep paratuberculosis (increased in paucibacillary and repressed in multibacillary diseased ileum) and it was also up-regulated in cattle PBMC stimulated by MAP (Coussens *et al.*, 2004). It binds with high affinity to IGF2 and inhibits many of its functions, including proliferation and differentiation (Kelley *et al.*, 1996). Its expression has previously been reported to be stimulated by IL-1β and TNFα but inhibited by the presence of TGFβ (Martin et al., 1994; Liu et al., 1999), and this might explain the significantly increased levels of IGFBP6 in paucibacillary lesions. In paucibacillary tissues, high IGFBP6 levels could play a role in controlling cellular proliferation by inhibiting IGF2 and inducing apoptosis. TLR2 was greatly up-regulated in multibacillary lesions, suggesting that the TLR2 signalling pathway is important in lepromatous pathogenesis. TLR2 is central to the immune response to mycobacteria, as knockout mice are more susceptible to infections. In human leprosy TLR2 is linked to differential pathology, as signalling through the mutated receptor leads to increased IL-10 production and a type 2 immune response (Kang *et al.*, 2004). Consistent with this are increased levels of IL-10 in multibacillary tissues and raised IFNγ in paucibacillary lesions (Smeed et al., 2007). Multibacillary tissues show high levels of both CXCR4 and TLR2 and, interestingly, cross-talk between these has been reported in *Porphyromonas gingivalis* infection that leads to the undermining of host defence (Hajishengallis *et al.*, 2008). The data on the linkage between TLR2 and MAP in ruminants is not clear. Analysis of TLR2 SNPs in sheep paratuberculosis failed to identify any linkage with pathological form (Nalubamba *et al.*, 2008) although in cattle there was significant association of a distinct SNP with MAP infection (Mucha *et al.*, 2009).

Due to their differential expression in multibacillary and paucibacillary sheep, CXCR4 and IGFBP6 were selected for preliminary SNP analysis and to assess whether these SNPs were associated with gene expression and pathology. All SNPs except IGFBP6 at 3743 were found in the three infected groups and there was no significant association with any disease form. The IGFBP6 SNP at 3743 is 3' to the stop codon and was not found in asymptomatic sheep; there was no statistically significant association with disease form possibly because of the small numbers of animals (n=10) in each group.

To summarize, this study identifies novel genes and pathways linked to mycobacterial pathologies. The two disease forms of sheep paratuberculosis are associated with distinct molecular profiles highlighted by the differential expression of chemokines, which are implicated in the different cellular infiltrates of the pathologies and the polarization of the immune response. Cells within the lesions show evidence of abnormal activation, with high levels of cytokine transcripts and a reduction in expression levels of transcripts for lymphocyte receptor molecules. Preliminary analysis could identify no significant link between CXCR4 and IGFBP6 SNPs and paratuberculosis pathologies.

Acknowledgements

We acknowledge Jill Sales (BioSS) for help with experimental and statistical design and the clinical staff at the Moredun Research Institute. This project was funded by the BBSRC and SEERAD (Project Grants 15/S19602 and 15/S13964). JAS was funded by BBSRC/Genesis-Faraday CASE studentship and sponsored by Moredun Scientific Ltd. Midlothian, UK.

References

Aho, A.D., McNulty, A.M., Coussens, P.M., 2003. Enhanced expression of interleukin-1α and tumor necrosis factor receptor-associated protein 1 in ileal tissues of cattle infected with Mycobacterium avium subsp. paratuberculosis. Infect. Immun. 71, 6486.

Andersen, C.L., Jensen, J.L., Orntoft, T.F., 2004. Normalization of real-rime quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64, 5245-5250.

Ben-Ali, M., Barbouche, M.R., Bousnina, S., Chabbou, A., Dellagi, K., 2004. Toll-like receptor 2 Arg677Trp polymorphism is associated with susceptibility to tuberculosis in Tunisian patients. Clin Diagn Lab Immunol 11, 625-626.

Bochud, P.Y., Hawn, T.R., Aderem, A., 2003. Cutting edge: a Toll-like receptor 2 polymorphism that is associated with lepromatous leprosy is unable to mediate mycobacterial signaling. J Immunol 170, 3451-3454.

Burrells, C., Clarke, C.J., Colston, A., Kay, J.A., Porter, J., Little, D., Sharp, J.M., 1998. 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.

