5-2018

Effects of Fescue Toxicosis on Whole Blood Gene Expression in Beef Cattle

Julie Atchley
University of Arkansas, Fayetteville

Effects of Fescue Toxicosis on Whole Blood Gene Expression in Beef Cattle

Julie A. Atchley

University of Arkansas

Dale Bumpers Honors College

KEY WORDS: fescue, gene expression, pathways, RNA-seq

TABLE OF CONTENTS

INTRODUCTION …… 3

LITERATURE REVIEW ……3-12

What is Fescue Toxicosis? …… 3-4

Utilization of RNA Sequencing Methods ……. 5-6

How Does RNAseq Work? ……..6-10

Gene Expression and Pathway Ontology Enrichment Analysis ……10-11

Previous Gene Expression Studies and Fescue Toxicosis ……11-12

MATERIALS AND METHODS ……...12-15

Population and Experimental Design ……12-13

Sample Collection ……....13

RNA Isolation, Quality Control, and Sequencing ……. 12-14

Data Analysis ……..... 14-15

RESULTS AND DISCUSSION ……... 15-19

Results ……... 15-17

Data Filtering ……15-16

DE Genes …….....16

GO Processes and KEGG Pathways Enriched by DAVID ……16-17

Discussion and Conclusion ……17-21

Tables …… 22-26

ACKNOWLEDGEMENTS ……27

BIBLIOGRAPHY..……28-30

Introduction

The United States beef industry is a large and profitable enterprise, incorporating approximately 913,246 beef cattle operations with a total of 93.5 million head of cattle with a predicted revenue of 67.56 billion dollars (NCBA, 2017). According to the USDA Census of Agriculture for 2012, Arkansas boasted a total of 23,422 cow/calf operations with an average of 40 head (USDA, 2018). The state of Arkansas has approximately 6 million acres of pasture land and the dominant forage types are cool season grasses, especially endophyte-infected tall fescue and Bermuda grasses (USDA, 2018). The effects of fescue toxicosis are a continuing concern for herd health and for the productivity and profitability of the beef industry because they are estimated to impact 20% of the US beef herd, resulting in an estimated $1 billion in annual losses to producers (Pratt et al., 2015).

Literature Review

What is Fescue Toxicosis?

Endophyte-infected tall fescue (toxic fescue) is one of the most dominant forage types in the Mid-Western United States (Stowe et al., 2014). Extremely hardy and drought resistant, toxic tall fescue grass owes these qualities to the symbiotic relationship it has developed with an endophyte fungus (*Epichloe coenophiala,* formerly *Neotyphodium*) that grows inside the plant (Coufal-Majewski et al., 2016; Leuchtmann et al., 2013). When it was first planted as a pasture grass in the United States in the early 1940s through late 1950s, toxic tall fescue was celebrated for its erosion control and drought resistant qualities; however, when ingested by cattle it has been proven to cause negative effects on growth, reproduction, longevity, and overall health of

beef cattle (Strickland et al., 2011; Bacon et al., 1995; Hoveland et al., 2009; Stowe et al., 2014). The endophytic fungus produces mycotoxins, called ergot-alkaloids, that bind to dopamine receptors and act as vasoconstrictors, among an array of other physiological effects (Strickland et al., 2011). The ergot-alkaloid that contributes primarily to the vasoconstrictive effects of this endophyte and is at highest levels is ergovaline (EV); however, there are other toxins produced, such as lysergic acid (Rattray et al., 2010; Strickland et al., 2011). The detrimental effects of this endophytic fungus on cattle have been termed "fescue toxicosis" by researchers and producers (Stowe et al., 2014). In mild to moderate cases of fescue toxicosis, symptoms may include reduced fertility, milk production, longevity, and rate of gain, as well as retention of hair coat (Moisa et al., 2015). In severe cases, fescue toxicosis can cause lameness and loss of extremities such as ears, tails, hooves, and sometimes entire limbs due to vasoconstriction in cold/wintry weather (Coufal-Majewski et al., 2016; Strickland et al., 2009; Rattray et al., 2010). However, many of the symptoms related to fescue toxicosis are less visible and production losses often go unrecognized by producers until an extreme case occurs (Strickland et al., 1993, 2009; Schmidt & Osborn, 1993; Oliver, 1997, 2005). In the United States alone, complications from fescue toxicosis cause the beef industry an average of over 1 billion dollars a year, making it one of the largest animal health related production loss for the grazing livestock industry (Strickland et al., 2009, 2011; Stowe et al., 2014). Since the majority of beef cattle operations in Arkansas are cow/calf producers (Smith & Cassidy, 2015), the impact of fescue toxicosis on cow reproduction and calf growth is extremely relevant to producers (Smith & Cassidy, 2015). However, as stated in the article by Strickland et al. 2011, "Although progress has been made in mitigating the problem [of fescue toxicosis], a comprehensive solution to the intoxication has not been realized" (Strickland et al., 2011).

