Final Technical Report

FCEB Project #51

Project Name: Predicting Fertility in Brahman Cows

Final Technical Report

Principal Investigator: Mario Binelli, PhD

1. Main Objective

The Main Objective of this proposal was to identify differences in uterine markers between cows that succeeded or failed to maintain the pregnancy. **This project addressed the following FCA Research and Education Priorities: Reproduction Efficiency (In utero losses) and Animal Genetics (Improving the Brahman Breed).**

.

2. Situation and Specific Objective

Early embryo mortality is elevated in cattle, and pregnancy success depends on the molecular composition of the uterus. Fertility of Brahman cows is rarely more than 60% after an opportunity of artificial insemination or natural service. It is unknown the extent to which cows showing estrus have distinct uterine compositions that result in successful or failed pregnancies. This project aimed to identify differences in the uterine composition between cows that succeeded or failed to maintain the pregnancy. The estrous cycles of multiparous, suckled Brahman cows (n=100) at the University of Florida Beef Unit were synchronized and cows were artificially inseminated based on estrus (D0). On D4, a sample from the uterine lumen was collected using a cytology brush (Cytobrush) and we compared the uterine composition of cows that were pregnant vs. open. We discovered and validated an original panel of uterine markers for fertility in Brahman cows, detectable 4 days after estrus. The specific objective of this proposal was to identify differences in the endometrial transcriptome between cows that succeeded or failed to maintain the pregnancy.

3. Approach

The estrous cycles of multiparous, suckled Brahman cows (n=100) at the University of Florida Beef Unit were synchronized using the 7-d CoSynch + CIDR protocol. On the day of CIDR removal, all cows received an injection of prostaglandin-F2α and were fitted with an estrus detection patch (Estrotect). Starting 36 h after CIDR removal, cows were observed for estrus behavior twice daily for 72 hours and artificially inseminated following the AM/PM rule. The day of detected heat was considered the experimental day 0 (DO). On D4, a sample from the

uterine lumen was collected using a cytology brush (Cytobrush) and stored at -80C for molecular analysis. On D30, cows were pregnancy-checked by ultrasound.

Cytobrush sample processing, RNA extraction and RNAseq: Cytobrush was used to probe the uterine luminal environment. Procedures were executed according to Cardoso *et al.*, (Reproduction in Domestic Animals, 56:1153-1157, 2017). Total RNA was extracted for transcriptomic analyses. Cell samples were subjected to total RNA isolation with Tryzol reagent following the manufacturer's guidelines. RNA extracts were submitted to RNAseq and we compared the transcriptome of endometrial epithelial cells lining the uterine lumen of cows that were pregnant vs. open. Samples were analyzed according to our previous reports (Silva et al., Biology of Reproduction 108:922-935; 2023).

.

4. Final Report

Samples were collected, submitted to transcriptomic analysis and bioinformatics analysis. Results from this project were turned into a chapter of the doctoral dissertation of Dr. Cecilia Rocha and are enclosed as an attachment to this report.

5. Conclusion

Project was concluded as planned. There is a clear difference in the uterine transcriptome of cows according to the pregnancy outcome.

CHAPTER 6
THE TRANSCRIPTOMIC SIGNATURE OF ENDOMETRIAL EPITHELIAL CELLS IS ASSOCIATED WITH THE PREGNANCY OUTCOME TO ARTIFICIAL INSEMINATION IN BRAHMAN COWS

Abstract

Crossbreeding strategies using *Bos indicus* genetics is a common management strategy in tropical and subtropical regions. However, there is a small number of studies describing the uterine biology of *Bos indicus* cows. We hypothesized that there is a transcriptional signature in endometrial epithelial cells on D4 after estrus that is associated with the ability of the cow to remain pregnant to AI. Pluriparous Brahman cows (N=78) were submitted to an estrous synchronization protocol, followed by AI 12 hours after estrus. Four days after estrus cows were submitted to endometrial cytology using a cytobrush for collection of endometrial epithelial cells. Pregnancy/AI was diagnosed on D30 and cells from endometrial cytology samples were submitted to RNA-seq (N=32 non-pregnant and N=32 pregnant). RNA-seq results were used for targeted and untargeted pathways analysis. Targeted analysis was based on pathways previously reported in the literature, untargeted analysis used Ingenuity Pathway Analysis (IPA) software. A total of 16,239 genes were detected in the RNA-seq, 6% (975/16,239) were significantly associated (P≤0.1) with pregnancy/AI, 64.9% of them (620/975) were associated negatively with pregnancy/AI. Target analysis showed a negative association of transcripts involved in cytoskeleton function, apoptosis, Th2 immune response, and cell-cell adhesion with pregnancy/AI. Untargeted analysis found activation of intracellular cholesterol biosynthesis and inactivation of Th2 immune response in pregnant cows. Receiver operating characteristic (ROC) curve analysis using single transcripts to predict the outcome of pregnancy harvested poor accuracy

(0.67-0.72). In conclusion the endometrial receptivity was predominantly marked by an overall reduction in the molecular response. Th2 response was associated negatively with pregnancy/AI, while cholesterol biosynthesis was associated positively. Both pathways were previously reported on D4 in the endometrium and involved with pregnancy/success in crossbred cows.

Introduction

Crossbreeding strategies using *Bos indicus* genetics (i.e. Brahman and Nelore breeds) improve adaptability of beef herds located in tropical and subtropical regions. However, the introduction of *Bos indicus* genetics comes with physiological changes in reproductive characteristics in the herd. For example, *Bos indicus* cattle, have prolonged post-partum anestrus (Short *et al.*, 1990; Day, 2004), delayed age at puberty (Sartori *et al.*, 2010), shorter estrus expression (Pinheiro *et al.*, 1998), reduced metabolization of sex steroid hormones (Sartori *et al.*, 2016), and reduced P4 concentrations during early gestation (Rocha et al., 2023, Chapter 7). Consequently, cows carrying more than 63% of Brahman genetics which were submitted to estrous synchronization programs, that were primarily designed to *Bos taurus* cows, had reduced follicular diameter at the end of the synchronization, reduced proportion of estrus expression, and reduced pregnancy/AI of up to 19.2%, 42.1% and 27.5% respectively when compared with cows carrying less than 19% of Brahman genetics, even the proportion of cyclicity being similar among them in the beginning of the synchronization protocol (Martins *et al.*, 2022). Follicular diameter and estrus expression are important programmers of the endometrial function during early gestation (i.e. first 30 days) for a successful pregnancy (Perry and Perry, 2008; Madsen *et al.*, 2015; Mesquita *et al.*, 2015; Ramos *et al.*, 2015; Northrop-Albrecht *et al.*, 2021).

Therefore, to characterize the endometrial function of *Bos indicus* cows during early gestation and associate with the outcome of pregnancy/AI is critical to improve reproductive efficiency of *Bos indicus* cattle.

The first 30 days of gestation involves the pre-attachment period (up to D20) and initial placentation (“attachment”) from D20 to 30 (Seo *et al.*, 2023). Pregnancy losses that happen in this window are consequences of a failed conceptus-maternal interaction, which might be caused by a developed compromised embryo, a dysfunctional endometrium, or environmental factors (Binelli *et al.*, 2022). As early as D4 of gestation, the embryonic development occurs in the confinement of the uterine lumen and relies on the uterine secretions for its development (Bazer, 1975; Hackett *et al.*, 1993). The uterine secretome is originated from the endometrial luminal and glandular epithelium (Forde *et al.*, 2014; Simintiras *et al.*, 2019a) and associated with the pregnancy outcome (Forde *et al.*, 2015, 2017; Simintiras *et al.*, 2021). Indeed, the messenger RNA (mRNA) expression of nutrient transporters in endometrial cells was highly correlated with their respective concentrations in uterine secretions (Forde *et al.*, 2014; França *et al.*, 2017). The endometrial environment of *Bos indicus* cows was previously studied on D7 of gestation and associated with the outcome of pregnancy/AI (Binelli *et al.*, 2015; Sá Filho *et al.*, 2017). The limitation of these studies was that the endometrial sample was obtained by biopsies, which usually contain a greater proportion of stromal cells than luminal and glandular epithelial (Maclean *et al.*, 2020). Thus, the characterization of cellular functions was a reflection of a cell pool rather than endometrial epithelial cells. Therefore, there is a need to characterize the transcriptomic

signature specific in endometrial epithelial cells to understand the endometrial biology of early pregnancy in *Bos indicus* cows.

