Transcriptome analysis reveals the importance of the immune system during early pregnancy in sheep

Kisun Pokharel[1], Jaana Peippo[1], Melak Weldenegodguad[1,2], Mervi Honkatukia[3], Meng-Hua Li[4*], Juha Kantanen[1*]

[1]Natural Resources Institute of Finland (Luke), Jokioinen, Finland

[2]Department of Environmental and Biological Sciences, Faculty of Science and Forestry, University of Eastern Finland, Kuopio, Finland

[3]NordGen – The Nordic Genetic Resources Center, Ås, Norway

[4]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences (CAS), Beijing, China

*** Correspondence:**
Prof. Meng-Hua Li
menghua.li@ioz.ac.cn
Prof. Juha Kantanen
juha.kantanen@luke.fi

Keywords: Finnsheep, Texel, progesterone, immunoglobulins, Siglec, endogenous retrovirus, endometrium, corpus luteum, prolificacy, preimplantation

Running title: Sheep preimplantation transcriptome

1 Abstract

The majority of pregnancy loss in ruminants occurs during the preimplantation stage, which is thus the most critical period determining reproductive success. While ovulation rate is the major determinant of litter size in sheep, interactions among the conceptus, corpus luteum and endometrium are essential for pregnancy success. To evaluate the role of reproductive tract function in sheep fertility, we performed a comparative transcriptome study by sequencing total RNA (mRNA and miRNA) from corpus luteum (CL) and endometrium tissues collected during the preimplantation stage of pregnancy in Finnsheep, Texel and F1 crosses. A total of 21,287 genes and 599 miRNAs were expressed in our dataset. Ten out of the top 25 most highly expressed genes were shared across tissues, indicating the complementary functions of the CL and endometrium. Moreover, highly expressed autosomal genes in the endometrium and CL were associated with biological processes such as progesterone formation (*STAR* and *HSD3B1*) in the CL and facilitation of maternal recognition of pregnancy, trophoblast elongation and implantation (*LGALS15*, *CST3*, *CST6*, and *EEF1A1*) in the endometrium. In the CL, a group of sialic acid-binding immunoglobulin (Ig)-like lectins (Siglecs), solute carriers (*SLC13A5*, *SLC15A2*, *SLC44A5*) and chemokines (*CCL5*, *CXCL13*, *CXCL9*) were upregulated in Finnsheep, while several multidrug resistance-associated proteins (MRPs) were upregulated in Texel ewes. We also identified a novel ERV gene located in a reduced FecL locus that is associated with sheep prolificacy and is upregulated in prolific Finnsheep. Moreover, we report, for the first time in any species, several genes that are active in the CL during early pregnancy (including *SIGLEC13*, *SIGLEC14*, *SIGLEC6*, *MRP4*, and *CA5A*). Importantly, functional analysis of differentially expressed genes suggested that Finnsheep have a better immune

system than Texel and that high prolificacy in Finnsheep might be governed by immune system regulation. Taken together, the findings of this study provide new insights into the interplay between the CL and the endometrium in gene expression dynamics during early pregnancy. The data and results will serve as a basis for studying this highly critical period of pregnancy, which has wide significance in mammalian fertility and reproduction.

2 Introduction

Litter size, a key determinant for the profitability of sheep production systems, is highly dependent on ovulation rate and embryo development in the uterus. Earlier studies have shown that the trait of high prolificacy can result due to the action of either a single gene with a major effect, as in the Chinese Hu, Boorola Merino, Lacaune and small-tailed Han breeds (Mulsant et al., 2001; Souza et al., 2001; Davis et al., 2002, 2006; Chu et al., 2007; Drouilhet et al., 2013), or different sets of genes, as in the Finnsheep and Romanov breeds (Ricordeau et al., 1990; Xu et al., 2018). The Finnsheep or Finnish landrace, one of the most highly prolific breeds, has been exported to more than 40 countries to improve local breeds, although the heritability of ovulation rate is low (Hanrahan and Quirke, 1984). In recent years, a $FecG^{F}$ (V371M) mutation in gene *GDF9* has been identified to be strongly associated with litter size in Finnsheep and breeds such as the Norwegian White Sheep, Cambridge and Belclare breeds, which were developed using Finnsheep (Hanrahan et al., 2004; Våge et al., 2013; Mullen and Hanrahan, 2014; Pokharel et al., 2018).

The success of pregnancy establishment in sheep and other domestic ruminants is determined at the preimplantation stage and involves coordination among pregnancy recognition, implantation and placentation, in which the corpus luteum (CL) and endometrium play vital roles (Geisert et al., 1992; Spencer et al., 2004b, 2007). The preimplantation stage of pregnancy is the most critical period in determining the litter size because of the high embryo mortality during this period. It has been shown that most embryonic deaths occur before day 18 of pregnancy in sheep (Quinlivan et al., 1966; Bolet, 1986; Rickard et al., 2017). However, due to the biological complexity of the process and to technical difficulties, embryo implantation is still not well understood.

The CL is an endocrine structure whose main function is to synthesize and secrete the hormone progesterone. Progesterone production is essential for the establishment of pregnancy. However, if pregnancy is not established, the CL will regress as a result of luteolysis, and a new cycle will begin. The endometrium is the site of blastocyst implantation, but its function is not limited to implantation. The outer lining of the endometrium secretes histotroph, a complex mixture of enzymes, growth factors, hormones, transport proteins and other substances that are key to conceptus survival and implantation, pregnancy recognition signal production and placentation (Spencer and Bazer, 2004; Forde et al., 2013). In addition, the endometrium also plays an important role in regulating the estrous cycle (Spencer et al., 2008).

The whole-transcriptome profiling approach enables a deeper understanding of the functions of both the CL and endometrium, which may allow the identification of genes and markers that are differentially expressed, for example, between breeds showing different litter size phenotypes. Although most of the studies associated with early pregnancy have been performed in sheep (Spencer et al., 2004b, 2007; Mamo et al., 2012; Bazer, 2013; Raheem, 2017), only a few studies have applied transcriptomic approaches to the endometrium and CL. A microarray-based transcriptomic study conducted by Gray et al. (2006) identified a number of endometrial genes regulated by progesterone (from the CL) and interferon tau (*IFNT*; from the conceptus) in pregnant vs uterine gland knockout (UGKO) ewes. In a more comprehensive study conducted by Brooks et al. (2016), transcriptome

analysis of uterine epithelial cells during the peri-implantation period of pregnancy identified various regulatory pathways and biological processes in sheep. Moore et al. (2016) combined gene expression data with genome-wide association studies (GWASs) to understand the roles of CL and endometrium transcriptomes in dairy cattle fertility. A study by (Kfir et al., 2018) identified differentially expressed genes (DEGs) between Day 4 and Day 11 in the CL in cattle. Recently, a study on endometrial gene expression differences between Finnsheep and European mouflon identified several genes associated with reproductive processes (Yang et al., 2018). Though these studies have certainly enhanced our understanding of the roles of the CL and endometrium during early pregnancy and in ruminant fertility in general, none of the studies have conducted specific comparisons between breeds with different reproductive potential. Thus, in this study, a comparison of transcriptome profiles between two breeds was conducted to provide insight into the similarities in developmental events in early pregnancy between the breeds. Using F1 crosses we were able to better understand the heritability of the genetic markers. Here, the main goal of this study was to build a global picture of transcriptional complexity in two tissues (Cl and endometrium) and examine differences in developmental profiles during early pregnancy in sheep breeds showing contrasting fertility phenotypes. Thus, this study has relevance to sheep breeding towards achieving improved reproductive capacity.

3 Materials and Methods

3.1 Experimental design

All procedures for the experiment and sheep sampling were approved by the Southern Finland Animal Experiment Committee (approval no. ESAVI/5027/04.10.03/2012). The animals were kept at Pusa Farm in Urjala, located in the province of Western Finland, during the experimental period. A total of 31 ewes representing three breed groups (Finnsheep (n=11), Texel (n=11) and F1 crosses (n=9) were included in the main experiment (please note that only 18 of the 31 ewes have been included in this study). Analyses were conducted for two different time points during the establishment of pregnancy: the follicular growth phase (Pokharel et al., 2018) and early pregnancy prior to implantation (current study). After ovary removal (Pokharel et al., 2018), the ewes were mated using two Finnsheep rams, and the pregnant ewes were slaughtered during the preimplantation phase of the pregnancy when the embryos were estimated to be one to three weeks old. At the slaughterhouse, a set of tissue samples (the pituitary gland, a CL, oviductal and uterine epithelial cells, and preimplantation embryos) were collected and stored in RNAlater reagent (Ambion/Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Of the collected tissue samples, CL and endometrium tissues were subjected to current study. Endometrial samples were collected from the uterine horns with a cytobrush, which was rinsed in a tube containing RNAprotect Cell Reagent (Qiagen, Valencia, CA, USA). One of the CLs was dissected from each ovary. For the present study, and particularly for the RNA-Seq of the endometrium and CL, six ewes each from the Finnsheep, Texel and F1 cross groups were included. Therefore, out of 31 ewes that were originally included in the main experiment, only 18 have been considered here. The experimental design have been described in more detail in an earlier study (Pokharel et al., 2018).

3.2 Library preparation and sequencing

Both mRNA and miRNA were extracted from the tissues using an RNeasy Plus Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. The details on RNA extraction have been described previously (Hu et al., 2015; Pokharel et al., 2018). RNA quality (RNA concentration and RNA integrity number) was measured using a Bioanalyzer 2100 (Agilent Technologies, Waldbronn, Germany) before sending the samples to the Finnish Microarray and Sequencing Center,

Turku, Finland, where library preparation and sequencing were performed. RNA libraries were prepared according to the Illumina TruSeq® Stranded mRNA Sample Preparation Guide (part # 15031047) which included poly-A selection step. Unique Illumina TruSeq indexing adapters were ligated to each sample during an adapter ligation step to enable pooling of multiple samples into one flow cell lane. The quality and concentrations of the libraries were assessed with an Agilent Bioanalyzer 2100 and by Qubit® Fluorometric Quantitation, Life Technologies, respectively. All samples were normalized and pooled for automated cluster preparation at an Illumina cBot station. High-quality libraries of mRNA and miRNA were sequenced with an Illumina HiSeq 2000 instrument using paired-end (2x100 bp) and single-end (1x50) sequencing strategies, respectively.

3.3 Data preprocessing and mapping for mRNA

The raw reads were assessed for errors and the presence of adapters using FastQC v0.11.6 (Simon Andrews). As we noticed the presence of adapters, Trim Galore v0.5.0 (Felix Krueger; Martin, 2011) was used to remove the adapters and low-quality reads and bases. The transcripts were quantified under the quasi-mapping-based mode in Salmon v0.11.2 (Patro et al., 2017). We extracted the FASTA sequences (oar31_87.fa) of the sheep transcriptome (oar31_87.gtf) using the gffread utility (Trapnell et al., 2010) and built the transcriptome index. The resulting index was used for transcript quantification (also known as pseudo alignment) of the RNA-Seq reads.

3.4 Data preprocessing and analysis for miRNA

The raw sequence data were initially screened to obtain an overview of the data quality, including the presence or absence of adapters, using FastQC v0.11.6 (Simon Andrews). Next, the Illumina adapters and low-quality bases were removed using Trim Galore v0.5.0 (Felix Krueger; Martin, 2011). In addition, reads that were too short (having fewer than 18 bases) after trimming were also discarded. To reduce downstream computational time, high-quality reads were collapsed using Seqcluster v1.2.4a7 (Pantano et al., 2011). The FASTQ output from Seqcluster was first converted into a FASTA file. The FASTA header was reformatted by including a sample-specific three letter code, which is also a requirement for miRDeep2 analysis. For instance, “>A01_1_x446 A01” represents sample C1033, whose first read was repeated 446 times.

The collapsed reads were mapped against the ovine reference genome (oar v3.1) using Bowtie (Langmead et al., 2009). The Bowtie parameters were adjusted so that (1.) the resulting alignments had no more than 1 mismatch (-v 1); (2.) the alignments for a given read were suppressed if more than 8 alignments existed for it (-m 8); and (3.) the best-aligned read was reported (--strata, --best). The alignment outputs (in SAM format) were coordinate-sorted and converted to BAM files. The sorted BAM files were converted to the miRDeep2 ARF format using the “bwa_sam_converter.pl” script.

miRDeep2 v2.0.0.5 (Friedländer et al., 2012) was used to identify known ovine miRNAs and to predict conserved (known in other species) and novel ovine miRNAs. Before running the miRDeep2 pipeline, we merged both the collapsed FASTA files and the mapped ARF files. Furthermore, hairpin and mature sequences of all species were extracted from miRBase v22 (Kozomara and Griffiths-Jones, 2011, 2014). The extracted sequences were grouped into mature ovine sequences, ovine hairpin sequences, and mature sequences for all species except sheep. The results from miRDeep2 were further processed to compile a list of all known and novel miRNAs. For novel and conserved miRNAs, we designated provisional IDs that included the genomic coordinates of the putative mature and star sequences.

3.5 Differential expression of RNA

For mRNA-Seq data, the gene expression estimates from Salmon were used to identify DEGs. The Salmon-based transcript counts were summarized to gene level estimates using tximport (Soneson et al., 2016).DESeq2 (Love et al., 2013) was used to compare gene expression differences. We started by considering both tissues, but after observing the high variation between the endometrial samples, we decided to analyze the two tissues separately. Furthermore, PCA plot (Fig. 2A) illustrated a high variation in gene expression estimates between the endometrial samples while tight clustering of the CL samples. Therefore, owing to sampling bias (explanation in results and discussion section), we did not proceed with differential gene expression analysis on endometrial samples. All replicates were collapsed before running DESeq. We set the filtering criteria for significant DEGs to an adjusted p-value of 0.1 (*padj <1*) and an absolute log2(fold change) of greater than 1 (*abs(log2Foldchange)>1*). All the significant DEGs were annotated with Bioconductor biomaRt (Durinck et al., 2005) to retrieve additional information (gene name, gene description, Entrez ID, human ortholog and chromosome number).

From the list of miRNAs discovered with miRDeep2, those with a minimum count of 10 reads across all samples were considered for expression analysis. We used DESeq2 for expression analysis, in which the technical replicates of three samples (C107, C4271 and C312) were collapsed prior to running the DESeq command. Because of the sampling bias in endometrium samples, we did not conduct breed wise differential expression analysis for endometrium samples. The differentially expressed miRNAs with adjusted p-values less than 0.1 were regarded as significant.