Utilization of RNA Sequencing Methods:

Whole transcriptome profiling is a powerful tool used to identify the abundance of all gene transcripts within a specific cell type or tissue (Kurkurba and Montgomery, 2015). Over the years, RNA sequencing has evolved from Northern Blot, qPCR, Q-RT-PCR, microarray, and SAGE, which all consume time and resources with lower throughput. Recently, RNA-sequencing (RNA-seq) has become the method of choice to determine the overall gene expression values for thousands of gene transcripts simultaneously. This whole transcriptome approach sequences all transcripts from isolated RNA within a tissue in short reads and allows researchers to count these short sequences as one "unit" of gene expression for a given gene transcript. In this way, gene expression can be analyzed as "counts" for each gene transcript and compared across treatment levels to make statistical comparisons. The number of counts per transcript is a function of the read length (typically 50-150 base pairs (bps) in length) and transcript length. Differences in the abundance of RNA expression across the transcriptome can be used to determine differences in transcript abundance as a function of the impact of treatments on the cells being studied (Illumina, Inc., 2018). This is especially useful in identifying differences in RNA expression due to disease or different environmental conditions, such as those caused by endophyte-infected tall fescue and fescue toxicosis (Illumina, Inc., 2018). However, in some cases the differences in transcript abundance are due to functional roles of the transcripts and in other cases they are reactive signatures to changes that have already occurred in the cells. In either case, transcript levels can be used as biomarkers of a physiological process. In this study, gene expression differences due to toxic fescue exposure may be a direct "causal" physiological impact on cells within an animal or they may be signaling a reactionary cellular response to detoxify the cell, mitigate cellular damage, or other "reactive" processes. Thus, gene

expression profiling by RNA-seq is only an association with the treatment and cannot be considered a causal factor without additional experimental and statistical evidence to support the claim.

The methodology for RNA-seq based transcriptional profiling used for this study was a relatively new method known as Quant-seq. Quant-seq methodology utilizes known genome sequencing information and uses proprietary technology to capture and sequence only the 3' poly-Adenylated (poly A) tail of the transcript, thus reducing the cost of sequencing (Lexicogen, Greenland, NH). In the context of this study, 150 base pairs of the 3' end were sequenced including a portion of the poly A tail of each transcript using Illumina technology (Illumina Inc., San Diego, CA). Since most transcripts are more than 1000 bps in length, this method results in at least a $1/10^{th}$ reduction in the amount of sequencing than would be necessary with other methods. The reduction in sequencing required to estimate the expression of a single gene transcript allows more samples to be combined, or multiplexed, into one lane within an Illumina technology-based sequencer flow cell. A flow cell contains eight, glass "lanes" imprinted with nano-wells that allow a multitude of single samples to be amplified individually in one run, which greatly reduces the cost of RNA-seq for whole transcriptome profiling experiments. The Quant-seq library kit includes bar code adaptors (i.e. short sequences) which act as signatures or fingerprints that are unique to a sample. Individual barcodes are attached to each amplified 3' end sequence generated to allow the identification of individual sample ID's based on the unique sequence added to the original copies of DNA or RNA, resulting in multiple amplicons (the length of sequence generated by an amplification method, such as PCR) for each sample. This approach to sequencing allows up to 96 unique samples to be multiplexed into a single sequencing lane without losing individual sample identification.