Campbell, D.J., Kim, C.H., Butcher, E.C., 2003. Chemokines in the systemic organization of immunity. Immunological Reviews 195, 58-71.

Clarke, C.J., 1997. The pathology and pathogenesis of paratuberculosis in ruminants and other species. J Comp Pathol 116, 217-261.

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, 5494-5502.

Coussens, P.M., Jeffers, A., Colvin, C.J., 2004. Rapid and transient activation of gene expression in peripheral blood mononuclear cells from Johne's disease positive cows exposed to Mycobacterium paratuberculosis in vitro. Microbial Pathogenesis 36, 93-108.

Dallas, P.B., Gottardo, N.G., Firth, M.J., Beesley, A.H., Hoffmann, K., Terry, P.A., Freitas, J.R., Boag, J.M., Cummings, A.J., Kees, U.R., 2005. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? BMC Genomics 6, 59.

de Silva, K., Begg, D., Carter, N., Taylor, D., Di Fiore, L., Whittington, R., 2009. The early lymphocyte proliferation response in sheep exposed to Mycobacterium avium subsp. paratuberculosis compared to infection status. Immunobiology.

Dustin, M.L., Springer, T.A., 1991. Role of lymphocyte adhesion receptors in transient interactions and cell locomotion. Annu Rev Immunol 9, 27-66.

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.

Ferwerda, G., Kullberg, B.J., de Jong, D.J., Girardin, S.E., Langenberg, D.M.L., van Crevel, R., Ottenhoff, T.H.M., Van der Meer, J.W.M., Netea, M.G., 2007. Mycobacterium paratuberculosis is recognized by Toll-like receptors and NOD2. J Leukoc Biol 82, 1011-1018.

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. Annual Review of Genomics and Human Genetics 8, 163-192.

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, 9063-9073.

Hajishengallis, G., Wang, M., Liang, S., Triantafilou, M., Triantafilou, K., 2008. Pathogen induction of CXCR4/TLR2 cross-talk impairs host defense function. Proc Natl Acad Sci U S A 105, 13532-13537.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., Vandesompele, J., 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol 8, R19.

Kang, T.J., Yeum, C.E., Kim, B.C., You, E.Y., Chae, G.T., 2004. Differential production of interleukin-10 and interleukin-12 in mononuclear cells from leprosy patients with a Toll-like receptor 2 mutation. Immunology 112, 674-680.

Kapsenberg, M.L., 2003. Dendritic cell control of pathogen-driven T-cell polarization. Nat Rev Immunol 3, 984-993.

Kelley, K.M., Oh, Y., Gargosky, S.E., Gucev, Z., Matsumoto, T., Hwa, V., Ng, L., Simpson, D.M., Rosenfeld, R.G., 1996. Insulin-like growth factor-binding proteins (IGFBPs) and their regulatory dynamics. The International Journal of Biochemistry & Cell Biology 28, 619-637.

Koets, A., Rutten, V., Hoek, A., van Mil, F., Müller, K., Bakker, D., Gruys, E., van Eden, W., 2002. Progressive bovine paratuberculosis is associated with local loss of CD4+ T

cells, increased frequency of γδ T cells, and related changes in T-cell function. Infect. Immun. 70, 3856-3864.

Kurade, N.P., Tripathis, B.N., Rajukumar, K., Parihar, N.S., 2004. Sequential development of histologic lesions and their relationship with bacterial isolation, faecal shedding and immune responses during progressive stages of experimental infection of lambs with Mycobacterium avium subsp. paratuberculosis. Vet Pathol 41, 378-387.

Liu, Y., Tsushima, T., Miyakawa, M., Isozaki, O., Yamada, H., Xu, Z.R., Iwamoto, Y., 1999. Effect of cytokines on production of insulin-like growth factor binding proteins (IGFBPs) from human fibroblasts in culture. Endocr J 46 Suppl, S63-S66.

Martin, J.L., Coverley, J.A., Baxter, R.C., 1994. Regulation of immunoreactive insulin-like growth factor binding protein-6 in normal and transformed human fibroblasts. J Biol Chem 269, 11470-11477.

Meisner, S.J., Mucklow, S., Warner, G., Sow, S.O., Lienhardt, C., Hill, A.V., 2001. Association of NRAMP1 polymorphism with leprosy type but not susceptibility to leprosy per se in west Africans. Am J Trop Med Hyg 65, 733-735.