3.6 Gene ontology and pathway analysis

The ClueGO v2.5.3 (Bindea et al., 2009) plugin in Cytoscape v3.7.0 (Shannon et al., 2003) was employed for gene functional analysis. Prior to performing the analyses, we downloaded the latest versions of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO) terms. In addition, we retrieved Entrez gene IDs for all expressed genes in our dataset using the biomaRt Bioconductor package. The enrichment analysis was based on a two-sided hypergeometric test with the Bonferroni step-down correction method. We used a custom reference set that included a list of all the expressed genes in our dataset. We also modified the default GO and pathway selection criteria in such a way that a minimum of three genes and three percent of genes from a given GO or KEGG pathway should be present in the query list. Furthermore, GO terms with a minimum level of three and a maximum level of 5 were retained.

3.7 Manual annotation of select genes

When we noticed that many genes lacked gene annotations, we manually annotated those among the top 25 most highly expressed genes in each tissue and the significant DEGs. First, we extracted the coding sequence of each novel gene using Ensembl BioMart. All genes that had coding sequences were BLASTed against the nonredundant (NR) nucleotide database. For the BLAST-based annotation, we chose the hit with the highest coverage and the highest percentage of sequence identity to the query sequence. Gene IDs that lacked coding sequences (CDSs) were queried back to the Ensembl database to retrieve existing information. Throughout the paper (including in the supplementary data files), the genes that were annotated based on the BLAST results have been marked with an asterisk (*), while those that were annotated based on information available in Ensembl are marked with a hash (#).

4 Results and Discussion

4.1 Phenotypic observations

After removal of the remaining ovary, we counted the number of CLs visually in each animal. With an average of 4.09, Finnsheep had the highest number of CLs, whereas Texel had an average of 1.7 CLs (Supplementary Table S1). F1 showed phenotypes closer to those of Finnsheep than those of Texel, having 3.75 CLs on average (Supplementary Table S1). We did not observe more than 2 CLs in the Texel group or fewer than 3 CLs in Finnsheep or F1 cross-bred. Similarly, on average, Finnsheep had the highest number of embryos (n=2.6), followed by F1 crosses (n=1.8) and Texel (n=1.5). The F1 crosses displayed phenotypes similar to those of Finnsheep; this was unsurprising, as we observed a similar pattern in an earlier study (Pokharel et al., 2018). Interestingly, the embryo survival rate in Texel where 1.5 embryos were present from 1.7 CLs (88%) on average. On the other hand, Finnsheep (63%) and F1 cross (48%) had remarkably low embryo survival rate. While these findings are based on fewer animals, the results are in line with earlier studies (Rhind et al., 1980; Silva et al., 2016) where typically higher litter size is associated with higher embryo mortality and vice versa. It would be of great interest to determine if productivity follows the same pattern in F2 (i.e., F1 x F1) crosses, backcrosses and presumably also in a reciprocal cross.

4.2 RNA-Seq data

From the 42 libraries (21 from each tissue, including three technical replicates), 4.4 billion raw reads were sequenced, of which 4.2 billion clean reads were retained after trimming. The summary statistics from Trim Galore revealed that up to 3.6% of the reads were trimmed, with reverse-strand reads having a comparatively higher percentage of trimmed bases. However, the percentage of reads that were excluded for being shorter than 18 bp was always less than 1% across all samples (Supplementary table S2). Up to 70% of the high-quality reads were mapped to the ovine reference transcriptome (Ensembl release 92).

4.3 Gene expression in CL and endometrium

A total of 21,287 gene transcripts were expressed in the whole data set, of which 1,019 and 959 were specific to the endometrium and CL, respectively. Genes such as cytochrome P450, family 11, subfamily A, polypeptide 1 (*CYP11A1*), C-C motif chemokines (*CCL21*, *CCL26*), serpin family A members (*SERPINA1*, *SERPINA5*), inhibin subunit alpha (*INHA*) and paternally expressed 10 (*PEG10*) were specific to CL while genes related with solute carriers (*SLC44A4*, *SLC7A9*, *SLC34A2*), ERVs were endometrium-specific (Table 1). Further grouping of the expressed genes showed that the most genes (n =19,440) were expressed in endometrium samples of Texel, while the fewest genes (n =19,305) were expressed in endometrium samples of F1 crosses indicating s high variation within the tissue. The cumulative difference in the number of genes in different samples and tissues might be due to transcriptional noise. Nevertheless, the total number of genes expressed in these tissues is comparatively higher than that in ovaries (Pokharel et al., 2018). As shown in Fig. 1, the highest number of breed-specific genes expressed in the CL was found in Finnsheep (n=254), followed by F1 crosses (n=204) and Texel (n=199). Similarly, from endometrium samples, we observed the highest number of unique genes in F1 crosses (n=284), followed by Finnsheep (n=244) and Texel (n=201). In a pairwise comparison, based on overall gene expression, Finnsheep and Texel shared a higher number of genes (n=260) in the CL than in the endometrium. Moreover, Finnsheep and F1 crosses were found to share relatively more common genes (n=278) than the other pairs (Fig. 1).

Several Igs were expressed in the endometrium samples. Igs are heterodimeric proteins that belong to the Ig superfamily (IgSF) (Williams and Barclay, 1988). Igs are composed of two heavy and two

light chains, and the light chain may further consist of a κ or λ chain (Williams and Barclay, 1988). Interestingly, the structure and organization of the genes enable Igs to be receptive to a virtually unlimited array of antigens rather than being limited to a fixed set of ligands (Honjo, 1983). This feature is particularly important for adaptation to changing environments and may have contributed to enabling Finnsheep, for example, to survive in the harsh Finnish climate. Studies on humans have shown that Igs, in general, improve pregnancy success (De Placido et al., 1994; Coulam and Goodman, 2000). In addition to 11 Ig genes representing both the light and heavy chains, the joining chain of multimeric IgA and IgM (*JCHAIN*) was also expressed. *JCHAIN* is a small polypeptide containing eight cysteine residues that makes disulfide (C-C) bonds with IgA and IgM to form multimers. Two of the eight cysteines are linked with cysteines available on the heavy chain of IgA or IgM to result in dimer or pentamer forms, respectively (Bastian et al., 1995). We also identified several genes associated with endogenous retroviruses (ERVs) in the endometrium samples. ERVs are copies of retroviral genomes that have been integrated into the host genome during evolution. Sheep ERVs share sequence similarity with exogenous and pathogenic Jaagsiekte sheep retrovirus (JSRV) (DeMartini et al., 2003). The genome of sheep contains at least 32 ERVs related to JSRV (Sistiaga-Poveda and Jugo, 2014), and these ERVs are essential during pregnancy, including during placental morphogenesis and conceptus elongation (Palmarini et al., 2001; Dunlap et al., 2006b; Spencer and Palmarini, 2012). A number of earlier studies have suggested critical roles of *enJSRV*s in uterine protection from viral infection, preimplantation conceptus development and placental morphogenesis (Dunlap et al., 2005, 2006b, 2006a; Denner, 2016). Interestingly, one of the novel genes (ENSOARG00000009959) predicted to be an ERV was part of reduced FecL locus which is linked to prolificacy in French Lacaune breed (Drouilhet et al., 2013). This gene is located on the reverse strand of chromosome X and has 24 paralogs. This gene is not listed for 162 (out of 184) species available in the Ensembl database. Although Ensembl lists 71 orthologs of this gene, none of them have even 50% sequence homology. A BLAST search against the NR database showed that 97% of the bases matched to the region of the reduced FecL locus (GenBank ID KC352617.1), which was recently characterized (Drouilhet et al., 2013). So far, only two genes, beta-1,4 N-acetylgalactosaminyltransferase 2 (*B4GALNT2*) and insulin-like growth factor 2 mRNA-binding protein 1 (*IGF2BP1*), and a pseudogene, ezrin-like protein, have been identified to exist in that region; our results have added one more gene. In addition to the finding that the best hit was related to the FecL locus, the gene appeared to be an ERV, as we noticed that the query gene had 98% sequence identity with a partial sequence of the endogenous-virus beta-2 pro/pol region (see also Fig. 4). Finally, several lincRNAs were also expressed in the dataset. LincRNAs are long ncRNAs (lncRNAs) that originate from intergenic regions and do not overlap a protein-coding transcript. LincRNAs have a wide array of functions, including transcriptional regulation, biogenesis, epigenetic regulation, tissue specificity and developmental patterning (see reviews by (Pauli et al., 2011; Ulitsky and Bartel, 2013; Deniz and Erman, 2017; Ransohoff et al., 2018).

Although we observed considerable overlap of genes between tissues, principal component analysis (PCA) of the 500 most highly expressed genes clearly indicated two distinct groups (Fig. 2A). Similarly, a heatmap plot based on the top 25 genes with the highest levels of gene expression variation across all samples showed a similar pattern (Fig. 2B). However, we did not observe any breed-specific clusters in either of the tissues, which was also the case in our earlier ovarian transcriptome study(Pokharel et al., 2018). In addition to distinctiveness in terms of gene expression, the PCA plot also revealed that the CL samples appeared to be more homogeneous than the endometrium samples. The two sub-clusters within the endometrium cluster is linked to the age of the embryo (i.e. days after mating) and indicated the experimental bias. More importantly, sampling bias owing to difference in days of collecting endometrium biopsies was apparent in the gene expression. In general, samples in the upper right were older (13-16 days) compared to those in the

lower right. Such difference in gene expression has been attributed to the effects of interferon-tau which causes massive changes in the endometrial gene expression starting day 13 of pregnancy (Gray et al., 2006; Spencer et al., 2007; Forde and Lonergan, 2017). Therefore, we did not proceed further with the differential gene expression comparisons of endometrium samples. However, similar bias was not observed in CL.

Despite sharing 15 of the top 25 highly expressed genes between the tissues there were 5393 differentially expressed genes between endometrium and CL (Supplementary table S3). We noticed that several genes belong to particular gene or protein family were upregulated exclusively in two tissues. Cilia and flagella associated proteins (*CFAP100*, *CFAP300*, *CFAP45*, *CFAP65*), desmosomes including two desmocollins (*DSC1*, *DSC2*), four desmogleins (*DSG1*, *DSG2*, *DSG3*) and desmoplakin (*DSP*), were upregulated in endometrium. Members of homeobox A and B were upregulated in endometrium while those from C and D were upregulated in CL. Thrombomodulin, thrombospondins (*THBS1*, *THBS2*, *THBS3*, *THBS4*), thrombospondin type 1 domain containing (*THSD1*, *THSD7A*, *THSD7B*), thromboxane A synthase 1 (*TBXAS1*) and thromboxane A2 receptor (*TBXA2R*) were all upregulated in CL. Transforming growth factors (*TGFB1*, *TGFB2*, *TGFB3*), *TGFB* receptors (*TGFBR1*, *TGFBR2*, *TGFBR3*), *TGFB1* induced transcript 1 (*TGFB1|1*) and *TGFB* induced (*TGFBI*) were upregulated in CL. Four genes associated with PDZ and LIM domain (*PDLIM2*, *PDLIM3*, *PDLIM4*) and PDZ and LIM domain protein 2-like were upregulated in CL. Guanylate binding proteins (n=7) were all upregulated in CL. EHD protein family comprises four members and three (*EHD2*, *EHD3*, *EHD4*) were exclusively upregulated in CL. Three transcripts (*ENSOARG00000017142*, *ENSOARG00000014427*, *ENSOARG00000019903*) belonging to endogenous retrovirus group K were significantly upregulated in endometrium. Similarly, Placenta-expressed transcript 1 protein and placenta-specific gene 8 protein-like were upregulated in endometrium. Plakophilin 2 (*PKP2*) and 3 (*PKP3*), and plakophilin-1-like were also upregulated in endometrium.

4.4 Most highly expressed genes

One of the most interesting findings was the significant interplay between the CL and endometrium during the preimplantation phase, as revealed by the most highly expressed genes. To obtain an overview of the most abundant genes in each tissue, we selected the top 25 genes (Table 2). We noticed that fifteen out of the top 25 genes were shared in both tissues, and the majority (9 out of 15) were mitochondrial genes. Mitochondrial genes play prominent roles during reproduction. We have also observed high levels of expression of mitochondrial genes in ovaries during the follicular growth phase (Pokharel et al., 2018). Six shared autosomal genes also appeared to play substantial roles during the preimplantation stage. Translationally controlled tumor protein (*TCTP*) is a highly conserved, multifunctional protein that plays essential roles in development and other biological processes in different species (Tuynder et al., 2002; Chen et al., 2007; Brioudes et al., 2010; Li et al., 2011; Branco and Masle, 2019). With a maximum level of expression on Day 5 of pregnancy, this protein has been shown to play a significant role in embryo implantation in mice (Li et al., 2011). Consistent with these earlier studies, *TCTP* appeared to have the highest level of expression during the embryo implantation period. Matrix Gla protein (*MGP*) is a vitamin K-dependent extracellular matrix protein whose expression has been shown to be correlated with development and maturation processes (Zhao and Nishimoto, 1996; Zhao and Warburton, 1997) and receptor-mediated adhesion to the extracellular matrix (Loeser and Wallin, 1992). Several studies have reported that *MGP* is highly expressed in the bovine endometrium (Spencer et al., 1999; Mamo et al., 2012; Forde et al., 2013). The high level of expression of *MGP* in our study is consistent with the results of earlier studies in which this gene was found to be elevated during the preimplantation stage in sheep

(Spencer et al., 1999; Gray et al., 2006) and cattle (Mamo et al., 2012). Similarly, (Casey et al., 2005) reported that *MGP* was significantly upregulated in nonregressed compared to regressed bovine CLs. Our data and supporting results from earlier studies on cattle show that *MGP* is highly expressed in both tissues during the preimplantation stage and plays important roles in superficial implantation and placentation in sheep. In summary, gene expression patterns in the CL and endometrium are similar.

Six genes (*NUPR1*, *BCL2L15*, *CST3*, *CST6*, *S100G*, and *OST4*; see Table 2 for descriptions) specific to the endometrium and one gene (*B2M*) common to both the CL and endometrium were also found to be highly abundant in a recent study in which the authors compared gene expression changes in the luteal epithelium and glandular epithelium during the peri-implantation stage in sheep (Brooks et al., 2016). Galectin 15 (*LGALS15*) is induced by *IFNT* and is involved in conceptus development and implantation (Kim et al., 2003; Gray et al., 2004; Lewis et al., 2007). *LGALS15* mRNA has been detected in ewes from Day 9 until Day 12 (Satterfield et al., 2006). *IFNT* is secreted by the ovine conceptus trophectoderm during the middle to late luteal phase and acts as the signal for maternal recognition of pregnancy. Furthermore, *LGALS15* is an important gene that facilitates adhesion of the trophectoderm to the endometrial luminal epithelium (Lewis et al., 2007; Spencer et al., 2007). Two cystatin (CST) family members, namely, cystatin C (*CST3*) and cystatin E/M (*CST6*), were highly expressed in the endometrium. Known for their importance during the elongation and implantation of the conceptus, CSTs are protease inhibitors that are initiated by progesterone, and their high expression levels are attributable to stimulation by *IFNT* (Spencer et al., 2008, 2015). Elongation factor 1-alpha (*EEF1A1*) is an important component of the protein synthesis machinery because it transports aminoacyl tRNA to the A sites of ribosomes in a GTP-dependent manner (Tatsuka et al., 1992; Mateyak and Kinzy, 2010). The high levels of expression of *EEF1A1* in the endometrium most likely correspond to the production and transport of progesterone and other molecules that are essential during the implantation stage. The exact function of BCL2-like 15 (*BCL2L15*) in the sheep endometrium is not known, nor has it been reported in the endometria of other species, but its high expression has been reported previously (Koch et al., 2010; Brooks et al., 2016; Romero et al., 2017).

