


RESEARCH ARTICLE

Open Access



Transcriptome profiles in peripheral white blood cells at the time of artificial insemination discriminate beef heifers with different fertility potential

Sarah E. Dickinson^{1,2}, Brock A. Griffin¹, Michelle F. Elmore^{1,2}, Lisa Kriese-Anderson¹, Joshua B. Elmore², Paul W. Dyce¹, Soren P. Rodning¹ and Fernando H. Biase^{1*} 

Abstract

Background: Infertility is a longstanding limitation in livestock production with important economic impact for the cattle industry. Female reproductive traits are polygenic and lowly heritable in nature, thus selection for fertility is challenging. Beef cattle operations leverage estrous synchronization in combination with artificial insemination (AI) to breed heifers and benefit from an early and uniform calving season. A couple of weeks following AI, heifers are exposed to bulls for an opportunity to become pregnant by natural breeding (NB), but they may also not become pregnant during this time period. Focusing on beef heifers, in their first breeding season, we hypothesized that: a- at the time of AI, the transcriptome of peripheral white blood cells (PWBC) differs between heifers that become pregnant to AI and heifers that become pregnant late in the breeding season by NB or do not become pregnant during the breeding season; and b- the ratio of transcript abundance between genes in PWBC classifies heifers according to pregnancy by AI, NB, or failure to become pregnant.

Results: We generated RNA-sequencing data from 23 heifers from two locations (A: six AI-pregnant and five NB-pregnant; and B: six AI-pregnant and six non-pregnant). After filtering out lowly expressed genes, we quantified transcript abundance for 12,538 genes. The comparison of gene expression levels between AI-pregnant and NB-pregnant heifers yielded 18 differentially expressed genes (DEGs) (*ADAM20*, *ALDH5A1*, *ANG*, *BOLA-DQB*, *DMBT1*, *FCER1A*, *GSTM3*, *KIR3DL1*, *LOC107131247*, *LOC618633*, *LYZ*, *MNS1*, *P2RY12*, *PPP1R1B*, *SIGLEC14*, *TPPP*, *TLL1*, *UGT8*, $eFDR \leq 0.02$). The comparison of gene expression levels between AI-pregnant and non-pregnant heifers yielded six DEGs (*ALAS2*, *CNKSR3*, *LOC522763*, *SAXO2*, *TAC3*, *TFF2*, $eFDR \leq 0.05$). We calculated the ratio of expression levels between all gene pairs and assessed their potential to classify samples according to experimental groups. Considering all samples, relative expression from two gene pairs correctly classified 10 out of 12 AI-pregnant heifers ($P = 0.0028$) separately from the other 11 heifers (NB-pregnant, or non-pregnant).

Conclusion: The transcriptome profile in PWBC, at the time of AI, is associated with the fertility potential of beef heifers. Transcript levels of specific genes may be further explored as potential classifiers, and thus selection tools, of heifer fertility.

Keywords: Biomarkers, Infertility, Pregnancy

* Correspondence: fbiase@auburn.edu

¹Department of Animal Sciences, Auburn University, 559 Devall Dr, Auburn, AL 36839, USA

Full list of author information is available at the end of the article



Background

Female infertility remains a limiting factor in cattle production systems. In beef heifers, pregnancy rates vary from 53% to 95% [1–10] under natural breeding (NB), and are reduced to the range of 48–69% [1, 4, 7, 9, 11, 12] if artificial insemination (AI) is the only breeding strategy utilized. Best management practices in heifer development have been used to increase the probability of reproductive success in a heifer's first breeding season [13]. For instance, heifers that reach 60% of their mature body weight [10], have a body conformation compatible with a healthy and well-nourished animal [3, 14], present reproductive structures indicative of cyclic animals [9, 15, 16], and are bred on their third estrus versus earlier cycles [8] may have a greater chance of becoming pregnant early in the breeding season [13]. Yet, under appropriate management, many of the heifers that are deemed reproductively mature according to morphological assessment and age criteria do not become pregnant. Unexplained infertility of otherwise healthy females impacts the cattle industry negatively and is a condition of significant importance in other livestock and humans [17].

In addition to the economic losses from infertile animals, heifers that conceive late in their first breeding season to NB are likely to cause losses to beef cattle operations. Following an unsuccessful AI, heifers that become pregnant to NB and calve after the first 21 days into their first calving season remain productive in the herd for a shorter period of time and wean less total pounds of calf than their early calving counterparts [18]. Therefore, improving the selection for heifers that become pregnant by AI at the beginning of the breeding season will reduce economic losses in beef cattle operations.

Genetic selection has been used extensively to improve production and reproductive traits in beef cattle operations. In heifers, fertility is assessed by first service conception and pregnancy rate. Nonetheless, low heritability estimates for pregnancy rate (0.07–0.13 [1, 4, 19]) and first service conception (0.02–0.18 [1, 4, 19, 20]) make it challenging to leverage statistical models to guide the decision making process for sire selection to improve female fertility in cattle. As a consequence, selection for fertility in beef heifers using traditional approaches has not achieved significant progress over generations.

Strategies leveraging molecular genetics biotechnology have added new perspective to understanding the genetic architecture of fertility. To that end, genomic polymorphisms [20–24], differential gene transcription in the hypothalamus [20], endometrium [25–29], and metabolites from follicular fluids [30] have been associated with fertility in heifers or cows. In women, investigation of circulating prognostic biomarkers have yielded promising

candidates that are predictive of infertility [31], in vitro fertilization [32], or pregnancy outcomes [32, 33]. These studies, and the physiological connection between reproduction and the immune system [34], support the rationale that peripheral white blood cells (PWBC) harbor invaluable molecular information predictive of the physiological state of beef heifers pertaining to their likelihood of pregnancy establishment.

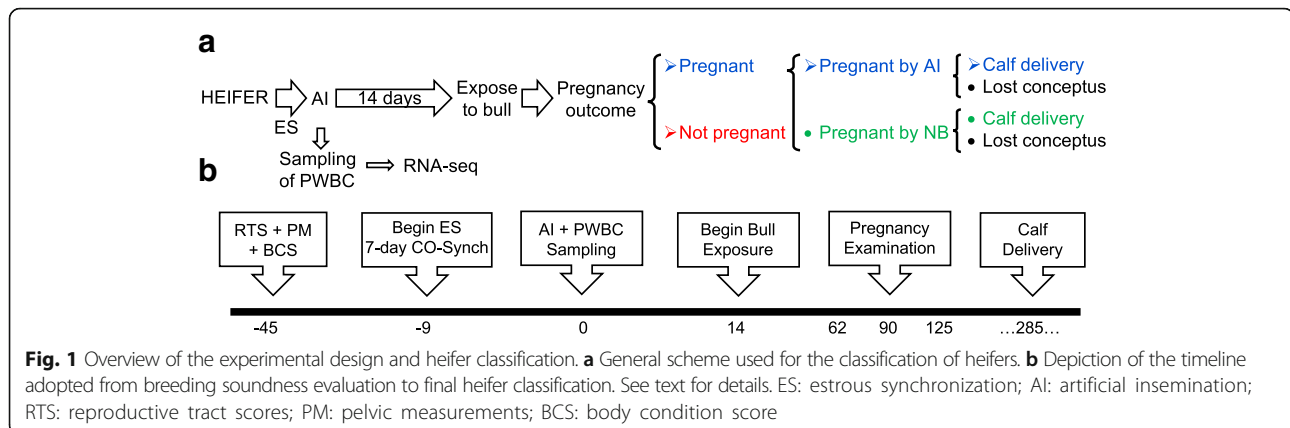
The molecular profile of circulating miRNAs [35] in the bloodstream and gene expression of PWBC [36] change during the early stages of pregnancy. Nonetheless, the molecular profiles of gene or protein expression in PWBC prior to fertilization have not been investigated as biomarkers for fertility in cattle. In this study, we tested the hypotheses that at the time of AI in beef heifers on their first breeding season: a- the transcriptome of PWBC differs between heifers that become pregnant to AI and heifers that become pregnant late in the breeding season by NB or do not become pregnant during the breeding season; and b- the ratio of transcript abundance between genes in PWBC classifies heifers according to pregnancy by AI, NB, or failure to become pregnant.

