Comprehensive transcriptional profiling of the gastrointestinal tract of ruminants from birth to adulthood reveals strong developmental stage specific gene expression

S. J. Bush[*1], M. E. B. McCulloch[*], C. Muriuki[*], M. Salavati[*], G. M. Davis[*2], I. L. Farquhar[*3], Z. M. Lisowski[*], A.L. Archibald[*], D. A. Hume[*4] and E. L. Clark[*]

[*]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, UK

Current Address:

[1]Nuffield Department of Clinical Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

[2]Faculty of Life Sciences, The University of Manchester, Manchester, UK

[3]Centre for Synthetic and Systems Biology, University of Edinburgh, Edinburgh, Scotland, UK

[4]Mater Research-University of Queensland, 37 Kent St, Woolloongabba, Qld 4104, Australia

Reference Numbers for Data in the Public Repositories

The raw RNA-Sequencing data are deposited in the European Nucleotide Archive (ENA) under study accessions PRJEB19199 (sheep) and PRJEB23196 (goat). Metadata for all samples is deposited in the EBI BioSamples database under group identifiers SAMEG317052 (sheep) and SAMEG330351 (goat).

Running Title: RNASeq analysis of the ruminant GI tract

Key Words: Gene expression, transcription, sheep, goat, ruminant, gastrointestinal tract, development, macrophage, immunity, RNA-Seq, FAANG.

Corresponding Author: Emily L. Clark, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, UK, emily.clark@roslin.ed.ac.uk

Abstract

One of the most significant physiological challenges to neonatal and juvenile ruminants is the development and establishment of the rumen. Using a subset of RNA-Seq data from our high-resolution atlas of gene expression in sheep (*Ovis aries*) we have provided the first comprehensive characterisation of transcription of the entire the gastrointestinal (GI) tract during the transition from pre-ruminant to ruminant. The dataset comprises 168 tissue samples from sheep at four different time points (birth, one week, 8 weeks and adult). Using network cluster analysis we illustrate how the complexity of the GI tract is reflected in tissue- and developmental stage-specific differences in gene expression. The most significant transcriptional differences between neonatal and adult sheep were observed in the rumen complex. Differences in transcription between neonatal and adult sheep were particularly evident in macrophage specific signatures indicating they might be driving the observed developmental stage-specific differences. Comparative analysis of gene expression in three GI tract tissues from age-matched sheep and goats revealed species-specific differences in genes involved in immunity and metabolism. This study improves our understanding of the transcriptomic mechanisms involved in the transition from pre-ruminant to ruminant. It highlights key genes involved in immunity, microbe recognition, metabolism and cellular differentiation in the GI tract. The results form a basis for future studies linking gene expression with microbial colonisation of the developing GI tract and will contribute towards identifying genes that underlie immunity in early development, which could be utilised to improve ruminant efficiency and productivity.

Introduction

Sheep are an important source of meat, milk and fibre for the global livestock sector and belong to one of the most successful species of herbivorous mammals, the ruminants. Adult sheep have four specialized four specialized chambers comprising their stomach: fermentative fore-stomachs encompassing the rumen, reticulum and omasum and the "true stomach", the abomasum (Dyce *et al.* 2010). The events surrounding the development of the rumen are among the most significant physiological challenges to young ruminants (Baldwin *et al.* 2004). As lambs transition from a milk diet to grass and dry pellet feed the gastrointestinal (GI) tract undergoes several major developmental changes. In neonatal lambs, feeding solely on milk, the fermentative fore-stomachs are not functional and the immature metabolic and digestive systems function similarly to that of a young monogastric mammal, with proteolytic digestion taking place inside the abomasum (Meale *et al.* 2017). At this stage the rumen has a smooth, stratified squamous epithelium with no prominent papillae (Baldwin *et al.* 2004). Suckling causes a reflex action that brings the walls of the reticulum together to form an 'oesophageal' or 'reticular' groove transferring milk and colostrum directly to the abomasum, where it is digested efficiently (Figure 1) (Dyce *et al.* 2010). In neonatal ruminants this is essential to ensure protective antibodies in the colostrum are transported intact to the abomasum.

The introduction of grass and dry feed into the diet (which usually occurs in very small amounts from one week of age) inoculates the rumen with microbes. These proliferate, facilitating the digestion of complex carbohydrates which the adult ruminant relies upon to meet its metabolic needs, and stimulating growth and development of the rumen and reticulum (Bryant *et al.* 1958). The transition from pre-ruminant to ruminant occurs gradually

from around 4 weeks of age. The rumen and reticulum are usually fully functional by the time the lamb reaches 8 weeks of age and has a completely grass and dry feed-based diet (Figure 1). The transition results in metabolic changes, as tissues shift from reliance on glucose supplied from milk to the metabolism of short-chain fatty acids as primary energy substrates. Whilst the most dramatic physical changes occurring during development are associated with the rumen epithelium, changes in intestinal mass, immunity and metabolism also occur in response to dietary changes (Baldwin *et al.* 2004).

Early development of the rumen complex and GI tract of sheep is of particular interest due to the transition from pre-ruminant to ruminant, which occurs over the period 8 weeks after birth. This period is crucial in the development of innate and acquired immunity, healthy rumen growth and establishment of the microbiome. These processes are likely to be intrinsically linked, as the GI tract protects the host from toxic or pathogenic luminal contents, while at the same time as supporting the absorption and metabolism of nutrients for growth and development (reviewed in (Steele *et al.* 2016; Meale *et al.* 2017)). Prior to the development of next generation sequencing technologies many studies used quantitative PCR to measure the expression of sets of candidate genes in ruminant GI tract tissues (reviewed in (Connor *et al.* 2010)). RNA-Sequencing (RNA-Seq) technology now provides a snapshot of the transcriptome in real-time to generate global gene expression profiles. This allows us to measure the expression of all protein coding genes throughout the development of the GI tract and associate these expression patterns with immunity, metabolism and other cellular processes at the gene/transcript level. The availability of high quality, highly contiguous, well annotated reference genomes for ruminant species, due in part to the efforts of the Functional Annotation of Animal Genomes Consortium (FAANG) (Andersson *et al.*

2015), has helped significantly in this task, particularly for sheep (Jiang *et al.* 2014) and goat (Bickhart *et al.* 2017; Worley 2017).

Previous studies characterising transcription in the GI tract throughout early development in ruminants examined links between feed intake and bacterial diversity and the development of the rumen (Connor *et al.* 2013; Wang *et al.* 2016; Xiang *et al.* 2016a). Another recent study characterised transcription in the adult rumen complex and GI tract of sheep, linking metabolic, epithelial and metabolic transcriptomic signatures (Xiang *et al.* 2016b). Similarly, Chao et al. 2017 performed transcriptional analysis of colon, caecum and duodenum from two breeds of sheep highlighting key genes involved in lipid metabolism (Chao *et al.* 2017). Others have linked diet to gene expression in the jejunal mucosa of young calves, suggesting that early feeding can have a profound effect on the expression of genes involved in metabolism and immunity (Hammon *et al.* 2018). The GI tract plays a significant role in defence against infection, because due to its large surface area it is often a site of primary infection via pathogen ingestion. This is particularly true in early development when the maturing lamb is exposed both to a range of pathogens from the environment and the rumen is colonised by commensal micro-organisms. Intestinal macrophages exhibit distinctive properties that reflect adaptation to a unique microenvironment (Bain & Mowat 2011). The microenvironment of the sheep GI tract changes dramatically during the transition from pre-ruminant to ruminant, and we hypothesise that the transcriptional signature of intestinal macrophages will reflect these physiological changes.

To characterise tissue specific transcription in the GI tract during early development we utilised a subset of RNA-Seq data, from Texel x Scottish Blackface (TxBF) lambs at birth, one week and 8 weeks of age and TxBF adult sheep, from our high resolution atlas of gene expression in sheep (Clark *et al.* 2017b). We characterise in detail the macrophage signature

126 in GI tract tissues at each developmental stage and link these to other key biological processes
127 ocurring as the lamb develops. We also perform comparative analysis of transcription in the
128 rumen, ileum and colon of one-week old sheep with age-matched goats. One of the most
129 significant physiological challenges to neonatal and juvenile ruminants is the development
130 and establishment of the rumen. A clearer understanding of the transcriptomic complexity
131 that occurs during the transition between pre-ruminant and ruminant will allow us to identify
132 key genes underlying immunity in neonatal ruminants and healthy growth and development
133 of the rumen and other tissues. These genes could then be utilised as novel targets for
134 therapeutics in sheep and other ruminants and as a foundation for understanding rumen
135 development, metabolism and microbial colonisation in young ruminants to improve
136 efficiency and productivity.

137 **Materials and Methods**

138 **Animals**

139 Approval was obtained from The Roslin Institute and the University of Edinburgh
140 Protocols and Ethics Committees. All animal work was carried out under the regulations of
141 the Animals (Scientific Procedures) Act 1986. Full details of all the sheep used in this study are
142 provided in the sheep gene expression atlas project manuscript (Clark *et al.* 2017b) and
143 summarised in Table 1. GI tract tissues were collected from three male and three female adult
144 Texel x Scottish Blackface (TxBF) sheep and nine Texel x Scottish Blackface lambs. Of these
145 nine lambs, three were observed at parturition and euthanised immediately prior to their first
146 feed. Three lambs were euthanised at one week of age pre-rumination (no grass was present
147 in their GI tract) and three at 8 weeks of age once rumination was fully established. All the
148 animals were fed *ad libitum* on a diet of hay and sheep concentrate nuts (16% dry matter

without water), with the exception of the lambs pre-weaning (birth and one week of age) who suckled milk from their mothers. Goat GI tract and alveolar macrophage samples from one-week old un-weaned male goat kids were obtained from an abattoir (Table 1). We also included available RNA-Seq data from a trio of Texel sheep (a ewe, a ram and their female lamb) that was released with the sheep genome paper (Jiang *et al.* 2014). By incorporating this data in the analysis, we were able to include tissues not included in the TxBF samples e.g. omentum and an additional developmental stage (6-10 months) (Table 1).

Tissue Collection

In total thirteen different regions of the sheep GI tract were sampled, as detailed in Table 1 and illustrated in Figure 1. All post mortems were undertaken by the same veterinary anatomist and tissue collection from the GI tract was standardised as much as possible. All GI tract tissue samples were washed twice in room temperature sterile 1x PBS (Mg^{2+} Ca^{2+} free) (P5493; Sigma Aldrich) then chopped into small pieces <0.5cm and transferred to RNAlater preservation solution (AM7021; Thermo Fisher Scientific, Waltham, USA). To maintain RNA integrity, all GI tract tissue samples were harvested within 30 minutes from the time of death. Alveolar macrophages, liver, rectum and oesophageal muscle and tissue from the adult sheep were also included as reference tissues for comparison in the analysis. Isolation of alveolar macrophages from the adult sheep is described in detail in the sheep gene expression atlas project manuscript (Clark *et al.* 2017b). The goat samples (Table 1) were collected using the same methods as described for sheep.

RNA Extraction and Library Preparation

We used a TRIzol® (15596018; Thermo Fisher Scientific) based RNA extraction method

which is described in detail in the sheep gene expression atlas project manuscript (Clark *et al.* 2017b). RNA quantity was measured using a Qubit RNA BR Assay kit (Q10210; Thermo Fisher Scientific) and RNA integrity estimated on an Agilent 2200 Tapestation System (Agilent Genomics, Santa Clara, USA) using the RNA Screentape (5067-5576; Agilent Genomics) to ensure RNA quality was of $RIN^e > 7$. RNA-Seq libraries were prepared by Edinburgh Genomics (Edinburgh Genomics, Edinburgh, UK) and run on the Illumina HiSeq 2500 (sheep) and Illumina HiSeq 4000 (goats) sequencing platform (Illumina, San Diego, USA). The GI tract tissues collected from the 9 TxBF lambs were sequenced at a depth of >25 million strand-specific 125bp paired-end reads per sample using the standard Illumina TruSeq mRNA library preparation protocol (poly-A selected) (Ilumina; Part: 15031047 Revision E). Depth refers to the number of paired end reads, therefore a depth of >25 million reads represents ~25M forward + ~25M reverse. The adult sheep GI tract tissues were also sequenced as above with the exception of the ileum, reticulum and liver which were sequenced using the Illumina TruSeq total RNA library preparation protocol (ribo-depleted) (Ilumina; Part: 15031048 Revision E) at a depth of >100 million reads per sample. The ileum and rumen samples from goats were sequenced at a depth of >30 million strand-specific 75p paired-end reads per sample using the standard Illumina TruSeq mRNA library preparation protocol (poly-A selected) (Ilumina; Part: 15031047 Revision E). The RNA-Seq libraries for the Texel dataset were also prepared by Edinburgh Genomics with RNA isolated using the method described in the sheep gene expression atlas project manuscript (Jiang *et al.* 2014; Clark *et al.* 2017b).

193

194 **Data Quality Control and Processing**

195 The raw data for sheep, in the form of .fastq files, was previously released with the
196 sheep gene expression atlas (Clark *et al.* 2017b) and is deposited in the European Nucleotide
197 Archive (ENA) under study accession number PRJEB19199
198 (http://www.ebi.ac.uk/ena/data/view/PRJEB19199). The goat data is also deposited in the
199 ENA under study accession number PRJEB23196
200 (http://www.ebi.ac.uk/ena/data/view/PRJEB23196). Both sets of data were submitted to the
201 ENA with experimental metadata prepared according to the FAANG Consortium metadata
202 and data sharing standards. Details of all the samples for sheep and goat, with associated data
203 and metadata can also be found on the FAANG Data Portal (http://data.faang.org/) (FAANG
204 2017). The raw read data from the Texel samples incorporated into this dataset and
205 previously published (Jiang *et al.* 2014) is located in the ENA under study accession PRJEB6169
206 (http://www.ebi.ac.uk/ena/data/view/PRJEB6169). The RNA-Seq data processing
207 methodology and pipelines used for this study are described in detail in (Clark *et al.* 2017b).
208 For each tissue a set of expression estimates, as transcripts per million (TPM), was obtained
209 using the high-speed transcript quantification tool Kallisto v0.43.0 (Bray *et al.* 2016) with the
210 Oar v3.1 reference transcriptome from Ensembl (Zerbino *et al.* 2018) as an index. Expression
211 estimates for the GI tract dataset were then filtered to remove low intensity signals (TPM<1)
212 and technical artefacts. To integrate expression estimates from the two different library types
213 we performed a ratio correction of the TPM values as described in (Bush *et al.* 2017).

214

215 **Network Cluster Analysis**

Network cluster analysis of the sheep GI tract dataset was performed using the network visualisation tool Graphia Professional (Kajeka Ltd, Edinburgh, UK) (Theocharidis *et al.* 2009; Livigni *et al.* 2018). To determine similarities between individual gene expression profiles a Pearson correlation matrix was calculated, for both sample-to-sample and gene-to-gene comparisons and filtered to remove relationships where $r < 0.81$ (sample-to-sample) and $r < 0.85$ (gene-to-gene). Network graphs were constructed by connecting nodes (genes or samples) with edges (where the correlation exceeded the threshold value). To interpret each graph a Markov Cluster algorithm (MCL) (van Dongen & Abreu-Goodger 2012) was applied at an inflation value (which determines cluster granularity) of 2.2. Visual examination was then used to interrogate the local structure of the graph. Similar samples/tissue types and genes with robust co-expression patterns formed clusters of highly interconnected nodes, implying related functions. To determine if genes within a cluster shared a similar biological function GO term enrichment based on gene ontology (Ashburner *et al.* 2000) was performed using the Bioconductor package 'topGO' (Alexa & Rahnenfuhrer 2010). For the gene-to-gene network analysis the top 20 largest clusters were assigned a functional class and sub-class based on GO term enrichment, gene function and the 'guilt-by-association' principle (Oliver 2000). Alveolar macrophage, oesophageal muscle and mucosa, rectum and liver from the adult sheep were included in the gene to gene network graph as 'reference' tissues for comparison, using a similar approach as described in (Xiang *et al.* 2016b). The Texel dataset (Jiang *et al.* 2014) was also included in the gene-to-gene network graph to add additional tissue samples and developmental stages not captured in the TxBF dataset. The sample-to-sample network analysis was used to illustrate transcriptional changes in tissues through each developmental stage (birth, one week, 8 weeks and adult) and included only the dataset from the TxBF sheep.