Processes required to conduct RNA-seq:

RNA sample preparation and isolation

Proper RNA isolation is the critical first step of RNA sequencing as it will determine the quantity and quality of RNA available for sequencing (Thermo Fisher Scientific, Inc., Waltham, MA). Ribonucleic acid can be isolated from any bodily tissue. There are several different methods used to isolate RNA including organic extraction, filter-based spin basket formats, magnetic particle methods, and direct lysis methods (Thermo Fisher Scientific, Inc., Waltham, MA). Each method of isolation has its benefits and drawbacks and is often chosen based on the sample type from which RNA will be isolated (Thermo Fisher Scientific, Inc., Waltham, MA). For this study, whole blood samples were collected using Tempus™ tubes that contain a specific reagent that simultaneously lyses whole blood cells, prevents RNase enzyme activity, and precipitates the RNA from the lysed blood cells (Thermo Fisher Scientific, Inc., Waltham, MA). Isolation of RNA can then be accomplished using the Tempus System RNA isolation kit (Thermo Fisher Scientific, Inc., Waltham, MA) that includes on-column DNase digestion to remove genomic DNA contamination. After isolation RNA quantity and purity are measured using a spectrophotometer. Then, RNA integrity (i.e. lack of RNA degradation) is measured to verify that full length transcripts are sequenced and not fragmented RNA pieces. The isolated and quality-checked RNA can then be prepared for sequencing.

How does RNA-seq work on Illumina technology?

RNA-seq using Illumina technology is composed on four basic steps as defined by Illumina, Inc., including library preparation, cluster generation, sequencing and data analysis. The next step in RNA sequencing is to create a library for analysis; this entails converting the

RNA to complementary DNA, or cDNA, and adding sequence adaptors that allow the amplification of genes (Kurkurba and Montgomery, 2015). This library enables the creation of multiplexes of 30 samples per flow cell lane with a total of 8 flow cell lanes available per run so that every sample can be run at one time and identified by their specific library, or barcode, that has been attached (Kurkurba and Montgomery, 2015). Hybridization of the RNA is conducted, which consists of adding reagents to resuspend the RNA and prepare the optimal concentration for sequencing. Molecules of RNA are bound into place in the nano-wells in the flow lanes and prepared for sequencing. Once sequencing is initialized, the RNA molecules are amplified, and clusters are formed by the addition of single base nucleotides that are marked by specific fluorescent dyes. Clusters contain approximately 500 copies of RNA and are 1 micron in diameter. Computer controlled imaging then reads the fluorescent dye markers and projects tens of millions of parallel clusters. Sequencing by synthesis allows one fluorescent base to be added per cycle, the fluorescence is read and flushed between each cycle, and the cycle is repeated 50 to 150 times on Illumina HiSeq machines resulting in a 50 to 150-base pair (bp) short sequences, or "reads" (Kurkurba and Montgomery, 2015). Sequences are then saved in FASTQ format files, which preserves sequence content, quality and other features descriptive of the sequencing process (Cock et al., 2009). Sequence reads are then summarized using a series of software to turn reads into gene expression values that can be analyzed using statistical methods to determine differential expression across treatments.