Results

Experiment overview

The experimental scheme of this study is outlined in Fig. 1a. Sixty pubertal, crossbred heifers (Angus x Simmental) were subjected to estrous synchronization followed by fixed-time AI with semen of proven fertility at two Auburn University Alabama Agricultural Experiment Stations. The heifers were then exposed to bulls for natural breeding and checked for pregnancy by rectal palpation. Figure 1b depicts the timeline of the experiment from breeding soundness to heifer classification. Heifers were identified as pregnant or not pregnant, and conceptus morphology was used to identify when conception occurred over the breeding season, for the classification of three groups (AI-pregnant, NB-pregnant, non-pregnant, Fig. 1a). We selected heifers from two experimental stations for transcriptome analyses. From station A, we carried out the experiment with AI-pregnant ($N=6$) and NB-pregnant ($N=5$) heifers. From station B, we conducted the experiment with AI-pregnant ($N=6$) and non-pregnant ($N=6$). Heifers presented similar averages for age, weaning weight, pelvic metrics, body condition score, and reproductive tract score within stations ($P>0.1$, Additional file 1: Table S1).

For each heifer, we collected peripheral blood at the time of AI, and assayed high throughput sequencing from PWBC. We generated over 557.2 million pairs of reads, averaging 20.9 million pairs of reads uniquely aligned to the bovine genome UMD3.1 [37] per sample (Additional file 1: Table S2).



Gene expression levels in PWBC associated with pregnancy outcome

We counted pairs of reads [38] according to the bovine Ensembl annotation [39] to estimate transcript abundance of expressed genes. In order to remove quantification uncertainty associated to lowly expressed genes and erroneous identification of differentially expressed genes [40, 41], we retained genes with more than one count per million (1 CPM) in six or more samples for downstream analyses, for each location independently. We quantified expression levels of 12,538 genes in all samples. Of these genes, 10,422 were expressed in PWBC of heifers located at both experimental stations. Furthermore, 1706 and 410 genes were exclusively expressed in PWBC of heifers located at experimental stations A or B, respectively (Fig. 2a). In order to strengthen the inferences of differentially expressed genes (DEG) between heifers of differential pregnancy classification, we analyzed the data from each station independently, and we adopted two algorithms implemented in the Bioconductor [42] packages edgeR [43] and DESeq2 [44]. The fold changes estimated by both algorithms were very similar ($r > 0.99$, $p < 0.0001$) and we used the output from edgeR package to report the fold changes of differential gene expression.

The comparison of gene expression profiles in PWBC between AI-pregnant and NB-pregnant heifers resulted in 18 DEGs (Fig. 2b, eFDR ≤ 0.02 , Additional file 1: Figure S1), of which *DMBT1*, *ADAM20*, *ALDH5A1*, *GSTM3*, *MNS1*, *P2RY12*, *TLL1*, *UGT8* showed greater and *ANG*, *BOLA-DQB*, *FCERIA*, *KIR3DL1*, *LOC107131247*, *LOC618633*, *LYZ*, *PPP1R1B*, *SIGLEC14*, *TPPP* displayed lower expression levels in NB-pregnant compared to AI-pregnant heifers (Table 1, Fig. 2d). Despite the low number of DEGs, we identified significant enrichment (FDR ≤ 0.002) for the GO biological process “metabolic process” (*ALDH5A1*, *GSTM3*, *LYZ*, *UGT8*).

The comparison of gene expression profiles in PWBC between AI-pregnant and non-pregnant heifers

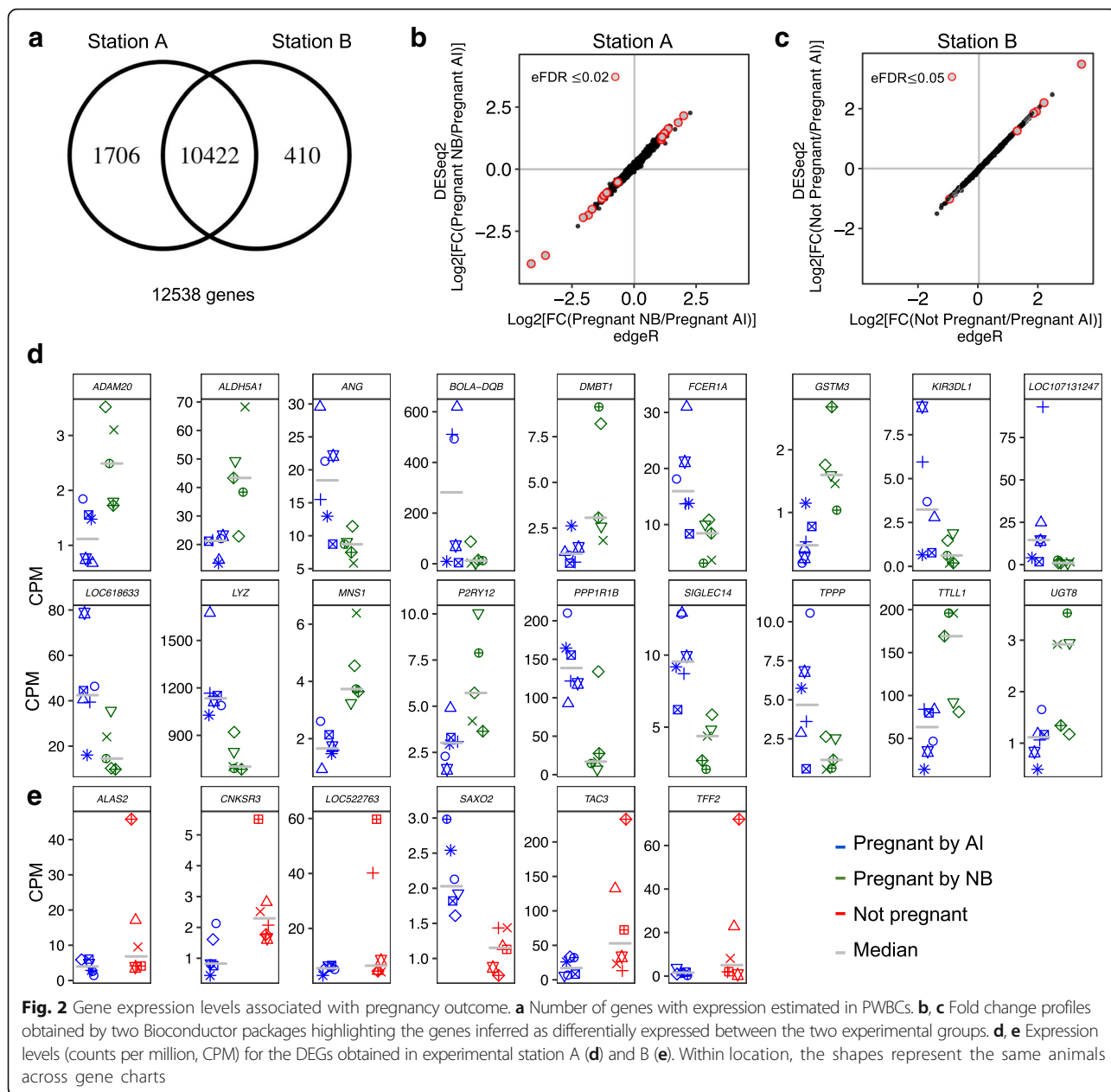
resulted in six DEGs (eFDR ≤ 0.05 , Fig. 2c, Additional file 1: Figure S1). The genes *ALAS2*, *CNKSR3*, *LOC522763*, *TAC3*, *TFE2* presented greater transcript abundance in non-pregnant heifers, whereas transcripts for *SAXO2* were less abundant in PWBC of non-pregnant heifers compared to heifers that became pregnant to AI (Table 2, Fig. 2e). No significant GO term was identified when these six DEGs were tested for enrichment of biological processes or molecular functions.