240

241 **Principal component analysis of transcriptional signatures in the developing GI tract**

242 All statistical analysis was carried out in R (v >= 3.0.0) (R Core Team 2014) unless stated
243 otherwise. We used Principal Component Analysis (PCA) to determine whether there was any
244 strong age or tissue related transcriptional patterns observed in the GI tract. PCA was
245 performed using FactoMineR v1.41 (Lê *et al.* 2008) with a subset of genes (n = 490) (extracted
246 from the alveolar macrophage clusters (clusters 7 and 10) from the gene-to-gene network
247 graph), centre scaled for computation of the principle components (PCs). The categorical data
248 was excluded from the Eigen vector calculation (passed as qualitative variables). The top 5
249 and 10 PCs explained 62.6% and 76.9% of variability in the data respectively. In order to
250 compare exploratory and discriminative power of PCs, the categorical data (age and tissue of
251 origin) was overlaid on the PCs coordinate maps afterwards using centroid lines and colouring
252 the observations by groups.

253 The expression levels as TPM of a small subset of macrophage and immune marker
254 genes (*CD14, CD68, CD163* and *IL10*) were plotted as a HeatMap using 'gplots'.

255

256 **Developmental Stage Specific Differential Expression Analysis for Sheep**

257 Differential expression analysis was used to compare gene-level expression estimates
258 from the Kallisto output as transcripts per million (TPM) across GI tract tissues and
259 developmental time points, and between age-matched sheep and goats. The R/Bioconductor
260 package tximport v1.0.3 was used to import and summarise the transcript-level abundance
261 estimates from Kallisto for gene-level differential expression analysis using edgeR v3.14.0
262 (Robinson *et al.* 2010), as described in (Soneson *et al.* 2015; Love *et al.* 2017). For RNA-Seq
263 experiments with less than 6 replicates per time point edgeR may be considered the optimal

differential expression analysis package (Schurch *et al.* 2016). We selected three tissues (abomasum, rumen and ileum) as representative samples from 3 of the major compartments of the GI tract: the stomach, rumen complex and small intestine, respectively. Gene expression patterns in each tissue were compared between birth and one week, and one week and 8 weeks of age. To determine whether there was any tissue- and developmental stage-specific overlap of differentially expressed genes we generated Venn diagrams using the software tool Venny (Oliveros 2007). To investigate the function of the differentially expressed genes we performed GO term enrichment (Ashburner *et al.* 2000) using 'topGO' (Alexa & Rahnenfuhrer 2010). Only GO terms with 10 or more associated genes were included.

Data Availability

All data analysed during this study are included in this published article and its additional files. The raw RNA-sequencing data are deposited in the European Nucleotide Archive (ENA) under study accessions PRJEB19199 (sheep) (https://www.ebi.ac.uk/ena/data/view/PRJEB19199) and PRJEB23196 (goat) (https://www.ebi.ac.uk/ena/data/view/PRJEB23196). Sheep data can also be viewed and downloaded via BioGPS (http://biogps.org/dataset/BDS_00015/sheep-atlas/) where the gene expression estimates for each tissue are searchable by gene name (http://biogps.org/sheepatlas). Sample metadata for all tissue and cell samples, prepared in accordance with FAANG consortium metadata standards, are deposited in the EBI BioSamples database under group identifiers SAMEG317052 (sheep) and SAMEG330351 (goat). All experimental protocols are available on the FAANG consortium website at http://www.ftp.faang.ebi.ac.uk/ftp/protocols. All supplementary material for this study has been deposited in Figshare.

Results and Discussion

Gene-to-gene network cluster analysis of the GI tract dataset

The dataset includes 168 RNA-Seq libraries in total from the TxBF described above sheep. Network cluster analysis of the GI tract data was performed using Graphia Professional (Kajeka Ltd, Edinburgh UK) (Theocharidis *et al.* 2009; Livigni *et al.* 2018). TPM estimates from Kallisto averaged across biological replicates (3 sheep per developmental stage) for the GI tract dataset were used to generate the network cluster graph. The full version of this averaged dataset was published with the sheep gene expression atlas and is available for download through the University of Edinburgh DataShare portal (http://dx.doi.org/10.7488/ds/2112). A version only including the TPM estimates for GI tract tissues, alveolar macrophages, thoracic oesophagus and liver is included here as Table S1.

The dataset was clustered using a Pearson correlation co-efficient threshold of $r = 0.85$ and MCL (Markov Cluster Algorithm (Gough *et al.* 2001)) inflation value of 2.2. The gene-to-gene network graph (Figure 2A) comprised 13,035 nodes (genes) and 696,618 edges (correlations above the threshold value). The network graph (Figure 2A) was highly structured comprising 349 clusters of varying size. Genes found in each cluster are listed in Table S2. Genes in Table S2 labelled 'assigned to' were annotated using an automated pipeline for the sheep gene expression atlas (Clark *et al.* 2017b). Clusters 1 to 20 (numbered in order of size; cluster 1 being the largest including 1724 genes in total) were annotated visually and assigned a broad functional 'class' (Table 2). Validation of functional classes was performed using GO term enrichment (Alexa & Rahnenfuhrer 2010) for molecular function, cellular component and biological process (Table S3). Figure 2B shows the network graph with the

312 nodes collapsed by class, and the largest clusters numbered 1 to 20, to illustrate the relative
313 number of genes proportional to the size of each cluster.
314
315 **Complexity of cell types within the GI tract is reflected in gene-to-gene network clustering**
316 The GI tract is a highly complex organ system in ruminants with region-specific cellular
317 composition. This complexity is illustrated by the highly structured gene-to-gene network
318 graph of GI tract tissues (Figure 2A). Other studies have characterised in detail the
319 transcriptional signatures in the GI tract of adult sheep (Xiang *et al.* 2016a; Xiang *et al.* 2016b)
320 and as such we will only broadly describe the clusters here and focus instead on
321 developmental stage-specific transcriptional patterns. As is typical of large gene expression
322 datasets from multiple tissues cluster 1, the largest cluster, was comprised largely of
323 ubiquitously expressed 'house-keeping' genes, encoding proteins that are functional in all cell
324 types. Enriched GO terms for genes within this cluster were for general cellular processes and
325 molecular functions performed by all cells including RNA-processing ($p=<1x10^{-30}$) and histone
326 binding ($p=3.9x10^{-8}$) (Table S3).
327 Although cluster 1 exhibited a general expression pattern the majority of the rest of
328 the clusters (with the exception of cluster 2, discussed below) exhibited some tissue/cellular
329 process specific gene expression pattern. The majority of clusters included genes associated
330 with more than one cell type/cellular process. This is because the lamina propria, one of the
331 three layers of the mucosa, which lies beneath the epithelium and lines the majority of the GI
332 tract, is made up of different cell types including endothelial, immune and connective tissues
333 (Takahashi-Iwanaga & Fujita 1985). As a consequence expression signatures of components
334 of the lamina propria were common across the network graph for sheep, as previously
335 reported from transcriptional analysis of the pig GI tract (Freeman *et al.* 2012).

The GI tract of the sheep is lined with epithelium, and the rumen in particular exhibits a strong epithelial signature, which differentiates it transcriptionally from other tissues (Xiang *et al.* 2016a; Xiang *et al.* 2016b). Genes in clusters 8 and 12, expressed in the rumen, exhibited a typical epithelial signature, and included genes in the KRT superfamily e.g. *KRT15*, with enriched GO terms for keratinocyte differentiation ($p=1.4x10^{-7}$ and $p=6.8x10^{-7}$, respectively). The rumen papillae help to increase the surface area available for the absorption of nutrients. The small intestine is also a major site of absorption and transport of nutrients. One small intestine specific cluster (cluster 11) was comprised of genes involved both in nutrient uptake and vesicle formation. Cluster 11 contained numerous genes in the solute carrier super family, a group of membrane transport proteins including *SLC10A4* and *SLC16A11*. Genes encoding water channel proteins were also found in cluster 11 including aquaporins *AQP7* and *AQP8*. Significant GO terms for cluster 11 included acrosomal vesicle (p=0.0033) and peptidyl-tyrosine autophosphorylation (p=0.00209).

Genes in cluster 5 were expressed in thoracic oesophageal skeletal muscle tissue and associated with muscle fibre development ($p=1.3x10^{-10}$) and skeletal muscle tissue development ($p=4.4x10^{-12}$). Several skeletal muscle specific genes were found within this cluster including, *ACTA1*, which is associated with skeletal muscle function and encodes a product belonging to the actin family of proteins (Clarke et al. 2007). Similarly, cluster 13 included genes associated with structural constituents of muscle (p=0.00108) but in this case the smooth muscle that surrounds the GI tract. Genes in cluster 13 were predominantly expressed in the rumen complex of adult and 8 week old sheep, and included several genes related to smooth muscle function and regulation (Table 2). *CALD1* for example encodes a calmodulin- and actin-binding protein that plays an essential role in the regulation of smooth

359 muscle and nonmuscle contraction (Huang et al. 2010). Intestinal motility, is the result of
360 contraction and relaxation of the smooth muscle, and a key function of the GI tract.

361 The numerous functions of the GI tract such as intestinal motility, exocrine and
362 endocrine secretion, immune function and circulation are under complex neural control.
363 Clusters 15 and 19 were associated with neuronal cells and included GO terms for synapse
364 ($p=0.000018$) and synaptic membrane ($p=4.9x10^{-7}$). However, overall expression of these
365 genes was low relative to other clusters, presumably because neurons comprise only a small
366 percentage of the cells that make up the GI tract, and therefore their expression level would
367 be reduced compared to other more abundant cell types. A similar trend was observed in
368 analysis of neuronal signatures in the pig GI tract (Freeman *et al.* 2012). This might also
369 account for an increased expression level at birth of genes in cluster 15 as there will be
370 comparatively fewer other cell types differentiated at this time point.

371 Several clusters in the top 20 largest clusters exhibited a strong immune signature and
372 are discussed in more detail below. Analysis of the gene-to-gene network graph revealed the
373 GI tract of the sheep comprises at least five major cell types: epithelia, immune cells, neuronal
374 cells and mesenchymal cells (muscle, connective tissue). Interrogation of the expression
375 profiles for each cluster in the gene-to-gene network graph revealed that cell type and tissue
376 specific expression signatures varied with developmental stage (discussed in more detail
377 below).

378

379 **Strong immune signatures highlight the role of the GI tract as an immune organ**

380 The GI tract has the largest surface area in the body in contact with the external
381 environment. The intestinal epithelium, a single-cell layer, is the barrier between the external
382 environment of the lumen (which contains pathogens, antigens and commensal bacteria) and

383 the body. As such it is not surprising that the second largest cluster of genes (cluster 2)
384 contained many genes associated with the immune response, their expression being two- to
385 three-fold higher in the ileum and Peyer's patches than other regions of the GI tract. This
386 pattern of expression was also observed in pig (Freeman *et al.* 2012). The lower small intestine
387 Peyer's patches (small masses of lymphatic tissue found throughout the ileum) form an
388 important part of the immune system by monitoring intestinal bacteria populations and
389 preventing the growth of pathogenic bacteria (Jung *et al.* 2010). Included in cluster 2 were
390 genes encoding many of the protein components of the B cell receptor complex
391 (*CD19*, *CD79B*, *CR2*) (Treanor 2012). Also evident in this cluster were many genes associated
392 with the cell cycle, for example cyclins, DNA polymerases and kinesins, which were identified
393 in the sheep gene expression atlas as a cell cycle specific cluster (Clark *et al.* 2017b). Significant
394 GO terms for cluster 2 include G2/M transition of mitotic cell cycle ($p=1.8x10^{-9}$) and meiotic
395 cell cycle process ($p=6.9x10^{-9}$). The high level of lymphocyte proliferation and replenishment
396 of intestinal macrophages by a high turnover of monocytes (Bain & Mowat 2011; Shaw *et al.*
397 2018), and therefore the frequency of cells undergoing mitosis in the Peyer's patches and
398 Ileum, might explain the association of cell cycle genes with an immune signature (David *et*
399 *al.* 2003).

400 Other GI clusters with an immune signature included clusters 3, 7, 9 and 10. Cluster 3
401 exhibited a strong immune signature, particularly associated with T-cells, with significant GO
402 terms for positive regulation of T-cell activation ($p=2.1x10^{-21}$) and cytokine receptor activity
403 ($p=5.4x10^{-5}$). Genes within cluster 3 included the T-cell marker genes *CD3, CD4* and *CD6,* as
404 well as Toll-like receptor genes *TLR1* and *TLR10* (Akira & Takeda 2004). Expression was highest
405 for this cluster in the small intestine particularly the ileum and Peyer's patches. Cluster 9
406 exhibited a similar immune related expression pattern including genes involved in T-cell-B-

cell interactions such as *CD37*, and cytokine production such as *IL10* and *CD80*. Significant GO terms for cluster 9 included B cell receptor signalling pathway ($p=4.9x10^{-6}$) and regulation of B-cell activation ($p=6.3x10^{-6}$) as well as Toll-like receptor 4 signalling pathway ($p=2.8x10^{-5}$).

Macrophages play an essential role in inflammation and protective immunity and in the GI tract are important for local homeostasis, maintaining a balance between microbiota and the host immune system (Mowat & Bain 2011). Clusters 7 and 10 contained many macrophage specific marker genes including *CD68* and *ADGRE1 (EMR1)* with enriched GO terms for the immune response ($p=7.3x10^{-6}$). Many of these genes are also commonly used as marker genes for intestinal macrophages including, *CCRL1* (synonym *CX3CR1*), *ITGAX* (synonym *CD11C*) and *ITGAM* (synonym *CD11B*) (Mowat and Bain 2011; Bain and Mowat 2011; Shaw et al. 2018). Other macrophage marker genes such as *CD14* and *CD163* were found in a much smaller macrophage-specific cluster, cluster 154, indicating some tissue specific differences in macrophage expression.

We included alveolar macrophages in the gene-to-gene network clustering of the GI tract dataset to look specifically at the expression of C-type lectins, and other receptors involved in bacterial recognition, in the GI tract relative to their expression in other populations of macrophages. Upon initial encounter with pathogens, the immune system needs to rapidly recognise these as potentially harmful. Innate immune cells, such as macrophages, use a limited number of pattern recognition receptors (PRRs), including C-type lectin receptors (CLRs), which activate immediate anti-microbial effectors or other mechanisms for defence against pathogens (reviewed in (Akira & Takeda 2004)). Analysis of the pig gene expression atlas revealed that there were a large number of genes for C-type lectins that were highly-expressed in alveolar macrophages but appeared down-regulated in the pig GI tract (Freeman *et al.* 2012). To determine if this was also the case in sheep we

examined the expression of six C-type lectin marker genes (*CD68*, *CLEC4D*, *CLEC4E, CLEC5A*, *CLEC7A* and *SIGLEC1*) in sheep using the gene expression profiles for the sheep atlas dataset on BioGPS (Clark *et al.* 2017a). In sheep, as in pig, the four C-type lectins and *SIGLEC1* all exhibited very weak expression in the GI tract tissues. *CLEC4D, CD68*, *CLEC4E* and *CLEC7A* were all highly expressed in sheep alveolar macrophages, as was observed in pig. *CLEC5A* and *SIGLEC1* did not show the same upregulation in sheep alveolar macrophages as they did in pigs indicating these genes exhibit a more species-specific expression pattern in alveolar macrophages. To investigate whether these trends also applied to other ruminants, we compared the expression of these genes in alveolar macrophages, ileum, large colon and rumen from one-week old sheep with age-matched goats (Figure 3). Similar expression patterns were observed for goat as for sheep, the only exception being *CLEC5A* which was upregulated in goat alveolar macrophages, unlike in sheep. Generally, the high expression of gene for C-type lectins in alveolar macrophages and down regulation in the GI tract appears to be conserved between both ruminants and monogastric mammals.

The expression pattern of *CD68, CLEC4D, CLEC4E* and *CLEC7A*, which are upregulated in alveolar macrophages and down regulated in the GI tract, indicate they are likely to be necessary for the elimination of inhaled pathogens, where such inflammatory responses to microbes would be undesirable in the intestine. *CLEC4E*, for example, has been shown to be important in protective immunity against pneumococcal pneumonia in lung macrophages in mice (Behler-Janbeck *et al.* 2016) and *CLEC7A* is involved in the innate immune respiratory response to fungal pathogens (Saijo & Iwakura 2011). Although ruminants have a larger and more diverse microflora in the GI tract compared to pigs, which are monogastric mammals (O'Donnell *et al.* 2017), expression patterns of the C-type lectin marker genes appear to be similar between the two species. Our findings generally support those observed in pig that

intestinal macrophages differ transcriptionally from lung and blood macrophage populations, which might be because they have adapted to be hypo-responsive to food-derived glycoproteins while alveolar macrophages need to utilise the receptors to both recognize and engulf potential pathogens (Freeman *et al.* 2012).