In a previous study, the transcriptomic signature of endometrial epithelial cells collected on D4 after estrus from crossbreed cows followed by ET 3 days later was associated with the pregnancy outcome measured on D30 after estrus (Martins *et al.*, 2021). Moreover, there were two different amino acid profiles in the ULF of non-inseminated *Bos indicus* heifers, on D4 after estrus (Silva *et al.*, 2021). Had those heifers being inseminated, it is possible that the contrasting amino acid profiles would be associated with different pregnancy outcomes. In addition, when endometrial epithelial cells from Brahman cows submitted to AI were collected on D4 after estrus and submitted to cell culture to access rbIFNT responsiveness and cellular proliferation, cells from cows who remained pregnant had 50% greater rbIFNT responsiveness and more intense cellular proliferation than the ones that did not remain pregnant (Chapter 4). Altogether, events that occur during early diestrus (D4) and their respective molecular signatures in endometrial cells, are likely decisive for the pregnancy outcome in *Bos indicus* cattle.

Here, for the first time, we interrogated the endometrial epithelial transcriptome of Brahman cows, collected on D4 after AI by endometrial cytology using a cytobrush. The endometrial cytology harvested samples that were 95% endometrial epithelial cells (Rocha et al., 2022, Chapter 3). Our overarching hypothesis is that there is a transcriptional signature in endometrial epithelial cells on D4 after estrus that is associated with the ability of the cow to remain pregnant to AI. Our objectives were to use RNA-seq results from D4 endometrial epithelial cells to perform, (1) target analysis

based on pathways from prior research conducted in the endometrium of cattle; (2) untargeted analysis to detect enriched pathways related to pregnancy/AI; and (3) prediction analysis to measure the accuracy of using genes from D4 to predict the outcome of pregnancy/AI on D30.

Material and Methods

Animals

All animal procedures were approved by the IACUC from the University of Florida under protocol No. IACUC-202200000055. Suckled multiparous (N=41) and 2-year-old primiparous (N=37) *Bos indicus* (Brahman) cows were used in the experiment. Cows had an average body weight: 503.5 ± 10kg, age: 4.06 ± 2.75 and were 115 ± 15 days postpartum. All cows had conceived and delivered a calf at least one time and were therefore considered to be of proven fertility. Cows were maintained in grazing conditions (Paspalum notatum), supplemented with hay, concentrate, and minerals to fulfill their maintenance requirements. The experiment was conducted over 3 years in 2021, 2022, and 2023 at the University of Florida Beef Research Unit, in Gainesville.

Experimental Design

Estrous cycles were synchronized using the Select Synch + CIDR protocol (Larson *et al.*, 2006) with a CIDR previously used once for 7 days, in 2021 and 2022 rounds. In 2023 round cows received the 7 & 7 Synch protocol (Andersen *et al.*, 2021). Briefly, for the Select Synch + CIDR protocol, on D−10 cows received an intravaginal P4-releasing insert containing 1.38g of P4 (CIDR; Eazi-Breed™ CIDR® Cattle insert, Zoetis Animal Health, Florham Park, USA) and were treated with GnRH analogue 100μg (Factrel®, Zoetis Animal Health). On D−3, the P4 insert was removed, cows

received a PGF2α injection, and were fitted with an estrus detection aid (ESTROTECT™, Rockway Inc, USA). For the 7 & 7 Synch protocol, a new CIDR was placed on D−17 and cows received a PGF2α injection, on D-10 and D-3 the same activities (GnRH and PGF2α injection, respectively) as the Select Synch protocol were performed. ESTROTECTs were evaluated three times daily (at 7am, 12pm and 6pm), from D−2 to D0 for estrus behavior and activation of the patch in all the rounds. Cows were considered in estrus when 50% or more of the gray lining on the patch was removed. Only cows that showed estrus were included in the study. The day of estrus was considered D0. Twelve hours after the beginning of estrus cows were AI by a single operator, using semen of 19 Brahman bulls which were assigned randomly. A total of 38 cows were used in 2021 (N=18 primiparous, N=20 multiparous), 20 in 2022 (N=8 primiparous, N=12 multiparous), and 20 in 2023 (N=11 primiparous, N=9 multiparous).

On D4, endometrial epithelial cells from the uterine body were collected transcervical from the cows using a cytobrush (Disposable cytology sampling brush; Viamed Ltd, UK) as described previously (Cardoso *et al.*, 2017) and blood samples were collected from the coccygeal vessels to measure P4 concentrations. Pregnancy per AI was determined by transrectal ultrasonography examination on D30 and was based on the presence or absence of a fetus heartbeat. In 2023, a subset of 32 cows were AI at estrus but not submitted to the endometrial cytology collections on D4. These cows served as a control group to determine if the uterine cytology caused pregnancy loss.

Collection of Endometrial Epithelial Cells

On D4 (i.e., 4 days after estrus and 3.5 days after AI), an endometrial cytology was performed to collect the luminal epithelial cells of the uterine body as described

previously by our group (Cardoso *et al.*, 2017). Briefly, for the collection, the cytobrush was connected to the tip of a conventional AI gun. The tip of the AI gun was positioned just past the cervix, and then the cytobrush was exposed and gently rotated to collect the first layer of luminal epithelial cells. After collection brushes were stored in a 1.5 mL microcentrifuge tube in liquid nitrogen, transported to the laboratory and stored in −80ºC for subsequent RNA extraction.

Plasma Progesterone Concentrations

Blood samples on D4 were collected in K2 EDTA tubes (BD Vacutainer, NJ, USA) and kept in ice until processing, within 1.5 hours of collection. Blood was centrifuged at 1,500x g for 10 minutes at 4°C. Plasma was recovered and stored in 1,5mL microcentrifuge tubes at -20°C until assayed. The concentration of P4 was assessed by a liquid-liquid phase, double-antibody radioimmunoassay using assay reagents supplied by MpBio (Santa Ana, USA), and validated previously (Pohler *et al.*, 2016). The minimum detectable P4 concentration was 0.04 ng/mL. The intra/inter-assay estimate of coefficient of variation was lower than 3%.

RNA Extraction

Total RNA extraction of the luminal epithelial cells collected from the cytobrush was performed using the RNeasy Mini Kit (Qiagen, MD, USA) following the manufacturer's instruction. The pellet was resuspended in 500 µL of buffer RLT supplemented with 1% of β-mercaptoethanol and disrupted by vortexing at maximum speed for 2 minutes. The lysate was transferred to a QIAshredder spin column (Qiagen) and centrifuged for 2 minutes at room temperature and full speed. The filtrated was mixed with 70% ethanol and the content was loaded in the RNeasy spin column. Samples were subjected to on-column DNase treatment using the RNase-free DNase

set (Qiagen). Concentration and purity of RNA in extracts were evaluated using spectrophotometry (NanoDrop, Microvolume Spectrophotometer and Fluorometer, Thermo Fisher Scientific, USA). RNA samples were stored at −80ºC for future library preparation and RNA sequencing.

Library generation and RNA Sequencing (RNA-seq)

RNA samples from pregnant and non-pregnant cows were sequenced by Novogene Corporation (https://en.novogene.com). The quality control of the RNA samples was performed before library generation, and RNA Integrity Number (RIN) ranged from 4.0 to 9.9. The sequencing library was constructed using the RNA-NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA). The sequencing was performed using NovaSeq6000 (Illumina,Sacramento,CA,USA). Fourteen samples, of which 10, two, and two were from 2021, 2022, and 2023 respectively, of which 11 and three were from multiparous and primiparous respectively, were excluded because they did not pass the quality control, prior to library preparation due to reduced RIN or reduced RNA concentration, or after sequencing due to low quality reads (N=78−14=64).