We selected the genes *ALDH5A1*, *FCERIA*, *LOC522763*, *SIGLEC14*, *TAC3*, and *TLL1* for independent assessment of differential gene expression by quantitative real-time polymerase chain reaction (qPCR). The averages of fold change calculated from the PCR data were correspondent to those obtained from RNA-seq (Spearman’s correlation = 0.94, $P < 0.02$, Additional file 1: Table S3). Therefore, we validated the results obtained by RNA-seq.

Detection of gene pairs to discriminate heifers pregnant by AI

Next, we used the top scoring pair (TSP) approach [45] to test whether the ratio between transcript levels of two genes within samples discriminated heifers presenting distinct pregnancy outcomes. According to this approach, a gene’s expression level is compared to the expression levels of all other genes. For instance, in station A, 12,128 genes formed 147,076,256 pairs, and 10,422 genes in station B formed 117,321,392 pairs.

The analysis of the transcriptome data from AI-pregnant and NB-pregnant heifers (station A) resulted in 1520 pairs of genes that discriminate most of the heifers according to their pregnancy outcome (Overall score = 1, $P < 0.0002$, 5000 randomizations). The pair of genes with the greatest discriminatory score was *DTX4* and ENSBTAG00000038233, whereby the transcript levels of *DTX4* are greater than the transcript levels of ENSBTAG00000038233 in NB-pregnant in contrast with AI-pregnant heifers (Fig. 3a). Clustering of the samples using the ratios of transcript levels of the top 20 gene



pairs (Additional file 1: Figure S2a) separated the heifers into two clusters that followed their pregnancy classification (Fig. 3b, $P \leq 0.01$, 5000 randomizations).

Analysis of the transcriptome data from 12 heifers sampled from station B, (AI-pregnant and non-pregnant) resulted in 88 gene pairs identified that separated most of the heifers in two groups (Overall score = 1, $P < 0.0002$, 5000 randomizations). The genes *U3* and *MMP19* formed the top scoring pair, in which the AI-pregnant heifers presented greater transcript abundance of *U3* compared to *MMP19*, and the opposite direction was observed for the non-pregnant heifers (Fig. 3c). Clustering of the samples using the ratios of transcript

levels of the top 20 gene pairs (Additional file 1: Figure S2b) resulted in the formation of two clusters that separated the samples by pregnancy outcome (Fig. 3d, $P < 0.01$, 5000 randomizations).

The TSP approach uses within subject transcript levels to calculate ratios between genes, and the analysis does not use variables that may create batch effects in animal experiments (i.e. time, genetic background, location of sampling). Thus, we interrogated the entire dataset (23 samples) under the binary classification of AI-pregnant ($N = 12$) and AI-not-pregnant ($N = 11$). There were four genes forming two pairs (*C11orf54*, *TAF1B*; *URB2*, *ENSTAG00000039129*) that discriminated 10 out of 12

Table 1 Differentially expressed genes associated with pregnancy originated from artificial insemination or natural breeding

Ensembl ID	Symbol	Description	LogFC(pregnant NB/pregnant AI) ^b
ENSBTAG00000022715	<i>DMBT1</i> ^a	Deleted in Malignant Brain Tumors 1	1.98
ENSBTAG00000001842	<i>GSTM3</i>	glutathione S-transferase Mu 3	1.76
ENSBTAG00000012030	<i>TTL1</i>	tubulin tyrosine ligase like 1	1.36
ENSBTAG00000000271	<i>MNS1</i>	meiosis specific nuclear structural 1	1.28
ENSBTAG00000021902	<i>ALDH5A1</i>	aldehyde dehydrogenase 5 family member A1	1.20
ENSBTAG00000004574	<i>UGT8</i>	UDP glycosyltransferase 8	1.13
ENSBTAG00000038377	<i>ADAM20</i>	ADAM metallopeptidase domain 20	1.11
ENSBTAG00000015837	<i>P2RY12</i>	purinergic receptor P2Y12	1.08
ENSBTAG00000026779	<i>LYZ</i>	Lysozyme C, non-stomach isozyme	-0.67
ENSBTAG00000045492	<i>ANG</i> ^a	angiogenin, ribonuclease, RNase A family, 5	-1.11
ENSBTAG00000040580	LOC618633 ^a	myeloid-associated differentiation marker-like	-1.22
ENSBTAG00000012887	<i>FCER1A</i>	Fc fragment of IgE receptor Ia	-1.28
ENSBTAG00000035868	<i>SIGLEC14</i> ^a	Sialic Acid Binding Ig Like Lectin 14	-1.31
ENSBTAG00000047116	<i>TPPP</i>	tubulin polymerization promoting protein	-1.70
ENSBTAG00000006035	<i>PPP1R1B</i>	protein phosphatase 1 regulatory inhibitor subunit 1B	-1.84
ENSBTAG00000047971	<i>KIR3DL1</i>	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 1 precursor	-2.05
ENSBTAG00000021077	<i>BOLA-DQB</i> ^a	major histocompatibility complex, class II, DQ beta	-3.57
ENSBTAG00000047764	LOC107131247	multidrug resistance-associated protein 4-like	-4.15

^aGenes annotated manually according to either Keeg pathways, Uniprot or NCBI Entrez databases

^bLogFC: log fold change output by edgeR package

heifers correctly (Fig. 3e, Overall score = 0.83). The clustering of 10 out of 12 AI-pregnant heifers independently from NB-pregnant and non-pregnant heifers, showed non-trivial ($P < 0.003$, hypergeometric test) patterns of ratios that identified heifers by pregnancy outcome, and clearly contrasted with ratio patterns obtained from random gene pairs (Fig. 3f).

Discussion

Our main goal was to identify differences in the transcriptome profile in PWBC at the time of AI in beef heifers with different pregnancy outcomes. In our experiment, we identified heifers that became pregnant to AI, to natural breeding, and heifers that failed to become pregnant during the breeding season. Sampling blood

from heifers of similar age and other phenotypic parameters within herd was central for us to work with pubertal heifers of similar nutritional status, and thus focus on the differences associated with the physiology of the reproduction driving the likelihood of pregnancy in beef heifers. Similar to other models of fertility and infertility in cattle [25–28], the categorical pregnancy outcomes adopted in our study identify heifers with distinct fertility potential. In the current study, we identified that variations in gene expression profiles of PWBC may be associated with the likelihood of a successful fertilization and pregnancy establishment.