Oxytocin (*OXT*) was one of the most highly expressed genes in the CL. In cyclic ewes, *OXT* secreted from the CL and posterior pituitary is widely known to bind with oxytocin receptor (*OTR*) from the endometrium to concomitantly release prostaglandin $F_{2\alpha}$ (*PGF*) pulses and induce luteolysis (Flint and Sheldrick, 2004; Spencer et al., 2004a, 2004b; Bazer, 2013). However, for noncyclic ewes, *OXT* plays an important role during peri-implantation and throughout pregnancy (Kendrick, 2000). *OXT* signaling is known to be influenced by progesterone, but the mechanism underlying the regulation is not yet clear due to conflicting findings (Grazzini et al., 1998; Gimpl et al., 2002; Fleming et al., 2006; Bishop, 2013). *OTR* expression in both the CL and endometrium was almost negligible compared to OXT expression. Steroidogenic acute regulatory protein (*STAR*) plays an important role in mediating the transfer of cholesterol to sites of steroid production (Stocco, 2000; Christenson and Devoto, 2003). Post ovulation, the expression of the majority of genes associated with progesterone synthesis starts to increase and peaks around the late luteal phase, when the CL has fully matured (Juengel et al., 1995; Devoto et al., 2001; Davis and LaVoie, 2018). *STAR*, together with the cytochrome P450 side chain cleavage (P450cc) complex and 3b-hydroxysteroid dehydrogenase/delta5 delta4-isomerase (*HSD3B1*), are the three most important actors involved in progesterone biosynthesis. *STAR* is involved in transporting free cholesterol to the inner mitochondrial membrane. The P450cc complex, composed of a cholesterol side chain cleavage enzyme (*CYP11A1*), ferredoxin reductase (*FDXR*) and ferredoxin (*FDX1*), converts the newly arrived cholesterol into pregnenolone (King and LaVoie, 2009). Finally, *HSD3B1* helps in converting pregnenolone to progesterone (Hu et al., 2010; Plant et al., 2015; Stouffer and Hennebold, 2015;

Davis and LaVoie, 2018). Two of these major genes involved in progesterone synthesis (*STAR* and *HSD3B1*) were ranked among the top 25 most highly expressed autosomal genes, while *CYP11A1* (TPM=2,080), *FDXR* (TPM=86) and *FDX1* (TPM=1,206) were also highly expressed.

4.5 Breed wise gene expression differences in the CL

Because of the sampling bias owing to the Overall, the CL appeared to display higher levels of gene expression differences between the breeds than the endometrium (Table 3). This tissue-specific difference is consistent with the findings of a recent study in cattle in which (Moore et al., 2016) identified nine and 560 DEGs in the endometrium and CL, respectively, between fertile and infertile cows. In the three possible pairwise comparisons for each tissue, the highest number of DEGs (n=199) was found between the pure breeds in the CL, while the fewest DEGs (n=2) were found between the Finnsheep and F1 crosses in the endometrium. In both tissues, the pure breed comparisons had the highest numbers of DEGs, but the two comparisons (for both the endometrium and CL) that involved the F1 crosses had the lowest. In other words, in the CL, Finnsheep had more DEGs (n=67) than F1 crosses, whereas in the endometrium, Texel had more DEGs (n=17) than F1 crosses.

We compared pair-wise (Finnsheep *vs* Texel, Finnsheep *vs* F1 and Texel *vs* F1) differential gene expression in the CL between three breeds. Out of the 199 significant DEGs in the CL of pure breeds (i.e Finnsheep *vs* Texel), 140 were upregulated in Finnsheep (Supplementary table S4) and the rest were downregulated. However, 91 out of the 199 genes lacked annotations (i.e., a gene name and gene description). We were able to retrieve the CDSs for 82 genes and employed a BLAST search against the NR database. Based on the Ensembl ncRNA prediction system, two out of nine genes that lacked CDSs were predicted to be miRNAs, and the rest were lincRNAs.

In the list of DEGs, we observed a few cases in which more than one gene from the same family was present. All eight genes related to multidrug resistance-associated proteins (*MRPs*) were upregulated in Texel ewes, of which seven genes were type 4, while one was type 1. Both *MRP1* and *MRP4* are lipophilic anion transporters. Earlier reports have suggested a role of *MRP4* in transporting prostaglandins in the endometrium (Lacroix-Pépin et al., 2011), and *MRP4* has been found to be upregulated in the endometrium in infertile cows compared to fertile cows (Moore et al., 2016). Although there are no reports regarding the existence and roles of both *MRP4* and *MRP1* in the CL, we speculate that the comparatively lower levels of these prostaglandin (PG) transporters in Finnsheep provide a luteoprotective effect. Six sialic acid-binding Ig-like lectins (Siglecs) were upregulated in Finnsheep. Based on a BLAST search and on the information available in Ensembl, the sequences were related to *SIGLEC-5* (*ENSOARG00000002701*), *SIGLEC-13* (*ENSOARG00000014846* and *ENSOARG00000014850*) and *SIGLEC-14* (*ENSOARG00000014875*, *ENSOARG00000002909*, and *ENSOARG00000001575*). Siglecs are transmembrane molecules that are expressed on immune cells and mediate inhibitory signaling (Varki and Angata, 2006). So far, *SIGLEC-13* has been reported only in nonhuman primates; it was deleted during the course of human evolution (Angata et al., 2004). The importance of Siglecs in immune system regulation has been reviewed elsewhere (Pillai et al., 2012). Siglecs constantly evolve through gene duplication events and may vary between species and even within a species (Cao and Crocker, 2011; Pillai et al., 2012; Bornhöfft et al., 2018). Here we have reported the expression Siglecs in CL which are known to play a role in the immune response during early pregnancy (preimplantation).

Similarly, three genes related to phospholipase A2 inhibitor and Ly6/PLAUR domain-containing protein-like (*PINLYP*) were upregulated in Finnsheep. Other genes with more than one member

included major histocompatibility complexes (MHCs) (*BOLA-DQA5*, *HLA-DMA*, *HLA-DRA*, *MICA*, *BOLA-DQB*2001*, etc.), chemokines (*CCL5*, *CXCL13*, and *CXCL9*), solute carriers (*SLC13A5*, *SLC44A5*, and *SLC15A2*), interleukin receptors (*IL2RG*, *IL12RB1*, and *IL12RB2*), cluster of differentiation factors (CDs) (*CD52*, *CD74*, and *CD300H*), granzymes (*GZMM* and *LOC114109030*/*GZMH*), calcium homeostasis modulators (*CALHM3* and *CALHM5*) and neurofilaments (*NEFL* and *NEFM*). Five out of seven significantly differentially expressed lincRNAs were upregulated in Finnsheep. *GZMM* and *GZMH* have been found to be upregulated in Yakutian cattle compared to Finncattle and Holstein cattle (Pokharel et al., 2019).

Out of 140 genes that were upregulated in Finnsheep, 50 genes (35.71%) were associated with 48 different GO terms (Fig. 3, Supplementary table S5) within the biological processes category, whereas 90 genes lacked GO annotations. The upregulated genes were associated with positive regulation of several processes such as “T cell migration”, “cytokine production”, “defense response”, “immune system process”, “interferon-gamma production”, “leukocyte chemotaxis”, “response to external stimulus”, , “cell–cell adhesion” and “cytokine-mediated signaling pathway”. Some biological processes potentially associated with implantation were “maintenance of location”, “plasma membrane invagination”, “import into cell”, “chemotaxis”, and “receptor internalization”. Other biological processes, such as “response to bacterium”, “response to lipopolysaccharide”, “lymphocyte-mediated immunity” and “chemotaxis”, could be associated with adaptation of Finnsheep to the rugged Finnish climate and with disease resistance. In summary, genes involved in the immune response were upregulated in Finnsheep CL during early pregnancy.

Similarly, only 40 of the 140 genes upregulated in Finnsheep were associated with 29 KEGG pathways (Supplementary fig. 1, Supplementary table S6). The majority of the pathways were associated with diseases; “tryptophan metabolism”, “cell adhesion molecules (CAMs)”, “Th1 and Th2 cell differentiation” and “Th17 cell differentiation” appeared to play roles in implantation. Out of 59 genes that were downregulated in Finnsheep, 17 and 14 genes were associated with GO IDs and KEGG pathways, respectively. However, after applying our selection criteria (GO terms with minimum level of 3 and maximum level of 5 whereby minimum of 3 genes and 3 % of genes from given GO terms, also see methods section), only one biological process, “negative regulation of endopeptidase activity” (associated genes: *COL28A1*, *LOC101104482*, and *SLPI*) and one KEGG pathway, “bile secretion” (associated genes: *LOC101106409*, *LOC101107772*, and *LOC101112460*), were identified.

Altogether, 67 genes were differentially expressed between Finnsheep and F1 crossbred ewes, of which 49 genes were upregulated in Finnsheep (Supplementary table S7). *CA5A* is a member of the carbonic anhydrase family of zinc-containing metalloenzymes, whose primary function is to catalyze the reversible conversion of carbon dioxide to bicarbonate. The mitochondrial enzyme *CA5A* plays an important role in supplying bicarbonate (HCO_{3-}) to numerous other mitochondrial enzymes. In a previous study, we observed downregulation of *CA5A* in the ovaries of Texel compared to F1 (Pokharel et al., 2018). More recently, *CA5A* was also shown to be expressed in the ovaries of the Pelibuey breed of sheep; the gene was upregulated in a subset of ewes that gave birth to two lambs compared to uniparous animals (Hernández-Montiel et al., 2019). However, there are no reports regarding the expression and function of *CA5A* in the CL. Based on the results from our current and earlier reports (Pokharel et al., 2018; Hernández-Montiel et al., 2019), *CA5A* appears to have an important function, at least until the preimplantation stage of reproduction. The level of expression in F1 crosses in the CL and endometrium followed the same pattern as that in the ovary, which led us to conclude that *CA5A* is heritable and potentially an imprinted gene. Further experiments are needed to determine whether the gene is associated with high prolificacy. Out of the 49 upregulated genes, 24

genes had available functional annotations and were associated with nine different GO terms (Fig. 5, Supplementary table S8). The majority of the GO terms were related to transport ("anion transport", "lipid transport", "organic anion transport", and "fatty acid transport") and regulation ("regulation of lipid transport", "regulation of homotypic cell–cell adhesion", "negative regulation of T cell activation", and "regulation of lipid localization").

Altogether, 20 out of the 49 genes upregulated in Finnsheep vs F1 crosses were linked to KEGG pathways. Based on the selection criteria, only two KEGG pathways, namely, "complement and coagulation cascades" (associated genes *C5AR1*, *F13A1*, and *VSIG4*) and "Fc gamma R-mediated phagocytosis" (associated genes: *FCGR1A*, *SCIN*, and *SYK*), were identified. The lowest number (n=22) of DEGs was observed between Texel and F1 crossbred ewes, with 13 genes being upregulated in Texel (Supplementary table S9).

We noticed that several genes were differentially expressed in more than one comparison, increasing our confidence in the identification of these DEGs. Few DEGs were exclusively up- or downregulated in particular breed compared to other two breeds. For example, a transcript encoding a miRNA (*ENSOARG00000022916*) was found to be always upregulated in Finnsheep compared to both Texel and F1 crosses. *HNRNPK* was always downregulated in Finnsheep, and *CA5A* was always upregulated in F1 crosses compared to the other two breeds. Similarly, coiled-coil domain-containing 73 (*CCDC73*) and a pseudogene (*ENSOARG00000020196*) were upregulated in Texel compared to F1 crosses. A lincRNA (*ENSOARG00000025875*) was downregulated in Finnsheep compared to Texel. *MICA** was upregulated in Finnsheep compared to Texel. Oxidized low-density lipoprotein receptor 1 (*OLR1*), NLR family apoptosis inhibitory protein (*NAIP*), macrophage scavenger receptor 1 (*MSR1*), high-affinity Ig gamma Fc receptor 1 precursor (*FCGR1A*), hemoglobin subunit alpha-1/2, folate receptor 3 (*FLOR3**), Fc gamma 2 receptor, chromogranin B (*CHGB*), Siglec-14, clavsin 2 (*CLVS2*), copine 4 (*CPNE4*), EPH receptor B6 (*EPBH6**) and *MICA** were exclusively upregulated in the CLs of Finnsheep. Gastrula zinc finger XICGF17.1-like (*LOC10562107*), crystallin mu (*CRYM*), myeloid-associated differentiation marker-like (*LOC101119079*) and tolloid-like 2 (*TLL2*) were downregulated in the CLs of Texel compared to the other two breeds. These results also indicated that F1 crosses were more similar to Finnsheep than Texel crosses. *CA5A* appeared to be upregulated in F1 crosses compared to both Finnsheep and Texel from both phases (i.e., in the CL in this study and in the ovary in our earlier study).

4.6 miRNAs expressed in the dataset

A total of 336.6 M reads were sequenced, of which approximately 42% contained adapters and/or low-quality bases. After trimming, more than 92% of the reads (n=311.3 M) were retained as high-quality clean reads. On average, collapsing of duplicate reads revealed 483,096 unique reads per sample, of which 54.4% of the unique sequences (collapsed reads) were mapped to the ovine reference genome. The detailed summary statistics for each sample are shown in Supplementary table S10. There were more collapsed reads and uniquely mapped reads for endometrial samples than CL samples despite the similar numbers of raw and clean reads in both tissues. After filtering out low-count (<10 reads) and ambiguous reads, a total of 599 miRNAs were included in the expression analysis. All the miRNAs quantified in this study are presented in Supplementary table S11 and have been sent to be considered for adding in the next release of miRBase. The majority of the expressed miRNAs (n=524) were shared by both tissues, with 43 and 32 miRNAs being unique to the CL and endometrium, respectively. Out of 599 miRNAs, 60 were conserved miRNAs in other species while 123 were known sheep miRNAs. Currently, 153 miRNAs are available in the miRBase database (Kozomara et al., 2019). The database was updated to the current version (22) from an earlier version (miRBase 21) after four years, and the overall number of miRNA sequences increased by over a

third. However, the number of sheep miRNAs remained the same. Moreover, studies that produce miRNA datasets have been scarce. As of April 2019, miRNA datasets from only three studies were available in the European Nucleotide Archive (ENA) database, with accession codes PRJNA308631 (n=3), PRJEB22101 (n=37) and PRJNA414087 (n=40); the PRJEB22101 dataset was from the first phase of this study (Pokharel et al., 2018). In the current study, we quantified over threefold more sheep miRNAs (n=599) than are available in miRBase. Therefore, these miRNAs will certainly improve the existing resources and will be valuable in future studies.