Network cluster analysis and principal component analysis of tissue samples reveals a strong effect of developmental stage on tissue specific transcription

Sample-to-sample network cluster analysis was used to visualise the effect of developmental stage on tissue specific gene expression. To perform the sample-to-sample network cluster analysis we used a version of the GI tract dataset that was not averaged across individuals and did not include the Texel dataset or the alveolar macrophage (AM), oesophagus and liver samples (Figure 4 A&B). The full version of this dataset was published with the sheep gene expression atlas and is available for download through the University of Edinburgh DataShare portal (http://dx.doi.org/10.7488/ds/2112). A version only including the TPM estimates used for this analysis is included here as Table S4.

Developmental stage specific transcriptional patterns were reflected particularly strongly in the rumen complex (reticulum, rumen and omasum) tissue samples (Figure 4A). Rumen complex samples from lambs at birth and one week of age formed a distinct cluster from rumen complex samples at 8 weeks of age and adult. This reflects the changes in tissue structure, function and morphology of the reticulum, rumen and omasum in the transition from pre-ruminant to ruminant (Figure 1). The close proximity of the adult oesophageal samples and the pre-ruminant rumen complex samples in the network cluster analysis indicate that tissue from the rumen complex at birth and one week of age is transcriptionally similar to adult oesophageal tissue in comparison with tissue from the rumen of adult sheep

(Figure 4A). At birth and one week of age the walls of the reticulum join together to form an 'oesophageal' or 'reticular' groove transferring milk and colostrum directly to the abomasum, where it is digested most efficiently bypassing the rumen (Figure 1) (Meale *et al.* 2017). Rumen tissue of lambs at birth and at one week of age will therefore be functionally and transcriptionally different to rumen tissue of adult sheep. The rumen of neonatal lambs, for example, is essentially non-functional in ketogenic capacity and does not exhibit the same high degree of keratinisation that is characteristic of the adult rumen (reviewed in (Baldwin *et al.* 2004)). Consequently, the strong epithelial and metabolic signatures of the adult rumen (Xiang *et al.* 2016b) are likely to be weaker at one week and birth driving some of the observed cluster separation between the developmental stages (Figure 4A).

The sample-to-sample graph also illustrates the strong tissue specific differences in transcription (Figure 4B). The pylorus and abomasum samples, for example, form distinct clusters from the other regions, reflecting the fact that they have unique functions. The pylorus and abomasum both have a glandular lining, with various glandular cell types involved with enzyme secretion, lubrication, and endocrine control and specialized structures, such as the pyloric and fundic glands of the stomach (Dyce *et al.* 2010). The large and small intestine and rumen complex also form large separate clusters highlighting the transcriptional similarities between tissues within and between these regions of the GI tract (Figure 4B).

Developmental stage- and tissue-specific transcriptional patterns observed in the sample to sample graph were also reflected in principal component analysis (PCA) of the dataset (Figure S1). Tissue specific clusters of samples were largely separated by organ system (rumen complex, small and large intestine, oesophagus and stomach) (Fig S1 A&B). The effect of age on transcription was particularly evident in PC2, with samples from newborn lambs clustered separately from the other 3 time points (Figure S1 C&D).

Developmental stage specific changes in immune gene signatures occur during the transition from pre-ruminant to ruminant

One of the largest macrophage populations in the body occupies the lamina propria of the GI tract (Bain & Mowat 2011). Macrophages are so numerous that the expression of macrophage-related genes can be detected within the total mRNA from GI tract samples. Constant exposure to food and environmental antigens and a wide diversity of commensal bacteria make the GI tract of developing sheep a primary site for pathogen infection (Mantovani & Marchesi 2014). As a consequence, GI immune cells are involved in a balanced immune response, focussed on controlling pathogen invasion while recognising commensal colonising microbes. This response will change as the lambs age reflecting changes in diet (the transition from milk to pellets and hay) and environment (transition from maternal transmission of pathogens to those from the surrounding environment). Intestinal mononuclear phagocytes, of which monocyte-derived macrophages form the most abundant population play a key role in the maintenance of the intestinal immune response to pathogens and homeostasis (Mantovani & Marchesi 2014). Macrophage specific genes are therefore likely to be driving some of the observed transcriptional differences observed between lambs at birth and one week of age and 8 week old and adult sheep, described in the previous section (See Figure 4 and Figure S1).

We used PCA to examine the clustering pattern of a set of 490 macrophage associated genes (extracted from the alveolar macrophage specific clusters 7 and 10 from the gene to gene network graph, detailed in Table S2) (Figure 5). We looked at tissue specific transcriptional patterns (Figure 5A) and found the small intestine and rumen complex distinguished by components PC1 and PC2, respectively, indicating differences in the

527 macrophage expression profile in these tissues. A similar effect was observed between
528 components PC1 and PC3 (Figure 5B). Developmental stage specific transcriptional patterns
529 in macrophage specific genes were evident when comparing components PC1 and PC3 (Figure
530 5C), and particularly when comparing PC3 and PC4 (Figure 5D), where the samples from
531 newborn lambs are clearly distinct from the other developmental stages. These differences
532 perhaps reflect that pathogen challenge in the newborn lambs will be reduced immediately
533 at birth and therefore the inflammatory response would be less than at later developmental
534 stages. In addition, the newborn lambs would not have received any colostrum which is key
535 in the homologous transfer of passive immunity between the mother and neonate, and the
536 initial source of acquired immunity by the newborn lamb (Stelwagen *et al.* 2009). A wide
537 variety of immune components linked to the innate immune response have been identified
538 in colostrum and milk including neutrophils and macrophages as well immunomodulatory
539 factors (including numerous pro- and anti-inflammatory cytokines) and peptides and proteins
540 with direct antimicrobial activity (Stelwagen *et al.* 2009). Another potential driver of these
541 differences is that the GI tract will gradually become colonised with gut macrophages, and at
542 birth there will be a comparatively higher proportion of monocytes (Shaw *et al.* 2018). The GI
543 tract harbours multiple distinct populations of macrophages, monocytes and other immune
544 cells that exhibit some level of stochasticity over time (Bain & Mowat 2011; Shaw *et al.* 2018)
545 which is reflected in the observed transcriptional patterns. The macrophage populations that
546 are most numerous in neonatal lambs may be quite different from those that are prevalent
547 in the GI tract of adult sheep and this will vary according to tissue.

548 To look in more detail at developmental stage specific differences in expression of
549 monocyte and macrophage marker genes we compared the expression of *CD14, CD68, CD163*
550 and *IL10* across GI tract tissues at each developmental stage (including liver and alveolar

551 macrophage samples for reference) (Figure 6). *CD14, CD68* and *CD163* were upregulated in
552 alveolar macrophages as expected (Gordon *et al.* 2014). This is not immediately obvious form
553 the heatmap however due to high level of expression of CD68 (2291 TPM) relative to the
554 levels of CD14 (171 TPM) and CD163 (168 TPM). Expression of *CD68* was highest in the ileum
555 relative to other GI tract tissues, which probably reflects the fact that ileum contains the
556 largest proportion of tissue macrophages relative to the other GI tract tissues sampled. The
557 expression of the monocyte marker gene *CD14* increased as the lambs aged across most of
558 the GI tract tissues sampled. This may reflect the lower number of monocytes in neonatal
559 lambs relative to juvenile and adult animals (Kramer *et al.* 2003) or the high turnover of blood
560 monocytes and continual replenishment of intestinal macrophages by monocytes that has
561 been reported in the GI tract (Bain *et al.* 2016). The effect was most obvious in the rumen
562 complex tissues in which *CD14* expression increased with development of the rumen. In
563 contrast *CD163* and *CD68* expression decreased as the lambs aged in the rumen complex but
564 increased in the developing small and large intestine. These differences in expression of
565 macrophage marker genes reflect tissue specific differences in macrophage colonisation with
566 age and the changing function of the rumen through these developmental stages. *CD163*, for
567 example, directly induces intracellular signaling leading to secretion of anti-inflammatory
568 cytokines (Moestrup & Moller 2004), and has been shown to be a macrophage receptor for
569 bacteria (Fabriek *et al.* 2009). During bacterial infection *CD163* on resident tissue
570 macrophages acts as an innate immune sensor, inducing local inflammation (Fabriek *et al.*
571 2009). Induction of *CD163* might therefore be less desirable in the rumen once colonisation
572 with intestinal microbes has been established, accounting for the decrease in expression of
573 *CD163* in the rumen complex. The increase in expression of *CD163* in the ileum and Peyer's

574 patches with age reflects the gradual colonisation of these tissues with intestinal
575 macrophages and their importance in the innate immune response.

576 The expression of *IL10* was very low across all the GI tract tissues which is in contrast
577 to other mammalian species where it has been shown to be constitutively expressed in
578 intestinal macrophages, reviewed in (Bain & Mowat 2011). For the majority of tissues TPM
579 was <3 with the exception of the ileum and Peyer's patch where expression ranged between
580 6 and 12 TPM. Interleukin 10 (*IL10*) is a key anti-inflammatory cytokine that can inhibit
581 proinflammatory responses of both innate and adaptive immune cells (Mantovani & Marchesi
582 2014; Shouval *et al.* 2014). A relative reduction in *IL10* expression in the GI tract could indicate
583 a suppressed macrophage response to intestinal microbes. This could be of potential benefit
584 for ruminant species, which have a different relationship with intestinal flora than
585 monogastric mammals and this would be interesting to explore further.

586

587 **Differential expression analysis reveals little overlap between stage-specific transcriptional**
588 **signatures**

589 To characterise which genes were driving transcriptional patterns, more broadly,
590 during the transition from pre-ruminant to ruminant we chose three regions of the GI tract
591 (rumen complex, stomach and small intestine) and selected one tissue per region (rumen,
592 abomasum and Ileum) to perform differential expression analysis using pairwise comparisons
593 between birth and one week and one week and 8 weeks of age (Table S5 – Rumen, Table S6
594 – Abomasum and Table S7 – Ileum). Differentially expressed genes are detailed in S5 Table
595 (rumen), S6 Table (abomasum) and S7 Table (ileum); those of particular interest are shown in
596 Table 3. To determine the main functions of genes exhibiting differential expression in each
597 comparison we used GO term enrichment (Table S8). GO terms related to immunity were

598 common across all three tissues but particularly the rumen and ileum between birth and one
599 week of age. In the ileum between birth and one week of age enriched GO terms were
600 predominantly associated with immunity and included 'immune response', 'chemokine
601 activity', 'chemokine production involved in the inflammatory response' as well as others
602 related to metabolism. In contrast between one and 8 weeks of age GO terms were associated
603 predominantly with metabolic processes, and vesicle formation. A similar but less
604 exaggerated trend was also observed in the rumen, with GO terms for 'defense response to
605 other organisms', 'immune response' and more generally for metabolism enriched between
606 birth and one week of age and then a shift towards metabolic and muscle and epithelial
607 differentiation between one and 8 weeks. In the abomasum enriched GO terms between both
608 sets of time points were more similar than for the other two tissues which probably reflects
609 the fact that the functional changes in the abomasum are comparatively less than the other
610 two tissues throughout the transition from pre-ruminant to ruminant.

611 Using Venn diagrams (Figure S2) we compared whether any differentially expressed
612 genes were shared across tissues and developmental stages. Relatively few genes were
613 shared between developmental stages (0-1 week and 1-8 weeks), fifty genes were shared
614 between the rumen time points, 19 between the ileum and 13 between the abomasum
615 (Figure S2). A greater number of differentially expressed genes were shared between the
616 rumen and ileum than either of these two tissues and the abomasum, which reflects the
617 different functional roles of these three tissues. Relatively few genes were differentially
618 expressed in two or more tissues and time points, the most notable being *HERC6* which is
619 involved in ubiquitin ligase activity and *ISG15* an IFN-gamma-inducing cytokine playing an
620 essential role in antimycobacterial immunity (Table 3).

621 The structural composition of the rumen and ileum changes significantly during the
622 transition from pre-ruminant to ruminant, with the most dramatic physical changes
623 associated with the rumen epithelium (Baldwin *et al.* 2004). Several genes associated with
624 connective tissue and collagen were differentially expressed between one week and 8 weeks
625 of age. Genes exhibiting greater than 2-fold up-regulation included *COL1A1* and *COL1A2*
626 (Table 3). Genes associated with keratinocytes and the epithelial signature of the rumen
627 exhibited differential expression patterns according to developmental stage. *KRT36*, for
628 example, showed a more than 4-fold increase in expression between one and 8 weeks of age
629 (Table 3). *KRT36* has been shown to exhibit significant transcriptional responses to changes
630 in diet in dairy cattle (Li *et al.* 2015) and the change in diet between one and 8 weeks of age
631 is likely to be driving at least a proportion of the observed expression patterns.
632 Similarly, the mitochondrial enzyme encoding gene *HMGCS2*, which belongs to the
633 HMG-Co synthase family and catalyses the first reaction in ketogenesis (Lane *et al.* 2002) was
634 8-fold up-regulated in the rumen between birth and one week of age (Table 3). *HMGCS2* has
635 also been shown to be differentially expressed in the calf rumen during the transition from
636 pre-ruminant to ruminant (Connor *et al.* 2013; Kato *et al.* 2016) and in the rumen of
637 developing lambs (Lane *et al.* 2002; Wang *et al.* 2016). It has been suggested that dietary
638 changes through early development are likely to promote ketogenesis in rumen epithelial
639 cells via PPAR-α-mediated activation of *HMGCS2* to promote papillary development, as well
640 as activation of genes promoting fatty acid beta-oxidation to support cellular differentiation
641 (Connor *et al.* 2013). *SLC14A1* has also been shown to be differentially expressed in the calf
642 rumen during the transition from pre-ruminant to ruminant (Connor *et al.* 2013) and was
643 differentially expressed in the rumen of lambs between birth and one week of age (Wang *et*
644 *al.* 2016). It encodes the protein SLC14A1 which mediates the basolateral cell membrane

transport of urea, a key process in nitrogen secretion into the ruminant GI tract (Abdoun *et al.* 2009). A characteristic of ruminants is a high level of nitrogen recycling in the GI tract. Recycling of nitrogen via urea secretion into the rumen allows ruminants to survive on low-protein diets whilst at the same time producing milk and meat for human consumption (Abdoun *et al.* 2009).

Large differences in developmental stage specific expression patterns are associated with the immune response

Several genes involved in the immune response were differentially expressed in the rumen, ileum and abomasum of developing lambs during the transition from pre-ruminant to ruminant. Many of these genes are likely to be part of the acute phase immune response, by regulating production of key cytokines such as *IL-6* and thus mediating activation of the NF-κB signaling pathways. Nuclear factor (NF)-κB and inhibitor of NF-κB kinase (IKK) proteins regulate innate- and adaptive-immune responses and inflammation (Perkins 2007). *IL36A* and *IL36B* are thought to influence the skin inflammatory response by acting directly on keratinocytes and macrophages and indirectly on T-lymphocytes to drive tissue infiltration, cell maturation and cell proliferation (Foster *et al.* 2014). *IL36A* is a cytokine that can activate the NF-kappa-B and MAPK signaling pathways to generate an inflammatory response and shows almost a 4-fold increase in expression between one and 8 weeks of age in the rumen. Other C-type lectins and genes involved in cytokine signalling also showed differential expression between the different developmental stages. *CLEC4E*, for example, was more than 2-fold upregulated in the Ileum between one and 8 weeks of age. *JAK3*, which encodes a member of the Janus kinase (JAK) family of tyrosine kinases is involved in cytokine receptor-

668 mediated intracellular signal transduction (Yeh & Pellegrini 1999), and was more than 5-fold
669 upregulated in the rumen between birth and one week of age.

670 Other immune genes, including *PIP* (Prolactin Induced Protein), were upregulated 3-
671 fold in the rumen of lambs between one and 8 weeks of age. *PIP* has been shown to be
672 preferentially expressed in the rumen of adult sheep (Xiang *et al.* 2016a), and is thought to
673 play a role in mucosal immunity in ruminants (Hassan *et al.* 2008). Multiple IFN-inducible
674 genes, including *IFIT2*, *IFIT3*, *MX1*, *MX2* and *ISG15*, were differentially expressed in the
675 abomasum, rumen and ileum during the transition from pre-ruminant to ruminant (Table 3).
676 Three of these genes - *IFIT2*, *IFIT3* and *MX1* - have recently been shown to be differentially
677 expressed in sheep fibroblast cells in a type I IFN-induced antiviral state (Shaw *et al.* 2017).
678 The expression pattern of some of these genes varied through development. *MX2* for
679 example, is up-regulated in the rumen between birth and one-week of age and down-
680 regulated between one and eight weeks of age (Table 3). The differential expression patterns
681 of these genes throughout the development of the rumen, abomasum and ileum highlights
682 both their importance in the innate immune response and the role of the GI tract as an
683 immunologically active site.