Melnick, M., Chen, H., Buckley, S., Warburton, D., Jaskoll, T., 1998. Insulin-like growth factor II receptor, transforming growth factor-beta, and Cdk4 expression and the developmental epigenetics of mouse palate morphogenesis and dysmorphogenesis. Dev Dyn 211, 11-25.

Morey, J.S., Ryan, J.C., Van Dolah, F.M., 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biol Proced Online 8, 175-193.

Mosmann, T.R., Coffman, R.L., 1989. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annu Rev Immunol 7, 145-173.

Mucha, R., Bhide, M.R., Chakurkar, E.B., Novak, M., Mikula, S., 2009. Toll-like receptors TLR1, TLR2 and TLR4 gene mutations and natural resistance to Mycobacterium avium subsp. paratuberculosis infection in cattle. Veterinary Immunology and Immunopathology 128, 381-388.

Murphy, J.T., Sommer, S., Kabara, E.A., Verman, N., Kuelbs, M.A., Saama, P., Halgren, R., Coussens, P.M., 2006. Gene Expression Profiling of Monocyte-Derived Macrophages Following Infection with Mycobacterium avium subspecies avium and Mycobacterium avium subspecies paratuberculosis. Physiol Genomics 28, 67-75.

Nalubamba, K.S., Smeed, J.A., Gossner, A.G., Watkins, C.A., Dalziel, R.G., Hopkins, J., 2008. Differential expression of pattern recognition receptors in the three pathological forms of sheep paratuberculosis. Microbes and Infection 10, 598-604.

Pulendran, B., Smith, J.L., Caspary, G., Brasel, K., Pettit, D., Maraskovsky, E., Maliszewski, C.R., 1999. Distinct dendritic cell subsets differentially regulate the class of immune response in vivo. Proc Natl Acad Sci U S A 96, 1036-1041.

Roupaka, S., 2004. Functional genomics of Johne's disease. MSc Thesis University of Edinburgh.

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.

Smith, I., 2003. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. Clin Microbiol Rev 16, 463-496.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric

averaging of multiple internal control genes. Genome Biology 3, research0034.1-research0034.11.

Veazey, R.S., Horohov, D.W., Krahenbuhl, J.L., Taylor, H.W., Oliver, J.L., Snider, T.G., 1995. Comparison of the resistance of C57BL/6 and C3H/He mice to infection with Mycobacterium paratuberculosis. Vet Microbiol 47, 79-87.

Vidal, S.M., Malo, D., Vogan, K., Skamene, E., Gros, P., 1993. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. Cell 73, 469-485.

Watkins, C.A., McKellar, A., Jensen, K., George, A., Jones, D., Sharp, J.M., Stevenson, K., Hopkins, J., 2008. Development and validation of an oligonucleotide microarray for immuno-inflammatory genes of ruminants. Veterinary Research Communications 32, 647-657.

Weiss, D.J., Evanson, O.A., McClenahan, D.J., Abrahamsen, M.S., Walcheck, B.K., 2001. Regulation of expression of major histocompatibility antigens by bovine macrophages infected with *Mycobacterium avium* subsp. *paratuberculosis* or *Mycobacterium avium* subsp. *avium*. Infect. Immun. 69, 1002-1008.

Figure Legends

Fig. 1. Venn diagram showing the relationships between genes differentially expressed in the ileum of sheep infected with MAP; comparisons between the three infected groups; A, asymptomatic; M, multibacillary; P, paucibacillary. The figures are the numbers of differentially expressed genes in the different comparisons with a fold change of ≥1.5 and $P \leq 0.05$. Figures in parentheses are the number of genes with fold change of ≥1.5 but P>0.05.

Fig. 2. Distribution of GO annotations of biological processes terms of genes that were differentially expressed (fold change of ≥1.5 and $P \leq 0.05$) in ileum tissue of MAP infected sheep. (a) Comparison of paucibacillary vs asymptomatic. (b) Comparison of multibacillary vs asymptomatic. (c) Comparison of multibacillary vs paucibacillary.

Fig. S1. Gene structures of bovine CXCR4 and IGFBP6 illustrating the positions of the cloning fragments and SNPs.