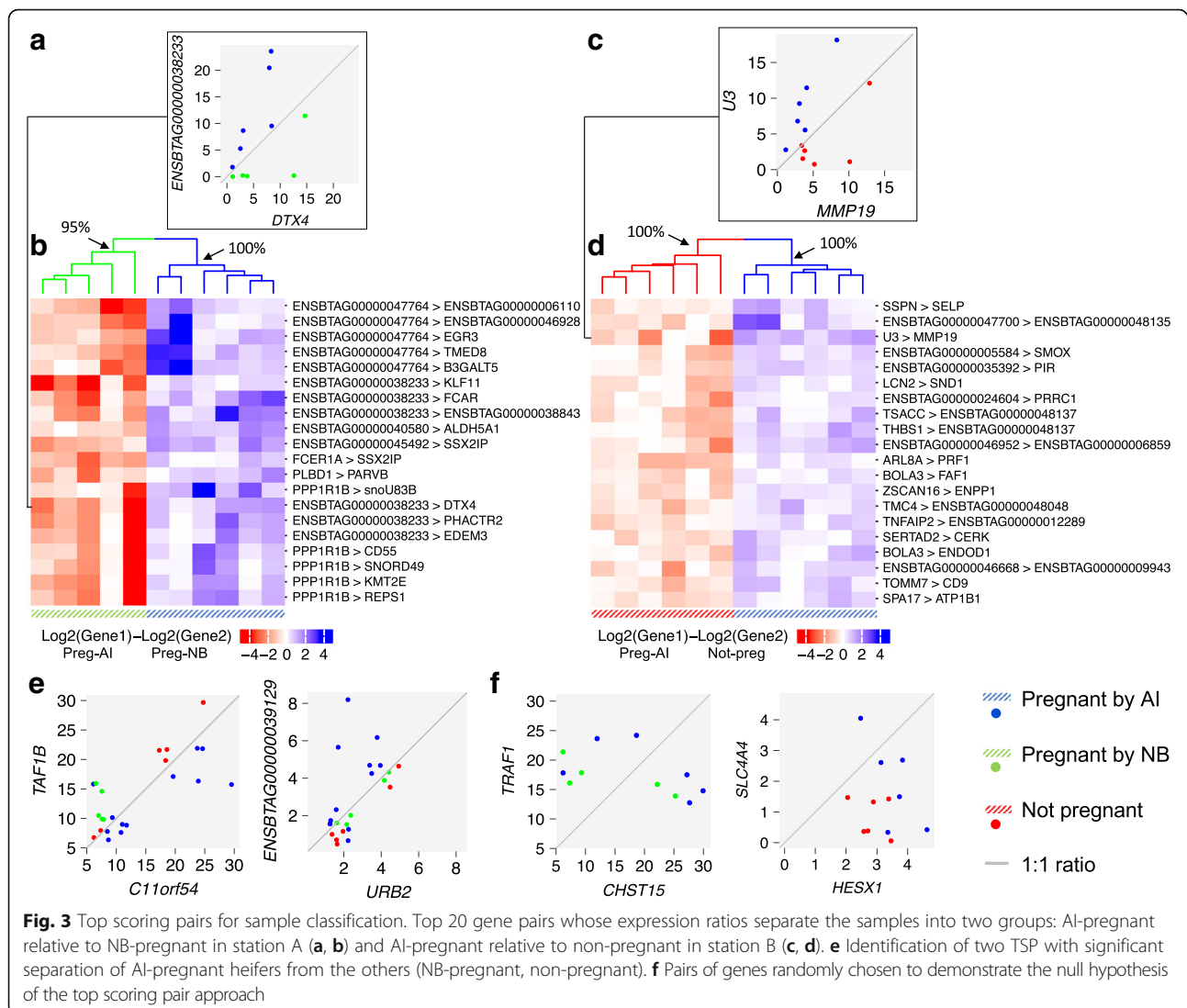
Considering the similarity of the heifers within location, as observed by age, phenotypic records (Additional file 1: Table S1), genetic background, reproductive, health, and

Table 2 Differentially expressed genes associated with pregnancy outcome in beef heifers

Ensembl ID	Gene symbol	Description	LogFC(not pregnant/pregnant AI) ^b
ENSBTAG00000030814	<i>TFF2</i>	trefoil factor 2	3.42
ENSBTAG00000021807	<i>TAC3</i>	tachykinin 3	2.17
ENSBTAG00000001308	LOC522763 ^a		1.92
ENSBTAG00000013178	<i>ALAS2</i>	5 ^l -aminolevulinic acid synthase 2	1.82
ENSBTAG00000012674	<i>CNKSR3</i>	CNKSR family member 3	1.28
ENSBTAG00000003414	<i>SAXO2</i>	stabilizer of axonemal microtubules 2	-0.98

^aGenes annotated manually according to NCBI Entrez databases

^bLogFC: log fold change output by edgeR package



nutritional management, and other environmental effects within station, one could anticipate the low number of DEGs inferred in this study. Our very stringent approach for inferring DEGs according to two independent algorithms was one reason for this observation. Nonetheless, this strategy [46] greatly reduces the chance of inferring false positives by leveraging the strengths of both algorithms [47]. Other transcriptome investigations of endometrial tissues of beef heifers [27–29] or dairy cows [48] of different fertility potential yielded DEGs in the order of few dozens. Of note, no previously identified DEGs have been found in more than one study. Furthermore, none of the DEGs identified in our study were observed in similar investigations focusing on women’s fertility [31–33]. This observation is not surprising given the polygenic and complex physiology involving fertilization and pregnancy in females.

We identified 18 DEGs associated with heifer pregnancy to AI compared to pregnancy from natural breeding. Gene ontology analysis showed significant enrichment of the biological process “metabolic process”, which included the genes “aldehyde dehydrogenase 5 family member A1” (*ALDH5A1*), “glutathione S-transferase Mu 3” (*GSTM3*), “UDP glycosyltransferase 8” (*UGT8*), and “Lysozyme C, non-stomach isozyme” (*LYZ*). The gene *ALDH5A1* is part of a family of aldehyde dehydrogenases that metabolizes aldehydes and reduces cellular toxicity. Additionally, there is evidence, in humans, that a functional *ALDH5A1* is associated with the concentration of glutathione in the bloodstream [49]. Also in humans, it has been hypothesized that up-regulation of *GSTM3* is a response to greater presence of cytotoxic products resultant of overabundance of reactive oxygen species (ROS) [50] and the need for the conjugation of ROS to glutathione [51] to mitigate the

toxic effects of ROS. As evidence supports the link between oxidative stress and female infertility in humans [51–53], the upregulation of *ALDH5A1* and *GSTM3* in PWBC suggests that a greater presence of ROS species in the blood stream may reduce the likelihood of pregnancy success to AI in beef heifers, but do not prevent the heifers from becoming pregnant to a bull later in the breeding season.

Although no significant enrichment was observed, it was noteworthy that four out of 18 DEGs were related to “cytoskeleton organization” (*MNS1*, *TLL1*, *TPPP*, *UGT8*). Interestingly, *UGT8* was down-regulated in the endometrium of women affected by implantation failure [54]. The gene *FCERIA* is associated with “positive regulation of granulocyte macrophage colony-stimulating factor production” was less expressed in NB-pregnant heifers. The down-regulation of the gene *FCERIA* in blood samples is associated with pre-term delivery in women [55]. Another gene related to the immune system, namely *KIR3DL1*, showed the lowest transcript abundance (4-fold) in NB-pregnant compared to AI-pregnant heifers. Interestingly, recurrent miscarriage patients presented lower occurrence of *KIR3DL1* in their blood compared to healthy women [56]. When expressed in natural killer (NK) cells, *KIR3DL1* inhibits [57] the cytotoxic function or the adhesive capacity of NK cells (reviewed by [58]).

We identified six DEGs in the PWBC of heifers associated with the pregnancy outcome of AI-pregnant or non-pregnant. It is critical to notice, however, that the inferences of *ALAS2*, *LOC52273*, *TAC3*, *TFE2* as DEGs, were mostly driven by some heifers that did not become pregnant, whereas others presented gene expression levels equivalent to the heifers that became pregnant to AI. The gene *TAC3* encodes the protein neurokinin B, whose expression is negatively regulated by ovarian derived steroids [59] and in turn stimulates the secretion of Gonadotropin-Releasing Hormone (GnRH) [60], which has central function on the release of follicle-stimulating hormone and luteinizing hormone. On the same note, expression of the gene *CNKSR3* was upregulated by luteinizing hormone in women’s endometrium [61] and follicular granulosa cells in buffalo cows [62]. In specific heifers, the dysregulation on these two genes is suggestive of an alteration in the hormonal feedback between the ovary the hypothalamic-pituitary axis in some of the heifers that did not become pregnant.