Bioinformatic Analysis of RNA-seq data to summarize gene expression levels

Summarization of FASTQ data from individual samples into gene expression data requires a series of bioinformatic software that remove poor-quality sequences, check sequence quality for minimum standards, align sequence to a reference genome, assign reads to genes

based on annotation information, and statistically analyze differences by treatment (Cock et al., 2009). Poor-quality sequences are removed with a trimming software (e.g. Sickle; Joshi et al., 2011) to make sure that reads with poor sequence quality are removed and that reads with poor sequences at their ends (a common problem during DNA sequencing) are trimmed to maintain only the high-quality sequences. Sequence quality can be determined using software such as FASTQC (Andrews, 2010) that check for poor-quality sequences that contain ambiguous DNA (i.e. Ns, not A, T, G or C sequences), or other problems. When samples have a reference genome (i.e. sequenced draft of the full genetic code of a species, often in FASTA sequence format) they are then aligned using an aligner (e.g. STAR software; Dobin et al., 2013). In cases where no reference is available, a reference transcriptome can be created to serve as the catalogue (i.e. reference) of all transcripts expressed within the tissue of study. In the case of this experiment, the bovine UMD 3.1 reference genome was used for alignment (Aleksey et al., 2009). In either case, only reads that align uniquely (i.e. one time in the genome) are often the only reads used for analysis to remove any doubt of incorrect assignment of a sequence to expression of an incorrect gene. This often removes a considerable number of reads (10-35%) from consideration as reflecting gene expression. A common issue in handling sequence data is the incompatibility of file formats across software. SAMtools (Li et al., 2009) provides options to format data for a variety of other software and can be used to change sequence data into a more usable format for software that summarize read counts into expression levels at the gene level. Software such as HTseq (Anders et al., 2015) are used to count the read alignments to specific gene annotations based on the position of read maps in comparison to previously defined gene coordinates (Anders et al., 2015). Many genomes now have fully annotated transcripts which are kept for easy use at databases such as ENSEMBL (Zerbino et al., 2018).

The viable sequence reads identified within a gene by HTSeq are then formatted and filtered for statistical analysis using commonly used computer programming languages (e.g. Perl, Python, R, Unix, etc.). Statistical analysis of read count data is typically performed using methods that account for the count-based nature of RNA-seq data using a negative binomial distribution, which also accounts for over-dispersion in the data. Additional filters are often developed upon visual inspection of data by plotting to remove samples with low or no read count or samples with low or inconsistent expression distributions. Of note, RNA-seq samples are notorious for highly expressed outliers. For this reason, data are normalized not only for read count based on the 75% quantile or median of the data but also for the overall library size to account for possible technical issues in library creation, imprecise pipetting (deviation in RNA quantity) and other within sample deviations in sample collection and preparation that could impact the total number of reads sequenced. Accounting for multiple testing (i.e. performing thousands of tests and thus identifying statistical differences by chance) is a final critical component of RNA-seq analysis. False discovery rate correction is a commonly used technique to control type I error.

Gene Expression and Pathway Ontology Enrichment Analysis:

Gene expression encompasses the entire process of transcribing DNA to mRNA and translating mRNA to proteins; a gene is considered expressed when the proteins it codes for are present and active in the function of the organism (NIH, 2018). Gene pathways are a biological series of actions (e.g. like an electrical circuit) within a cell that lead to a certain biological process or change in the cell (NHGRI, 2018). Such biological processes could be related to cellular growth, death, immune response or any number of other very detailed processes that are often specific to a specific cell type. Often in experiments concerning gene expression, so many

genes are identified as differentially expressed that it is difficult to interpret the biological meaning of so many transcripts at once. Because gene pathways, or biological processes, are often affected by tens, hundreds or thousands of gene transcripts, a common approach to summarizing information from the data is to sort the gene transcripts into pathways. Gene ontologies are major bioinformatics initiatives, with the main goal of connecting information to genes and gene products (i.e. such as pathway membership) across all species. The gene ontology initiative is especially relevant to the study of gene expression pathways, because it attempts to annotate the functions and products of all known genes to a specific sequence, creating a consortium of information that is easily accessed from a single location. The Kyoto Encyclopedia of Genes and Genomes, KEGG, is an integrated database that aids in the high-level understanding of biological systems, such as what genes, gene functions, and products are assigned to specific sequence and pathways (KEGG, 2018). KEGG analysis can help to explain and connect the biological processes associated with certain genes and gene effects, including abnormal function due to disease or illness (KEGG, 2018). In addition, KEGG connects pathways of genes that work together in common processes (i.e. TGF-B signaling genes) that effect a coordinated network of changes within a cell that have been experimental documented.