Table 1
Mean fold–changes of significantly differentially expressed genes.

NCBI Official Symbol (Gene)	Accession Number	P v A	M v A	M v P
ANPEP (CD13)	X98240	–2.32	–2.50	
C3 (complement C3)	AF038130	1.53	1.75	
C7 (complement C7)	AF162274		1.5	
C9 (complement C9)	F14494			–1.61
CFB (complement Factor B)	NM_001040526	1.52		
CCL2 (MCP-1)	M84602	1.65*	1.64	
CCL5 (RANTES)	AJ007043	–1.52		
CD3E	Z12969		–1.59	
CD3G	X52994		–1.67	
CD63 (LAMP3)	**AJ012589**		**1.87**	**1.62**
CLIC5	AF109199			–1.54
CXCL10 (IP-10)	AB070717	2.25		–1.72*
CXCR3 (CD183)	X95876		1.54	
CXCR4 (CD185)	**AF399642**		**1.82**	**1.9**
ICAM1 (CD54)	AF110984	*1.45*	1.9	*1.31*
IGF2R (CD222)	**AF327649**		**1.74**	**1.51**
IGFBP3	AF327651	–1.61	–1.56	
IGHA1	AF024645	1.96	1.85	
IGHM	X59994	2.31	1.83	
IGJ (J chain)	L20311	1.89		–1.61
IL-18RAP (CD218b)	NM_010553	1.68		
ITGA4 (CD49D)	AB055841	1.52		–1.59
ITGAL (CD11A/LFA-1)	BF654844	1.9	3.58	1.95
ITGAM (CD11B/MAC-1)	AW353414			1.51
ITGB2 (CD18)	M81233		2.03	1.95
MHC-DYA	AJ251357	1.67		
MMP7 (matrilysin)	AB031323			1.84
MMP9 (gelatinase B)	**X78324**		**1.98**	**1.76**
pIGR (secretory component)	L04797	1.56		
SELL (CD62L)	X62882	1.58*	1.85	1.83*
SIRPA (CD172A)	Y11045	*1.41*	1.65	
SLC11A1 (NRAMP1)	AF005380		1.57	1.63*
TFRC (CD71)	M11507	1.78	2.79	1.81*
TLR2 (CD282)	**AF310951**		**1.52**	**1.52**
TNFRSF1B (CD120b)	NM_001040490		1.62	1.5
TYROBP (DAP12)	**AJ419226**		**1.95**	**1.68**

A – asymptomatic; M – multibacillary; P – paucibacillary.

Bold denotes genes selected for validation by qRT–PCR. Changes less than 1.5 fold (and $P \leq 0.05$) are in italics to show trends in the data. *Changes not statistically significant, but >1.5 fold. The blank cells are fold change <1.5 and P>0.05.

Table 2
Relative fold-change of genes analysed by qRT-PCR.

	P v A	M v A	M v P
CD63	1.1 (0.7)[a]	**3.2 (0.00)**	**3.0 (0.00)**
CXCR4	**2.89 (0.05)**	**6.13 (0.01)**	**2.12 (0.05)**
IGFBP6	**1.86 (0.05)**	-1.37 (0.5)	**-2.54 (0.02)**
IGF2R	**2.0 (0.02)**	**4.2 (0.00)**	**2.1 (0.00)**
ITGB2	**2.1 (0.02)**	**4.3 (0.00)**	**2.1 (0.00)**
MMP9	**1.6 (0.02)**	**14.9 (0.00)**	**9.3 (0.00)**
TLR2	1.6 (0.4)	**6.8 (0.00)**	**4.3 (0.00)**
TYROBP	**3.1 (0.01)**	**9.4 (0.00)**	**3.0 (0.00)**

A – asymptomatic; M – multibacillary; P – paucibacillary.

[a] mean fold change (P value). Bold, $P \leq 0.05$.