684 Several genes involved in both metabolism and immunity were differentially
685 expressed during the transition from pre-ruminant to ruminant. *DUOXA2*, for example is
686 involved in thyroid hormone synthesis and lactoperoxidase-mediated antimicrobial defense
687 at the surface of mucosa (Bae *et al.* 2010). The rumen is the main site of colonisation by micro-
688 organisms as the lamb develops. *DUOX2* and *DUOXA2*, which encode subunits of dual oxidase
689 have previously been shown to be upregulated in the rumen of adult sheep (Xiang *et al.*
690 2016a; Xiang *et al.* 2016b) and here we found *e*xpression of *DUOXA2* 4-fold up-regulated in
691 the rumen between birth and one week of age. *DUOXA2* might be involved in controlling

692 microbial colonization as the lamb transitions from pre-ruminant to ruminant. Similarly, *IDO1*
693 was more than 4-fold upregulated in the ileum of lambs between birth and one week of age
694 and has been implicated in immune modulation through its ability to limit T-cell function and
695 engage mechanisms of immune tolerance (Grohmann *et al.* 2003; Plain *et al.* 2011). *IDO1*
696 encodes indoleamine 2,3-dioxygenase (IDO) which is a heme enzyme that catalyzes the first
697 and rate-limiting step in tryptophan catabolism to N-formyl-kynurenine (Grohmann *et al.*
698 2003). Through its expression in monocytes and macrophages this enzyme modulates T-cell
699 behaviour by its peri-cellular catabolization of the essential amino acid tryptophan (Munn &
700 Mellor 2013). It has also been shown to be highly expressed in the jejunal mucosa of pre-
701 weaning calves (Hammon *et al.* 2018).
702
703 **Comparative analysis of the rumen, colon and ileum of one week old age-matched sheep**
704 **and goats reveals differences in expression of genes involved in metabolism and immunity**
705 We performed a comparative analysis of the gene expression estimates for age-
706 matched one week old sheep and goats for three GI tract tissues: rumen, ileum and colon.
707 The goat gene expression estimates as transcripts per million (TPM) both for these tissues
708 and alveolar macrophages are included in Table S9. Full lists of all differentially expressed
709 genes between sheep and goat are included in Table S10 (rumen), Table S11 (ileum) and Table
710 S12 (colon). The top 25 differentially expressed genes between sheep and goat in either
711 direction for each of the three tissues are shown in Figure 7. Differentially expressed genes
712 between the rumen of sheep and goats included several solute carrier genes (*SLC5A8*, *SLC9A4,*
713 *SLC14A1*, and *SLC27A6*), a myosin V heavy chain gene *MYO5A* associated with connective
714 tissue, and *KRT3* which is involved in the simple and stratified differentiation of epithelial
715 tissues. Other genes within these super-families were also differentially expressed between

716 the ileum and colon of sheep and goats. Similarly, genes associated with immunity and
717 metabolism were differentially expressed between sheep and goat GI tract tissues. For
718 example, C-C motif chemokine ligands *CCL5* and *CCL20* were upregulated, respectively, in the
719 goat ileum and colon. Chemokines form a superfamily of genes involved in immunoregulatory
720 and inflammatory processes (Griffith *et al.* 2014). *CCL5* is one of the predominant cytokines
721 expressed during damage and inflammation of epithelial keratinocytes (Wetzler *et al.* 2000).
722 In contrast, *CSF3R* was down regulated in the colon of goats relative to sheep (Figure 7).
723 *CSF3R* (also known as *GCSF*) encodes a protein that controls the production, differentiation
724 and production of granulocytes (Nagata & Fukunaga 1991). Differences in expression
725 between the two species might therefore reflect variation in the density of granulocyte cells.
726 *CSF3R* has been shown to undergo changes in anatomical distribution during human fetal
727 development in the intestine (Calhoun *et al.* 1999), and similar trends could also occur in
728 neonatal animals.

729 Other genes with key roles in metabolic pathways and immune regulation were
730 differentially expressed between goats and sheep, including *IDO2*, which was upregulated in
731 the colon of one-week old goats relative to age-matched sheep. *IDO2*, with its paralogue
732 *IDO1,* mentioned above, catalyses the first and rate-limiting step of the catabolism of the
733 essential amino acid tryptophan along the kynurenine pathway (Grohmann *et al.* 2003) and
734 is involved in the metabolic control of immune responses (Munn & Mellor 2013). IDO has
735 been implicated in downregulating immune responses to *Mycobacterium*
736 *avium* subsp. *paratuberculosis*, the causative agent of Johne's disease (Plain *et al.* 2011). In
737 sheep and cattle an increase in IDO expression correlates with progression to clinical
738 mycobacterial disease (Plain *et al.* 2011) with the small intestine the primary site of infection
739 (Koets *et al.* 2015). Since the GI tract contains important regions related to immunity

differences in the expression of genes with immune and metabolic function in the tissues analysed might underlie differences in disease susceptibility and response to pathogens between the two species. Interestingly, *SLC14A1* which was differentially expressed in the rumen of sheep between birth and one week of age, was also up-regulated in the rumen of goats relative to sheep (Figure 7). As mentioned above the protein encoded by *SLC14A1* mediates the basolateral cell membrane transport of urea, a key process in nitrogen secretion into the ruminant GI tract (Abdoun *et al.* 2009). Nitrogen recycling is an important topic in ruminant management because movements of nitrogen across the GI tract can more than double nitrogen intake and have major effects on nitrogen metabolism (Lapierre & Lobley 2001). Nitrogenous content of urea between sheep and goats has been shown to vary (Bristow Andrew *et al.* 1992), which may be influenced by species specific differences in expression of key genes involved in iodine and nitrogen recycling pathways in ruminants.

The results from our differential expression analysis of the three GI tract tissues in age-matched sheep and goats indicate that there are basal differences in the expression of some key genes involved in pathways related to immunity and metabolism between the two species. These genes would be candidates for further analysis to elucidate the functional significance of species specific differences in gene expression e.g. (Young *et al.* 2018).

Conclusions

By characterising tissue specific transcription in the GI tract through the transition from pre-ruminant to ruminant we have shown that there are significant developmental stage specific differences in gene expression particularly between neonatal lambs and 8 week and adult sheep. These differences were most obvious in the rumen complex where significant morphological and physiological changes occur as the lamb transitions from a milk-based to

764 a grass and pellet diet. Differences in the expression of protein coding genes with age were
765 observed both when the whole transcriptome was included in the analysis, and also when
766 only a subset of macrophage specific genes were analysed.

767 We focused only on protein-coding genes in this study, and have presented a wider
768 characterisation of non-coding transcripts elsewhere (Bush *et al.* 2018). A low level, highly
769 tissue specific expression pattern is characteristic of lncRNAs in sheep and goats (Bush *et al.*
770 2018). Further characterisation of the lncRNAs and other non-coding transcripts expressed
771 specifically in the GI tract through development, might help to infer something about their
772 function and would be an interesting direction for future work. Similarly, analysis of isoform
773 regulation using RNA-Seq data, as in (Katz *et al.* 2010), in the GI tract tissues would also be
774 interesting, although both this and analysis of the non-coding transcriptome would require
775 further sequencing to generate high depth total RNA-Seq libraries.

776 Other studies have examined links between feeding regime, the host transcriptome
777 and bacterial diversity in sheep using sequencing analysis of 16S rRNA genes (Wang *et al.*
778 2016). A full characterisation linking tissue- and developmental stage-specific microbial
779 colonisation of the GI tract would require a detailed 16S sequencing metagenomic approach
780 e.g. (Wallace *et al.* 2015). Further work would ideally examine this link further to profile in
781 parallel transcription in the GI tract and microbial colonisation to provide information as to
782 how microbial colonisation influences transcription during development. An additional area
783 to explore further would be the relative *IL10* insufficiency in the GI tract of the lambs – which
784 suppresses the macrophage response to intestinal microbes – as this could potentially be
785 beneficial to ruminant species which, compared to monogastric mammals, have a quite
786 different relationship with gut microbes.

The results we present in this study lay a foundation for further work, providing baseline estimates of gene expression in the GI tract at the whole tissue level from healthy lambs. In their 2002 review of gene expression in the ruminant GI tract Connor *et al.* indicated that the use of techniques such as laser-capture microdissection will be needed to further characterize expression profiles of individual cell types within the GI tract, and to remove expression biases that may occur in studies evaluating whole tissue samples (Connor *et al.* 2010). Cutting edge single cell RNA-Seq technology provides the solution to this, allowing a cell specific level of resolution of gene expression profiles. Single cell messenger RNA-Seq has already been applied to cells from mouse GI tract organoids revealing rare cell types (Grün *et al.* 2015) and the technology will hopefully be applied in the future to cells from the GI tract of ruminants, since *in vitro* systems are now available (Hamilton *et al.* 2018).

One of the most significant physiological challenges to neonatal and juvenile ruminants is the development and establishment of the rumen. The transition from pre-ruminant to ruminant involves not only growth of the rumen and cellular differentiation but also has a major effect on metabolism, immunity and physiological processes in other GI tract tissues (Baldwin *et al.* 2004), which is reflected in the extensive transcriptional complexity observed in this study. Using this sub-set of RNA-Seq data, from our high resolution of atlas of gene expression in sheep (Clark *et al.* 2017b) we have improved our understanding of the genetic and genomic mechanisms involved in the transition between pre-ruminant and ruminant in sheep and highlighted key genes underlying healthy growth and development that could be utilised to improve productivity in sheep and other ruminants.

Author's Contributions

ELC coordinated and designed the study with assistance from MEBM, ALA and DAH. DAH acquired the funding with ALA. MEBM, ILF and ELC performed sample collection from sheep. CM, ELC and ZML performed sample collection from goats. GMD and CM performed the RNA extractions. ELC performed the network cluster and gene expression analysis. MS performed the Principal Component Analysis and created Figure 1. CM performed the comparative analysis of sheep and goat. DAH and ZML provided guidance on interpretation of the gut macrophage transcriptional signatures. SJB performed all bioinformatic analyses. ELC wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the farm staff at Dryden farm and members of the sheep tissue collection team from The Roslin Institute and R(D)SVS who were involved in tissue collections for the sheep and goat gene expression atlas projects. The authors are also grateful for the support of the FAANG Data Coordination Centre in the upload and archiving of the sample data and metadata and BioGPS for hosting the sheep atlas dataset on their annotation portal.

Funding

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC; www.bbsrc.ac.uk) grant BB/L001209/1 ('Functional Annotation of the Sheep Genome') and Institute Strategic Program grants 'Farm Animal Genomics' (BBS/E/D/2021550), 'Blueprints for Healthy Animals' (BB/P013732/1) and 'Transcriptomes, Networks and Systems' (BBS/E/D/20211552). The goat RNA-Seqdata was funded by the Roslin Foundation, which also supported SJB. CM was supported by a Newton Fund Ph.D.

832 studentship. Zofia Lisowski was supported by a PhD studentship from the Horse Racing and
833 Betting Levy Board. Edinburgh Genomics is partly supported through core grants from the
834 BBSRC (BB/J004243/1), National Research Council
835 (NERC; www.nationalacademies.org.uk/nrc) (R8/H10/56), and Medical Research Council
836 (MRC; www.mrc.ac.uk) (MR/K001744/1). The funders had no role in study design, data
837 collection and analysis, decision to publish, or preparation of the manuscript.

838 **Literature Cited**

839

840 Abdoun K., Stumpff F., Rabbani I. & Martens H. (2009) Modulation of urea transport across
841 sheep rumen epithelium in vitro by SCFA and CO2. *Am J Physiol Gastrointest Liver Physiol* **298**,
842 G190-G202.

843 Akira S. & Takeda K. (2004) Toll-like receptor signalling. *Nat Rev Immunol* **4**, 499-511.

844 Alexa A. & Rahnenfuhrer J. (2010) topGO: Enrichment analysis for Gene Ontology. URL
845 http://www.bioconductor.org/packages/release/bioc/html/topGO.html.

846 Andersson L., Archibald A.L., Bottema C.D., Brauning R., Burgess S.C., Burt D.W., Casas E.,
847 Cheng H.H., Clarke L., Couldrey C., Dalrymple B.P., Elsik C.G., Foissac S., Giuffra E., Groenen
848 M.A., Hayes B.J., Huang L.S., Khatib H., Kijas J.W., Kim H., Lunney J.K., McCarthy F.M., McEwan
849 J.C., Moore S., Nanduri B., Notredame C., Palti Y., Plastow G.S., Reecy J.M., Rohrer G.A.,
850 Sarropoulou E., Schmidt C.J., Silverstein J., Tellam R.L., Tixier-Boichard M., Tosser-Klopp G.,
851 Tuggle C.K., Vilkki J., White S.N., Zhao S. & Zhou H. (2015) Coordinated international action to
852 accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes
853 project. *Genome Biol* **16**, 57.

854 Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K.,
855 Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C.,

856 Richardson J.E., Ringwald M., Rubin G.M. & Sherlock G. (2000) Gene ontology: tool for the
857 unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**.
858 Bae Y.S., Choi M.K. & Lee W.-J. (2010) Dual oxidase in mucosal immunity and host–microbe
859 homeostasis. *Trends Immunol* **31**, 278-87.
860 Bain C.C., Hawley C.A., Garner H., Scott C.L., Schridde A., Steers N.J., Mack M., Joshi A.,
861 Guilliams M., Mowat A.M.I., Geissmann F. & Jenkins S.J. (2016) Long-lived self-renewing bone
862 marrow-derived macrophages displace embryo-derived cells to inhabit adult serous cavities.
863 *Nature Commun* **7**, 11852.
864 Bain C.C. & Mowat A.M. (2011) Intestinal macrophages – specialised adaptation to a unique
865 environment. *Eur Journal Immunol* **41**, 2494-8.
866 Baldwin R.L.V., McLeod K.R., Klotz J.L. & Heitmann R.N. (2004) Rumen Development,
867 Intestinal Growth and Hepatic Metabolism In The Pre- and Postweaning Ruminant. *J Dairy Sci*
868 **87**, E55-E65.
869 Behler-Janbeck F., Takano T., Maus R., Stolper J., Jonigk D., Tort Tarrés M., Fuehner T., Prasse
870 A., Welte T., Timmer M.S.M., Stocker B.L., Nakanishi Y., Miyamoto T., Yamasaki S. & Maus
871 U.A. (2016) C-type Lectin Mincle Recognizes Glucosyl-diacylglycerol of Streptococcus
872 pneumoniae and Plays a Protective Role in Pneumococcal Pneumonia. *PLoS Pathog* **12**,
873 e1006038.
874 Bickhart D.M., Rosen B.D., Koren S., Sayre B.L., Hastie A.R., Chan S., Lee J., Lam E.T., Liachko
875 I., Sullivan S.T., Burton J.N., Huson H.J., Nystrom J.C., Kelley C.M., Hutchison J.L., Zhou Y., Sun
876 J., Crisa A., Ponce de Leon F.A., Schwartz J.C., Hammond J.A., Waldbieser G.C., Schroeder S.G.,
877 Liu G.E., Dunham M.J., Shendure J., Sonstegard T.S., Phillippy A.M., Van Tassell C.P. & Smith
878 T.P.L. (2017) Single-molecule sequencing and chromatin conformation capture enable de
879 novo reference assembly of the domestic goat genome. *Nat Genet* **49**, 643-50.