We did not perform differential gene expression analysis on the endometrial samples because of the sampling bias. Two miRNAs, both upregulated in Finnsheep, were significantly differentially expressed between the pure breeds in the CL, while the other comparisons did not reveal any significantly differentially expressed miRNAs. Of these two significantly differentially expressed miRNAs, rno-miR-451-5p is a conserved miRNA similar to one found in rats (*Rattus norvegicus*). The other, oar-18_757_mt, is a novel miRNA expressed on chromosome 18. Chromosomal placement of the quantified miRNAs revealed a large cluster of miRNAs on chromosome 9 that we also observed in the ovaries (Fig.5).

4.7 Limitations and thoughts for future studies

We acknowledge certain limitations of this study. We believe that with the availability of a better annotated reference genome, the data from this study will reveal additional information that we may have missed in this analysis. With sequencing costs becoming increasingly inexpensive, increasing the sample size of each breed group would certainly add statistical power. Given that time-series experiments are not feasible with the same animal, sampling could be performed with a larger group of animals at different stages of pregnancy to obtain an overview of gene expression changes. One ovary from each ewe was removed earlier (Pokharel et al., 2018) and all the CLs for this study was collected from the remaining ovary. Therefore, there might be some impact due to possible negative feedback effects on overall gene expression. It should be noted that overall gene expression and, more specifically, differential expression between breeds is inherently a stochastic process; thus, there is always some level of bias caused by individual variation (Hansen et al., 2011). After noticing the bias caused by pregnancy length on endometrium, we were unable to proceed further with breed-wise differential expression analyses. Including more individuals in future experiments as well as more controlled tissue sampling will minimize such bias. We are also aware that the overall gene expression profiles may have been affected by the absence of one CL that was removed for earlier study, (Pokharel et al., 2018). Having said that, as CL was removed from all ewes in current study, we do not expect any bias in the gene expression comparison. The results from breeding experiments have shown that productivity traits such as litter sizes may not carry on to F2 crosses (F1 x F1) and/or backcrosses. Therefore, future experiments that involve F2 crosses and backcrosses would provide more valuable findings related to prolificacy. Moreover, by doing a reciprocal cross experiment, we might be able to get insight into POE and measure the potential contribution of the Texel and Finnsheep in each cross. In addition, replicating such experiments in different environments would be relevant for breeding strategies to mitigate the effects of climate change. To minimize or alleviate noise from tissue heterogeneity, single-cell experiments may prove beneficial in future studies. While we observed interplay between the endometrium and the CL, it would be equally interesting to measure the transcriptional patterns in the embryo. Finally, the application of gene-modifying technologies such as CRISPR/Cas9 to edit certain regions (such as the region in the FecL locus homologous to the partial retrovirus sequence) may provide important insights into phenotypes associated with infertility, prolificacy and other traits of interest.

5 Conclusion

We compiled the most comprehensive list thus far of genes (n=21,287) and miRNAs (n=599) expressed in the CL and endometrium, which are the most important tissues during the preimplantation stage and therefore determine the success of pregnancy in sheep. Our results agree well with the (limited) existing reports, which are mainly focused on the interplay of the endometrium and conceptus, but we have shown that the CL plays an equally important role. The relative scarcity of transcriptomic information about the CL means that its functional importance is underrated. We identified several key transcripts, including coding genes (producing mRNA) and noncoding genes (miRNAs, snoRNAs, and lincRNAs), that are essential during early pregnancy. Functional analysis primarily based on literature searches and earlier studies revealed the significant roles of the most highly expressed genes in pregnancy recognition, implantation and placentation. F1 crosses were more closely related to Finnsheep than to Texel, as indicated by phenotypic and gene expression results that need to be validated with additional experiments (with F2 crosses and backcrosses). Several genes with potential importance during early pregnancy (including *SIGLEC13*, *SIGLEC14*, *SIGLEC6*, *MRP4*, and *CA5A*) were reported in the CL for the first time in any species. The roles of retroviruses during early pregnancy and in breed-specific phenotypes were indicated by the observed gene expression dynamics, especially in the endometrium. A novel gene sharing similarity with an ERV was identified in the FecL locus. The results from this study show the importance of the immune system during early pregnancy. We also highlight the need for improved annotation of the sheep genome and emphasize that our data will certainly contribute to such improvement. We observed a cluster of miRNAs on chromosome 18 homologous to that found on chromosome 14 in humans. Taken together, our data provide new information to aid in understanding the complex reproductive events during the preimplantation period in sheep and may also have implications for other ruminants (such as goats and cattle) and mammals, including humans.

6 Abbreviations

Ig (immunoglobulin), Siglec (sialic acid-binding Ig-like lectin), ERV (endogenous retrovirus), CDS (coding sequence), OXT (oxytocin), MRP (multidrug resistance-associated protein), CL (corpus luteum), lincRNA (long intergenic noncoding RNA), TPM (Transcripts Per Million)

7 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

8 Author Contributions

J.K. and M.H.L. conceived and designed the project. J.K., M.H. and J.P. collected the samples. K.P. analyzed the data and wrote the manuscript. J.P. contributed substantially to revising the manuscript. M.W. and J.K. contributed to the data analysis and manuscript writing, respectively. All authors revised and approved the final manuscript.

8.1 Funding

This study was funded by the Academy of Finland (decisions 250633, 250677 and 285017). This study is part of the ClimGen (“Climate Genomics for Farm Animal Adaptation”) project funded by FACCE-JPI ERA-NET Plus on Climate Smart Agriculture. K.P. acknowledges financial support from the Niemi Foundation.

9 Acknowledgments

We are grateful to Anu Tuomola for providing the experimental facilities on her farm in Urjala and to Kati Kaisaioki from Lallin Lammas Ltd. for providing the slaughtering facilities. We thank Johanna Rautiainen, Arja Seppälä, Annukka Numminen, Kalle Saastamoinen, Ilma Tapio, Tuula Marjatta Hamama, Anneli Virta and Magnus Andersson for their valuable assistance during this study. The authors wish to acknowledge the CSC – IT Center for Science, Finland, for computational resources. This study was supported by the Finnish Functional Genomics Centre of the University of Turku, Åbo Akademi and Biocenter Finland.

10 References

Angata, T., Margulies, E. H., Green, E. D., and Varki, A. (2004). Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mechanisms. *Proceedings of the National Academy of Sciences* 101, 13251–13256. doi:10.1073/pnas.0404833101.

Bastian, A., Kratzin, H., Fallgren-Gebauer, E., Eckart, K., and Hilschmann, N. (1995). "Intra- and Inter-Chain Disulfide Bridges of J Chain in Human S-IgA," in (Springer, Boston, MA), 581–583. doi:10.1007/978-1-4615-1941-6_122.

Bazer, F. W. (2013). Pregnancy recognition signaling mechanisms in ruminants and pigs. *Journal of Animal Science and Biotechnology*. doi:10.1186/2049-1891-4-23.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics (Oxford, England)* 25, 1091–3. doi:10.1093/bioinformatics/btp101.

Bishop, C. V. (2013). Progesterone inhibition of oxytocin signaling in endometrium. *Frontiers in Neuroscience*. doi:10.3389/fnins.2013.00138.

Bolet, G. (1986). "Timing and Extent of Embryonic Mortality in Pigs Sheep and Goats: Genetic Variability," in *Embryonic Mortality in Farm Animals* (Dordrecht: Springer Netherlands), 12–43. doi:10.1007/978-94-009-5038-2_2.

Bornhöfft, K. F., Goldammer, T., Rebl, A., and Galuska, S. P. (2018). Siglecs: A journey through the evolution of sialic acid-binding immunoglobulin-type lectins. *Developmental & Comparative Immunology* 86, 219–231. doi:10.1016/J.DCI.2018.05.008.

Branco, R., and Masle, J. (2019). Systemic signalling through TCTP1 controls lateral root formation in Arabidopsis. *Journal of Experimental Botany*. doi:10.1093/jxb/erz204.

Brioudes, F., Thierry, A.-M., Chambrier, P., Mollereau, B., and Bendahmane, M. (2010). Translationally controlled tumor protein is a conserved mitotic growth integrator in animals and plants. *Proceedings of the National Academy of Sciences* 107, 16384–16389. doi:10.1073/pnas.1007926107.

Brooks, K., Burns, G. W., Moraes, J. G. N., and Spencer, T. E. (2016). Analysis of the Uterine Epithelial and Conceptus Transcriptome and Luminal Fluid Proteome During the Peri-Implantation Period of Pregnancy in Sheep. *Biology of Reproduction* 95, 88–88. doi:10.1095/biolreprod.116.141945.

Cao, H., and Crocker, P. R. (2011). Evolution of CD33-related siglecs: regulating host immune functions and escaping pathogen exploitation? *Immunology* 132, 18–26. doi:10.1111/j.1365-2567.2010.03368.x.

Casey, O. M., Morris, D. G., Powell, R., Sreenan, J. M., and Fitzpatrick, R. (2005). Analysis of gene expression in non-regressed and regressed bovine corpus luteum tissue using a customized ovarian cDNA array. *Theriogenology* 64, 1963–1976. doi:10.1016/j.theriogenology.2005.04.015.

Chen, S. H., Wu, P.-S., Chou, C.-H., Yan, Y.-T., Liu, H., Weng, S.-Y., et al. (2007). A Knockout Mouse Approach Reveals that TCTP Functions as an Essential Factor for Cell Proliferation and Survival in a Tissue- or Cell Type–specific Manner. *Molecular Biology of the Cell* 18, 2525–2532. doi:10.1091/mbc.e07-02-0188.

Christenson, L. K., and Devoto, L. (2003). Cholesterol transport and steroidogenesis by the corpus luteum. *Reproductive biology and endocrinology : RB&E* 1, 90. doi:10.1186/1477-7827-1-90.

Chu, M. X., Mu, Y. L., Fang, L., Ye, S. C., and Sun, S. H. (2007). Prolactin receptor as a candidate gene for prolificacy of small tail Han sheep. *Animal Biotechnology* 18, 65–73. doi:10.1080/10495390601090950.

Coulam, C. B., and Goodman, C. (2000). Increased pregnancy rates after IVF/ET with intravenous immunoglobulin treatment in women with elevated circulating C56+ cells. *Early pregnancy (Online)* 4, 90–8.

Davis, G. H., Farquhar, P. A., O'Connell, A. R., Everett-Hincks, J. M., Wishart, P. J., Galloway, S. M., et al. (2006). A putative autosomal gene increasing ovulation rate in Romney sheep. *Animal reproduction science* 92, 65–73. doi:10.1016/j.anireprosci.2005.05.015.

Davis, G. H., Galloway, S. M., Ross, I. K., Gregan, S. M., Ward, J., Nimbkar, B. V, et al. (2002). DNA tests in prolific sheep from eight countries provide new evidence on origin of the Booroola (FecB) mutation. *Biology of reproduction* 66, 1869–74.

Davis, J. S., and LaVoie, H. A. (2018). *Molecular Regulation of Progesterone Production in the Corpus Luteum*. 3rd ed. Elsevier Inc. doi:10.1016/b978-0-12-813209-8.00015-7.

De Placido, G., Zullo, F., Mollo, A., Cappiello, F., Nazzaro, A., Colacurci, N., et al. (1994). Intravenous immunoglobulin (IVIG) in the prevention of implantation failures. *Annals of the New York Academy of Sciences* 734, 232–4.

DeMartini, J. C., Carlson, J. O., Leroux, C., Spencer, T., and Palmarini, M. (2003). Endogenous retroviruses related to jaagsiekte sheep retrovirus. *Current topics in microbiology and immunology* 275, 117–37.

Deniz, E., and Erman, B. (2017). Long noncoding RNA (lincRNA), a new paradigm in gene expression control. *Functional & Integrative Genomics* 17, 135–143. doi:10.1007/s10142-016-0524-x.

698 Denner, J. (2016). Expression and function of endogenous retroviruses in the placenta. *APMIS* 124,
699 31–43. doi:10.1111/apm.12474.

700 Devoto, L., Kohen, P., Gonzalez, R. R., Castro, O., Retamales, I., Vega, M., et al. (2001). Expression
701 of Steroidogenic Acute Regulatory Protein in the Human Corpus Luteum throughout the
702 Luteal Phase. *The Journal of Clinical Endocrinology & Metabolism* 86, 5633–5639.
703 doi:10.1210/jcem.86.11.7982.

704 Drouilhet, L., Mansanet, C., Sarry, J., Tabet, K., Bardou, P., Woloszyn, F., et al. (2013). The Highly
705 Prolific Phenotype of Lacaune Sheep Is Associated with an Ectopic Expression of the
706 B4GALNT2 Gene within the Ovary. *PLoS genetics* 9, e1003809.
707 doi:10.1371/journal.pgen.1003809.

708 Dunlap, K. A., Palmarini, M., Adelson, D. L., and Spencer, T. E. (2005). Sheep Endogenous
709 Betaretroviruses (enJSRVs) and the Hyaluronidase 2 (HYAL2) Receptor in the Ovine Uterus
710 and Conceptus. *Biology of Reproduction* 73, 271–279. doi:10.1095/biolreprod.105.039776.

711 Dunlap, K. A., Palmarini, M., and Spencer, T. E. (2006a). Ovine Endogenous Betaretroviruses
712 (enJSRVs) and Placental Morphogenesis. *Placenta* 27, 135–140.
713 doi:10.1016/j.placenta.2005.12.009.

714 Dunlap, K. A., Palmarini, M., Varela, M., Burghardt, R. C., Hayashi, K., Farmer, J. L., et al. (2006b).
715 Endogenous retroviruses regulate periimplantation placental growth and differentiation.
716 *Proceedings of the National Academy of Sciences of the United States of America* 103,
717 14390–5. doi:10.1073/pnas.0603836103.

718 Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., et al. (2005). BioMart
719 and Bioconductor: a powerful link between biological databases and microarray data analysis.
720 *Bioinformatics (Oxford, England)* 21, 3439–40. doi:10.1093/bioinformatics/bti525.

721 Felix Krueger Trim Galore. Available at:
722 http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [Accessed January 9, 2019].