The TSP approach compares the levels of transcript abundance for each possible pair of genes expressed within a sample [45], and it has been used as a classification or prediction tool in biomedicine [45, 63–65]. We employed this approach to evaluate the usefulness of gene expression levels in PWBC at the time of AI for classification of heifers with different pregnancy outcome.

For each experimental station, the use of transcript levels for the top 20 pairs of genes clustered AI-pregnant heifers separately from the others with 100% confidence of cluster formation. Because this approach is parameter free [66, 67] with the exception of the binary variable that separates subjects into two categories, we used the algorithm to identify TSPs in all 23 samples that could identify AI-pregnant heifers. The ratio between the expression levels for four gene pairs misclassified only two out of the 12 AI-pregnant heifers.

Our investigation focused on PWBC, which are mostly composed of circulating immune cells. The immune system and female fertility are connected at many levels with the reproductive function in cattle (reviewed by Fair [34]), and circulating cells of the immune system respond to reproductive hormones [68, 69]. Our results show that specific genes have transcript abundance correlated with whether a heifer became pregnant to AI, could become pregnant later by natural breeding, or failed to become pregnant. We hypothesize that PWBC change their transcriptome as the heifers undergo the follicular phase of their estrous cycle. These changes most likely reflect the heifer’s readiness for fertilization.

The physiological relationship between the immune system of healthy heifers and their likelihood of becoming pregnant by AI is yet to be studied. In addition, further investigation is required to assess how our results may translate to other herds, especially when accounting for different management strategies, breeds, and genetic background. Although further work is needed to develop robust approaches to identify molecular markers in the transcriptome of PWBC, taken together, our results suggest a window of opportunity for the use of gene expression data as source of prognostic molecular markers of pregnancy likelihood in beef heifers.

Conclusions

At the time of AI, specific genes expressed in PWBC displayed differential transcript abundance in heifers classified according to their pregnancy outcome (AI-, NB-, non-pregnant). This variable expression is likely associated with the heifers’ physiological condition that relates to their fertility at the time of AI. The data suggest that the heifer’s metabolic status may be critical for the AI success, and impaired hormonal regulation is among the multiple factors that may hinge the chances of pregnancy in beef heifers. Further investigation is needed to confirm these hypotheses. Using a parameter free approach, the transcript abundance of specific gene pairs distinguished most AI-pregnant, relative to NB- or non-pregnant heifers. This result showed that the transcriptome of PWBC has a promising potential to be used as a source of data to classify heifers of distinct potential to become pregnant.

Methods

Animal handling and heifer classification according to pregnancy outcome

Crossbred beef heifers (Angus-Simmental cross) from two Auburn University research stations (Station A: Wiregrass Research and Extension Center, $n = 27$; and Station B: Black Belt Research and Extension Center, $n = 33$) were developed to reach a target weight of 60% of their mature body weight by 13.5 months of age [13, 70]. Pre-breeding examinations were performed approximately 45 days before breeding to evaluate the pubertal status of each heifer. Reproductive tract scores (scale of 1–5; 1 = pre-pubertal, 5 = pubertal, luteal phase [16]), pelvic width, and pelvic height were determined through transrectal palpation by a single, experienced veterinarian. Additionally, heifers were evaluated for body condition score (BCS; scale of 1–9 with 1 = emaciated and 9 = obese [3]).

Heifers were then subjected to estrous synchronization for fixed-time artificial insemination with the 7-Day CO-Synch protocol. Briefly, heifers received an injection of GnRH (i.m.; 100 μ g; Cystorelin[®]; Merial, Duluth, GA) and insertion of a CIDR (intravaginal insert; 1.38 g progesterone; Eazi-Breed[®] CIDR[®]; Zoetis Inc., Kalamazoo, MI) on day – 9, followed by CIDR removal and an injection of prostaglandin F₂ α (PGF; i.m.; 25 mg; Lutalyse[®]; Zoetis Inc., Kalamazoo, MI) on day – 2. All heifers then received a second GnRH injection (i.m.; 100 μ g; Cystorelin[®]; Merial, Duluth, GA) and were inseminated with a dose of semen of proven fertility on day 0, 54 \pm 2 h after CIDR removal and PGF injection. Two professionals were responsible for insemination procedures in both experimental stations, taking turns on random heifers.

Immediately after AI, 10 ml of blood was drawn from the jugular vein into vacutainers containing 18 mg K2 EDTA (Becton, Dickinson and Company, Franklin, NJ). The tubes were inverted for 8–10 times and immersed in ice. Upon arrival in the laboratory, the tubes were sprayed with 10% bleach and rinsed to eliminate contamination from the field. The tubes were centrifuged for 10 min at 2000 \times g at 4 $^{\circ}$ C. The buffy coat was removed and deposited into 14 ml of red blood cell lysis solution (0.15 M ammonium chloride, 10 mM potassium bicarbonate, 0.1 mM EDTA, Cold Spring Harbor Protocols) for 10 min at room temperature (24–25 $^{\circ}$ C). The solution was then centrifuged for 5 min at 800 \times g at 4C to pellet the PWBCs. The aqueous layer was discarded and the pellet was re-suspended in 200 μ l of RNeasy[®] (Lifetechnologies[™], Carlsbad, CA). The PWBCs were then stored at – 80 $^{\circ}$ C prior to RNA extraction. This procedure was reproduced for both experimental stations.

Fourteen days after insemination, heifers were exposed to two fertile bulls for natural breeding for the remainder of the 86 day breeding season on station A or 42 day

breeding season on station B. An experienced veterinarian performed pregnancy evaluation by transrectal palpation on day 62 and 125 post insemination at station A, and on day 95 post insemination at station B. Presence or absence of a conceptus, alongside morphological features indicating fetal age were recorded, and heifers were classified as pregnant to AI, pregnant to natural service, or non-pregnant. Heifers that became pregnant after the first 21 days of breeding were identified as late breeding for the purpose of this study.

Selection of heifers for RNA-sequencing of PWBC

Eleven heifers (six AI-pregnant and five NB-pregnant) were selected from station A, and twelve heifers (six AI-pregnant and six non-pregnant) were selected from station B for RNA-sequencing. Within station, heifers were selected according to their similarities of age and phenotypic parameters. Data for age, weaning weight, pelvic height, pelvic width, and pelvic area were compared between groups using Krustal-Wallis rank sum test. Body condition and reproductive tract scores were tested using Fisher's exact test. Tests were conducted in R software. Selected heifers did not differ for phenotypic traits associated with puberty (Additional file 1: Table S1), and all heifers were of pubertal status at the time of breeding. The selection of heifers from different groups that were phenotypically similar, according to trait average and standard deviation, avoided the addition of covariates in the analysis of differential gene expression.