RNA Sequencing (RNA-seq) Analysis: Quality Control, Mapping, and Gene Expression Estimation

The quality of the RNA-seq reads was evaluated using the software FastQC (version 0.11.7, Babraham Bioinformatics, UK). Read trimming and adapter removal were performed using the software Trim Galore (version 0.4.4, Babraham Bioinformatics, UK) with the following parameters: --paired, --quality 20, --length 50, --clip_R1 15, --clip_R2 15, --three_prime_clip_R1 5, and --three_prime_clip_R2 5. After editing, sequencing reads were mapped to the latest bovine reference genome (ARS-

UCD1.2) using the software Hisat2 (v2.1.0) (Kim *et al.*, 2015). The number of reads that mapped to each annotated gene in the bovine genome was estimated using htseq-count (v0.6.1p1) with the option intersection-nonempty (Anders *et al.*, 2015).

Bioinformatics and Statistical Analysis

This RNA-seq study was designed to evaluate the impact that endometrial gene expression has on the outcome of pregnancy/AI. After removing lowly expressed genes (i.e., genes with counts per million < 0.5 in at least half of the samples), a total of 16,239 genes from 64 (N=32 non-pregnant; N=32 pregnant) samples were analyzed. Gene expression was normalized using the trimmed mean of M-values method implemented in the R package EdgeR (Robinson *et al.*, 2010). The impact that the expression level of gene g has on pregnancy outcome was evaluated using the following logistic regression: $\log\left(\frac{P_P}{1-P_P}\right) = \beta_0 + year + \ parity \ + \ \beta_{ge} \,.\log\left(ge_{gi}\right)$

where P_P represents the probability of pregnancy, $year$ represents the round of trial (3 levels), $parity$ represents the parity of the cow (2 levels; primiparous or multiparous), and $log\left(ge_{gi}\right)$ represents the log-transform normalized gene expression level of gene g in sample i. Note that β_{GE} represents the change in the log-odds when log-transform normalized gene expression level increases one unit with the influence of year and parity held constant. For each gene, the null hypothesis H_0: $\beta_{GE} = 0$ was tested using a likelihood ratio test. Therefore, we have taken a unique approach of using gene expression to predict a binary dependent variable (pregnant or non-pregnant), which has intrinsically lower statistical power than using a linear dependent variable. The magnitude and direction of the effect of gene expression on the pregnancy outcome were estimated by the regression coefficient (β_{ge}) of each gene. Specifically, a

negative coefficient means that there was a negative association between an increase in gene expression and the probability of pregnancy success, a positive coefficient means that there was a positive association between an increase in gene expression and the probability of pregnancy.

Multivariate Analysis

Multivariate analyses were run using the MetaboAnalyst software 6.0 (Xia *et al.*, 2009), by entering the expression values of the 16,239 genes for each cow. The data was normalized using log transformation, mean-centered, submitted to a principal component analysis (PCA), and a supervised method analysis (sparse partial least squares discriminant; SPLS-DA). With the SPLS-DA results we established a set of genes that most contributed to the separation between pregnant and non-pregnant cows. This contribution was based on the parameter “variable importance in the projection” (VIP), which displays the overall importance of each variable to the projection of component 1 and component 2 in the SPLS-DA.

Targeted Analyses of Differentially Expressed Genes

To study specific pathways that were regulated in the endometrial epithelial cells on D4, and that were associated with the outcome of pregnancy/AI, a targeted approach was used. To avoid bias, a list of genes was generated prior to differential gene expression analysis of the current RNA-seq analysis. A list of target cell and tissue functions and pathways was generated based on prior research conducted in the endometrium of cattle, during early diestrus, and that had a physiological association with receptivity to the embryo. (Supplementary Table C-1). We selected the following cell and tissue functions and pathways: cellular proliferation (Mesquita *et al.*, 2015), immune system (Martins *et al.*, 2021; Silva *et al.*, 2023b), cell-cell adhesion (Binelli *et*

al., 2015; Sponchiado *et al.*, 2017; Johnson *et al.*, 2023), nutrient transporters (Gao *et al.*, 2009a, b; Satterfield *et al.*, 2010; Forde *et al.*, 2014b; França *et al.*, 2015, 2017), endometrial responses to sperm (Akthar *et al.*, 2021, 2023), genes regulated by endometrial region (Hoorn *et al.*, 2023), and blood vessel development (De Bem *et al.*, 2021). Then, within each cell and tissue function and pathway, we selected target transcripts. Specifically, for cellular proliferation, we selected 117 genes which were the top 50 genes of cellular proliferation in KEEG pathways dataset (Kanehisa, 2000), 13 specific cellular proliferation markers of epithelial cells (Tompkins *et al.*, 2011; Locard-Paulet *et al.*, 2022), 19 cell cycle regulators (Whitfield *et al.*, 2002), 12 cytoskeleton genes (Hall, 2009), and 23 apoptosis genes (Li *et al.*, 2014). For the immune system we selected 126 genes which were the top 50 chemokine signaling in KEEG pathways that were previously reported expressed in epithelial cells (Kanehisa, 2000), 22 genes from IL10 signaling (Verma *et al.*, 2016; Silva *et al.*, 2023b), 37 genes from overall interleukin signaling, which included interleukins and its receptors published previously in the literature and expressed in epithelial cells (Akdis *et al.*, 2016), and 17 genes from the pattern recognition receptors pathway expressed in the endometrium (Sheldon *et al.*, 2017, 2019). For cell-cell adhesion we selected 96 genes, which were 28 genes from the metalloproteinases family, expressed in epithelial cells (NCBI), 43 integrins previously reported to be regulated in the endometrium (Johnson *et al.*, 2023), or that were found when the key word “ITG” was searched in the present RNA-seq dataset, 29 cadherins previously reported as regulated in the endometrium (Sağsöz *et al.*, 2022), or that were found when the key word “CDH” was searched in the present RNA-seq dataset, and 12 specific targets that were studied previously in the endometrium as

markers of cellular adhesion (Sponchiado *et al.*, 2017). For nutrient transporters we selected 305 genes, which were 296 solute carrier transporters (SLC) reported previously to be expressed in the endometrium (Gao *et al.*, 2009a, b; Satterfield *et al.*, 2010; Forde *et al.*, 2014b; França *et al.*, 2015, 2017), or that were revealed when the key word “SLC” was searched in the present RNA-seq dataset, and nine genes from the aquaporin family that were previously reported to be expressed in the endometrium (Lee *et al.*, 1997). For the endometrial response to sperm, we selected nine transcripts recently reported to be regulated in the endometrium (Akthar *et al.*, 2021, 2023). For the genes regulated by the endometrial region, we selected 51 genes reported previously as differentially regulated in the caudal region of the endometrium compared to the cranial region (Sponchiado *et al.*, 2017; Hoorn *et al.*, 2023). Lastly, for the blood vessels development (De Bem *et al.*, 2021) we selected the top 50 genes in the KEEG pathway (Kanehisa, 2000).

To determine which genes were differentially expressed within each pathway, we calculated the adjusted p-value based on the total number of genes. For example, for genes related to cellular proliferation, we adjusted the p-value for the total number of comparisons (N=117). For this, we used the optimized false discovery rate (FDR) approach to obtain q-values. After determining the significant genes, we calculated the fold change between pregnant and non-pregnant cows and ranked the top 10 genes from all the selected pathways with greatest increased fold change and the top 10 genes with greatest decreased fold change to include in a heatmap. Heatmap analysis was performed using MetaboAnalyst using the log normalization of each transcript and the Ward criteria for clustering.