723 Fleming, J. A. G. W., Spencer, T. E., Safe, S. H., and Bazer, F. W. (2006). Estrogen regulates
724 transcription of the ovine oxytocin receptor gene through GC-rich SP1 promoter elements.
725 *Endocrinology*. doi:10.1210/en.2005-1120.

726 Flint, A. P. F., and Sheldrick, E. L. (2004). Ovarian oxytocin and the maternal recognition of
727 pregnancy. *Reproduction*. doi:10.1530/jrf.0.0760831.

728 Forde, N., and Lonergan, P. (2017). Interferon-tau and fertility in ruminants. *Reproduction* 154, F33–
729 F43. doi:10.1530/REP-17-0432.

730 Forde, N., Mehta, J. P., McGettigan, P. A., Mamo, S., Bazer, F. W., Spencer, T. E., et al. (2013).
731 Alterations in expression of endometrial genes coding for proteins secreted into the uterine
732 lumen during conceptus elongation in cattle. *BMC Genomics* 14, 321. doi:10.1186/1471-
733 2164-14-321.

734 Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2
735 accurately identifies known and hundreds of novel microRNA genes in seven animal clades.
736 *Nucleic acids research* 40, 37–52. doi:10.1093/nar/gkr688.

737 Geisert, R. D., Morgan, G. L., Short, E. C., and Zavy, M. T. (1992). Endocrine events associated with
738 endometrial function and conceptus development in cattle. *Reproduction, fertility, and*
739 *development* 4, 301–5.

740 Gimpl, G., Wiegand, V., Burger, K., and Fahrenholz, F. (2002). Cholesterol and steroid hormones:
741 Modulators of oxytocin receptor function. in *Progress in Brain Research* doi:10.1016/S0079-
742 6123(02)39006-X.

743 Gray, C. A., Abbey, C. A., Beremand, P. D., Choi, Y., Farmer, J. L., Adelson, D. L., et al. (2006).
744 Identification of Endometrial Genes Regulated by Early Pregnancy, Progesterone, and
745 Interferon Tau in the Ovine Uterus. *Biology of Reproduction* 74, 383–394.
746 doi:10.1095/biolreprod.105.046656.

747 Gray, C. A., Adelson, D. L., Bazer, F. W., Burghardt, R. C., Meeusen, E. N. T., and Spencer, T. E.
748 (2004). Discovery and characterization of an epithelial-specific galectin in the endometrium
749 that forms crystals in the trophectoderm. *Proceedings of the National Academy of Sciences*
750 101, 7982–7987. doi:10.1073/pnas.0402669101.

751 Grazzini, E., Guillon, G., Mouillac, B., and Zingg, H. H. (1998). Inhibition of oxytocin receptor
752 function by direct binding of progesterone. *Nature*. doi:10.1038/33176.

753 Hanrahan, J. P., Gregan, S. M., Mulsant, P., Mullen, M., Davis, G. H., Powell, R., et al. (2004).
754 Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated
755 with both increased ovulation rate and sterility in Cambridge and Belclare sheep (Ovis aries).
756 *Biology of reproduction* 70, 900–9. doi:10.1095/biolreprod.103.023093.

757 Hanrahan, J. P., and Quirke, J. F. (1984). Contribution of variation in ovulation rate and embryo
758 survival to within breed variation in litter size. *Genetics of reproduction in sheep / edited by*
759 *R.B. Land, D.W. Robinson*.

760 Hansen, K. D., Wu, Z., Irizarry, R. A., and Leek, J. T. (2011). Sequencing technology does not
761 eliminate biological variability. *Nature Biotechnology* 29, 572–573. doi:10.1038/nbt.1910.

762 Hernández-Montiel, W., Collí-Dula, R. C., Ramón-Ugalde, J. P., Martínez-Núñez, M. A., and
763 Zamora-Bustillos, R. (2019). RNA-seq Transcriptome Analysis in Ovarian Tissue of Pelibuey
764 Breed to Explore the Regulation of Prolificacy. *Genes* 10, 358. doi:10.3390/genes10050358.

765 Honjo, T. (1983). Immunoglobulin Genes. *Annual Review of Immunology* 1, 499–528.
766 doi:10.1146/annurev.iy.01.040183.002435.

767 Hu, J., Zhang, Z., Shen, W.-J., and Azhar, S. (2010). Cellular cholesterol delivery, intracellular
768 processing and utilization for biosynthesis of steroid hormones. *Nutrition & metabolism* 7, 47.
769 doi:10.1186/1743-7075-7-47.

770 Hu, X., Pokharel, K., Peippo, J., Ghanem, N., Zhaboyev, I., Kantanen, J., et al. (2015). Identification
771 and characterization of miRNAs in the ovaries of a highly prolific sheep breed. *Animal*
772 *genetics*, n/a-n/a. doi:10.1111/age.12385.

773 Juengel, J. L., Meberg, B. M., Turzillo, A. M., Nett, T. M., and Niswender, G. D. (1995). Hormonal
774 regulation of messenger ribonucleic acid encoding steroidogenic acute regulatory protein in
775 ovine corpora lutea. *Endocrinology* 136, 5423–5429. doi:10.1210/endo.136.12.7588291.

776 Kendrick, Keith. M. (2000). Oxytocin, motherhood and bonding. *Experimental Physiology* 85, 111s–
777 124s. doi:10.1111/j.1469-445x.2000.tb00014.x.

778 Kfir, S., Basavaraja, R., Wigoda, N., Ben-Dor, S., Orr, I., and Meidan, R. (2018). Genomic profiling
779 of bovine corpus luteum maturation. *PLOS ONE* 13, e0194456.
780 doi:10.1371/journal.pone.0194456.

781 Kim, S., Choi, Y., Bazer, F. W., and Spencer, T. E. (2003). Identification of Genes in the Ovine
782 Endometrium Regulated by Interferon τ Independent of Signal Transducer and Activator of
783 Transcription. *Endocrinology* 144, 5203–5214. doi:10.1210/en.2003-0665.

784 King, S. R., and LaVoie, H. A. (2009). "Regulation of the Early Steps in Gonadal Steroidogenesis,"
785 in *Reproductive Endocrinology* (Boston, MA: Springer US), 175–193. doi:10.1007/978-0-
786 387-88186-7_16.

787 Koch, J. M., Ramadoss, J., and Magness, R. R. (2010). Proteomic profile of uterine luminal fluid
788 from early pregnant ewes. *Journal of Proteome Research* 9, 3878–3885.
789 doi:10.1021/pr100096b.

790 Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences
791 to function. *Nucleic Acids Research* 47, D155–D162. doi:10.1093/nar/gky1141.

792 Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-
793 sequencing data. *Nucleic Acids Research* 39, D152–D157. doi:10.1093/nar/gkq1027.

794 Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs
795 using deep sequencing data. *Nucleic Acids Research* 42, D68–D73. doi:10.1093/nar/gkt1181.

796 Lacroix-Pépin, N., Danyod, G., Krishnaswamy, N., Mondal, S., Rong, P.-M., Chapdelaine, P., et al.
797 (2011). The Multidrug Resistance-Associated Protein 4 (MRP4) Appears as a Functional
798 Carrier of Prostaglandins Regulated by Oxytocin in the Bovine Endometrium. *Endocrinology*
799 152, 4993–5004. doi:10.1210/en.2011-1406.

800 Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient
801 alignment of short DNA sequences to the human genome. *Genome biology* 10, R25.
802 doi:10.1186/gb-2009-10-3-r25.

803 Lewis, S. K., Farmer, J. L., Burghardt, R. C., Newton, G. R., Johnson, G. A., Adelson, D. L., et al.
804 (2007). Galectin 15 (LGALS15): A Gene Uniquely Expressed in the Uteri of Sheep and
805 Goats that Functions in Trophoblast Attachment1. *Biology of Reproduction* 77, 1027–1036.
806 doi:10.1095/biolreprod.107.063594.

Li, S., Chen, X., Ding, Y., Liu, X., Wang, Y., and He, J. (2011). Expression of translationally controlled tumor protein (TCTP) in the uterus of mice of early pregnancy and its possible significance during embryo implantation. *Human Reproduction* 26, 2972–2980. doi:10.1093/humrep/der275.

Loeser, R. F., and Wallin, R. (1992). Cell adhesion to matrix Gla protein and its inhibition by an Arg-Gly-Asp-containing peptide. *The Journal of biological chemistry* 267, 9459–62.

Love, M., Anders, S., and Huber, W. (2013). Differential expression of RNA-Seq data at the gene level – the DESeq2 package.

Madeira, F., Park, Y, M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 2–4.

Mamo, S., Mehta, J. P., Forde, N., McGettigan, P., and Lonergan, P. (2012). Conceptus-Endometrium Crosstalk During Maternal Recognition of Pregnancy in Cattle. *Biology of Reproduction* 87, 6, 1–9. doi:10.1095/biolreprod.112.099945.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12.

Mateyak, M. K., and Kinzy, T. G. (2010). eEF1A: thinking outside the ribosome. *The Journal of biological chemistry* 285, 21209–13. doi:10.1074/jbc.R110.113795.

Moore, S. G., Pryce, J. E., Hayes, B. J., Chamberlain, A. J., Kemper, K. E., Berry, D. P., et al. (2016). Differentially Expressed Genes in Endometrium and Corpus Luteum of Holstein Cows Selected for High and Low Fertility Are Enriched for Sequence Variants Associated with Fertility1. *Biology of Reproduction* 94. doi:10.1095/biolreprod.115.132951.

Mullen, M. P., and Hanrahan, J. P. (2014). Direct evidence on the contribution of a missense mutation in GDF9 to variation in ovulation rate of Finnsheep. *PloS one* 9, e95251. doi:10.1371/journal.pone.0095251.

Mulsant, P., Lecerf, F., Fabre, S., Schibler, L., Monget, P., Lanneluc, I., et al. (2001). Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Mérino ewes. *Proceedings of the National Academy of Sciences of the United States of America* 98, 5104–9. doi:10.1073/pnas.091577598.

Palmarini, M., Gray, C. A., Carpenter, K., Fan, H., Bazer, F. W., and Spencer, T. E. (2001). Expression of Endogenous Betaretroviruses in the Ovine Uterus: Effects of Neonatal Age, Estrous Cycle, Pregnancy, and Progesterone. *Journal of Virology* 75, 11319–11327. doi:10.1128/JVI.75.23.11319-11327.2001.

Pantano, L., Estivill, X., and Martí, E. (2011). A non-biased framework for the annotation and classification of the non-miRNA small RNA transcriptome. *Bioinformatics* 27, 3202–3203. doi:10.1093/bioinformatics/btr527.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* 14, 417–419. doi:10.1038/nmeth.4197.

845 Pauli, A., Rinn, J. L., and Schier, A. F. (2011). Non-coding RNAs as regulators of embryogenesis.
846 *Nature reviews. Genetics* 12, 136–49. doi:10.1038/nrg2904.

847 Pillai, S., Netravali, I. A., Cariappa, A., and Mattoo, H. (2012). Siglecs and immune regulation.
848 *Annual review of immunology* 30, 357–92. doi:10.1146/annurev-immunol-020711-075018.

849 Plant, T. M., Zeleznik, A. J., and Auchus, R. J. (2015). "Chapter 8 – Human Steroid Biosynthesis," in
850 *Knobil and Neill's Physiology of Reproduction*, 295–312. doi:10.1016/B978-0-12-397175-
851 3.00008-9.

852 Pokharel, K., Peippo, J., Honkatukia, M., Seppälä, A., Rautiainen, J., Ghanem, N., et al. (2018).
853 Integrated ovarian mRNA and miRNA transcriptome profiling characterizes the genetic basis
854 of prolificacy traits in sheep (Ovis aries). *BMC Genomics* 19. doi:10.1186/s12864-017-4400-
855 4.

856 Pokharel, K., Weldenegodguad, M., Popov, R., Honkatukia, M., Huuki, H., Lindeberg, H., et al.
857 (2019). Whole blood transcriptome analysis reveals footprints of cattle adaptation to
858 sub□arctic conditions. *Animal Genetics* 50, 217–227. doi:10.1111/age.12783.

859 Quinlivan, T. D., Martin, C. A., Taylor, W. B., and Cairney, I. M. (1966). Estimates of pre- and
860 perinatal mortality in the New Zealand Romney Marsh ewe. I. Pre- and perinatal mortality in
861 those ewes that conceived to one service. *Journal of reproduction and fertility* 11, 379–90.

862 Raheem, K. A. (2017). An insight into maternal recognition of pregnancy in mammalian species.
863 *Journal of the Saudi Society of Agricultural Sciences* 16, 1–6.
864 doi:10.1016/j.jssas.2015.01.002.

865 Ransohoff, J. D., Wei, Y., and Khavari, P. A. (2018). The functions and unique features of long
866 intergenic non-coding RNA. *Nature reviews. Molecular cell biology* 19, 143–157.
867 doi:10.1038/nrm.2017.104.

868 Rhind, S. M., Robinson, J. J., Fraser, C., and McHattie, I. (1980). Ovulation and embryo survival
869 rates and plasma progesterone concentrations of prolific ewes treated with PMSG. *J. Reprod.*
870 *Fertil.* 58, 139–144. doi:10.1530/jrf.0.0580139.

871 Rickard, J. P., Ryan, G., Hall, E., Graaf, S. P. de, and Hermes, R. (2017). Using transrectal
872 ultrasound to examine the effect of exogenous progesterone on early embryonic loss in sheep.
873 *PLOS ONE* 12, e0183659. doi:10.1371/journal.pone.0183659.

874 Ricordeau, G., Thimonier, J., Poivey, J. P., Driancourt, M. A., Hochereau-De-Reviers, M. T., and
875 Tchamitchian, L. (1990). I.N.R.A. research on the Romanov sheep breed in France: A review.
876 *Livestock Production Science* 24, 305–332. doi:10.1016/0301-6226(90)90009-U.

877 Romero, J. J., Liebig, B. E., Broeckling, C. D., Prenni, J. E., and Hansen, T. R. (2017). Pregnancy-
878 induced changes in metabolome and proteome in ovine uterine flushings. *Biology of*
879 *Reproduction* 97, 273–287. doi:10.1093/biolre/iox078.

880 Satterfield, M. C., Bazer, F. W., and Spencer, T. E. (2006). Progesterone Regulation of
881 Preimplantation Conceptus Growth and Galectin 15 (LGALS15) in the Ovine Uterus. *Biology*
882 *of Reproduction* 75, 289–296. doi:10.1095/biolreprod.106.052944.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498–504. doi:10.1101/gr.1239303.

Silva, C. L. A. D., Brand, H. van den, Laurenssen, B. F. A., Broekhuijse, M. L. W. J., Knol, E. F., Kemp, B., et al. (2016). Relationships between ovulation rate and embryonic and placental characteristics in multiparous sows at 35 days of pregnancy. *animal* 10, 1192–1199. doi:10.1017/S175173111600015X.