RNA extraction, library preparation, and RNA sequencing

Total RNA was then isolated from PWBCs of 23 heifers using TRIzol[™] reagent (Invitrogen, Carlsbad, CA) following the manufacturer's protocol. RNA yield was quantified using the Qubit[™] RNA Broad Range Assay Kit (Eurogene, OR) on a Qubit[®] Fluorometer, and integrity was assessed on Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA) using an Agilent RNA 6000 Nano kit (Agilent, Santa Clara, CA). We obtained RIN values ranging between 7.7 and 8.8. Furthermore, samples with rRNA ratios (28S:18S) greater than 1.5 were further processed for library construction (range 1.5–1.8). Libraries were prepared with the TruSeq Stranded mRNA Library Prep kit (Illumina, Inc., San Diego, CA) following manufacturer's instructions. Libraries were quantified with Qubit[™] dsDNA High Sensitivity Assay Kit (Eurogene, OR) and quality was evaluated using the High Sensitivity DNA chip (Agilent, Santa Clara, CA) on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). Libraries were sequenced in a HiSeq 2500 system at the Genomic Services Laboratory at HudsonAlpha, Huntsville, AL to generate 125 nucleotide long pair-end reads.

RNA sequencing data processing

Sequences were trimmed of their adapters and submitted to a custom build bioinformatics pipeline [71]. Reads were aligned to the bovine genome (UMD3.1 [37]), and sequences aligning to multiple places on the genome, with 5 or more mismatches were filtered out. The sequences were then marked for duplicates, and non-duplicated pairs of reads were used for gene expression study. The read-pairs were counted against the Ensembl gene annotation [39] (version 1.87) using HTSeq [72].

Differentially expressed genes

Differences of transcript levels between samples at each experimental station were determined from fragment counts [38] using the Bioconductor packages “edgeR” [73] and “DESeq2” [44] in R software [74]. Genes were considered detected if the counts per million was greater than one in six or more samples. For each experimental station, a gene was inferred as differentially expressed if the nominal P value was ≤ 0.01 . This nominal P value corresponded to empirical false discovery rate (eFDR) of 0.02 for station A and 0.05 for station B (Additional file 1: Figure S1), as calculated according to the procedure outlined elsewhere [75], using 10,000 randomizations of sample classification.

Validation of DEGs

We used RNA extracted from the PWBC of the 23 heifers from station A and B whose PWBC transcriptome was evaluated through RNA sequencing to confirm the DEGs by RT-qPCR. We synthesized complementary DNA from 500 ng of total RNA and using oligodT₁₅ (Promega, Madison, WI). Reverse transcription was carried out with SuperScriptII (Invitrogen, Carlsbad, CA) following manufacturer’s recommendations. The final RT reaction was diluted 1:2 (v:v) and 1 μ l was used as template for each PCR reaction using Perfecta SYBR Green FastMix (Quanta Biosciences), and 100 nM of each primer (Additional file 1: Table S3, IDT) in a final volume of 10 μ l. Primers were designed using Primer-Blast application following the recommendations for obtaining target-specific primers for PCR [76]. The reactions were assayed in a Roche Light Cycler 480 equipment (Roche) equipment with pre-incubation at 95 °C for 1 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 45 s. A melting curve was generated using the thermocycler’s default parameters to validate the presence of a unique amplicon. The identification of unique amplicon is a proxy of primer specificity. Primer efficiency and cycle threshold (CT) was determined for all reactions using the LinRegPCR program [77].

We used *GAPDH* as a reference gene, which presented similar Ct values across all samples (Additional file 1: Figure S4) and showed no difference of transcript

abundance between the groups tested ($P > 0.9$, t-test, Additional file 1: Table S4). The Δ CT was calculated for each corresponding target gene relative to the reference gene, and the values of Δ CT were used as input for a t-test to assess the significance of differences between the two groups [78]. We inferred that the averages of gene expression levels were statistically different when $P \leq 0.1$. We adopted alpha = 0.1 for qPCR analysis because comparing normalized gene expression levels between groups with six samples in each group presents the power of 0.65 to detect an effect of 1 at the significance level of 0.1.

Pairs of genes with expression ratios indicating fertility categorization

Fragments per kilobase per million reads (FPKM) were calculated using the function “rpkm()” from “edgeR”. FPKM was used as input for the calculation of TSP using the package “tspair” [66]. The TSP approach [45] identifies genes whose transcript abundance ratios within each individual can classify subjects into binary categories. The ratios of the 20 TSP were used as input for hierarchical clustering of the samples, and the robustness of the clusters was calculated using 5000 randomizations with the R package “pvclust” [79].

Additional files

Additional file 1: Supplementary figures and tables. (DOCX 639 kb)

Additional file 2: R Code that reproduces the RNA-seq analyses performed in this study. (PDF 1676 kb)

Abbreviations

AI: Artificial insemination; BCS: Body condition score; CIDR: Controlled internal drug release; CPM: Count per million; CT: Cycle threshold; DEG: Differentially expressed gene; FDR: False discovery rate; FPKM: Fragment per kilobase per million reads; GnRH: Gonadotropin releasing hormone; N: Number; P: Probability; PGF: Prostaglandin F2 α ; PM: Pelvic measurements; PWBC: Peripheral white blood cells; qPCR: Quantitative real-time polymerase chain reaction; RNA: Ribonucleic acid; RT: Reverse transcription; RTS: Reproductive tract score; TSP: Top scoring pairs

Acknowledgements

The authors would like to thank the director and associates of Auburn University Wiregrass Research and Extension Center and Black Belt Research and Extension Center. We are thankful for the animal records obtained from the Alabama Beef Cattle Improvement Association. We are thankful for the comments from two anonymous reviewers.

Funding

This work was partially funded by Alabama Cattlemen’s Association, the Alabama Agricultural Experiment Station, and the Hatch program of the National Institute of Food and Agriculture, U.S. Department of Agriculture. Funding sources did not influence the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

Sequences and metadata were deposited in the NCBI GEO databank under the accession GSE103628. The R code to reproduce the RNA-sequencing analysis is deposited as Additional file 2.

Authors' contributions

FB, SR, PD conceived the study; SR, JE supervised the heifer development program, and artificially inseminated the heifers; ME, LK-A received, curated and maintained animal records; BG analyzed animal records; FB and SD performed experiments, analyses, and wrote the paper. All authors read and approved the final manuscript.

Ethics approval

All animals sourced in this study belonged to Auburn University, and had no relationship with private owners or commercial sources. All procedures with animals were performed in accordance with the protocols approved by Institutional Animal and Care and Use Committee in Auburn University.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Animal Sciences, Auburn University, 559 Devall Dr, Auburn, AL 36839, USA. ²Alabama Cooperative Extension System, Auburn, AL, USA.