880 Bray N.L., Pimentel H., Melsted P. & Pachter L. (2016) Near-optimal probabilistic RNA-seq
881 quantification. *Nat Biotech* **34**.
882 Bristow Andrew W., Whitehead David C. & Cockburn John E. (1992) Nitrogenous constituents
883 in the urine of cattle, sheep and goats. *J Sci Food Agric* **59**, 387-94.
884 Bryant M.P., Small N., Bouma C. & Robinson I. (1958) Studies on the Composition of the
885 Ruminal Flora and Fauna of Young Calves. *J Dairy Sci* **41**, 1747-67.
886 Bush S.J., McCulloch M.E.B., Summers K.M., Hume D.A. & Clark E.L. (2017) Integration of
887 quantitated expression estimates from polyA-selected and rRNA-depleted RNA-seq libraries.
888 *BMC Bioinformatics* **18**, 301.
889 Bush S.J., Muriuki C., McCulloch M.E.B., Farquhar I.L., Clark E.L. & Hume D.A. (2018) Cross-
890 species inference of long non-coding RNAs greatly expands the ruminant transcriptome. *GSE*
891 **50**, 20.
892 Calhoun D.A., Donnelly Jr W.H., Du Y., Dame J.B., Li Y. & Christensen R.D. (1999) Distribution
893 of Granulocyte Colony-Stimulating Factor (G-CSF) and G-CSF-Receptor mRNA and Protein in
894 the Human Fetus. *Pediatric Res* **46**, 333.
895 Chao T., Wang G., Ji Z., Liu Z., Hou L., Wang J. & Wang J. (2017) Transcriptome Analysis of
896 Three Sheep Intestinal Regions reveals Key Pathways and Hub Regulatory Genes of Large
897 Intestinal Lipid Metabolism. *Sci Rep* **7**, 5345.
898 Clark E.L., Bush S.J., McCulloch M.E.B., Farquhar I.L., Lefevre L., Young R., Tsang H., Whitelaw
899 C.B., Watson M., Summers K.M., Archibald A.L. & Hume D.A. (2017a) The Roslin Sheep Gene
900 Expression Atlas Dataset on BioGPS. URL http://biogps.org/dataset/BDS_00015/sheep-atlas/
901 Clark E.L., Bush S.J., McCulloch M.E.B., Farquhar I.L., Young R., Lefevre L., Pridans C., Tsang H.,
902 Wu C., Afrasiabi C., Watson M., Whitelaw C.B., Freeman T.C., Summers K.M., Archibald A.L. &

903 Hume D.A. (2017b) A high resolution atlas of gene expression in the domestic sheep (Ovis
904 aries). *PLoS Genet* **13**, e1006997.
905 Clarke N.F., Ilkovski B., Cooper S., Valova V.A., Robinson P.J., Nonaka I., Feng J.J., Marston S.,
906 and North K. (2007) The pathogenesis of ACTA-1 related congenital fibre type disproportion.
907 Ann Neurol **61,** 552-561.
908 Connor E.E., Baldwin R.L., Li C.-j., Li R.W. & Chung H. (2013) Gene expression in bovine rumen
909 epithelium during weaning identifies molecular regulators of rumen development and
910 growth. *Funct Integr Genomics* **13**, 133-42.
911 Connor E.E., Li R.W., Baldwin R.L. & Li C. (2010) Gene expression in the digestive tissues of
912 ruminants and their relationships with feeding and digestive processes. *Animal* **4**, 993-1007.
913 David C.W., Norrman J., Hammon H.M., Davis W.C. & Blum J.W. (2003) Cell Proliferation,
914 Apoptosis, and B- and T-Lymphocytes in Peyer's Patches of the Ileum, in Thymus and in Lymph
915 nodes of Preterm Calves, and in Full-Term Calves at Birth and on Day 5 of Life. *J Dairy Sci* **86**,
916 3321-9.
917 Dyce K., Sack W. & Wensing C. (2010) *Textbook of Veterinary Anatomy*. Saunders.
918 FAANG (2017) FAANG Data Portal. URL http://data.faang.org/home.
919 Fabriek B.O., van Bruggen R., Deng D.M., Ligtenberg A.J.M., Nazmi K., Schornagel K., Vloet
920 R.P.M., Dijkstra C.D. & van den Berg T.K. (2009) The macrophage scavenger receptor CD163
921 functions as an innate immune sensor for bacteria. *Blood* **113**, 887.
922 Foster A.M., Baliwag J., Chen C.S., Guzman A.M., Stoll S.W., Gudjonsson J.E., Ward N.L. &
923 Johnston A. (2014) IL-36 Promotes Myeloid Cell Infiltration, Activation, and Inflammatory
924 Activity in Skin. *J Immunol* **192**, 6053.
925 Freeman T.C., Ivens A., Baillie J.K., Beraldi D., Barnett M.W., Dorward D., Downing A., Fairbairn
926 L., Kapetanovic R., Raza S., Tomoiu A., Alberio R., Wu C., Su A.I., Summers K.M., Tuggle C.K.,

927 Archibald A.L. & Hume D.A. (2012) A gene expression atlas of the domestic pig. *BMC Biology*
928 **10**, 90.

929 Gordon S., Plüddemann A. & Martinez Estrada F. (2014) Macrophage heterogeneity in tissues:
930 phenotypic diversity and functions. *Immunol Rev* **262**, 36-55.

931 Gough J., Karplus K., Hughey R. & Chothia C. (2001) Assignment of homology to genome
932 sequences using a library of hidden Markov models that represent all proteins of known
933 structure. *J Mol Biol* **313**, 903-19.

934 Griffith J.W., Sokol C.L. & Luster A.D. (2014) Chemokines and Chemokine Receptors:
935 Positioning Cells for Host Defense and Immunity. *Ann Rev Immunol* **32**, 659-702.

936 Grohmann U., Fallarino F. & Puccetti P. (2003) Tolerance, DCs and tryptophan: much ado
937 about IDO. *Trends Immunol* **24**, 242-8.

938 Grün D., Lyubimova A., Kester L., Wiebrands K., Basak O., Sasaki N., Clevers H. & van
939 Oudenaarden A. (2015) Single-cell messenger RNA sequencing reveals rare intestinal cell
940 types. *Nature* **525**, 251.

941 Hamilton C., Young R., Jayaraman S., Sehjal A., Paxton E., Thomson S., Katzer F., Hope J., Innes
942 E., Morrison L.J. and Mabbott N.A. (2018) Development of in vitro enteroids derived from
943 bovine small intestinal crypts. *BMC Vet Res* **49**, 54.

944 Hammon H.M., Frieten D., Gerbert C., Koch C., Dusel G., Weikard R. & Kühn C. (2018) Different
945 milk diets have substantial effects on the jejunal mucosal immune system of pre-weaning
946 calves, as demonstrated by whole transcriptome sequencing. *Sci Rep* **8**, 1693.

947 Hassan M.I., Waheed A., Yadav S., Singh T.P. & Ahmad F. (2008) Prolactin inducible protein in
948 cancer, fertility and immunoregulation: structure, function and its clinical implications. *Cell*
949 *Mol Life Sci* **66**, 447.

950 Huang R., Grabarek Z. & Wang C.L. (2010) Differential effects on caldeson on the intermediate
951 conformational states on polymerizing actin. *J Biol Chem* **285**, 71-79.
952 Jiang Y., Xie M., Chen W., Talbot R., Maddox J.F., Faraut T., Wu C., Muzny D.M., Li Y., Zhang
953 W., Stanton J.-A., Brauning R., Barris W.C., Hourlier T., Aken B.L., Searle S.M.J., Adelson D.L.,
954 Bian C., Cam G.R., Chen Y., Cheng S., DeSilva U., Dixen K., Dong Y., Fan G., Franklin I.R., Fu S.,
955 Fuentes-Utrilla P., Guan R., Highland M.A., Holder M.E., Huang G., Ingham A.B., Jhangiani S.N.,
956 Kalra D., Kovar C.L., Lee S.L., Liu W., Liu X., Lu C., Lv T., Mathew T., McWilliam S., Menzies M.,
957 Pan S., Robelin D., Servin B., Townley D., Wang W., Wei B., White S.N., Yang X., Ye C., Yue Y.,
958 Zeng P., Zhou Q., Hansen J.B., Kristiansen K., Gibbs R.A., Flicek P., Warkup C.C., Jones H.E.,
959 Oddy V.H., Nicholas F.W., McEwan J.C., Kijas J.W., Wang J., Worley K.C., Archibald A.L.,
960 Cockett N., Xu X., Wang W. & Dalrymple B.P. (2014) The sheep genome illuminates biology of
961 the rumen and lipid metabolism. *Science* **344**, 1168.
962 Jung C., Hugot J.-P. & Barreau F. (2010) Peyer's Patches: The Immune Sensors of the Intestine.
963 *Int J Inflam* **2010**, 823710.
964 Kato D., Suzuki Y., Haga S., So K., Yamauchi E., Nakano M., Ishizaki H., Choi K., Katoh K. & Roh
965 S.-G. (2016) Utilization of digital differential display to identify differentially expressed genes
966 related to rumen development. J *Anim Sci* **87**, 584-90.
967 Katz Y., Wang E.T., Airoldi E.M. & Burge C.B. (2010) Analysis and design of RNA sequencing
968 experiments for identifying isoform regulation. *Nat Methods* **7**, 1009-15.
969 Koets A.P., Eda S. & Sreevatsan S. (2015) The within host dynamics of Mycobacterium avium
970 ssp. paratuberculosis infection in cattle: where time and place matter. *Vet Res* **46**, 61.
971 Kramer B.W., Jobe A.H. & Ikegami M. (2003) Monocyte Function in Preterm, Term, and Adult
972 Sheep. *Pediatric Res* **54**, 52.

973 Lane M.A., Baldwin R.L. & Jesse B.W. (2002) Developmental changes in ketogenic enzyme
974 gene expression during sheep rumen development. *J Anim Sci* **80**, 1538-44.
975 Lapierre H. & Lobley G.E. (2001) Nitrogen Recycling in the Ruminant: A Review. *J Dairy Sci* **84**,
976 E223-E36.
977 Lê S., Josse J. & F. H. (2008) FactoMineR : An R Package for Multivariate Analysis. *J Stat Softw*
978 **25**, 253-8.
979 Li Y., Carrillo J.A., Ding Y., He Y., Zhao C., Zan L. & Song J. (2015) Ruminal Transcriptomic
980 Analysis of Grass-Fed and Grain-Fed Angus Beef Cattle. *PLoS One* **10**, e0116437.
981 Livigni A., O'Hara L., Polak M.E., Angus T., Wright D.W., Smith L.B. & Freeman T.C. (2018) A
982 graphical and computational modeling platform for biological pathways. *Nat Protoc* **13**, 705.
983 Love M.I., Soneson C. & Robinson M.D. (2017) Importing transcript abundance datasets with
984 tximport.
985 https://bioconductor.org/packages/release/bioc/vignettes/tximport/inst/doc/tximport.html
986 Mantovani A. & Marchesi F. (2014) IL-10 and Macrophages Orchestrate Gut Homeostasis.
987 *Immunity* **40**, 637-9.
988 Meale S.J., Chaucheyras-Durand F., Berends H., Guan L.L. & Steele M.A. (2017) From pre- to
989 postweaning: Transformation of the young calf's gastrointestinal tract. *J Dairy Sci* **100**, 5984-
990 95.
991 Moestrup S.K. & Moller H.J. (2004) CD163: a regulated hemoglobin scavenger receptor with
992 a role in the anti-inflammatory response. *Ann Med* **36**, 347-54.
993 Munn D.H. & Mellor A.L. (2013) Indoleamine 2,3 dioxygenase and metabolic control of
994 immune responses. *Trends Immunol* **34**, 137-43.
995 Nagata S. & Fukunaga R. (1991) Granulocyte colony-stimulating factor and its receptor. *Prog*
996 *Growth Factor Res* **3**, 131-41.

997 O'Donnell M.M., Harris H.M.B., Ross R.P. & O'Toole P.W. (2017) Core fecal microbiota of
998 domesticated herbivorous ruminant, hindgut fermenters, and monogastric animals. *Open*
999 *Microbiol J* **6**, e00509.
1000 Oliver S. (2000) Proteomics: Guilt-by-association goes global. *Nature* **403**, 601-3.
1001 Oliveros J.C. (2007) Venny. An interactive tool for comparing lists with Venn's diagrams.
1002 http://bioinfogp.cnb.csic.es/tools/venny/index.html.
1003 Perkins N.D. (2007) Integrating cell-signalling pathways with NF-[kappa]B and IKK function.
1004 *Nat Rev Mol Cell Biol* **8**, 49-62.
1005 Plain K.M., de Silva K., Earl J., Begg D.J., Purdie A.C. & Whittington R.J. (2011) Indoleamine 2,3-
1006 Dioxygenase, Tryptophan Catabolism, and Mycobacterium avium subsp. paratuberculosis: a
1007 Model for Chronic Mycobacterial Infections. *Infect Immun* **79**, 3821-32.
1008 R Core Team (2014) R: A Language and Environment for Statistical Computing. URL
1009 http://www.R-project.org.
1010 Robinson M.D., McCarthy D.J. & Smyth G.K. (2010) edgeR: a Bioconductor package for
1011 differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-40.
1012 Saijo S. & Iwakura Y. (2011) Dectin-1 and Dectin-2 in innate immunity against fungi. *Int*
1013 *Immunol* **23**, 467-72.
1014 Schurch N.J., Schofield P., Gierliński M., Cole C., Sherstnev A., Singh V., Wrobel N., Gharbi K.,
1015 Simpson G.G., Owen-Hughes T., Blaxter M. & Barton G.J. (2016) How many biological
1016 replicates are needed in an RNA-seq experiment and which differential expression tool should
1017 you use? *RNA* **22**, 839-51.
1018 Shaw A.E., Hughes J., Gu Q., Behdenna A., Singer J.B., Dennis T., Orton R.J., Varela M., Gifford
1019 R.J., Wilson S.J. & Palmarini M. (2017) Fundamental properties of the mammalian innate

1020 immune system revealed by multispecies comparison of type I interferon responses. *PLoS*
1021 *Biology* **15**, e2004086.
1022 Shaw T.N., Houston S.A., Wemyss K., Bridgeman H.M., Barbera T.A., Zangerle-Murray T.,
1023 Strangward P., Ridley A.J.L., Wang P., Tamoutounour S., Allen J.E., Konkel J.E. & Grainger J.R.
1024 (2018) Tissue-resident macrophages in the intestine are long lived and defined by Tim-4 and
1025 CD4 expression. *J Exp Med,* 215(6):1507-1518.
1026 Shouval D.S., Ouahed J., Biswas A., Goettel J.A., Horwitz B.H., Klein C., Muise A.M. & Snapper
1027 S.B. (2014) Interleukin 10 Receptor Signaling: Master Regulator of Intestinal Mucosal
1028 Homeostasis in Mice and Humans. *Adv Immunol* **122**, 177-210.
1029 Soneson C., Love M.I. & Robinson M.D. (2015) Differential analyses for RNA-seq: transcript-
1030 level estimates improve gene-level inferences [version 1; referees: 2 approved].
1031 *F1000Research* **4**.
1032 Steele M.A., Penner G.B., Chaucheyras-Durand F. & Guan L.L. (2016) Development and
1033 physiology of the rumen and the lower gut: Targets for improving gut health. *J Dairy Sci* **99**,
1034 4955-66.
1035 Stelwagen K., Carpenter E., Haigh B., Hodgkinson A. & Wheeler T.T. (2009) Immune
1036 components of bovine colostrum and milk1. *J Anim Sci* **87**, 3-9.
1037 Takahashi-Iwanaga H. & Fujita T. (1985) Lamina propria of intestinal mucosa as a typical
1038 reticular tissue. A scanning electron-microscopic study of the rat jejunum. *Cell Tissue Res*
1039 242(1), 57-66.
1040 Theocharidis A., van Dongen S., Enright A.J. & Freeman T.C. (2009) Network visualisation and
1041 analysis of gene expression data using BioLayout Express3D. *Nat Protocols* **4**.
1042 Treanor B. (2012) B-cell receptor: from resting state to activate. *Immunology* **136**, 21-7.