Simon Andrews FastQC A Quality Control tool for High Throughput Sequence Data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed January 18, 2013].

Sistiaga-Poveda, M., and Jugo, B. M. (2014). Evolutionary dynamics of endogenous Jaagsiekte sheep retroviruses proliferation in the domestic sheep, mouflon and Pyrenean chamois. *Heredity* 112, 571–8. doi:10.1038/hdy.2013.136.

Soneson, C., Love, M. I., and Robinson, M. D. (2016). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* 4, 1521. doi:10.12688/f1000research.7563.2.

Souza, C. J., MacDougall, C., MacDougall, C., Campbell, B. K., McNeilly, A. S., and Baird, D. T. (2001). The Booroola (FecB) phenotype is associated with a mutation in the bone morphogenetic receptor type 1 B (BMPR1B) gene. *The Journal of endocrinology* 169, R1-6.

Spencer, T. E., and Bazer, F. W. (2004). Conceptus signals for establishment and maintenance of pregnancy. *Reproductive biology and endocrinology : RB&E* 2, 49. doi:10.1186/1477-7827-2-49.

Spencer, T. E., Burghardt, R. C., Johnson, G. A., and Bazer, F. W. (2004a). Conceptus signals for establishment and maintenance of pregnancy. *Animal Reproduction Science* 82–83, 537–550. doi:10.1016/j.anireprosci.2004.04.014.

Spencer, T. E., Forde, N., and Lonergan, P. (2015). The role of progesterone and conceptus-derived factors in uterine biology during early pregnancy in ruminants. *Journal of Dairy Science* 99, 5941–5950. doi:10.3168/jds.2015-10070.

Spencer, T. E., Johnson, G. A., Bazer, F. W., and Burghardt, R. C. (2004b). Implantation mechanisms: insights from the sheep. *Reproduction (Cambridge, England)* 128, 657–68. doi:10.1530/rep.1.00398.

Spencer, T. E., Johnson, G. A., Bazer, F. W., Burghardt, R. C., and Palmarini, M. (2007). Pregnancy recognition and conceptus implantation in domestic ruminants: roles of progesterone, interferons and endogenous retroviruses. *Reproduction, fertility, and development* 19, 65–78.

Spencer, T. E., and Palmarini, M. (2012). Endogenous retroviruses of sheep: a model system for understanding physiological adaptation to an evolving ruminant genome. *The Journal of reproduction and development* 58, 33–7.

Spencer, T. E., Sandra, O., and Wolf, E. (2008). Genes involved in conceptus–endometrial interactions in ruminants: insights from reductionism and thoughts on holistic approaches. *REPRODUCTION* 135, 165–179. doi:10.1530/REP-07-0327.

Spencer, T. E., Stagg, A. G., Joyce, M. M., Jenster, G., Wood, C. G., Bazer, F. W., et al. (1999). Discovery and Characterization of Endometrial Epithelial Messenger Ribonucleic Acids Using the Ovine Uterine Gland Knockout Model [1]. *Endocrinology* 140, 4070–4080. doi:10.1210/endo.140.9.6981.

Stocco, D. M. (2000). The role of the StAR protein in steroidogenesis: challenges for the future. *The Journal of endocrinology* 164, 247–53.

Stouffer, R. L., and Hennebold, J. D. (2015). "Structure, Function, and Regulation of the Corpus Luteum," in *Knobil and Neill's Physiology of Reproduction*, 1023–1076. doi:10.1016/B978-0-12-397175-3.00023-5.

Tatsuka, M., Mitsui, H., Wada, M., Nagata, A., Nojima, H., and Okayama, H. (1992). Elongation factor-1α gene determines susceptibility to transformation. *Nature* 359, 333–336. doi:10.1038/359333a0.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* 28, 511–5. doi:10.1038/nbt.1621.

Tuynder, M., Susini, L., Prieur, S., Besse, S., Fiucci, G., Amson, R., et al. (2002). Biological models and genes of tumor reversion: cellular reprogramming through tpt1/TCTP and SIAH-1. *Proceedings of the National Academy of Sciences of the United States of America* 99, 14976–81. doi:10.1073/pnas.222470799.

Ulitsky, I., and Bartel, D. P. (2013). lincRNAs: genomics, evolution, and mechanisms. *Cell* 154, 26–46. doi:10.1016/j.cell.2013.06.020.

Våge, D. I., Husdal, M., Kent, M. P., Klemetsdal, G., and Boman, I. a (2013). A missense mutation in growth differentiation factor 9 (GDF9) is strongly associated with litter size in sheep. *BMC genetics* 14, 1. doi:10.1186/1471-2156-14-1.

Varki, A., and Angata, T. (2006). Siglecs—the major subfamily of I-type lectins. *Glycobiology* 16, 1R-27R. doi:10.1093/glycob/cwj008.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi:10.1093/bioinformatics/btp033.

Williams, A. F., and Barclay, A. N. (1988). The Immunoglobulin Superfamily—Domains for Cell Surface Recognition. *Annual Review of Immunology* 6, 381–405. doi:10.1146/annurev.iy.06.040188.002121.

Xu, S.-S., Gao, L., Xie, X.-L., Ren, Y.-L., Shen, Z.-Q., Wang, F., et al. (2018). Genome-Wide Association Analyses Highlight the Potential for Different Genetic Mechanisms for Litter Size Among Sheep Breeds. *Frontiers in Genetics* 9, 118. doi:10.3389/fgene.2018.00118.

Yang, J., Li, X., Cao, Y.-H., Pokharel, K., Hu, X.-J., Chen, Z.-H., et al. (2018). Comparative mRNA and miRNA expression in European mouflon (Ovis musimon) and sheep (Ovis aries) provides novel insights into the genetic mechanisms for female reproductive success. *Heredity*. doi:10.1038/s41437-018-0090-1.

Zhao, J., and Nishimoto, S. K. (1996). Matrix Gla protein gene expression is elevated during postnatal development. *Matrix biology : journal of the International Society for Matrix Biology* 15, 131–40.

Zhao, J., and Warburton, D. (1997). Matrix Gla protein gene expression is induced by transforming growth factor-beta in embryonic lung culture. *American Journal of Physiology-Lung Cellular and Molecular Physiology* 273, L282–L287. doi:10.1152/ajplung.1997.273.1.L282.

11 FIGURE LEGENDS

Figure 1 Venn diagram showing the distribution of genes expressed in the A) CL and B) endometrium of Finnsheep, Texel and F1 crosses.

Figure 2 Sample relatedness. (A) PCA plot of the top 500 expressed genes in the CL (left) and endometrium (right) and (B) heatmap of the top 25 most variable genes across all samples. Tissue-specific samples are denoted with a trailing c (for CL) or e (for endometrium). Legend: FS – Finnsheep; TX – Texel ewes, F1 – F1 crosses of Finnsheep and Texel sheep

Figure 3 GO terms associated with the list of genes that were upregulated in the CLs of Finnsheep compared to Texel

Figure 4 Multiple sequence alignment (partial) of the novel endogenous retrovirus gene. (A) The novel ERV identified in this study belongs to FecL locus and is located between *B4GALNT2* and Enzrin-like protein. (B) The multiple sequence alignment was prepared using Clustal Omega (Madeira et al., 2019) based on novel ERV (nERV, ENSOARG00000009959), ovine endogenous-virus beta-2 pro/pol region, partial sequence (kERV, AY193894.1), *Ovis canadensis canadensis* isolate 43U chromosome 17 sequence (OC43U, CP011902.1) and the reverse complement of reduced FecL locus (RFecL, KC352617). The bases are colored based on the nucleotide coloring scheme in Jalview (Waterhouse et al., 2009).

Figure 5 miRNA clusters in sheep (Chr. 18, top) and humans (Chr. 14, bottom). Only three (marked in black font color) out of 46 miRNAs in this cluster were not expressed in our data.

12 TABLES

Table 1: Top 25 genes (ranked by TPM) exclusively expressed in CL and endometrium.

	TPM	Gene name	Chr.	Gene description
CL				
ENSOARG00000003867	2080.2	CYP11A1	18	cytochrome P450, family 11, subfamily A, polypepti

ENSOARG00000009107	1233.5	CCL21	2	C-C motif chemokine ligand 21
ENSOARG00000013402	1004.3	PTGFR	1	prostaglandin F receptor
ENSOARG00000013340	622.3	CCL26	24	C-C motif chemokine ligand 26
ENSOARG00000019663	424.4	PTHLH	3	parathyroid hormone like hormone
ENSOARG00000015144	273.6	SERPINA5	18	serpin family A member 5
ENSOARG00000004774	248.0	LOC101114790*	11	C-C motif chemokine 15
ENSOARG00000014882	225.6	SERPINA1	18	serpin family A member 1
ENSOARG00000009230	201.1	DPT	12	dermatopontin
ENSOARG00000016052	188.9	AOX1	2	aldehyde oxidase 1
ENSOARG00000008189	177.9	PKIB	8	cAMP-dependent protein kinase inhibitor beta
ENSOARG00000004455	177.6	LHCGR	3	luteinizing hormone/choriogonadotropin receptor
ENSOARG00000020188	160.0	FAM110B*	9	Ovis aries family with sequence similarity 110 meml
ENSOARG00000009597	151.2	HS6ST2*	X	heparan sulfate 6-O-sulfotransferase 2
ENSOARG00000000474	143.3	GJA4	1	gap junction protein alpha 4
ENSOARG00000002475	131.9	PEG10	4	paternally expressed 10
ENSOARG00000010344	130.1	LTBP1	3	latent transforming growth factor beta binding protei
ENSOARG00000020976	118.9	GLDN	7	gliomedin
ENSOARG00000014966	116.3	INSL3	5	insulin like 3
ENSOARG00000009887	111.7	CHST15	22	carbohydrate sulfotransferase 15
ENSOARG00000017273	108.7	PROS1	1	protein S
ENSOARG00000020243	104.9	INHA	2	Ovis aries inhibin subunit alpha (INHA), mRNA.
ENSOARG00000019971	99.6	PLN	8	phospholamban
ENSOARG00000005118	92.8	KCNK12	3	potassium two pore domain channel subfamily K me
ENSOARG00000016448	90.1	TFR2	24	transferrin receptor 2
Endometrium				
ENSOARG00000004013	2378.2	FXYD4	25	FXYD domain containing ion transport regulator 4
ENSOARG00000012900	1429.8	IGFBP1	4	insulin like growth factor binding protein 1
ENSOARG00000002255	1191.2		19	acyl-coenzyme A thioesterase THEM4-like
ENSOARG00000019845	1078.1		3	enJSRV-20 endogenous virus Jaagsiekte sheep retro
ENSOARG00000015448	654.7	LOC114115259*	1	endogenous retrovirus group K member 6 Pro protei
ENSOARG00000014187	589.7	HamM*	19	endogenous virus Jaagsiekte sheep retrovirus
ENSOARG00000011311	544.3	CLDN7	11	claudin 7
ENSOARG00000010828	436.5	HAVCR1*	5	hepatitis A virus cellular receptor 1
ENSOARG00000004021	420.7	GRP*	23	gastrin releasing peptide
ENSOARG00000003992	389.0	SLC44A4	20	solute carrier family 44 member 4
ENSOARG00000019903	367.3		14	endogenous retrovirus group K member 9 Pol protei
ENSOARG00000004279	333.7	SLC7A9	14	solute carrier family 7 member 9
ENSOARG00000018755	327.1	TNNI1	12	troponin I1, slow skeletal type
ENSOARG00000013020	316.0	HSD11B1	12	hydroxysteroid 11-beta dehydrogenase 1
ENSOARG00000003224	300.3	SMPDL3B	2	sphingomyelin phosphodiesterase acid like 3B
ENSOARG00000012115	266.9		15	Jaagsiekte sheep retrovirus-like element
ENSOARG00000013158	263.8	SPDYC	21	speedy/RINGO cell cycle regulator family member C
ENSOARG00000006119	257.5	MMP7	15	matrix metallopeptidase 7
ENSOARG00000004751	246.5	TMEM92	11	transmembrane protein 92
ENSOARG00000010191	234.5	PEBP4	2	phosphatidylethanolamine binding protein 4
ENSOARG00000007680	232.5	SLC34A2	6	solute carrier family 34 member 2
ENSOARG00000004297	216.7	ALOX12*	11	arachidonate 12-lipoxygenase, 12S type

ENSOARG00000007592	211.8	PTGS2	12	prostaglandin-endoperoxide synthase 2
ENSOARG00000008009	211.1	CDH17	9	cadherin 17
ENSOARG00000014451	204.7	IFI27L2	18	interferon alpha-inducible protein 27-like protein 2

990

991 **Table 2: List of the 25 most abundant genes in the CL and endometrium**. Fifteen of the top 25
992 genes were shared by both tissues and were dominated by mitochondrial genes. The table includes
993 the Ensembl gene ID, chromosome number (Chr.), gene ID (GeneID) and gene description. The table
994 is divided into three sections; the first section lists the 25 genes that were shared by the two tissues,
995 and the other two list the remaining 10 genes in the endometrium and CL. Gene IDs and annotations
996 that were not available in BioMart were retrieved based on a homology search using the nucleotide
997 BLAST (marked with an asterisk, "*") or on information available in Ensembl (marked with a hash,
998 "#").