Untargeted Analyses of Differentially Expressed Genes Using Ingenuity Pathway (IPA)

Untargeted pathway analyses were performed using IPA (QIAGEN) for the transcriptomic dataset. The IPA platform successfully mapped 14,031 out of the 16,239 genes analyzed. The set of DEGs used in each analysis was determined based on raw $P \leq 0.1$ (N=807 genes mapped). Emphasis was given to the significant canonical pathways and bio-functions after multiple comparisons correction (i.e., −log (Benjamini–Hochberg p-value)) ≥1.3 that were predicted to be increased or decreased based on a z-score ≥ 2.0 and ≤ −2.0, respectively. Furthermore, IPA was used to connect the data set molecules with predicted upstream regulators and downstream functions (i.e., regulator effects analysis) using the IPA algorithms. The predictions were deemed to be activated or inhibited based on the aforementioned z-score criteria. We performed a secondary enrichment analysis using DAVID Bioinformatics Resources (Huang *et al.*, 2008) relying on *Bos taurus* annotation. KEEG pathways were also accessed through DAVID and contrasted to canonical pathways from IPA results. The top 10 enriched pathways were related to similar processes when comparing DAVID and IPA, thus only IPA analysis was reported in the results.

Univariate Analysis

The effect of P4 concentrations on pregnancy/AI was determined by logistic regression using the GLIMMIX procedure of SAS (version 9.4, SAS Institute Inc., Cary, NC) and a model fitted to binomial distribution. The model included P4 concentrations as a linear explanatory variable, with year and parity as categorical explanatory variables. We tested the linear and quadratic effects of P4 in the model. For the effect of

endometrial cytology on the pregnancy/AI in the 2023 year, we used the same binomial distribution model but as fixed effect only the endometrial cytology (yes or not) variable.

Receiver operating characteristic curve analyses were performed using the 10 most significant DEGs (raw p-value) in JMP Pro software (SAS Institute Inc., NC, USA). We tested the ability of the selected genes to predict the outcome of pregnancy. We used the raw count per million of each cow for each transcript as the explanatory variable. Specificity, sensitivity, and accuracy were calculated as reported previously (Pugliesi *et al.*, 2014).

Results

The overall pregnancy/AI was 48.7% (38/78). There was no significant effect (P>0.1) of year in the pregnancy/AI (2021 47.3% ([18/38], 2022 40% [8/20], and 2023 60% [12/20]). The samples used for RNA-seq analysis were distributed among years (2021, N=28; 2022, N=18; 2023, N=18). There was no significant effect of parity (P<0.1) in the pregnancy/AI (primiparous 56.7% [21/37] and multiparous 41.4% [17/41]). Samples used for RNA-seq were distributed among primiparous and multiparous cows (N=34 and 30, respectively). The average plasma P4 concentrations on D4 was 1.08 ± 0.08 ng/mL and no significant effect of P4 (P>0.1) in the pregnancy/AI was observed. For the 2023 data, endometrial cytology did not affect pregnancy/AI (P>0.1). Cows that displayed estrus and were submitted to endometrial cytology had similar pregnancy/AI (60% 12/20) than cows that displayed estrus and were not submitted to endometrial cytology (68% 22/32).

Overall RNA Sequencing (RNA-seq) Results

Supplementary Table C-2 shows the 16,239 genes detected in the RNA-seq, their respective ensemble ID, name, regression coefficient (β), and raw p-value. Among all detected genes, there were 6% (975/16,239) significantly associated ($P \leq 0.1$) with pregnancy/AI, 64.9% (620/975) were associated negatively with pregnancy/AI and 34.3% (335/975) were associated positively with pregnancy/AI. No genes were significant when adjusted p-values (q-value) were calculated.

For the multivariate analysis, in the PCA plot, component 1 explained 15.8% of the variability between pregnant and non-pregnant cows, while component 2 explained 11.8% of the variability (Figure 6-1, Panel A). There was an overlap between pregnant and non-pregnant cows, most of the overlap happened from -5 to 5 distance where 65.6% (21/32) and 75% (24/32) of the pregnant and non-pregnant cows respectively were clustered. For the SPLS-DA analysis component 1 contributed to 13.8% of the variability between pregnant and non-pregnant cows and component 2 contributed to 6.6% of the variability (Figure 6-1, Panel B). There was a discrete overlap between pregnant and non-pregnant cows, which happened from distance 0 to 5 with 6.3% (2/32) and 3.12% (1/32) of pregnant and non-pregnant cows respectively. A secondary result from the SPLS-DA analysis was the VIP plot (Figure 6-1, Panel C). The VIP analysis ranked the top 10 genes which contributed the most for the separation of pregnant and non-pregnant cows, UL16-binding protein 27 (*ULBP27*), BRICHOS domain containing 5 (*BRICD5*), and Aldo-keto reductase family 1 member C4 (*AKR1C4*) were the top three genes with a VIP score from 0.4 to 0.5. They were

associated positively with the odds of pregnancy (β coefficients: 0.64, 1.77, and 1.18, respectively).

Targeted Analysis

For the selected 754 genes in the targeted analysis, 9.41% (71/754) were considered significant after applying the q-value correction (adjusted p-value ≤0.1). Out of the 71 significant genes, 19 were from the cellular proliferation pathways, 15 from immune system, 11 from cell-cell adhesion, 22 from nutrient transporters, two from endometrial response to sperm, two from endometrial region, and two from blood vessel development (Supplementary Table C-1; Figure 6-2, Panel A). The significant genes from endometrial response to sperm overlapped with immune response.

For cellular proliferation pathways, 16.2% of the selected genes (19/117) were associated with pregnancy/AI, 18 were associated negatively and one positively (Figure 6-2, Panel A). Whitin the significant genes, two were from the cellular proliferation pathway, cyclin dependent kinase 4 (*CDK14*) and cyclin dependent kinase 6 (*CDK6*) and had a negative association with pregnancy/AI. Nine genes were from cytoskeleton regulators, such as, ras homolog family member H (*RHOH*), Rac/Cdc42 guanine nucleotide exchange factor 6 (*ARHGEF6*), rhophilin Rho GTPase binding protein 1 (*RHPN1*), Rho guanine nucleotide exchange factor 38 (*ARHGEF38*), Rho guanine nucleotide exchange factor 3 (*ARHGEF3*), Rho GTPase activating protein 17 (*ARHGAP17*), alsin Rho guanine nucleotide exchange factor 2 (*ALS2*), Rho/Rac guanine nucleotide exchange factor 2 (*ARHGEF2*), and Rho GTPase activating protein 28 (*ARHGAP28*), and were associated negatively with pregnancy/AI. Seven of the significant genes were part of apoptosis pathway, CASP8 and FADD like apoptosis regulator (*CFLAR*), TNF receptor superfamily member 21 (*TNFRSF21*), TNF receptor

superfamily member 1B (*TNFRSF1B*), TNF receptor superfamily member 19 (*TNFRSF19*), TNF receptor superfamily member 1A (*TNFRSF1A*), BCL2 related protein A1 (*BCL2A1*), and caspase 4 (*CASP4*), and were associated negatively with pregnancy/AI. The single gene associated positively with pregnancy/AI was from apoptosis pathway (caspase 8; *CASP8*). No significant genes were observed for the "proliferation markers" and "cell cycle regulators" in the cellular proliferation pathways.

For the immune system functions, 11.9% of genes (15/126) were associated negatively with pregnancy/AI (Figure 6-2, Panel A). Five were involved in chemokine signaling, such as C-X-C motif chemokine ligand 8 (*CXCL8*), C-X-C motif chemokine ligand 9 (*CXCL9*), X-C motif chemokine receptor 1 (*XCR1*), C-C motif chemokine receptor 4 (*CCR4*), and C-C motif chemokine receptor 1 (*CCR1*). Five were part of IL10 signaling, interleukin 10 receptor subunit alpha (*IL10RA*), intercellular adhesion molecule 1 (*ICAM1*), inhibitor of nuclear factor kappa B kinase subunit beta (*IKBKB*), interleukin 1 alpha (*IL1A*), and interleukin 10 (*IL10*). Five genes were involved in overall interleukin signaling, two being exclusively from the pathway, interleukin 7 receptor (*IL7R)* and interleukin 20 receptor (*IL20R)*, and three overlapped with IL10 signaling (*IL10*, *IL1A*, and *IL10RA*). Lastly, three genes were from pattern recognition receptors, toll like receptor 4 (*TLR4*), toll like receptor 3 (*TLR3*), and NLR family pyrin domain containing 3 (*NLRP3*). No positive association was found for immune system genes and pregnancy/AI.