Received: 13 September 2017 Accepted: 29 January 2018

Published online: 09 February 2018

References

- Bormann JM, Totir LR, Kachman SD, Fernando RL, Wilson DE. Pregnancy rate and first-service conception rate in Angus heifers. *J Anim Sci*. 2006;84:2022–5.
- Roberts AJ, Geary TW, Grings EE, Waterman RC, MacNeil MD. Reproductive performance of heifers offered ad libitum or restricted access to feed for a one hundred forty-day period after weaning. *J Anim Sci*. 2009;87(9):3043–52.
- Rae DO, Kunkle WE, Chenoweth PJ, Sand RS, Tran T. Relationship of parity and body condition score to pregnancy rates in Florida beef-cattle. *Theriogenology*. 1993;39(5):1143–52.
- Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaver RL, Silver GA, Thomas MG. Heritability and Bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. *J Anim Sci*. 2013;91(2):605–12.
- Grings EE, Geary TW, Short RE, MacNeil MD. Beef heifer development within three calving systems. *J Anim Sci*. 2007;85(8):2048–58.
- Funston RN, Deutscher GH. Comparison of target breeding weight and breeding date for replacement beef heifers and effects on subsequent reproduction and calf performance. *J Anim Sci*. 2004;82(10):3094–9.
- Funston RN, Larson DM. Heifer development systems: dry-lot feeding compared with grazing dormant winter forage. *J Anim Sci*. 2011;89(5):1595–602.
- Byerley DJ, Staigmiller RB, Berardinelli JG, Short RE. Pregnancy rates of beef heifers bred either on Puberal or 3rd estrus. *J Anim Sci*. 1987;65(3):645–50.
- Gutierrez K, Kasimanickam R, Tibary A, Gay JM, Kastelic JP, Hall JB, Whittier WD. Effect of reproductive tract scoring on reproductive efficiency in beef heifers bred by timed insemination and natural service versus only natural service. *Theriogenology*. 2014;81(7):918–24.
- Martin JL, Creighton KW, Musgrave JA, Klopfenstein TJ, Clark RT, Adams DC, Funston RN. Effect of prebreeding body weight or progesterin exposure before breeding on beef heifer performance through the second breeding season. *J Anim Sci*. 2008;86(2):451–9.
- Diskin MG, Sreenan JM. Fertilization and embryonic mortality rates in beef heifers after artificial insemination. *J Reprod Fertil*. 1980;59(2):463–8.
- Rorie RW, Lester TD, Lindsey BR, McNew RW. Effect of timing of artificial insemination on gender ratio in beef cattle. *Theriogenology*. 1999;52(6):1035–41.
- Larson RL, White BJ, Laflin S. Beef Heifer Development. *Vet Clin North Am Food Anim Pract*. 2016;32(2):285–302.
- Arango JA, Cundiff LV, Van Vleck LD. Genetic parameters for weight, weight adjusted for body condition score, height, and body condition score in beef cows. *J Anim Sci*. 2002;80(12):3112–22.
- Holm DE, Nielen M, Jorritsma R, Irons PC, Thompson PN. Evaluation of pre-breeding reproductive tract scoring as a predictor of long term reproductive performance in beef heifers. *Prev Vet Med*. 2015;118(1):56–63.
- Holm DE, Thompson PN, Irons PC. The value of reproductive tract scoring as a predictor of fertility and production outcomes in beef heifers. *J Anim Sci*. 2009;87(6):1934–40.
- Quaas A, Dokras A. Diagnosis and treatment of unexplained infertility. *Rev Obstet Gynecol*. 2008;1(2):69–76.
- Cushman RA, Kill LK, Funston RN, Mousel EM, Perry GA. Heifer calving date positively influences calf weaning weights through six parturitions. *J Anim Sci*. 2013;91(9):4486–91.
- MacNeil MD, Geary TW, Perry GA, Roberts AJ, Alexander LJ. Genetic partitioning of variation in ovulatory follicle size and probability of pregnancy in beef cattle. *J Anim Sci*. 2006;84:1646–50.
- Fortes MRS, Snelling WM, Reverter A, Nagaraj SH, Lehnert SA, Hawken RJ, DeAtley KL, Peters SO, Silver GA, Rincon G, et al. Gene network analyses of first service conception in Brangus heifers: use of genome and trait associations, hypothalamic-transcriptome information, and transcription factors. *J Anim Sci*. 2012;90:2894–906.
- Hawken RJ, Zhang YD, Fortes MRS, Collis E, Barris WC, Corbet NJ, Williams PJ, Fordyce G, Holroyd RG, Walkley JRW, et al. Genome-wide association studies of female reproduction in tropically adapted beef cattle. *J Anim Sci*. 2012;90:1398–410.
- McDanel TG, Kuehn LA, Thomas MG, Snelling WM, Smith TPL, Pollak EJ, Cole JB, Keele JW. Genomewide association study of reproductive efficiency in female cattle. *J Anim Sci*. 2014;92:1945–57.
- McDanel TG, Kuehn LA, Thomas MG, Snelling WM, Sonstegard TS, Matukumalli LK, Smith TPL, Pollak EJ, Keele JW. Y are you not pregnant: identification of Y chromosome segments in female cattle with decreased reproductive efficiency. *J Anim Sci*. 2012;90:2142–51.
- Fortes MRS, Deatley KL, Lehnert SA, Burns BM, Reverter A, Hawken RJ, Boehansen G, Moore SS, Thomas MG. Genomic regions associated with fertility traits in male and female cattle: advances from microsatellites to high-density chips and beyond. *Anim Reprod Sci*. 2013;141:1–19.
- Peterson AJ, Donnison MJ, Pearson S, McMillan WH. Contrasting early embryo development in a herd of recipient cattle with previously high or low pregnancy rates. *Theriogenology*. 1999;51(1):229.
- McMillan WH, Donnison MJ. Understanding maternal contributions to fertility in recipient cattle: development of herds with contrasting pregnancy rates. *Anim Reprod Sci*. 1999;57(3–4):127–40.
- Minten MA, Bilby TR, Bruno RG, Allen CC, Madsen CA, Wang Z, Sawyer JE, Tibary A, Neibergs HL, Geary TW, et al. Effects of fertility on gene expression and function of the bovine endometrium. *PLoS One*. 2013;8(8):e69444.
- Geary TW, Burns GW, Moraes JG, Moss JI, Denicol AC, Dobbs KB, Ortega MS, Hansen PJ, Wehrman ME, Neibergs H, et al. Identification of beef heifers with superior uterine capacity for pregnancy. *Biol Reprod*. 2016;95(2):47.
- Salliew-Wondim D, Holker M, Rings F, Ghanem N, Ulas-Cinar M, Peippo J, Tholen E, Looft C, Schellander K, Tesfaye D. Bovine pretransfer endometrium and embryo transcriptome fingerprints as predictors of pregnancy success after embryo transfer. *Physiol Genomics*. 2010;42(2):201–18.
- Bender K, Walsh S, Evans ACO, Fair T, Brennan L. Metabolite concentrations in follicular fluid may explain differences in fertility between heifers and lactating cows. *Reproduction*. 2010;139:1047–55.
- Michou VI, Kanavaros P, Athanassiou V, Chronis GB, Stabamas S, Tsilivakos V. Fraction of the peripheral blood concentration of CD56(+)/CD16(-)/CD3(-) cells in total natural killer cells as an indication of fertility and infertility. *Fertil Steril*. 2003;80:691–7.
- Thum MY, Bhaskaran S, Abdalla HI, Ford B, Sumar N, Shehata H, Bansal AS. An increase in the absolute count of CD56(dim)CD16(+)-CD69(+) NK cells in the peripheral blood is associated with a poorer IVF treatment and pregnancy outcome. *Hum Reprod*. 2004;19(10):2395–400.
- Dons'koi B, Chernyshov VP, Osypchuk DV. Peripheral blood natural killer cells activation status determined by CD69 upregulation predicts implantation outcome in IVF. *J Reprod Immunol*. 2014;101:45.
- Fair T. The contribution of the maternal immune system to the establishment of pregnancy in cattle. *Front Immunol*. 2015;6(7)
- Pohler KG, Green JA, Moley LA, Gunewardena S, Hung WT, Payton RR, Hong X, Christenson LK, Geary TW, Smith MF. Circulating microRNA as candidates for early embryonic viability in cattle. *Mol Reprod Dev*. 2017;84(8):731–43.
- Pugliesi G, Miagawa BT, Paiva YN, Franca MR, Silva LA, Binelli M. Conceptus-induced changes in the gene expression of blood immune cells and the