	Chr.	GeneID	Description
Common			
ENSOARG00000007815	6	LOC105580399*	Cercocebus atys 60S ribosomal protein L41-like
ENSOARG00000000019	MT	COX2	cytochrome c oxidase subunit II
ENSOARG00000018666	7	RPLP1	ribosomal protein lateral stalk subunit P1
ENSOARG00000000035	MT	CYTB	cytochrome b
ENSOARG00000000021	MT	ATP8	ATP synthase F0 subunit 8
ENSOARG00000007617	10	TCTP	tumor protein, translationally controlled 1
ENSOARG00000000023	MT	COX3	cytochrome c oxidase subunit III
ENSOARG00000003793	23	TSMB4X	thymosin beta-4
ENSOARG00000000037	MT	Mt tRNA#	mitochondrial tRNA
ENSOARG00000020724	3	MGP	matrix Gla protein
ENSOARG00000000016	MT	COX1	cytochrome c oxidase subunit I
ENSOARG00000000022	MT	ATP6	ATP synthase F0 subunit 6
ENSOARG00000000028	MT	ND4	NADH dehydrogenase subunit 4
ENSOARG00000003782	7	B2M	beta-2-microglobulin
ENSOARG00000000006	MT	ND1	NADH dehydrogenase subunit 1
Endometrium			
ENSOARG00000019088	AMGL01125506.1	LGALS15	lectin, galactoside-binding, soluble, 15
ENSOARG00000003184	24	NUPR1	nuclear protein 1, transcriptional regulator
ENSOARG00000019924	1	BCL2L15	BCL2-like 15
ENSOARG00000006202	13	CST3	cystatin C
ENSOARG00000021079	1	S100A11	S100 calcium-binding protein A11
ENSOARG00000016080	13	ATP5F1E	PRELI domain-containing 3B
ENSOARG00000013018	X	S100G	S100 calcium-binding protein G
ENSOARG00000019491	3	OST4	oligosaccharyltransferase complex subunit 4, noncatalytic
ENSOARG00000001346	21	CST6	cystatin E/M
ENSOARG00000006149	8	EEF1A1	eukaryotic translation elongation factor 1 alpha 1
CL			
ENSOARG00000004595	13	OXT	oxytocin/neurophysin I prepropeptide
ENSOARG00000022293	13	RF02216#	misc. RNA
ENSOARG00000000027	MT	ND4L	NADH dehydrogenase subunit 4L
ENSOARG00000002586	15	APOA1	apolipoprotein A1

ENSOARG00000002472	25	MSMB	microseminoprotein beta
ENSOARG00000001269	26	STAR	steroidogenic acute regulatory protein
ENSOARG00000000010	MT	ND2	NADH dehydrogenase subunit 2
ENSOARG00000013157	X	TIMP1	TIMP metallopeptidase inhibitor 1
ENSOARG00000020402	1	HSD3B1	hydroxy-delta-5-steroid dehydrogenase, 3 beta- a steroid delta-isomerase 1
ENSOARG00000000033	MT	ND6	NADH dehydrogenase subunit 6

Table 3: Numerical summary of differentially expressed genes in the CL and endometrium. Legend: FS – Finnsheep, TX – Texel, F1 – F1-cross

Comparison	**CL**		**Endometrium**	
	Upregulated	Downregulated	Upregulated	Downregulated
FS vs TX	140	59	22	21
FS vs F1	49	18	2	0
TX vs F1	13	9	5	12

13 Supplementary Material

Fig. S1: KEGG pathways associated with the DEGs upregulated in Finnsheep compared to Texel in the CL

Fig. S2: GO terms associated with DEGs upregulated in Finnsheep compared to F1 crosses

Fig. S3: PCA of the top 500 expressed miRNAs in the CL (left) and endometrium (right)

Table S1: Phenotype data of the samples

Table S2: Summary of the samples included in mRNA-Seq

Table S3: List of DEGs between the CLs of Finnsheep and Texel ewes

Table S4: List of DEGs between the CL and endometrium

Table S5: List of GO terms associated with the upregulated genes in the CLs of Finnsheep compared to Texel

Table S6: List of KEGG pathways associated with the upregulated genes in the CLs of Finnsheep compared to Texel

Table S7: List of DEGs between the CLs of Finnsheep and those of F1 crosses

Table S8: List of GO terms associated with the upregulated genes in the CLs of Finnsheep compared to F1 crosses

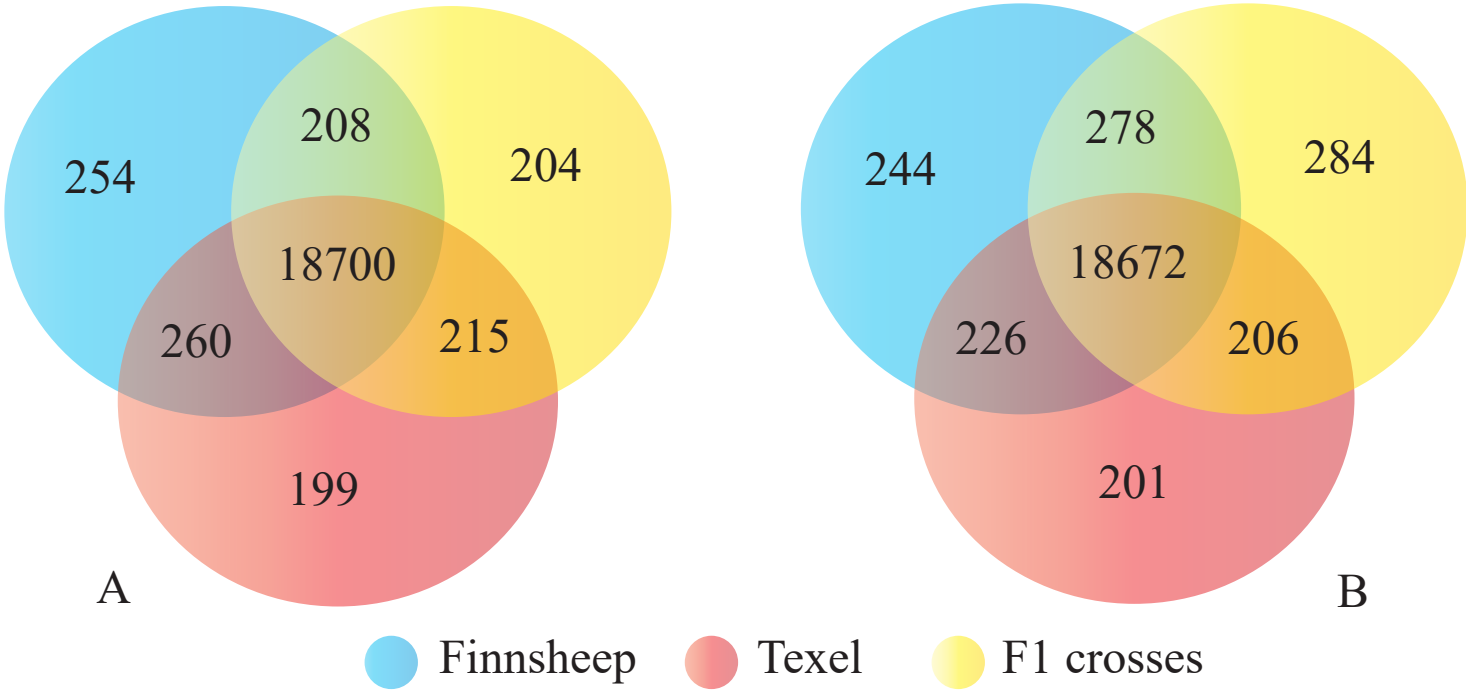
Table S9: List of DEGs between the CLs of Texel and F1 crosses

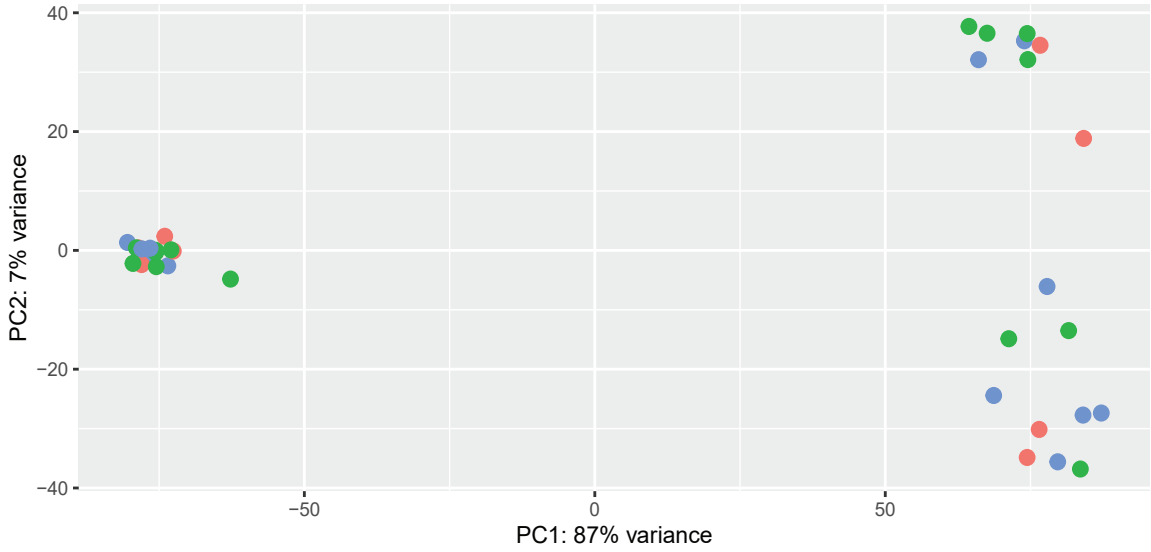
Table S10: Summary of the miRNA-Seq data

Table S11: List of miRNAs quantified in this study

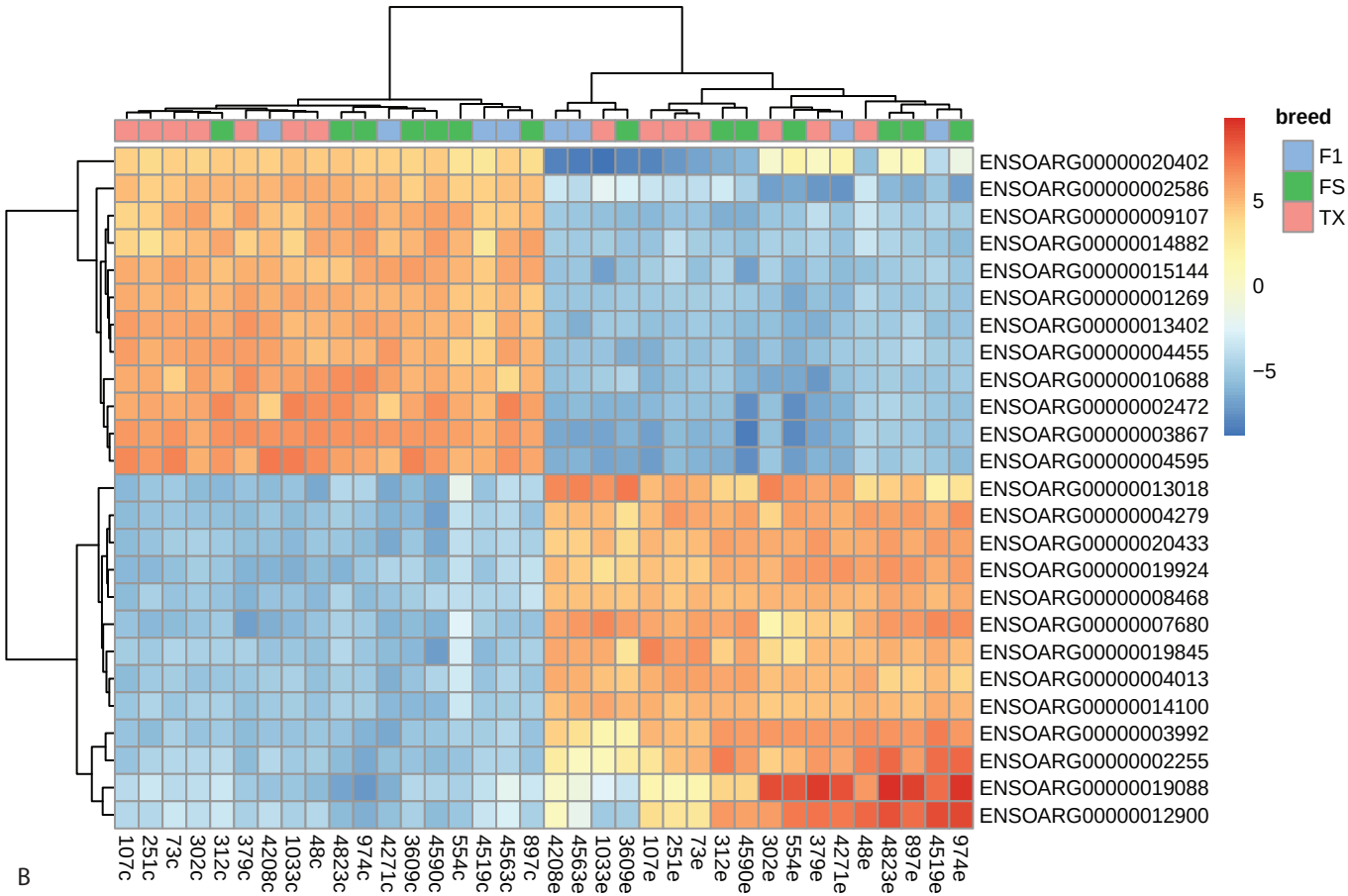
Data Availability Statement

The raw FASTQ sequence data (for both mRNAs and miRNAs) from this study have been deposited in the European Nucleotide Archive (ENA) database under accession code PRJEB32852. The accession codes for each sample are included in the sample summary tables (Supplementary table S3 and S10 for mRNA and miRNA, respectively).

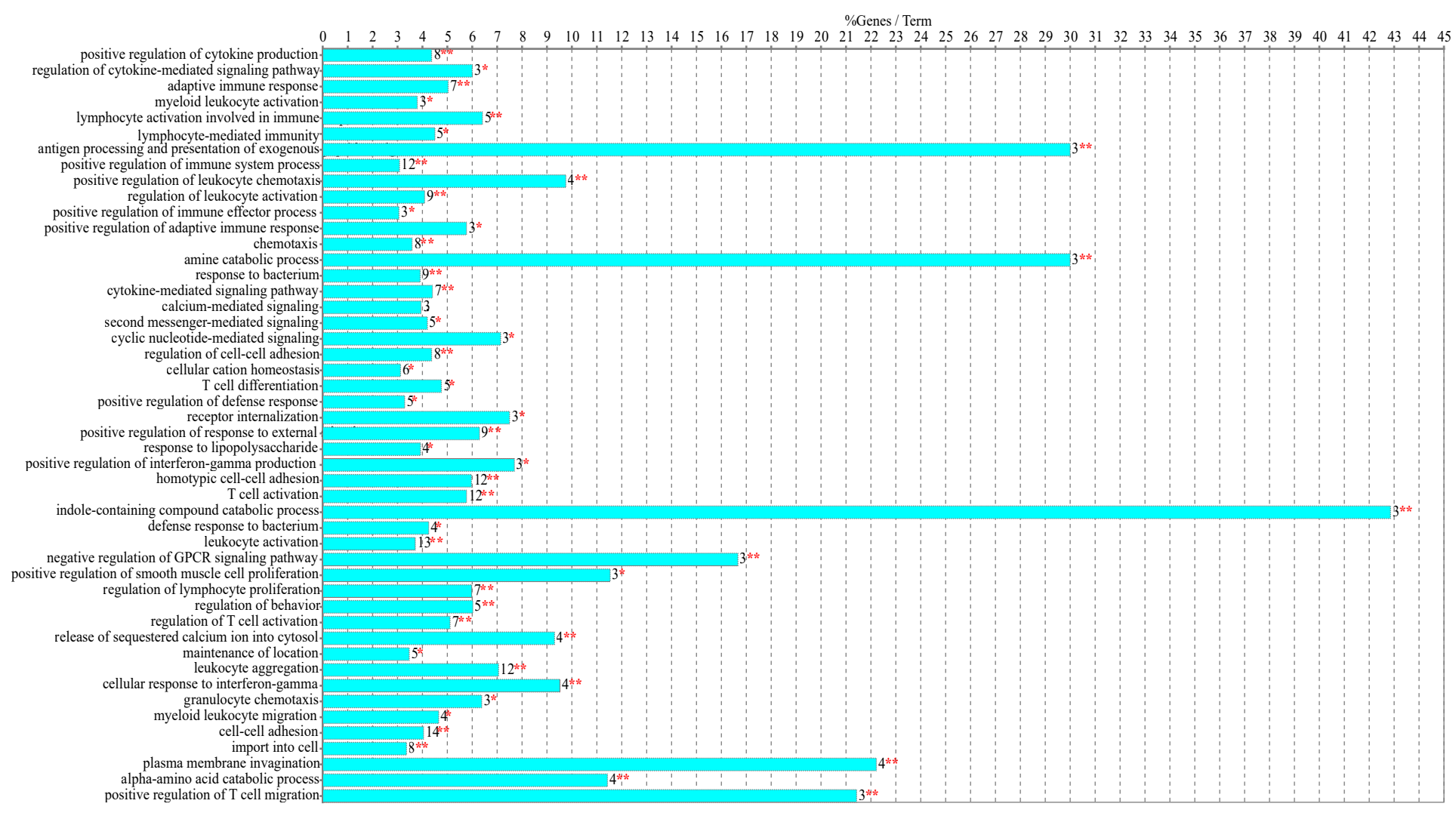

254
208
204
18700
260
215
199
A
244
278
284
18672
226
206
201
B
Finnsheep
Texel
F1 crosses