Table 3
Distribution of SNPs

	Frequency (10 in each group)			
Genotype	Multi	Pauci	Asympto	χ^2 P value
CXCR4 C ^{2080} T	2	4	4	0.33
CXCR4 G ^{2275} A	2	2	4	0.51
CXCR4 G ^{3361} A	4	2	1	0.27
IGFBP-6 C ^{12} A	1	1	1	1.00
IGFBP-6 G ^{2730} A	4	5	5	0.88
IGFBP-6 G ^{3743} A	2	1	0	0.33

Fig 1

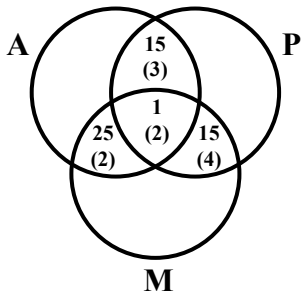


Fig 2

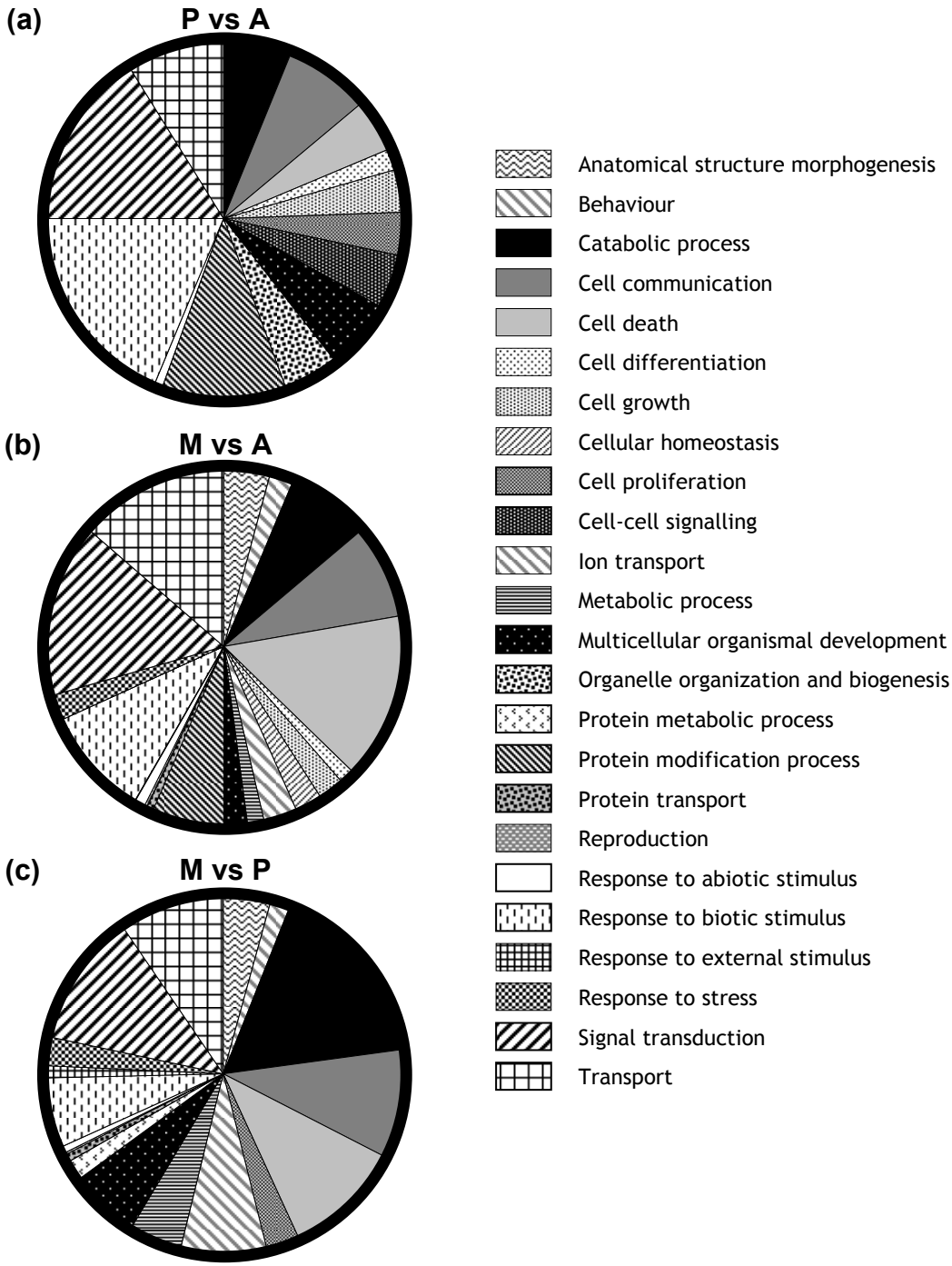
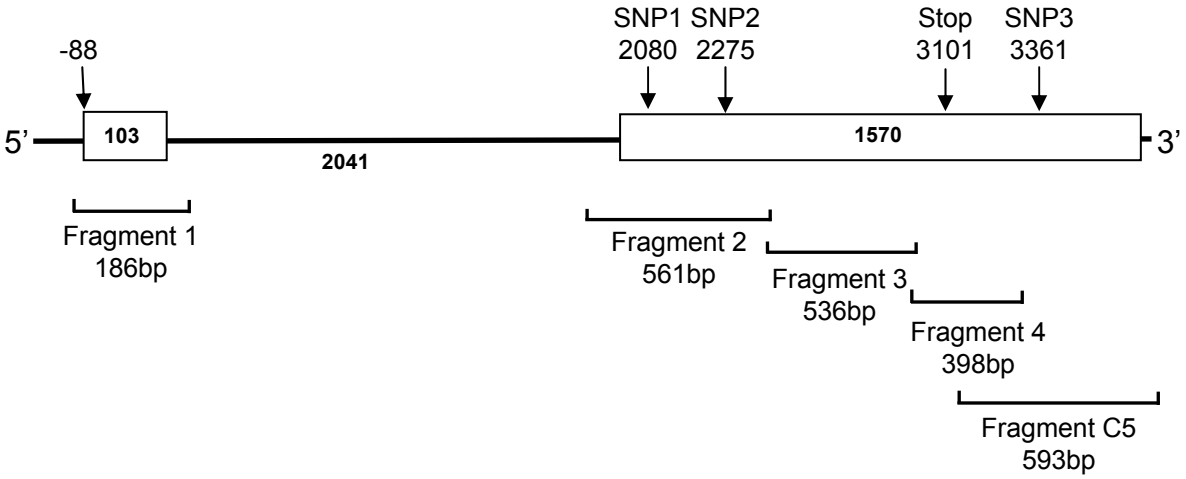