1043 van Dongen S. & Abreu-Goodger C. (2012) Using MCL to extract clusters from networks.
1044 *Methods Mol Biol* **804**, 281-95.
1045 Wallace R.J., Rooke J.A., McKain N., Duthie C.-A., Hyslop J.J., Ross D.W., Waterhouse A.,
1046 Watson M. & Roehe R. (2015) The rumen microbial metagenome associated with high
1047 methane production in cattle. *BMC Genomics* **16**, 839.
1048 Wang W., Li C., Li F., Wang X., Zhang X., Liu T., Nian F., Yue X., Li F., Pan X., La Y., Mo F., Wang
1049 F. & Li B. (2016) Effects of early feeding on the host rumen transcriptome and bacterial
1050 diversity in lambs. *Sci Rep* **6**, 32479.
1051 Wetzler C., Kämpfer H., Pfeilschifter J. & Frank S. (2000) Keratinocyte-Derived Chemotactic
1052 Cytokines: Expressional Modulation by Nitric Oxide in Vitro and during Cutaneous Wound
1053 Repair in Vivo. *Biochem Biophys Res Commun* **274**, 689-96.
1054 Worley K.C. (2017) A golden goat genome. *Nat Genet* **49**, 485-6.
1055 Xiang R., McNally J., Rowe S., Jonker A., Pinares-Patino C.S., Oddy V.H., Vercoe P.E., McEwan
1056 J.C. & Dalrymple B.P. (2016a) Gene network analysis identifies rumen epithelial cell
1057 proliferation, differentiation and metabolic pathways perturbed by diet and correlated with
1058 methane production. *Sci Rep* **6**, 39022.
1059 Xiang R., Oddy V.H., Archibald A.L., Vercoe P.E. & Dalrymple B.P. (2016b) Epithelial, metabolic
1060 and innate immunity transcriptomic signatures differentiating the rumen from other sheep
1061 and mammalian gastrointestinal tract tissues. *PeerJ* **4**, e1762.
1062 Yeh T.C. & Pellegrini S. (1999) The Janus kinase family of protein tyrosine kinases and their
1063 role in signaling. *Cell Mol Life Sci CMLS* **55**, 1523-34.
1064 Young R., Bush S.J., Lefevre L., McCulloch M.E.B., Lisowski Z.M., Muriuki C., Waddell L.A.,
1065 Sauter K.A., Pridans C., Clark E.L. & Hume D.A. (2018) Species-Specific Transcriptional

1066 Regulation of Genes Involved in Nitric Oxide Production and Arginine Metabolism in
1067 Macrophages. *ImmunoHorizons* **2**, 27.
1068 Zerbino D.R., Achuthan P., Akanni W., Amode M R., Barrell D., Bhai J., Billis K., Cummins C.,
1069 Gall A., Girón C.G., Gil L., Gordon L., Haggerty L., Haskell E., Hourlier T., Izuogu O.G., Janacek
1070 S.H., Juettemann T., To J.K., Laird M.R., Lavidas I., Liu Z., Loveland J.E., Maurel T., McLaren W.,
1071 Moore B., Mudge J., Murphy D.N., Newman V., Nuhn M., Ogeh D., Ong C.K., Parker A., Patricio
1072 M., Riat H.S., Schuilenburg H., Sheppard D., Sparrow H., Taylor K., Thormann A., Vullo A.,
1073 Walts B., Zadissa A., Frankish A., Hunt S.E., Kostadima M., Langridge N., Martin F.J., Muffato
1074 M., Perry E., Ruffier M., Staines D.M., Trevanion S.J., Aken B.L., Cunningham F., Yates A. &
1075 Flicek P. (2018) Ensembl 2018. *Nucleic Acids Res* **46**, D754-D61.
1076

1077 **Table 1: Details of animals and samples from the GI tract included in this study**

Species	Breed	Developmental Stage	Sex	GI Tract Tissues	Reference & ENA Study Accession
Sheep	Texel x Scottish Blackface	Birth (neonate/ pre-ruminant)	1 male, 2 females	**Stomach** Abomasum, Pylorus **Large Intestine** Caecum, Colon Spiral, Colon Large **Small Intestine** Duodenum, Ileum, Jejunum, Peyer's Patch **Rumen Complex** Omasum, Reticulum, Rumen	Clark et al. 2017 (PRJEB19199)
Sheep	Texel x Scottish Blackface	One Week (transition from pre-ruminant to ruminant)	1 male, 2 females		
Sheep	Texel x Scottish Blackface	8 Weeks (ruminant)	2 males, 1 female		
Sheep	Texel x Scottish Blackface	Adult (GI tract)	3 males, 3 females		
Sheep	Texel x Scottish Blackface	Adult (reference tissues)	3 males, 3 females	Oesophageal mucosa, Oesophageal muscle, Rectum, Liver, Alveolar Macrophages	Clark et al. 2017 (PRJEB19199)
Sheep	Texel	6 – 10 Months	1 female	Abomasum, Colon, Caecum, Rectum, Rumen, Omentum, Peyer's Patch	Jiang et al. 2014 (PRJEB6169)
Sheep	Texel	Adult	1 male, 1 female		
Goat	Crossbred	One Week	3 male	Ileum, Rumen, Colon Large, Alveolar Macrophages	This Study (PRJEB23196)

1078

Table 2: Functional annotation of the top 20 clusters for the sheep GI tract dataset including liver, oesophagus and alveolar macrophages as reference samples.

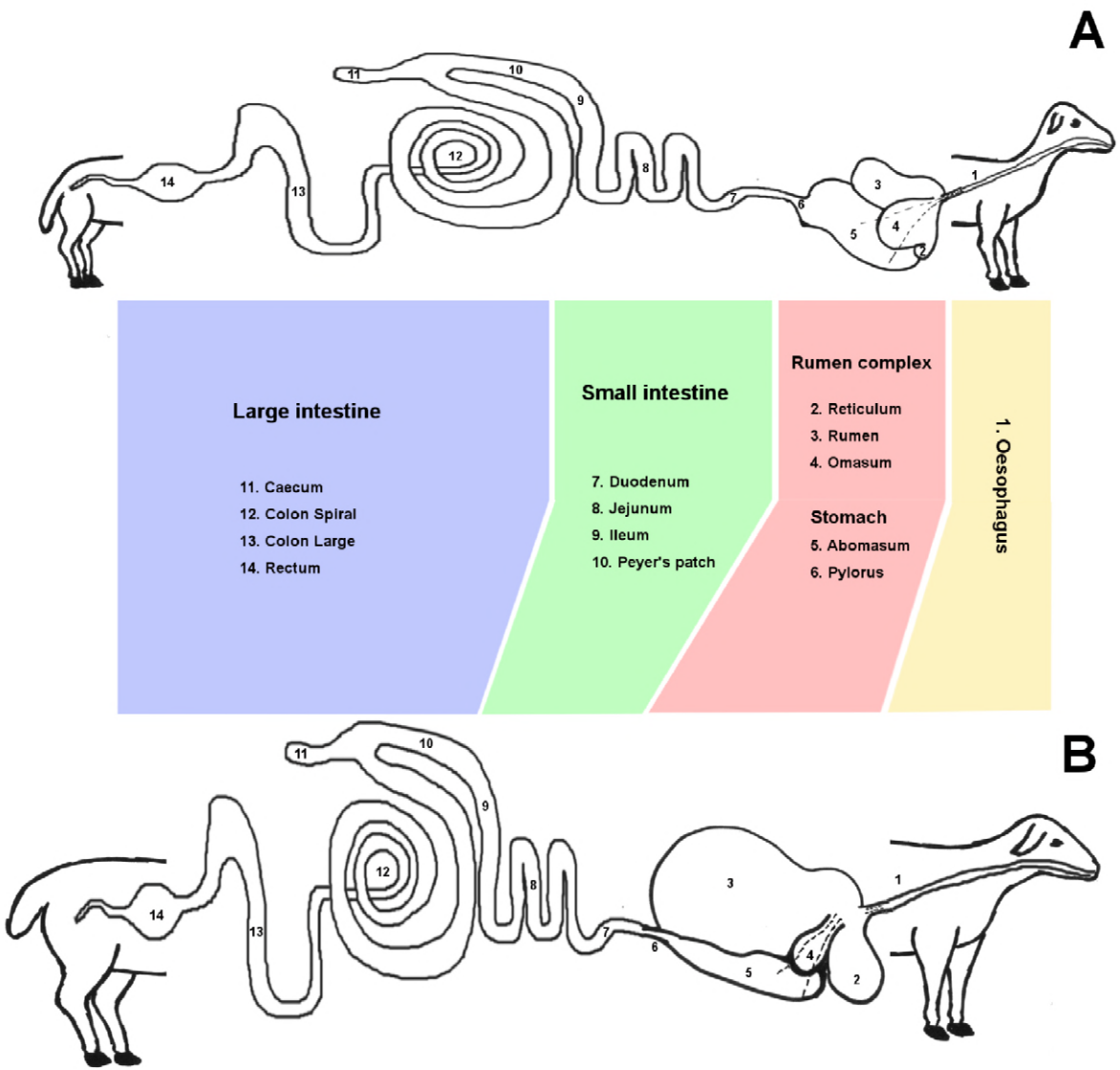
Cluster	Functional Class	Tissue Specific Expression	Marker Genes
1	House-Keeping (1)	Ubiquitous	*ATM, CDK9, CEP63*
2	Cell-Cycle & Immune	Peyer's Patch, Ileum	*CBX2, CD19, CDK1, CKLF*
3	Immune (T-Cells)	Small and Large Intestine	*CD4, CD8A, IL16, IL2RA, PARP8*
4	Metabolism	Liver	*SERPINA10, SLC10A1, FGF12, IGF1, IGF2*
5	Muscle - skeletal muscle	Thoracic Oesophagus	*MYH4, MYL2, OXCT1, ACTA1*
6	Lipids	Omentum, Rectum	*PLIN1, ADIPOQ*
7	Immune - Macrophages	Alveolar Macrophages	*CLEC4E, CD68, CCL4, CSF1, TLR4, CCR1, ITGAX, ITGAM, ADGRE1, MERTK, ICAM-1*
8	Epithelial - Rumen	Rumen Complex - General	*KRT15, N4BP3, NOD1, SOX2*
9	Immune	Ileum, Peyer's Patch, Alveolar Macrophages	*IL10, CD73, CLEC10A*
10	Immune - Macrophages	Alveolar Macrophages	*CD83, CD44, TNF-alpha, VSIG4, P2RY12, JUNB*
11	Vesicle Formation/Nutrient Uptake	Small Intestine (at birth)	*SLC10A4, SLC16A11, AQP7, AQP8*
12	Epithelial/Immune - Rumen	Rumen Complex (increasing with age)	*KRT5, PIP, IL19, VSIG8, IL36A, IL36B*
13	Smooth Muscle	Adult Reticulum and Rumen	*FIBIN, MYL, FMOD, CALD1, MYH11, ACTG2, CNN1*
14	Ribosomal	Ubiquitous	*RPS3, RPL32*
15	Neuronal	Large and small intestine (up-regulation at birth)	*IGF2BP1, MAPK10, FBXL16, VIP*
16	Oxidative Phosphorylation	Ubiquitous	*ATP5B, COX10, NDUFA3*
17	Mitochondria	Ubiquitous	*PLAG1, F3, ESR2*
18	Connective Tissue	Rectum	*COL6A6, GDF6*
19	Neuronal	True stomach, large and small intestine	*CPS1, DEPCR2*
20	Ribosomal	Ubiquitous	*RPL5, RPL15*

1081 **Table 3: Examples of differentially expressed genes identified using pairwise comparison of abomasum, rumen and ileum samples from sheep**

1082 **at birth vs one week and one-week vs 8 weeks.**

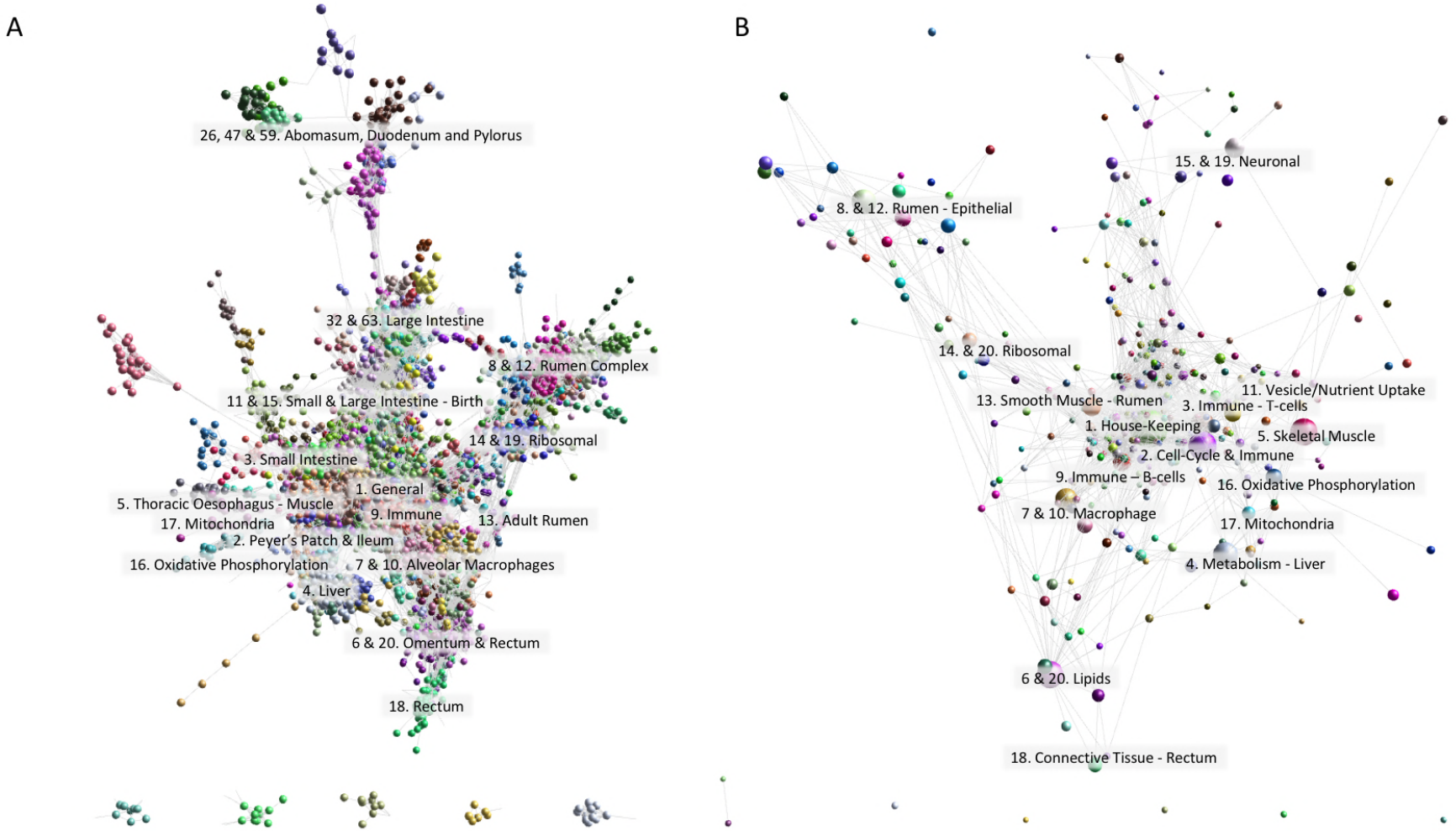
1083

Gene	Tissue	Developmental Stage	Role	Fold Change	P-Value	Direction	TPM 1	TPM 2
PIP	Rumen	1 vs 8 Weeks	Mucosal Immunity	3.1	$5.1x10^{-7}$	Up	375	2113
CLEC4E	Ileum	1 vs 8 Weeks	Inflammatory and Immune Response	2.4	$2.7x10^{-4}$	Up	15	76
IL36A	Rumen	1 vs 8 Weeks	Inflammatory Response	3.7	$1.4x10^{-5}$	Up	275	2647
JAK3	Rumen	0 vs 1 Week	Cytokine signalling	5.5	$2.1x10^{-83}$	Up	115	4311
ISG15	Abomasum	1 vs 8 Weeks	IFN response	3.1	$1.1x10^{-67}$	Up	524	602
	Abomasum	0 vs 1 Week	IFN response	2.6	$1.2x10^{-36}$	Up	15	99
	Rumen	0 vs 1 Week	IFN response	6.2	$1.4x10^{-9}$	Up	10	625
	Ileum	0 vs 1 Week	IFN response	3.4	$5.5x10^{-8}$	Up	18	167
IFIT3	Abomasum	1 vs 8 Weeks	IFN-induced antiviral protein	3.1	$5x10^{-20}$	Up	337	439
IFIT2	Ileum	0 vs 1 Week	IFN-induced antiviral protein	2.0	$7.7x10^{-31}$	Up	76	275
MX2	Abomasum	0 vs 1 Week	IFN-induced protein	2.1	$7x10^{-7}$	Up	19	86
	Rumen	0 vs 1 Week	IFN-induced protein	5.4	$8 x10^{-7}$	Up	13	443
	Rumen	1 vs 8 Weeks	IFN-induced protein	-2.8	$4.3x10^{-3}$	Down	443	47
MX1	Abomasum	1 vs 8 Weeks	IFN-induced antiviral protein	2.1	$5.5x10^{-21}$	Up	1048	1307
	Rumen	0 vs 1 Week	IFN-induced antiviral protein	3.6	$2x10^{-5}$	Up	220	2256
IDO1	Ileum	0 vs 1 Week	Amino acid catabolism	4.2	$2x10^{-8}$	Up	88	1467
COL1A1	Rumen	1 vs 8 Weeks	Collagen - connective tissue	2.1	$3.3x10^{-5}$	Up	29269	92501
	Ileum	1 vs 8 Weeks	Collagen - connective tissue	-2.3	$8.5x10^{-15}$	Down	35506	6919
COL1A2	Rumen	1 vs 8 Weeks	Collagen - connective tissue	2.1	$7.3x10^{-5}$	Up	25919	80206
	Ileum	1 vs 8 Weeks	Collagen - connective tissue	-2.1	$5.6x10^{-14}$	Down	33384	7541
HERC6	Rumen	1 vs 8 Weeks	Ubiquitin ligase activity	-3.9	$4.3x10^{-3}$	Down	319	16
	Ileum	0 vs 1 Week	Ubiquitin ligase activity	3.2	$7.7x10^{-11}$	Up	25	211
	Abomasum	1 vs 8 Weeks	Ubiquitin ligase activity	3.0	$1.1x10^{-11}$	Up	357	636
KRT36	Rumen	1 vs 8 Weeks	Keratin	3.9	$6.5x10^{-8}$	Up	6384	69503
HMGCS2	Rumen	0 vs 1 Week	Ruminal ketone body synthesis	8.4	$2.3x10^{-50}$	Up	66	19467
DUOXA2	Rumen	0 vs 1 Week	Thyroid hormone synthesis	4.1	$3.2x10^{-5}$	Up	12	168
SLC14A1	Rumen	0 vs 1 Week	Nitrogen and Iodine Recycling	2.4	$9.3x10^{-16}$	Up	2181	10158