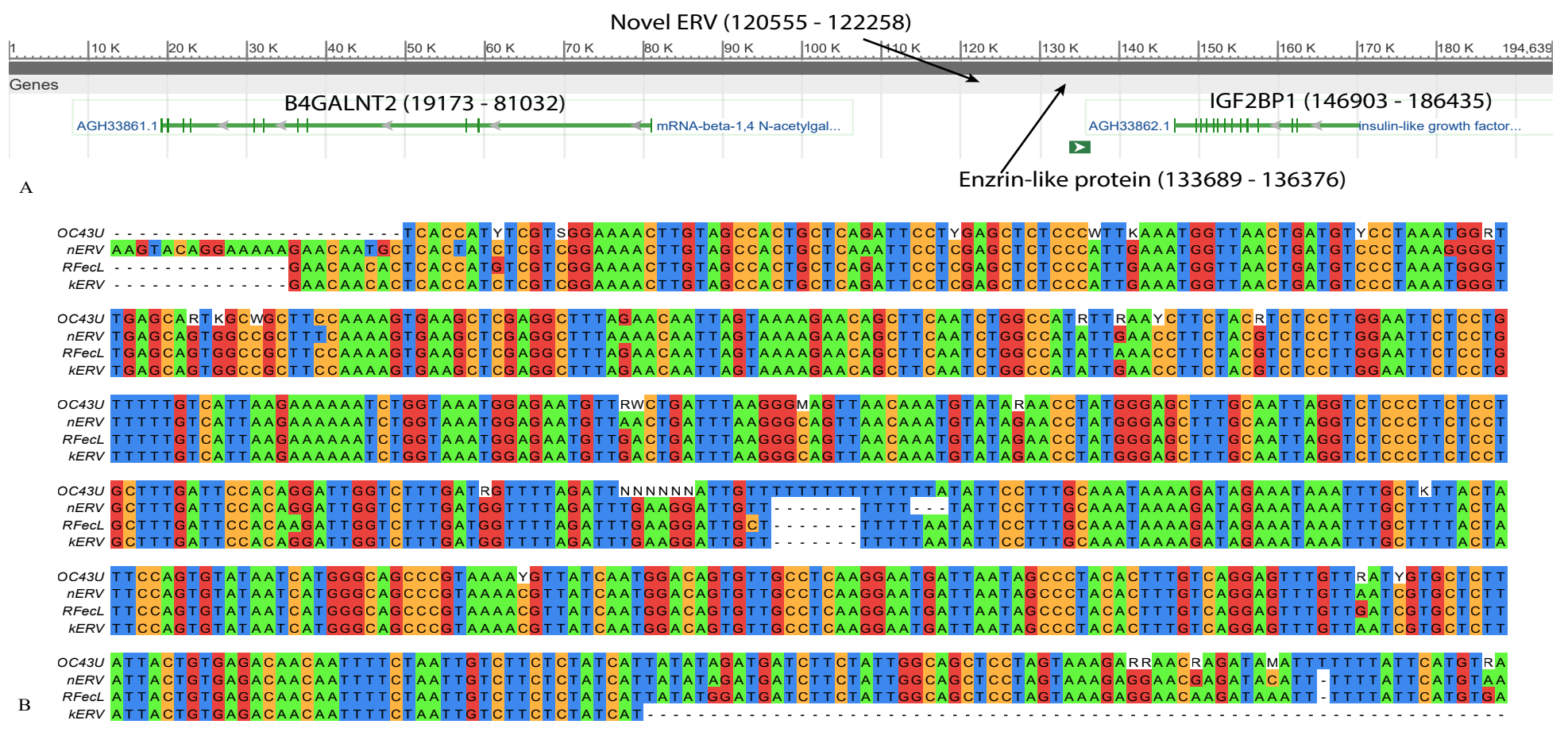

PC2: 7% variance
PC1: 87% variance
40
20
0
−20
−40
−50
0
50
group
F1
FS
TX

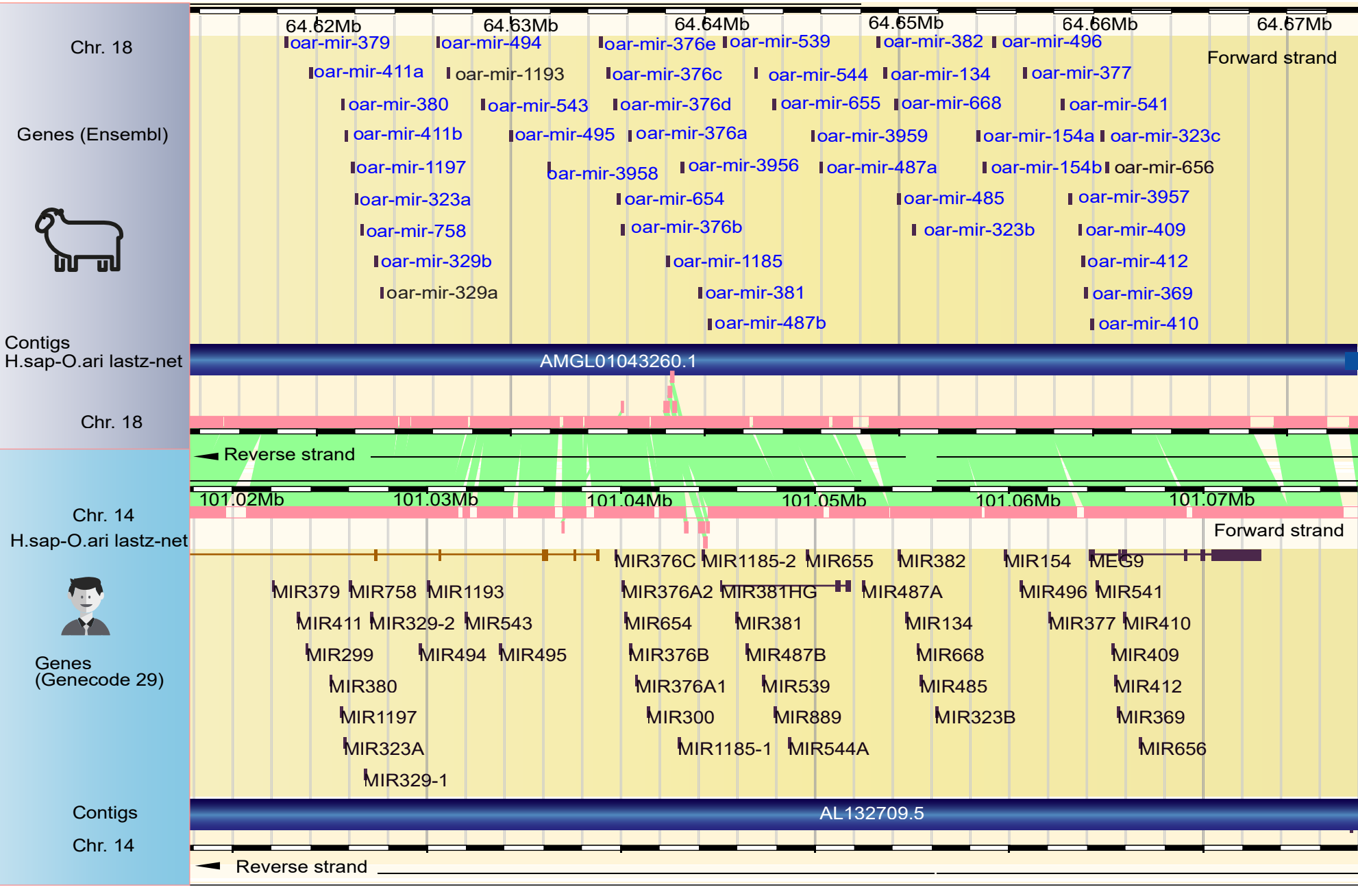
A


breed
F1
FS
TX
5
0
−5
ENSOARG00000020402
ENSOARG00000002586
ENSOARG00000009107
ENSOARG00000014882
ENSOARG00000015144
ENSOARG00000001269
ENSOARG00000013402
ENSOARG00000004455
ENSOARG00000010688
ENSOARG00000002472
ENSOARG00000003867
ENSOARG00000004595
ENSOARG00000013018
ENSOARG00000004279
ENSOARG00000020433
ENSOARG00000019924
ENSOARG00000008468
ENSOARG00000007680
ENSOARG00000019845
ENSOARG00000004013
ENSOARG00000014100
ENSOARG00000003992
ENSOARG00000002255
ENSOARG00000019088
ENSOARG00000012900
107c
251c
73c
302c
312c
379c
4208c
1033c
48c
4823c
974c
4271c
3609c
4590c
554c
4519c
4563c
897c
4208e
4563e
1033e
3609e
107e
251e
73e
312e
4590e
302e
554e
379e
4271e
48e
4823e
897e
4519e
974e

B


%Genes / Term
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
positive regulation of cytokine production 8**
regulation of cytokine-mediated signaling pathway 3*
adaptive immune response 7**
myeloid leukocyte activation 3*
lymphocyte activation involved in immune 5**
lymphocyte-mediated immunity 5*
antigen processing and presentation of exogenous 3**
positive regulation of immune system process 12**
positive regulation of leukocyte chemotaxis 4**
regulation of leukocyte activation 9**
positive regulation of immune effector process 3*
positive regulation of adaptive immune response 3*
chemotaxis 8**
amine catabolic process 3**
response to bacterium 9**
cytokine-mediated signaling pathway 7**
calcium-mediated signaling 3
second messenger-mediated signaling 5*
cyclic nucleotide-mediated signaling 3*
regulation of cell-cell adhesion 8**
cellular cation homeostasis 6*
T cell differentiation 5*
positive regulation of defense response 5*
receptor internalization 3*
positive regulation of response to external 9**
response to lipopolysaccharide 4*
positive regulation of interferon-gamma production 3*
homotypic cell-cell adhesion 12**
T cell activation 12**
indole-containing compound catabolic process 3**
defense response to bacterium 4*
leukocyte activation 13**
negative regulation of GPCR signaling pathway 3**
positive regulation of smooth muscle cell proliferation 3*
regulation of lymphocyte proliferation 7**
regulation of behavior 5**
regulation of T cell activation 7**
release of sequestered calcium ion into cytosol 4**
maintenance of location 5*
leukocyte aggregation 12**
cellular response to interferon-gamma 4**
granulocyte chemotaxis 3*
myeloid leukocyte migration 4*
cell-cell adhesion 14**
import into cell 8**
plasma membrane invagination 4**
alpha-amino acid catabolic process 4**
positive regulation of T cell migration 3**


Novel ERV (120555 - 122258)
1
10 K
20 K
30 K
40 K
50 K
60 K
70 K
80 K
90 K
100 K
110 K
120 K
130 K
140 K
150 K
160 K
170 K
180 K
194,639
Genes
B4GALNT2 (19173 - 81032)
AGH33861.1
mRNA-beta-1,4 N-acetylgal...
IGF2BP1 (146903 - 186435)
AGH33862.1
insulin-like growth factor...
Enzrin-like protein (133689 - 136376)
A
B
OC43U
nERV
RFecL
kERV


Chr. 18
Genes (Ensembl)
Contigs
H.sap-O.ari lastz-net
Chr. 18
Chr. 14
H.sap-O.ari lastz-net
Genes
(Genecode 29)
Contigs
Chr. 14
64.62Mb
64.63Mb
64.64Mb
64.65Mb
64.66Mb
64.67Mb
Forward strand
oar-mir-379
oar-mir-494
oar-mir-376e
oar-mir-539
oar-mir-382
oar-mir-496
oar-mir-411a
oar-mir-1193
oar-mir-376c
oar-mir-544
oar-mir-134
oar-mir-377
oar-mir-380
oar-mir-543
oar-mir-376d
oar-mir-655
oar-mir-668
oar-mir-541
oar-mir-411b
oar-mir-495
oar-mir-376a
oar-mir-3959
oar-mir-154a
oar-mir-323c
oar-mir-1197
bar-mir-3958
oar-mir-3956
oar-mir-487a
oar-mir-154b
oar-mir-656
oar-mir-323a
oar-mir-654
oar-mir-485
oar-mir-3957
oar-mir-758
oar-mir-376b
oar-mir-323b
oar-mir-409
oar-mir-329b
oar-mir-1185
oar-mir-412
oar-mir-329a
oar-mir-381
oar-mir-369
oar-mir-487b
oar-mir-410
AMGL01043260.1
Reverse strand
101.02Mb
101.03Mb
101.04Mb
101.05Mb
101.06Mb
101.07Mb
Forward strand
MIR376C
MIR1185-2
MIR655
MIR382
MIR154
MEG9
MIR379
MIR758
MIR1193
MIR376A2
MIR381HG
MIR487A
MIR496
MIR541
MIR411
MIR329-2
MIR543
MIR654
MIR381
MIR134
MIR377
MIR410
MIR299
MIR494
MIR495
MIR376B
MIR487B
MIR668
MIR409
MIR380
MIR376A1
MIR539
MIR485
MIR412
MIR1197
MIR300
MIR889
MIR323B
MIR369
MIR323A
MIR1185-1
MIR544A
MIR656
MIR329-1
AL132709.5
Reverse strand