For the cell-cell adhesion 11.4% (11/96) of the genes were associated negatively with pregnancy/AI (Figure 6-2, Panel A). Whitin them, three were metalloproteinases, from the matrix metallopeptidase (MMP) family (*MMP9*, *MMP12*, and *MMP25*), five

were integrins, such as, integrin subunit alpha D (*ITGAD*), integrin subunit beta 6 (*ITGB6*), integrin subunit alpha M (*ITGAM*) integrin subunit alpha X (*ITGAX*), and integrin subunit alpha V (*ITGAV*). There was one cadherin gene, protocadherin gamma subfamily C, 3 (*PCDHGC3*), and the last two were specific targets previously reported in the literature, intercellular adhesion molecule 1 (*ICAM1*) and galectin 9 (*LGALS9*). No positive association between cell-cell adhesion genes and pregnancy/AI was observed.

For nutrient transporters, 7.21% (22/305) of the genes were associated with pregnancy/AI, three were associated positively and 19 were associated negatively (Figure 6-2, Panel A). Within the significant genes, 21 were from the SLC family, including eight amino acids SLCs, which were associated negatively (*SLC16A6*, *SLC66A1*, *SLC43A2, SLC36A1*, *SLC7A7*, *SLC38A5*, *SLC13A3,* and *SLC1A1*) and one amino acids SLC associated positively (*SLC22A15*) with pregnancy/AI. Three sodium SLCs were associated negatively with pregnancy/AI (*SLC9A2*, *SLC9B1,* and *SLC9A9)* and one sodium SLC associated positively with pregnancy/AI (*SLC24A3*). Four sugar SLCs were associated negatively with pregnancy/AI (*SLC2A9*, *SLC2A3, SLC5A2*, and *SLC35F6*). Other nutrients SLCs related with vitamins (*SLC46A1*), minerals (*SLC39A11)*, and ions (*SLC34A2*, and *SLC46A2)* were associated negatively with pregnancy/AI, while the water transporter aquaporin 5 (*AQP5*) was associated positively with pregnancy/AI.

For the genes regulated by endometrial response to sperm, 22.2% (2/9) were associated with pregnancy/AI. The association was negative, and both genes (*TLR4* and *IL10*) were also part of the immune system pathway. For the genes regulated by endometrial region, 3.9% (2/51) were associated with pregnancy/AI. The *AKR1C4 gene*

was associated positively and insulin like growth factor 2 receptor (*IGF2R*) negatively. For the genes related to blood vessel development, 4% (2/50) were associated with pregnancy/AI (alcohol dehydrogenase 5; *ADH5* and adrenoceptor beta 1; *ADRB1*), and the association was negative.

Figure 6-2, Panel B shows the heatmap for the clustering analysis using the top 10 greatest and top 10 smallest fold change, genes are in order of significance. The top five genes are associated negatively with pregnancy/AI and four out of five are from the immune system.

Untargeted Analysis

Enrichment analysis for the 975 significant genes (raw $P \leq 0.1$) resulted in 167 significant canonical pathways (Supplementary Table C-3), of which the 10 most significant (adjusted $P<1.3$) are represented in Figure 6-3, Panel A. There was an intimate functional association among the three activated pathways, which were "activation of gene expression by SREBF" (12 DEGs), "Superpathway of cholesterol biosynthesis" (10 DEGs), and "Cholesterol biosynthesis" (9 DEGs). They all shared eight common genes, cytochrome P450 family 51 subfamily A member 1 (*CYP51A1*), farnesyl-diphosphate farnesyltransferase 1 (*FDFT1*), farnesyl diphosphate synthase (*FDPS*), 3-hydroxy-3-methylglutaryl-CoA reductase (*HMGCR*), 3-hydroxy-3-methylglutaryl-CoA synthase 1 (*HMGCS1*), isopentenyl-diphosphate delta isomerase 1 (*IDI1*), lanosterol synthase (*LSS*), and squalene epoxidase (*SQLE)*. In each case, gene expression was associated positively with the odds of pregnancy/AI. These transcripts are primarily associated with de novo synthesis of cholesterol. Here, activation of such pathways was associated with increased pregnancy/AI.

For the inhibited canonical pathways with Z-score $\leq$ -2, a similar intimate relationship was observed. In this case the “Interleukin-10 signaling” (12 DEGs), “Macrophage classical activation signaling pathway” (25 DEGs), “Neuroinflammation signaling pathway” (33 DEGs), and “Interleukin-4 and Interleukin-13 signaling” (17 DEGs) also shared common genes. They all shared six common genes, *CXCL8*, *IL10*, *IL1A*, *TNFRSF1B*, *ICAM1*, and tyrosine kinase 2 (*TYK2;* Supplementary Table C-3). In each case, gene expression was associated negatively with the odds of pregnancy/AI. These transcripts are primarily associated with the Th2 anti-inflammatory response. Here, inactivation of such pathways was associated with increased pregnancy/AI.

Lastly, the algorithm of IPA enrichment analysis calculated regulator effects which explains how predicted activated or inhibited upstream regulators (e.g., transcription factors, receptors, or signaling proteins) might cause increase or decrease in phenotypic or functional outcomes downstream. There were 1,022 networks, of which 27 had a consistency score greater than 3 (i.e., a measure of how causally consistent and densely connected a network is; Supplementary Table C-4). The top 1 regulatory network had an inhibitory profile which was associated with one upstream regulator, KLF transcription factor 6 (*KLF6*), and the downstream outcome of cell movement of tumor cell lines (Figure 6-3, Panel B). Similarly, the top 2 regulatory network had an inhibitory profile, which was associated with one upstream regulator phosphatidylinositol 3-kinase (PI3K) complex and the downstream outcome of invasion of cells (Figure 6-3, Panel C). Both networks shared seven regulatory genes associated negatively with pregnancy/AI, which were either part of the immune system and chemotaxis of immune cells (*CCR1*, CD274 molecule; *CD274*, *CXCL8*, *ICAM1*, and cytochrome b-245 beta

chain; *CYBB*) or metalloproteinases important for cellular adhesion (*MMP9* and *MMP12*). Collectively both regulatory networks corroborate with the idea that inhibition of migratory capacity of cells in general is positively associated with the odds of pregnancy/AI.

Predictability of Selected Transcripts

ROC curve analysis was generated to the top 10 most significant genes (Table 6-1). The accuracy of prediction was poor and ranged from 0.67 to 0.72. The greatest sensitivity (0.97) was observed for *A0A3Q1NBF5_BOVIN*, and protein tyrosine phosphatase receptor type T (*PTPRT*), while greatest specificity (0.97) was observed for *LOC101903758*. Overall, genes with a sensitivity greater than 0.8 had a specificity smaller than 0.55.

Discussion

Introduction of *Bos indicus* genetics in beef herds for greater adaptability of increased ambient temperatures comes with the consequence of reduced reproductive efficiency (Baruselli *et al.*, 2004). Brahman cows have up to 27.5% reduction in pregnancy/AI measured on D30 when compared to *Bos taurus* cows (Martins *et al.*, 2022). During early pregnancy, the conceptus relies on the biochemical composition of the uterine lumen for its development, which is a reflection of the endometrial epithelial cells function (Forde *et al.*, 2014). In this paper we investigated for the first time the relationship of gene expression on D4 in endometrial epithelial cells with the ability of Brahman cows to remain pregnant by AI. We hypothesized that there is a transcriptional signature in endometrial epithelial cells on D4 after estrus that is associated with the ability of the cow to remain pregnant to AI. Cows were submitted to AI, 3.5 days later endometrial epithelial cells were harvested by endometrial cytology (Rocha et al., 2022,

Chapter 3), submitted to RNA-seq analysis, and results were associated retrospectively with the outcome of pregnancy, measured on D30. Here, we curated a list of genes involved in processes relevant to uterine receptivity to the embryo and determined the extent to which their expression predicted the pregnancy outcome. We established that pregnancy success on D30 was associated with a down regulation of genes involved in cytoskeleton, apoptosis, cellular adhesion, nutrient transporters, and immune response functions on D4. Pathway analysis corroborated these findings and indicated that Th2 anti-inflammatory response was inactivated, and cholesterol biosynthesis activated on D4. The ability of selected individual transcripts to predict the pregnancy outcome was poor. Overall, pregnancy success to AI coincided with a general down regulation of transcripts, appropriate regulation of the immune response, and is most likely controlled by a set of genes rather than a single transcript.