- ultrasound-accessed luteal function in beef cattle: how early can we detect pregnancy? *Biol Reprod.* 2014;91(4):95.
37. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, Hanrahan F, Pertea G, Van Tassell CP, Sonstegard TS, et al. A whole-genome assembly of the domestic cow, *Bos Taurus*. *Genome Biol.* 2009;10:R42.
 38. Anders S, McCarthy DJ, Chen YS, Okoniewski M, Smyth GK, Huber W, Robinson MD. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc.* 2013;8(9):1765–86.
 39. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, et al. Ensembl 2014. *Nucleic Acids Res.* 2013;42:D749–55.
 40. Bullard JH, Purdom E, Hansen KD, Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics.* 2010;11:94.
 41. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11:R25.
 42. Reimers M, Carey VJ. Bioconductor: an open source framework for bioinformatics and computational biology. *Methods in enzymol.* 2006;411:119–34.
 43. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 2012;40:4288–97.
 44. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
 45. Geman D, d'Avignon C, Naiman DQ, Winslow RL. Classifying gene expression profiles from pairwise mRNA comparisons. *Stat Appl Genet Mol Biol.* 2004;3: Article19.
 46. Suarez-Vega A, Gutierrez-Gil B, Klopp C, Robert-Granie C, Tosser-Klopp G, Aranz JJ. Characterization and comparative analysis of the milk transcriptome in two dairy sheep breeds using RNA sequencing. *Sci Rep-Uk.* 2015;5:18399.
 47. Schurch NJ, Schofield P, Gierlinski M, Cole C, Sherstnev A, Singh V, Wrobel N, Gharbi K, Simpson GG, Owen-Hughes T, et al. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA.* 2016;22(6):839–51.
 48. Moore SG, Pryce JE, Hayes BJ, Chamberlain AJ, Kemper KE, Berry DP, McCabe M, Cormican P, Lonergan P, Fair T, et al. Differentially expressed genes in endometrium and corpus luteum of Holstein cows selected for high and low fertility are enriched for sequence variants associated with fertility. *Biol Reprod.* 2016;94(1):19.
 49. Niemi AK, Brown C, Moore T, Enns GM, Cowan TM. Evidence of redox imbalance in a patient with succinic semialdehyde dehydrogenase deficiency. *Mol Genet Metab Rep.* 2014;1:129–32.
 50. Corton M, Botella-Carretero JI, Lopez JA, Camafeita E, Millan JLS, Escobar-Morreale HF, Peral B. Proteomic analysis of human omental adipose tissue in the polycystic ovary syndrome using two-dimensional difference gel electrophoresis and mass spectrometry. *Hum Reprod.* 2008;23(3):651–61.
 51. Lamoureux GL, Rusness DG. The mechanism of action of Bas-145-138 as a Safener for Chlorimuron ethyl in corn - effect on hydroxylation, glutathione conjugation, glucoside conjugation, and Acetolactate synthase. *Pestic Biochem Phys.* 1992;42(2):128–39.
 52. Gupta S, Ghulmiyyah J, Sharma R, Halabi J, Agarwal A. Power of proteomics in linking oxidative stress and female infertility. *Biomed Res Int.* 2014;2014: 916212.
 53. Ruder EH, Hartman TJ, Goldman MB. Impact of oxidative stress on female fertility. *Curr Opin Obstet Gyn.* 2009;21(3):219–22.
 54. Maekawa R, Taketani T, Mihara Y, Sato S, Okada M, Tamura I, Jozaki K, Kajimura T, Asada H, Tamura H, et al. Thin endometrium transcriptome analysis reveals a potential mechanism of implantation failure. *Reprod Med Biol.* 2017;16(2):206–27.
 55. Enquobahrie DA, Williams MA, Qiu CF, Muhie SY, Slentz-Kesler K, Ge ZP, Sorenson T. Early pregnancy peripheral blood gene expression and risk of preterm delivery: a nested case control study. *BMC Pregnancy Childb.* 2009;9:56.
 56. Faridi RM, Das V, Tripathi G, Talwar S, Parveen F, Agrawal S. Influence of activating and inhibitory killer immunoglobulin-like receptors on predisposition to recurrent miscarriages. *Hum Reprod.* 2009;24(7):1758–64.
 57. Rajagopalan S, Long EO. Understanding how combinations of HLA and KIR genes influence disease. *J Exp Med.* 2005;201(7):1025–9.
 58. Pegram HJ, Andrews DM, Smyth MJ, Darcy PK, Kershaw MH. Activating and inhibitory receptors of natural killer cells. *Immunol Cell Biol.* 2011;89(2):216–24.
 59. Rance NE, Bruce TR. Neurokinin-B gene-expression is increased in the arcuate nucleus of Ovariectomized rats. *Neuroendocrinology.* 1994;60(4):337–45.
 60. Lehman MN, Coolen LM, Goodman RL. Minireview: Kisspeptin/neurokinin B/Dynorphin (KNDy) cells of the arcuate nucleus: a central node in the control of gonadotropin-releasing hormone secretion. *Endocrinology.* 2010;151(8): 3479–89.
 61. Humaidan P, Van Vaerenbergh I, Bourgain C, Alsjberg B, Blockeel C, Schuit F, Van Lommel L, Devroey P, Fatemi H. Endometrial gene expression in the early luteal phase is impacted by mode of triggering final oocyte maturation in recFSH stimulated and GnRH antagonist co-treated IVF cycles. *Hum Reprod.* 2012;27(11):3259–72.
 62. Rao JU, Shah KB, Puttaiah J, Rudraiah M. Gene expression profiling of preovulatory follicle in the buffalo cow: effects of increased IGF-I concentration on periovulatory events. *PLoS One.* 2011;6(6):e20754.
 63. Zhao H, Logothetis CJ, Gorlov IP. Usefulness of the top-scoring pairs of genes for prediction of prostate cancer progression. *Prostate Cancer Prostatic Dis.* 2010;13(3):252–9.
 64. Shi P, Ray S, Zhu Q, Kon MA. Top scoring pairs for feature selection in machine learning and applications to cancer outcome prediction. *BMC Bioinformatics.* 2011;12:375.
 65. Czajkowski M, Kretowski M. Top scoring pair decision tree for gene expression data analysis. *Adv Exp Med Biol.* 2011;696:27–35.
 66. Leek JT. The tspan package for finding top scoring pair classifiers in R. *Bioinformatics.* 2009;25(9):1203–4.
 67. Magis AT, Price ND. The top-scoring 'N' algorithm: a generalized relative expression classification method from small numbers of biomolecules. *BMC Bioinformatics.* 2012;13:227.
 68. Athreya BH, Pletcher J, Zulian F, Weiner DB, Williams WV. Subset-specific effects of sex hormones and pituitary gonadotropins on human lymphocyte proliferation in vitro. *Clin Immunol Immunopathol.* 1993;66(3):201–11.
 69. Giglio T, Imro MA, Filaci G, Scudeletti M, Puppo F, De Cecco L, Indiveri F, Costantini S. Immune cell circulating subsets are affected by gonadal function. *Life Sci.* 1994;54(18):1305–12.
 70. Patterson DJ, Smith MF. Management considerations in beef heifer development and puberty. *Vet Clin North Am Food Anim Pract.* 2013; 29(3):xiii–xiv.
 71. Biase FH, Rabel C, Guillomot M, Hue I, Andropolis K, Oliveira R, Wallace R, Le Bourhis D, Richard C, et al. Massive dysregulation of genes involved in cell signaling and placental development in cloned cattle conceptus and maternal endometrium. *P Natl Acad Sci USA.* 2016;113(51):14492–501.
 72. Anders S, Pyl PT, Huber W. HTSeq - a python framework to work with high-throughput sequencing data. *Bioinformatics.* 2014;31:166–9.
 73. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26:139–40.
 74. Ihaka R, Gentleman R: a language and environment for statistical computing. *J Comput Graph Stat.* 2012;5:299–14.
 75. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *P Natl Acad Sci USA.* 2003;100:9440–5.
 76. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics.* 2012;13:134.
 77. Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett.* 2003;339(1):62–6.
 78. Yuan JS, Reed A, Chen F, Stewart CN Jr. Statistical analysis of real-time PCR data. *BMC Bioinformatics.* 2006;7:85.
 79. Suzuki R, Shimodaira H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics.* 2006;22:1540–2.