A
Large intestine
11. Caecum
12. Colon Spiral
13. Colon Large
14. Rectum
Small intestine
7. Duodenum
8. Jejunum
9. Ileum
10. Peyer's patch
Rumen complex
2. Reticulum
3. Rumen
4. Omasum
Stomach
5. Abomasum
6. Pylorus
1. Oesophagus
B

1084

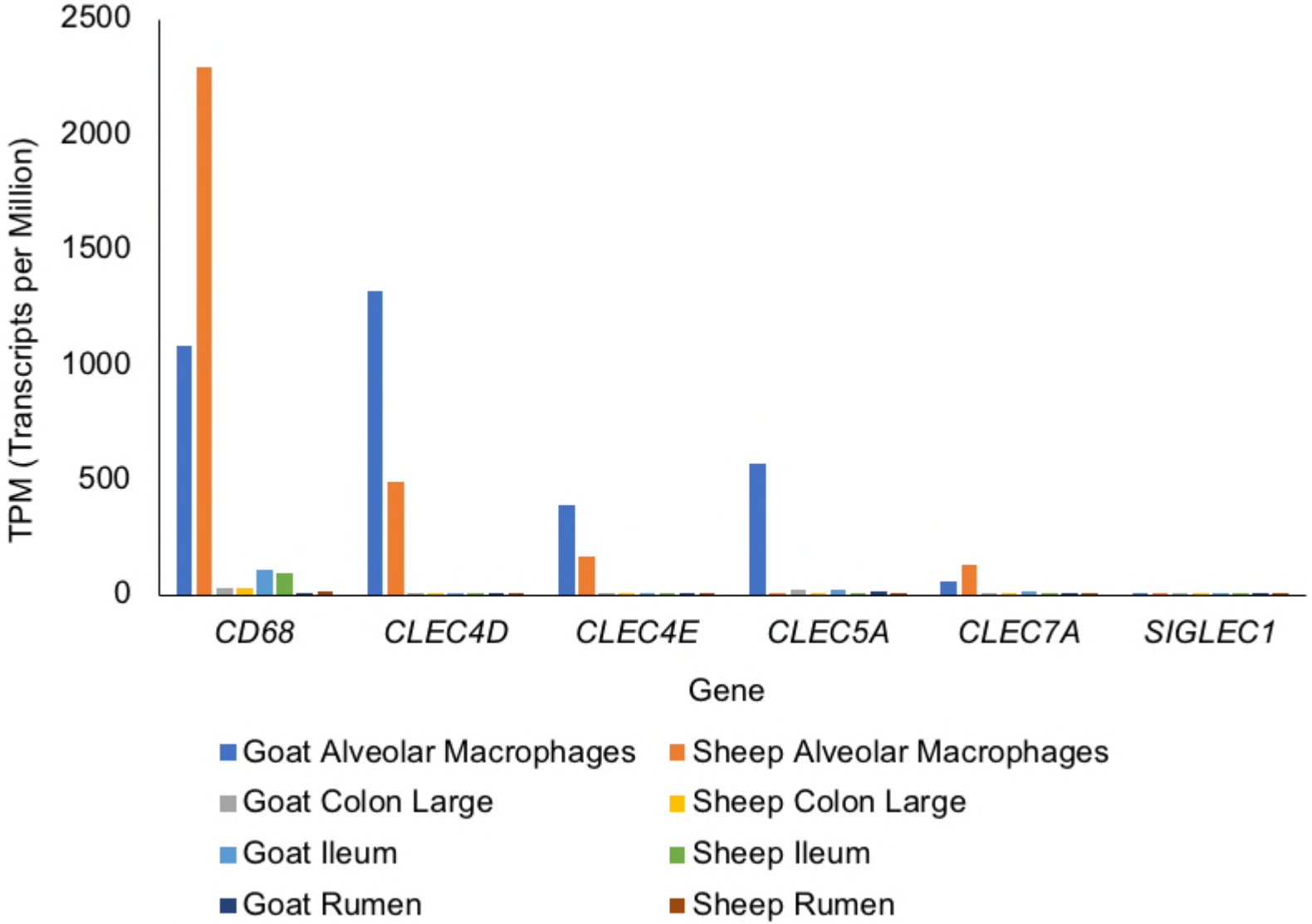
Figure 1: Diagrammatic representation of the morphological changes that occur in the gastrointestinal tract of a sheep during the transition from (A) pre-ruminant to (B) ruminant. The 14 regions sampled for this study are numbered. The oesophageal groove is indicated with dotted lines.


A
26, 47 & 59. Abomasum, Duodenum and Pylorus
32 & 63. Large Intestine
8 & 12. Rumen Complex
11 & 15. Small & Large Intestine - Birth
14 & 19. Ribosomal
3. Small Intestine
1. General
5. Thoracic Oesophagus - Muscle
9. Immune
13. Adult Rumen
17. Mitochondria
2. Peyer's Patch & Ileum
16. Oxidative Phosphorylation
7 & 10. Alveolar Macrophages
4. Liver
6 & 20. Omentum & Rectum
18. Rectum
B
15. & 19. Neuronal
8. & 12. Rumen - Epithelial
14. & 20. Ribosomal
11. Vesicle/Nutrient Uptake
13. Smooth Muscle - Rumen
3. Immune - T-cells
1. House-Keeping
5. Skeletal Muscle
2. Cell-Cycle & Immune
9. Immune – B-cells
16. Oxidative Phosphorylation
7 & 10. Macrophage
17. Mitochondria
4. Metabolism - Liver
6 & 20. Lipids
18. Connective Tissue - Rectum

1089

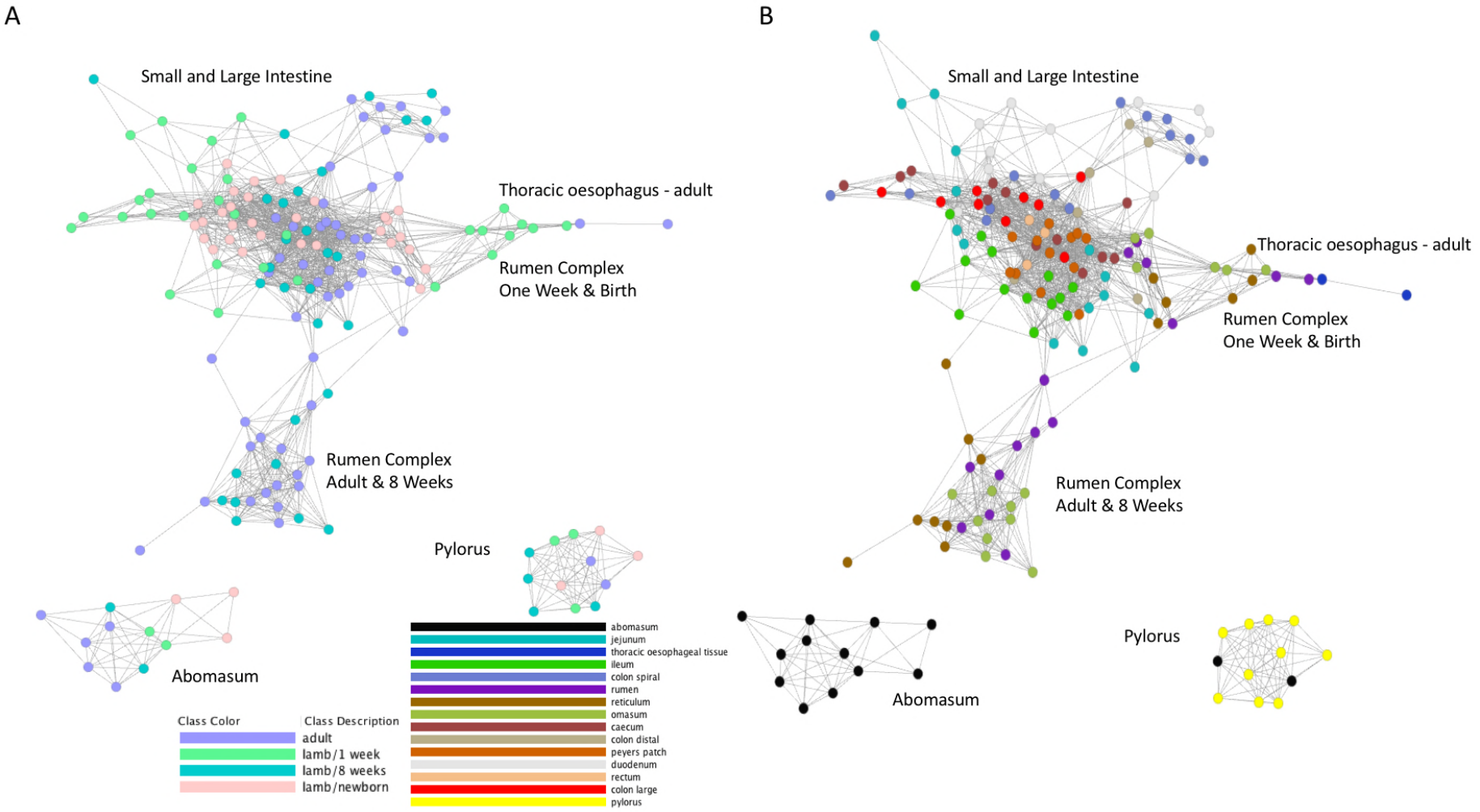
1090

Figure 2 A: Gene-to-gene network graph of sheep GI tract tissues and including alveolar macrophages, oesophageal tissue and liver. The top 20 largest clusters are annotated by functional class. B: Gene-to-gene network graph with the nodes collapsed by class to illustrate the relative size of each cluster. Created using Graphia Professional with parameters Pearson's R=0.85, MCLi=2.2, Minimum Component Size=2, Minimum Cluster Size=2.


2500
2000
1500
1000
500
0
TPM (Transcripts per Million)
CD68
CLEC4D
CLEC4E
CLEC5A
CLEC7A
SIGLEC1
Gene
Goat Alveolar Macrophages
Sheep Alveolar Macrophages
Goat Colon Large
Sheep Colon Large
Goat Ileum
Sheep Ileum
Goat Rumen
Sheep Rumen

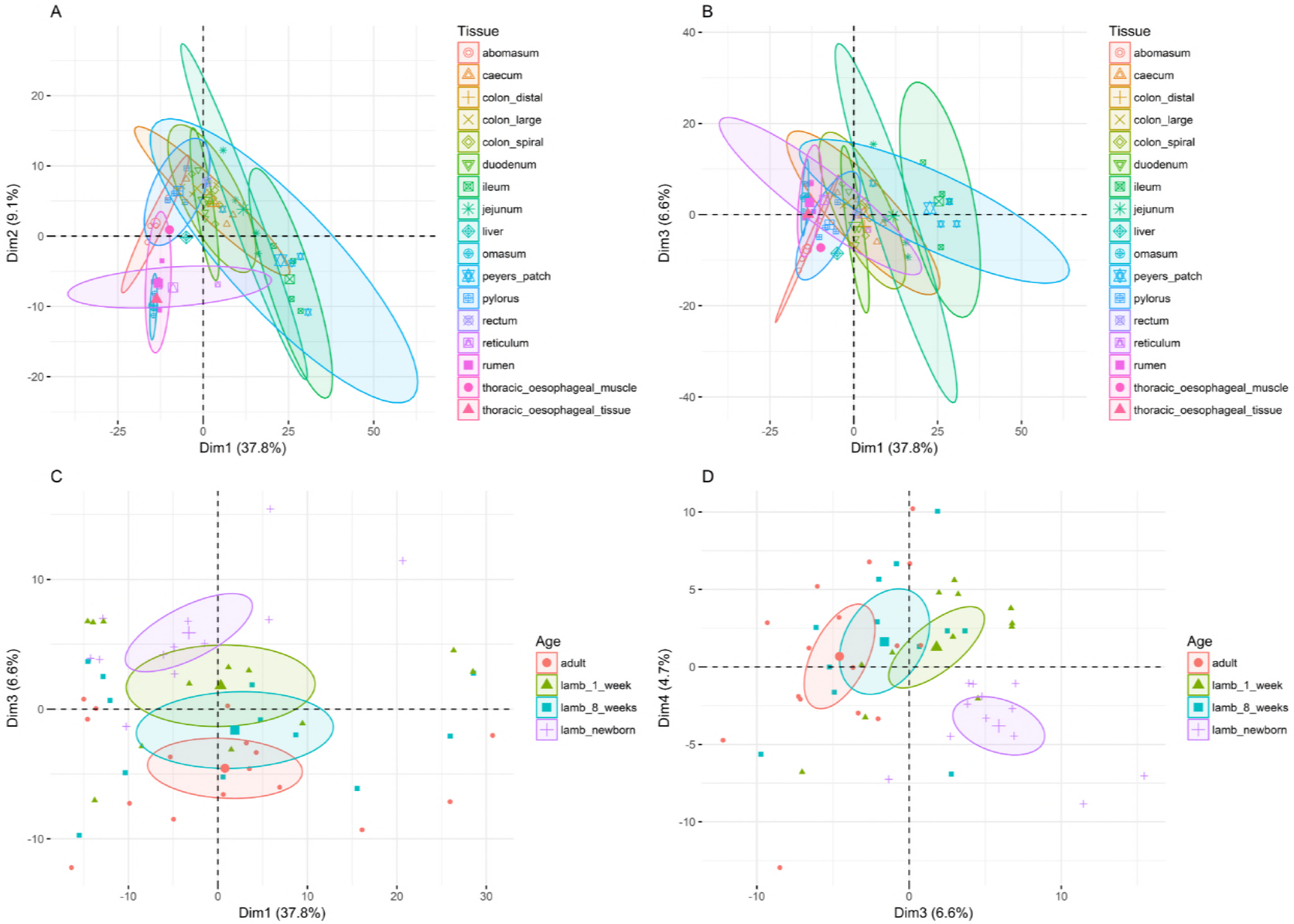
1096

Figure 3: Comparative analysis of the expression of five C-type lectin genes, measured as transcripts per million (TPM), across three GI tract tissues (Large Colon, Ileum and Rumen) and alveolar macrophages in sheep and goats.