Our in vivo model allowed to associate transcriptomic changes in endometrial epithelial cells that were related from the same cycle that pregnancy was taking place. We associated the changes in transcription with pregnancy by AI, which is the most common reproductive technology used for breeding in beef cattle (Baruselli *et al.*, 2017). The sample collection was performed in the uterine body that is physically distant from the uterine-tubal junction where the embryo is expected to be located on D4. It is plausible that patterns of gene expression will differ according to the uterine region (Araújo *et al.*, 2016; Sponchiado *et al.*, 2017; Hoorn *et al.*, 2023). However, we would like to use an approach that would be relevant for practical use in the future. It is expected that at some degree changes in gene expression of key transcripts are similar in the whole organ. Another key aspect of our experimental design was including in the

study only cows that displayed estrus, where ovulation is expected in 97% (Sá Filho *et al.*, 2010; Fricke *et al.*, 2014). Therefore, pregnancy failure was likely not due to synchronization problems. To increase the relevance of our transcriptomic studies we performed a targeted and untargeted analysis of our data. By using this approach, we were able to draw more concrete conclusions in regards of pathways critical for pregnancy establishment in cattle, that is because most of the software available for enrichment analysis use algorithms that rely in the human genome to identify pathologic cellular process, which does not align well with the bovine endometrium physiological mechanism.

Genes related to cell-cycle, cellular adhesion, and nutrient transport were mostly associated negatively with pregnancy/AI. We selected these processes in our target analysis because of their previous involvement with fertility from D4 to D30 of gestation (Gao *et al.*, 2009a–c; França *et al.*, 2015, 2017; Mesquita *et al.*, 2015; D'Occhio *et al.*, 2020). Regarding cell-cycle (cellular proliferation), the down regulated genes in our study were linked to cytoskeleton function, apoptosis, and only two were part of cellular proliferation pathway. Because all of them were associated with pregnancy in the same direction, it is difficult to draw concrete conclusions of the proliferative state in the endometrial epithelial cells associated with pregnancy. Similarly, for the cellular adhesion molecules, most of the regulated genes were integrins, which not only control cellular adhesion but are also required for the cell cycle (Kamranvar *et al.*, 2022). Genes coding for nutrient transporters, such as SLCs, were also downregulated in cows that remained pregnant to AI. Amino acid transporters were the most abundant class down regulated. Regulation of amino acid transporters is consistent with data from Silva *et al*.,

2021. These authors analyzed the metabolomic composition of the ULF in *Bos indicus* heifers on D4 after estrus and reported that heifers clustered in two groups of contrasting luminal amino acid concentrations. Had those heifers being AI, one of the profiles most likely would be associated with greater pregnancy/AI. The results of the current work suggest that low amino acid concentrations would be more likely associated with pregnancy success. However, if the down regulation of the SLCs transporters reflects an effective modulation of nutrient transport to the lumen remains to be determined. Not only the presence of the transporter protein but also the basal vs. apical location on the cells is unknown. Collectively, the findings of our targeted analysis suggested a down regulation of cytoskeleton function, cellular apoptosis, cellular adhesion, and amino acids transporters. Although down regulated, genes were still being expressed in the endometrium, which likely is reflecting a greater control of the intracellular machinery events in pregnant cows and exaggerated endometrial responses in non-pregnant cows. A balance of cellular processes that require energy need to be well controlled in this period, as the uterus will experience a high metabolic demand in the subsequent weeks, to support conceptus elongation (Binelli *et al.*, 2015; Simintiras *et al.*, 2019a, 2021).

Targeted and untargeted analysis of the luminal epithelial transcriptome revealed down regulation of the Th2 anti-inflammatory system in pregnant cows. Molecules responsible for the Th2 system activation such as IL10, IL10R, IL4R, and chemokines involved in the recruitment of immune cells were previously reported to be regulated on D4 in endometrial epithelial cells by the P4 concentrations of the previous cycle (Silva *et al.*, 2023b) and negatively associated with pregnancy by ET (Martins *et al.*, 2021). Both

studies used crossbred animals in the models, rather than *Bos indicus* cattle. Consistently, with the down regulation of Th2 response, in our study, pathways related to the migration of macrophages were also down regulated in pregnant cows. We observed that such down regulation in immune cells migration was consistent with the top two regulatory networks provided by the IPA algorithm, which downstream effects were decreased cell movement and invasion. The appropriate balance between Th1 and Th2 immune responses during early pregnancy is critical for the pregnancy establishment (Oliveira *et al.*, 2013). Pregnancy is characterized by a predominant Th2 immune response in the preimplantation period (Oliveira *et al.*, 2013). Although Th2 is considered an anti-inflammatory response, exacerbated Th2 inhibits uterine response to pathogens, likely increasing susceptibility to infections (Berger, 2000). We interpret these results to mean that pregnant cows have a regular expression of Th2 response, while non-pregnant cows have an exacerbated Th2 response, which limits pregnancy success. Such exacerbated response could be due to a greater immune response to the allogenic semen used for AI (Akthar *et al.*, 2021, 2023), or an intrinsic characteristic of the cow in terms of immune response. For example, it has been shown previously that cows that undergo consecutives pregnancy loss have increased in vivo and in vitro expression of *IL6,* a Th2-modulated cytokine (Chapter 5 and Herath et al., 2009).

Increased de novo synthesis of cholesterol on D4 in endometrial epithelial cells was associated with pregnancy success on D30. The IPA algorithm selected three canonical pathways that were activated and contained similar transcripts related to cholesterol biosynthesis. Genes from these specific pathways were involved in the intracellular production of the cholesterol molecule from the acetyl-CoA precursor (Shi

et al., 2022). Similarly, when the transcriptomic signature of endometrial epithelial cells of crossbred recipient cows was investigated on D4, the activation of gene expression by SREBP pathway was up regulated in pregnant recipients (Martins *et al.*, 2021). Possible explanations for the increased intracellular biosynthesis of cholesterol include: (1) involvement of cholesterol in the cell cycle. Cholesterol is a primary lipid that is essential for membrane biogenesis, cellular proliferation, and cellular differentiation (Singh *et al.*, 2013; Yang *et al.*, 2020). Most of the intracellular cholesterol is contained within the plasma membrane. Cellular proliferation is physiologically increased during early diestrus (Johnson *et al.*, 1997), which would require increased cholesterol biosynthesis; and (2) cholesterol is being produced and accumulated in intracellular lipids droplets (van Meer, 2001; Tauchi-Sato *et al.*, 2002). It has been suggested that endometrial cells accumulate lipids in early diestrus to supply the subsequent elevated demand for lipids to the elongating conceptus (Ribeiro *et al.*, 2016b). The dynamics of lipids droplets accumulation linearly increase from D5 to D15 of gestation in the endometrium of pregnant cows (King *et al.*, 2021). Altogether our observations suggest that increase in intracellular cholesterol biosynthesis are associated with pregnancy success by either contributing to the progress of the cell cycle or being reserved as a source of energy to the increased metabolic demand during the elongation period.