(a)
P vs A
(b)
M vs A
(c)
M vs P
Anatomical structure morphogenesis
Behaviour
Catabolic process
Cell communication
Cell death
Cell differentiation
Cell growth
Cellular homeostasis
Cell proliferation
Cell-cell signalling
Ion transport
Metabolic process
Multicellular organismal development
Organelle organization and biogenesis
Protein metabolic process
Protein modification process
Protein transport
Reproduction
Response to abiotic stimulus
Response to biotic stimulus
Response to external stimulus
Response to stress
Signal transduction
Transport

Fig S1

CXCR4



IGFBP- 6

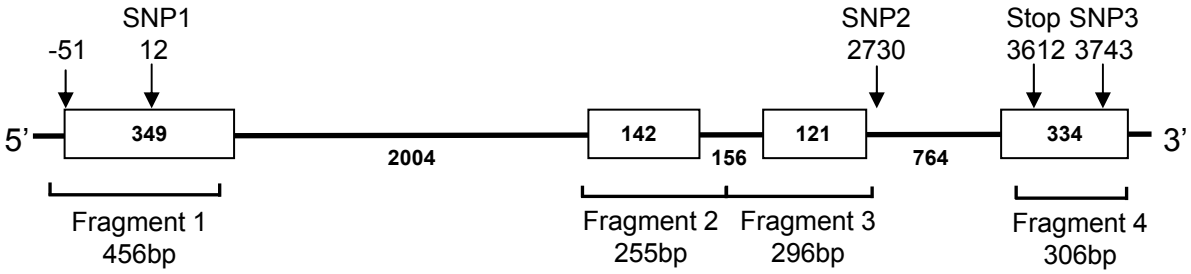


Table S1
Pathology, breed and age of sheep

Pathology	Breed	Age (y)
Pauci[a]	Blackface × Bleu du Maine	4.5
Pauci[a]	Texel	5.5
Pauci[a]	Blackface × Bleu du Maine × Lleyn × Roussin	2.5
Pauci[a]	Lleyn × Roussin	3.5
Pauci[a]	Blackface × Bleu du Maine	6
Pauci[a]	Texel	2.5
Pauci[b]	Blackface × Bleu du Maine	4.5
Pauci[b]	Blackface	2
Pauci	Blackface × Bleu du Maine	4
Pauci	Blackface	3.5
Multi[a]	Blackface	3
Multi[a]	Blackface	4
Multi[a]	Blackface	3
Multi[a]	Blackface	3
Multi[a]	Blackface	5
Multi[a]	Blackface	5
Multi[b]	Blackface	2
Multi[b]	Blackface × Bleu du Maine	4.5
Multi	Blackface × Bleu du Maine	4
Multi	Blackface	4
Asympto[a]	Blackface × Bleu du Maine	7
Asympto[a]	Greyface	4
Asympto[a]	Texel	1
Asympto[a]	Blackface	2.5
Asympto[a]	Greyface	4
Asympto[a]	Greyface	4
Asympto[b]	Blackface × Bleu du Maine	7
Asympto[b]	Blackface × Bleu du Maine	4.5
Asympto	Blackface × Bleu du Maine	3.5
Asympto	Blackface	3

[a] animals used in the microarray experiments

[b] additional animals used in qRT-PCR experiments

Table S2
Primer sequences for qRT-PCR

Gene Accession no	Primers	Annealing Temp °C	Product size (bp)	Primer conc (nM)
CD63 BC151412.1	F - GGGCTGTGTGGAGAAGATTG R- GATGAGGGGGCTGAAGAGAC	61	178	F: 700 R: 1000
CXCR4 AF399642	F - ACCTCCTGTTTGTCCTACG R - AATGTCCACCTCGCTTGC	62	163	F: 500 R: 500
IGFBP-6 NM_001134308	F - AAGGAGAGTAAGCCCCAAGC R - CGGGAAGGAGTGGTAGAGGT	60	95	F: 500 R: 500
ITGB2 NM_001009485	F - CTCACCGACAACTCCAAACA R - AAAGTGGAACCCATCGTCTG	57	180	F: 400 R: 400
IGF2R AF353513	F - GACGACCTGAAGACCCTGAA R - GCAAATGAAGCGGATGATG	60	150	F: 500 R: 500
MMP9 X78324	F - GAGGGTAAGGTGCTGCTGTTC R - AAGATGTCGTGCGTGCTAATG	62	133	F: 400 R: 500
TLR2 AM183218	F - GCACTTCAACCCTCCCTTTTA R - TCTCCGAAAGCACAAAGATG	57	125	F: 500 R: 500
TYROBP AJ419228.1	F - GACCTGATGCTGACCCTCC R - CTGTCTCCGTGATGTGCTGT	56	112	F: 400 R: 500
SDHA NM_174178	F - ACCTGATGCTTTGTGCTCTGC R - CCTGGATGGGCTTGGAGTAA	62	126	F: 500 R: 500
YWHAZ BC102382	F - TGTAGGAGCCCGTAGGTCATC R - TCTCTCTGTATTCTCGAGCCAT	58	101	F: 500 R: 500

Table S3. Primer sets used to sequence CXCR4 and IGFBP-6

Gene Accession Number	Primer Sequence	Size (bp)	Start position[a]
CXCR4	F1 - GAAACTTCAGTGTGTTGGCT	186	-96
NM_174301	R1 - CCACTCAGAGAGGCGGTTG		
	F2 - ATTCCCTTGCCTGTTTTTCA	561	2035
	R2 - GTACCTCTCATCCACCTCCTTG		
	F3 - GTCTGGCTACCTGCTGTCCT	536	2521
	R3 - GAATGTCCACCTCGCTTGC		
	F4 - CCCATCCTCTATGCCTTCCT	398	2938
	R4 - CGGTCCTACAATGACACACAGC		
	F5 - GAATCTTCAAGTTTTCACTCCAG	593	3076
	R5 - GCCAAAGGAATGCCAATAG		
IGFBP-6	F1 - CTGGGAAGGGAGGAGGTAGAG	456	-227
NM_001040495	R1 - CGCAGTTGGGAGTGTAGACC		
	F2 - CTTGCGGCTGTAGGAGAGAATC	255	2291
	R2 - CAGGACTGGGATGGGAGACTT		
	F3 - AAGTCTCCCATCCCAGTCCT	296	2525
	R3 - GAAGGGCAGCGTCTCATTTA		
	F4 - GAGGTCCTTGCTGGTGTG	306	3514
	R4 - GAACACAGAGACATAAACCAGAG		

[a] numbered relative to start codon.