A
Small and Large Intestine
Thoracic oesophagus - adult
Rumen Complex
One Week & Birth
Rumen Complex
Adult & 8 Weeks
Pylorus
Abomasum
Class Color
Class Description
adult
lamb/1 week
lamb/8 weeks
lamb/newborn
abomasum
jejunum
thoracic oesophageal tissue
ileum
colon spiral
rumen
reticulum
omasum
caecum
colon distal
peyers patch
duodenum
rectum
colon large
pylorus
B
Small and Large Intestine
Thoracic oesophagus - adult
Rumen Complex
One Week & Birth
Rumen Complex
Adult & 8 Weeks
Abomasum
Pylorus

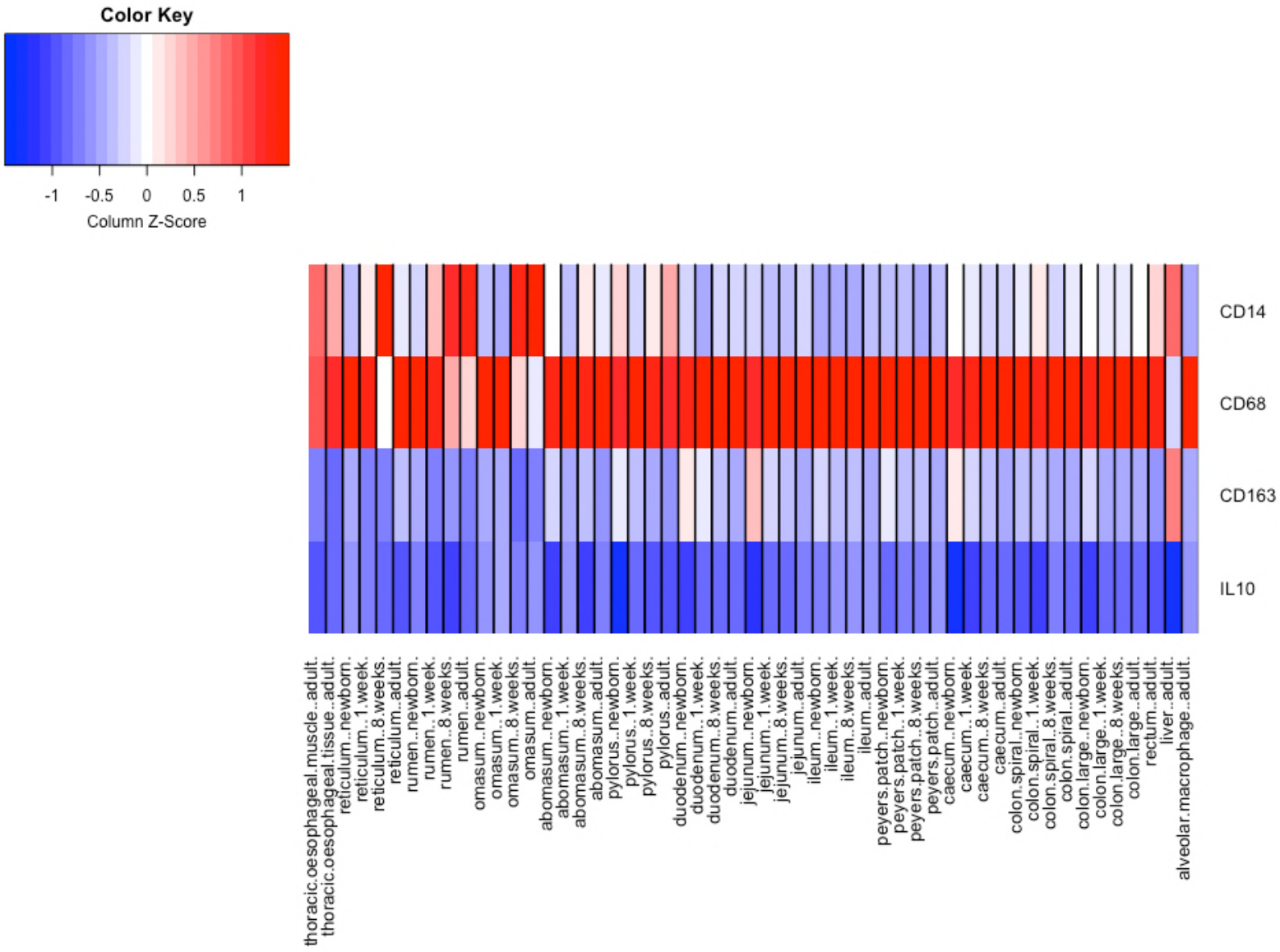
1100

Figure 4: Sample-to-sample network graph of sheep GI tract samples coloured according to developmental stage (A) and tissue (B). Created using Graphia Professional with parameters Pearson's R=0.85, MCLi=2.2, Minimum Component Size=2, Minimum Cluster Size=2.


A
Dim2 (9.1%)
Dim1 (37.8%)
B
Dim3 (6.6%)
Dim1 (37.8%)
Tissue
abomasum
caecum
colon_distal
colon_large
colon_spiral
duodenum
ileum
jejunum
liver
omasum
peyers_patch
pylorus
rectum
reticulum
rumen
thoracic_oesophageal_muscle
thoracic_oesophageal_tissue
C
Dim3 (6.6%)
Dim1 (37.8%)
D
Dim4 (4.7%)
Dim3 (6.6%)
Age
adult
lamb_1_week
lamb_8_weeks
lamb_newborn

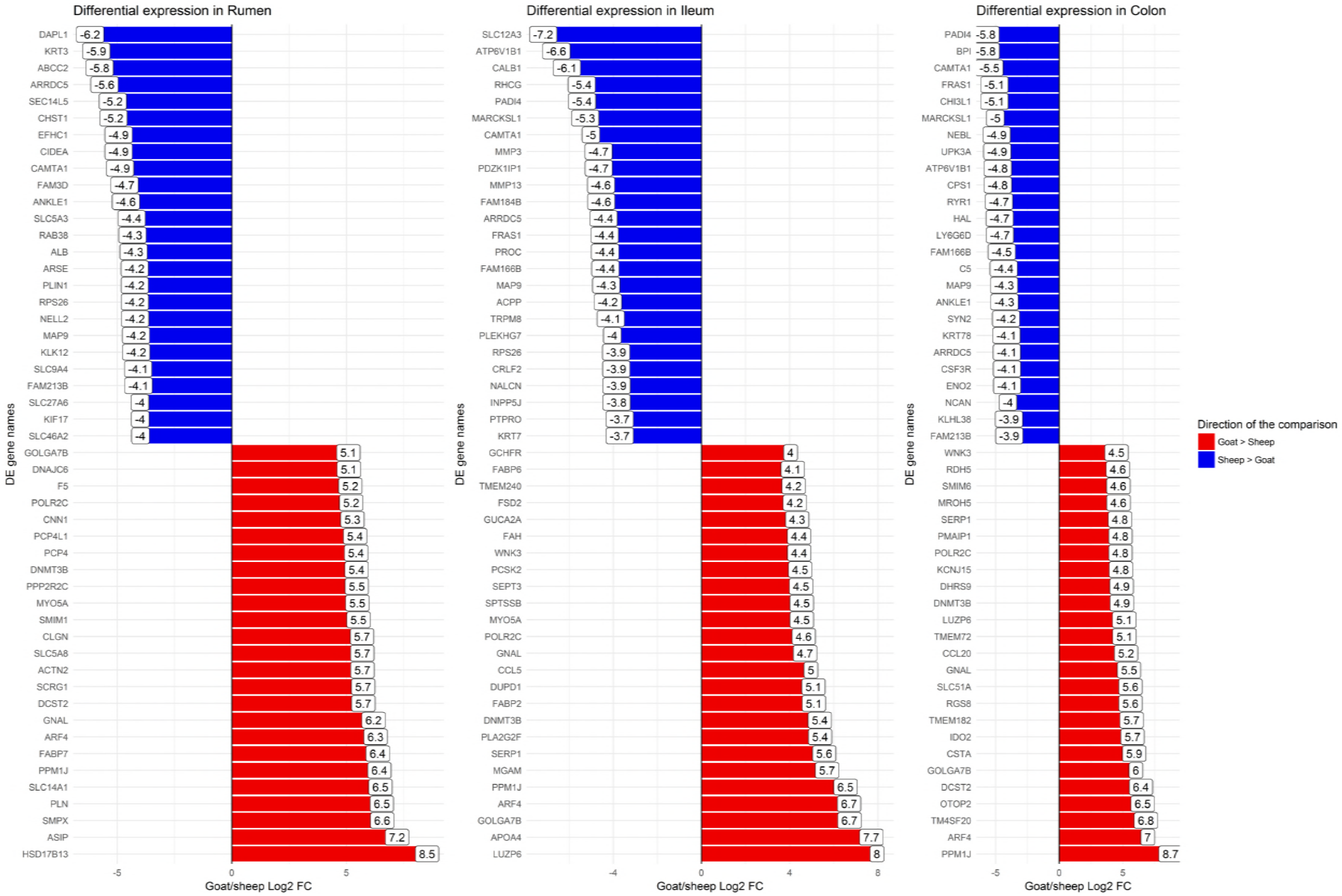
1104

Figure 5: Principal Component Analysis of macrophage specific signatures in the sheep GI tract illustrating separation by tissue (A & B) and developmental stage (C & D) in three and four components, respectively.


Color Key
-1
-0.5
0
0.5
1
Column Z-Score
CD14
CD68
CD163
IL10
thoracic.oesophageal.muscle..adult.
thoracic.oesophageal.tissue..adult.
reticulum..newborn.
reticulum..1.week.
reticulum..8.weeks.
reticulum..adult.
rumen..newborn.
rumen..1.week.
rumen..8.weeks.
rumen..adult.
omasum..newborn.
omasum..1.week.
omasum..8.weeks.
omasum..adult.
abomasum..newborn.
abomasum..1.week.
abomasum..8.weeks.
abomasum..adult.
pylorus..newborn.
pylorus..1.week.
pylorus..8.weeks.
pylorus..adult.
duodenum..newborn.
duodenum..1.week.
duodenum..8.weeks.
duodenum..adult.
jejunum..newborn.
jejunum..1.week.
jejunum..8.weeks.
jejunum..adult.
ileum..newborn.
ileum..1.week.
ileum..8.weeks.
ileum..adult.
peyers.patch..newborn.
peyers.patch..1.week.
peyers.patch..8.weeks.
peyers.patch..adult.
caecum..newborn.
caecum..1.week.
caecum..8.weeks.
caecum..adult.
colon.spiral..newborn.
colon.spiral..1.week.
colon.spiral..8.weeks.
colon.spiral..adult.
colon.large..newborn.
colon.large..1.week.
colon.large..8.weeks.
colon.large..adult.
rectum..adult.
liver..adult.
alveolar.macrophage..adult.

1108

Figure 6: Red/blue heat map visualisation of the expression of a set of four immune marker genes (*CD14, CD68, CD163* and *IL10*) across sheep GI tract tissues at each developmental stage (including thoracic oesophagus, liver and alveolar macrophage samples for reference). Red indicates up-regulation of genes and blue down-regulation using z-score normalisation of TPM values for each gene.


Differential expression in Rumen
DE gene names
DAPL1 -6.2
KRT3 -5.9
ABCC2 -5.8
ARRDC5 -5.6
SEC14L5 -5.2
CHST1 -5.2
EFHC1 -4.9
CIDEA -4.9
CAMTA1 -4.9
FAM3D -4.7
ANKLE1 -4.6
SLC5A3 -4.4
RAB38 -4.3
ALB -4.3
ARSE -4.2
PLIN1 -4.2
RPS26 -4.2
NELL2 -4.2
MAP9 -4.2
KLK12 -4.2
SLC9A4 -4.1
FAM213B -4.1
SLC27A6 -4
KIF17 -4
SLC46A2 -4
GOLGA7B 5.1
DNAJC6 5.1
F5 5.2
POLR2C 5.2
CNN1 5.3
PCP4L1 5.4
PCP4 5.4
DNMT3B 5.4
PPP2R2C 5.5
MYO5A 5.5
SMIM1 5.5
CLGN 5.7
SLC5A8 5.7
ACTN2 5.7
SCRG1 5.7
DCST2 5.7
GNAL 6.2
ARF4 6.3
FABP7 6.4
PPM1J 6.4
SLC14A1 6.5
PLN 6.5
SMPX 6.6
ASIP 7.2
HSD17B13 8.5
-5
0
5
Goat/sheep Log2 FC
Differential expression in Ileum
DE gene names
SLC12A3 -7.2
ATP6V1B1 -6.6
CALB1 -6.1
RHCG -5.4
PADI4 -5.4
MARCKSL1 -5.3
CAMTA1 -5
MMP3 -4.7
PDZK1IP1 -4.7
MMP13 -4.6
FAM184B -4.6
ARRDC5 -4.4
FRAS1 -4.4
PROC -4.4
FAM166B -4.4
MAP9 -4.3
ACPP -4.2
TRPM8 -4.1
PLEKHG7 -4
RPS26 -3.9
CRLF2 -3.9
NALCN -3.9
INPP5J -3.8
PTPRO -3.7
KRT7 -3.7
GCHFR 4
FABP6 4.1
TMEM240 4.2
FSD2 4.2
GUCA2A 4.3
FAH 4.4
WNK3 4.4
PCSK2 4.5
SEPT3 4.5
SPTSSB 4.5
MYO5A 4.5
POLR2C 4.6
GNAL 4.7
CCL5 5
DUPD1 5.1
FABP2 5.1
DNMT3B 5.4
PLA2G2F 5.4
SERP1 5.6
MGAM 5.7
PPM1J 6.5
ARF4 6.7
GOLGA7B 6.7
APOA4 7.7
LUZP6 8
-4
0
4
8
Goat/sheep Log2 FC
Differential expression in Colon
DE gene names
PADI4 -5.8
BPI -5.8
CAMTA1 -5.5
FRAS1 -5.1
CHI3L1 -5.1
MARCKSL1 -5
NEBL -4.9
UPK3A -4.9
ATP6V1B1 -4.8
CPS1 -4.8
RYR1 -4.7
HAL -4.7
LY6G6D -4.7
FAM166B -4.5
C5 -4.4
MAP9 -4.3
ANKLE1 -4.3
SYN2 -4.2
KRT78 -4.1
ARRDC5 -4.1
CSF3R -4.1
ENO2 -4.1
NCAN -4
KLHL38 -3.9
FAM213B -3.9
WNK3 4.5
RDH5 4.6
SMIM6 4.6
MROH5 4.6
SERP1 4.8
PMAIP1 4.8
POLR2C 4.8
KCNJ15 4.8
DHRS9 4.9
DNMT3B 4.9
LUZP6 5.1
TMEM72 5.1
CCL20 5.2
GNAL 5.5
SLC51A 5.6
RGS8 5.6
TMEM182 5.7
IDO2 5.7
CSTA 5.9
GOLGA7B 6
DCST2 6.4
OTOP2 6.5
TM4SF20 6.8
ARF4 7
PPM1J 8.7
-5
0
5
Goat/sheep Log2 FC
Direction of the comparison
Goat > Sheep
Sheep > Goat

1113

Figure 7: The top 25 differentially expressed genes up-regulated and top 25 down-regulated genes between goat and sheep in the rumen, ileum and colon of age-matched one-week old animals.

Supplemental Material

Table S1: Averaged TPM expression estimates for sheep GI tract tissues, alveolar macrophages, thoracic oesophagus and liver, including the TxBF and Texel data.

Table S2: List of genes within each cluster from the gene-to-gene network graph (Figure 2) including sheep GI tract tissues, alveolar macrophages, thoracic oesophagus and liver.

Table S3: GO term enrichment for molecular function, cellular component and biological process for each cluster from the gene-to-gene network graph for sheep (Figure 2).

Table S4: Individual sheep GI tract TPM expression estimates used for this analysis, which is a subset of data from the TxBF sheep atlas.

Table S5: Differentially expressed genes in the sheep rumen using pairwise comparisons between birth and one week and one week and 8 weeks of age.

Table S6: Differentially expressed genes in the sheep ileum using pairwise comparisons between birth and one week and one week and 8 weeks of age.

Table S7: Differentially expressed genes in the sheep abomasum using pairwise comparisons between birth and one week and one week and 8 weeks of age.

Table S8: GO term enrichment of differentially expressed genes in the rumen, ileum and abomasum of sheep between two developmental stages birth and one week and one week and 8 weeks of age.

Table S9: Individual TPM expression estimates for rumen, ileum and colon from one-week old goats.

Table S10: Differentially expressed genes in rumen samples from one-week old sheep and goats.

Table S11: Differentially expressed genes in ileum samples from one-week old sheep and goats.

Table S12: Differentially expressed genes in colon samples from one-week old sheep and goats.

Figure S1: Principal Component Analysis of sheep GI tract samples illustrating separation by tissue (A & B) and developmental stage (C & D) in three components.

Figure S2: Visualisation of differentially expressed genes within and between ileum, rumen and abomasum tissues at two developmental stages (0-1 week) and (1-8 weeks) in sheep using Venn diagrams.