The use of individual transcripts on D4 to predict the outcome of pregnancy on D30 resulted in poor accuracy. The ROC curve analysis and VIP results have shown that using the expression of a single gene to separate pregnant and non-pregnant cows on D30 did not yield acceptable accuracy or VIP scores above 1. The results reinforce the idea that success in pregnancy is influenced by multifactorial maternal, embryonic,

and environmental components. Therefore, a combination of genes and machine learning algorithms to predict the outcome of the pregnancy should be the focus of future studies. This approach was previously done on D0 of dairy cows with an accuracy of 0.88 (Hoorn *et al.*, 2024).

In conclusion, the transcriptional signature of endometrial epithelial cells on D4 after estrus is related to the ability of Brahman cows to remain pregnant on D30. Endometrial receptivity was predominantly marked by a reduction in the endometrial expression of genes involved in the cell cycle, nutrient transport, and Th2 response. One exception was intracellular cholesterol biosynthesis which was up regulated in pregnancy. It appears that the pregnant endometrium focusses the cellular efforts in the initial period of pregnancy to avoid endometrial exacerbated responses, which most likely save energy for the future increased metabolic demand in the uterus during elongation. The present study reiterates the importance of an appropriate regulation of the Th2 immune system and cholesterol biosynthesis for pregnancy establishment. Both pathways were reported previously to be regulated on D4 in crossbred cows, and now were also reported in Brahman cows.

Table 6-1. Receiver operating characteristic (ROC) curve results for the top 10 most significant genes in the RNA-seq analysis of day 4 endometrial epithelial cells between pregnant and non-pregnant cows on day 30.

	ARHGEF6	LOC508153	BRICD5	AKR1C4	A0A3Q1NBF5_BOVIN	TAF3	ADPGK	PTPRT	A0A3Q1LKW9_BOVIN	LOC101903758
Number of cows	64	64	64	64	64	64	64	64	64	64
True positive	26	22	24	17	31	26	22	31	29	13
True negative	20	22	22	26	13	17	22	12	15	31
False positive	12	10	10	6	19	15	10	20	17	1
False negative	6	10	8	15	1	6	10	1	3	19
Sensitivity[1]	0.81	0.69	0.75	0.53	0.97	0.81	0.69	0.97	0.91	0.41
Specificity[2]	0.63	0.69	0.69	0.81	0.41	0.53	0.69	0.38	0.47	0.97
Accuracy[3]	0.72	0.69	0.72	0.67	0.69	0.67	0.69	0.67	0.69	0.69

[1]Sensitivity: Probability that a test result will be positive when the cow is pregnant. [2]Specificity: Probability that a test results will be negative when the cow is non-pregnant. [3] Accuracy: Probability of correct predictions.

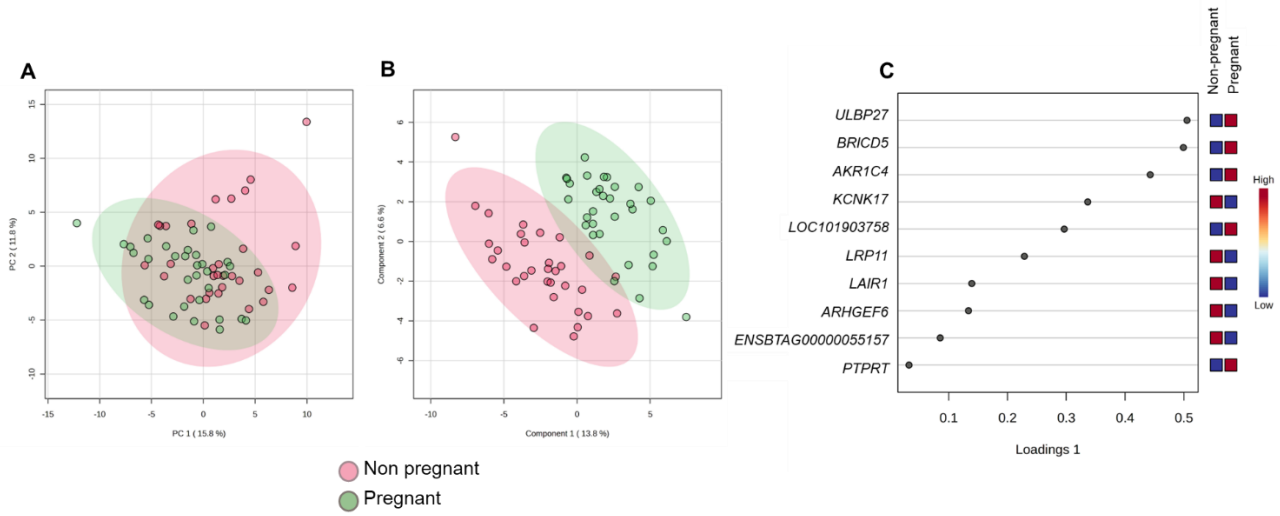

Figure 6-1. Multivariate analysis considering the 16,239 genes from the RNA-seq results and their contribution to the variability between pregnant and non-pregnant cows. (A) Principal Component Analysis (PCA) plot showing the separation between pregnant and non-pregnant cows. (B) Sparse partial least squares discriminant analysis (SPLS-DA) plot. Each datapoint is one cow, pink and green circular shades represent the 95% confidence interval. (C) Variable importance in the projection (VIP) plot, showing the top 10 genes that contributed the most for the pregnant and non-pregnant separation in the SPLS-DA analysis. The X axis shows the VIP score range, each gene correspondent VIP score is marked by the gray circle. Red and blue reflect the β coefficient for that transcript in the logistic equation to predict the pregnancy/AI. Red color means the association is positive, blue colors means the association is negative.

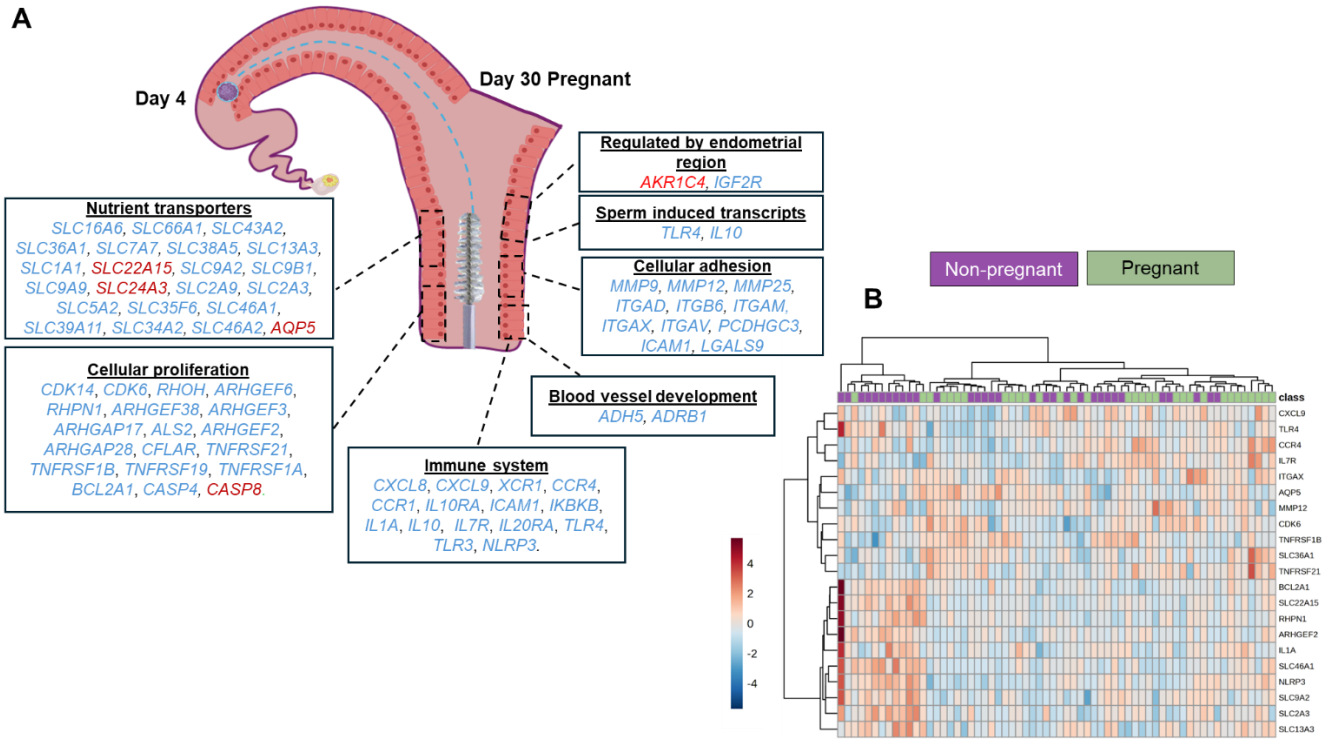

Figure 6-2. Results for targeted analysis. (A) graphical representation showing the genes for the targeted analysis which were associated positively (red) and negatively (blue) with pregnancy/AI in specific cellular processes. (B) Heatmap of the top 10 genes with greatest fold change and top 10 smallest fold change, and that were significant in the targeted analysis. Cows were clustered based on the Ward method. Pregnant cows are green and non-pregnant purple. Red means a positive association, and blue means a negative association.

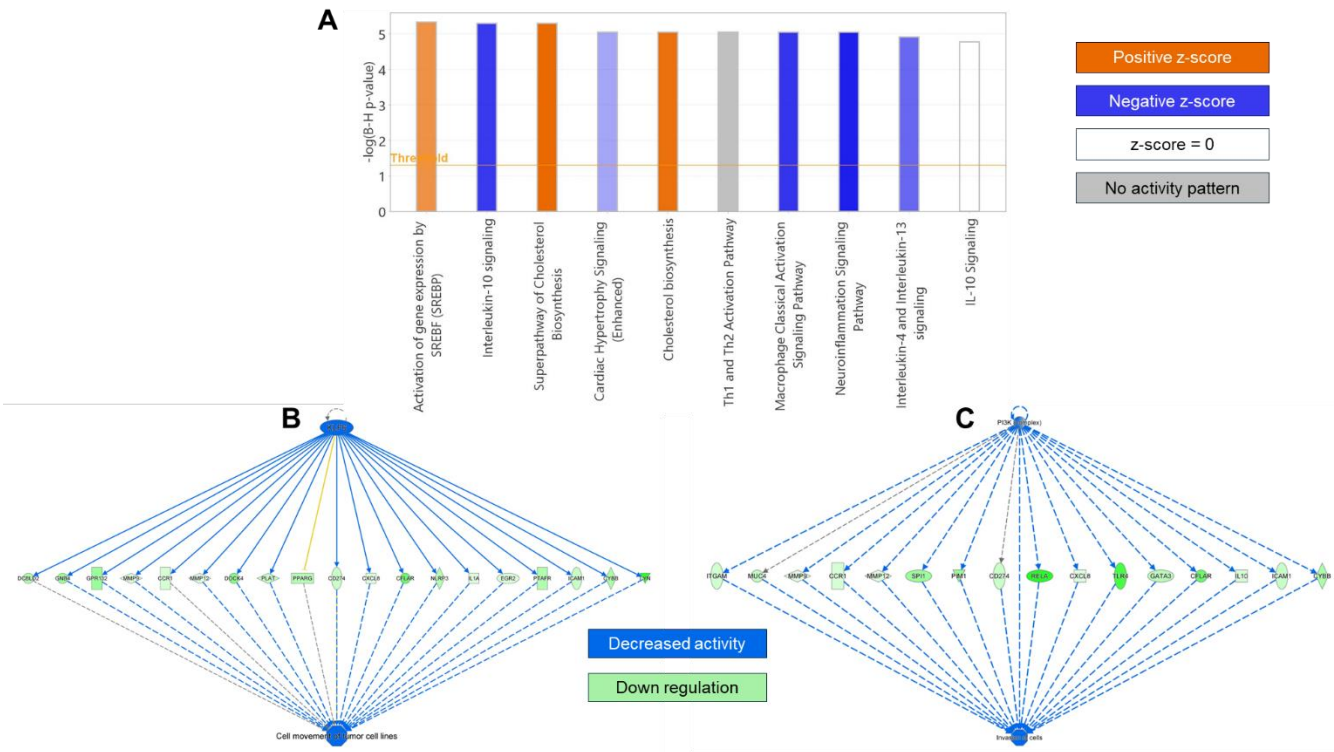

Figure 6-3. Untargeted analysis results. (A) top 10 enriched canonical pathways in the IPA analysis. Orange and blue-colored bars indicated predicted pathways activation or inactivation in relationship with pregnant cows. The color intensity reflects the Z-score of the pathway. Darker colors are more significant Z-score pathways. The Y axis shows the adjusted p-value of each pathway. Orange line crossing the graph is the threshold for the pathway to be considered significant in the adjusted p-value (1.3). (B-C) Network representation of the top 1 (B) and top 2 (C) upstream regulators, target genes in the dataset, and the consequent downstream effects. Upstream regulators colored in blue represent a decreased predicted activity in pregnant cows. Genes colored in green denote a negative relationship with the odds of pregnancy. Collor intensity indicates the degree of up or down regulation in the dataset. Connective lines colored in blue, represent a relationship that leads to inhibition. Gray and yellow connective lines indicate that the effect could not be predicted or was inconsistent, respectively. Solid and dash lines imply direct or indirect relationships, respectively. The shape of each regulated genes represents its function, which can be accessed in the IPA webpage (https://qiagen.my.salesforce-sites.com/KnowledgeBase/articles/Basic_Technical_Q_A/Legend).

UF | UNIVERSITY of FLORIDA

PLEASE REMIT TO:
UNIVERSITY OF FLORIDA BOARD OF TRUSTEES
Contracts & Grants
PO Box 931297
Atlanta, GA 31193-1297

SPONSOR:
FL CATTLE ENHANCEMENT BOARD
P.O. Box 421929
Kissimmee FL 34742-1929
United States

Invoice Date:	08/15/2024
Invoice Period:	05/01/2024 - 07/31/2024
Principal Investigator:	Binelli,Mario
Award Begin Date:	10/30/2023
Award End Date:	07/31/2024
UF FEIN:	59-6002052

Sponsor Award ID:	51
Award Title:	Predicting Fertility in Brahman Cows
Award Amount:	$38,106.00

Invoice #	I000130495
UF Award #	AWD15801
Primary Project #	P0324615
Primary Department:	60090000
Current Invoice Amount:	$13,924.89

Description	Current	Cumulative
Personnel - Salary	$60.13	$7,294.11
Personnel - Fringe Benefits	($530.03)	$3,124.89
Materials and Supplies	$2,765.85	$3,029.75
Contractual Services	$10,062.00	$10,878.79
Other Expenses	$75.00	$75.00
Direct Cost	$12,432.95	$24,402.54
Facilities and Administrative Costs	$1,491.94	$2,928.30
Total	$13,924.89	$27,330.84

For billing questions, please call 352.392.1235
Torres,Kannika S kannika@ufl.edu
Please reference the UF Award Number and Invoice Number in all correspondence

By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate, and the expenditures, disbursements and cash receipts are for the purposes and objectives set forth in the terms and conditions of the federal award. I am aware that any false, fictitious, or fraudulent information, or the omission of any material fact, may subject me to criminal, civil, or administrative penalties for fraud, false statements, false claims or otherwise. (U.S Code Title 18, Section 1001 and Title 31, Sections 3729-3730 and 3801-3812).

Kannika Torres

Certifying Official

Payment History	
Cumulative Invoices:	$27,330.84
Payments Received:	$13,405.95
Outstanding Balance:	$13,924.89
Note: Outstanding balance includes current invoice amount	

FOR UF USE ONLY Additional Projects: N

Project ID	Deptid	Department Name	Current	Cumulative
P0324615	60090000	AG-ANIMAL SCIENCES	$13,544.21	$27,330.84