Results from Previous Gene Expression Studies Concerning Fescue Toxicosis:

Research has been conducted concerning the effects of fescue toxicosis on whole gene expression and gene expression pathways (Li et al., 2017; Liao et al., 2015). A study conducted by Li et al. investigated the difference in gene expression profiles in the pituitaries of predominantly Angus steers. Steers grazed pastures with higher amounts of endophyte-infected (toxic) tall fescue (n = 10) and lower amounts of toxic tall fescue grass (n = 9). After a grazing period of 89 days on their respective pasture types, the steers were humanely slaughtered and

pituitary glands were removed and flash frozen (Li et al., 2017). Three pituitaries were excluded from the analysis due to tissue damage upon removal, resulting in 8 samples for each pasture type. Ribonucleic acid (RNA) was isolated from the pituitary tissue and microarray analysis of the RNA was conducted (Li et al., 2017). Differentially expressed genes (n = 542) were identified by micro-array analysis. Upon further cluster enrichment analysis, pathways related to prolactin and dopamine signaling were affected showing a clear differentiation between steers fed low amounts of toxic tall fescue and those fed higher amounts of toxic tall fescue (Li et al., 2017). The study conducted by Liao et al. resulted in a transcriptome analysis of differentially expressed genes involved in hepatic metabolic pathways in beef steers. Two groups of beef steers (also predominantly Angus) grazed pastures with higher amounts of toxic fescue (n = 10) and lower amounts of toxic fescue (n = 9) for a period of 89 days (Liao et al., 2015). Steers were humanely slaughtered, liver samples were retrieved; RNA was then isolated from the hepatic tissue and microarray analysis was conducted, resulting in the identification of 427 differentially expressed (DE) genes (Liao et al., 2015). These DE genes were analyzed through cluster enrichment and found to be involved in cell-mediated immune response pathways (Liao et al., 2015).

Materials and Methods

Population and Experimental Design

One hundred Hereford and Charolais cross-bred cows of varying age and parity were used for this study. Cows were housed at the University of Arkansas Livestock and Forestry Research Station in Batesville, Arkansas. They were allocated as evenly as possible by sire breed and parity to two pasture types: toxic endophyte-infected tall fescue (P1; n = 50) and non-

endophyte infected tall fescue (P2; n = 50). In both groups, animals grazed on two different pastures of the same treatment type, with 25 animals per pasture that were rotationally grazed every other week on pastures of the same forage type. Pastures were tested approximately once a month for ergovaline to verify toxicity at > 300 ppb (University of Missouri Veterinary Diagnostic Lab). One novel fescue pasture was identified as toxic, resulting in a total of three pasture types used within this experiment. The third pasture type was called "alternating" as it resulted in 25 cows grazing toxic fescue for two weeks and non-toxic fescue for two weeks. The groups by pasture type were then identified as: toxic (P1; n = 50), non-toxic (P2; n = 25), and alternating (P3; n = 25). Cows were on a yearly program that included vaccinations, dewormer, and topical parasite spray. All cows were given ample access to water and free-choice mineral mixture.

Sample Collection

Whole blood samples were collected via jugular venipuncture in Tempus™ tubes for RNA isolation in August 2016 according to manufacturer's instructions (Applied Biosystems®, Foster City, CA). Tempus™ tubes were immediately stored on ice after sampling for transport. Blood tubes were then refrigerated at 4℃ for approximately two days and then stored at -80℃ until used for RNA isolation (approximately 7 months).

RNA Isolation, Quality Control, and Sequencing

RNA from the Tempus™ tubes was isolated using the Tempus™ Spin RNA Isolation Kit according to manufacturers' instructions (Applied Biosystems®, Foster City, CA). Concentration and purity of RNA were analyzed using a Nanodrop™ One (Thermo Fisher Scientific, Inc. Waltham, MA); the goal concentration of samples was > 20 ng/uL and the goal purity

(A260/280) was 1.9 – 2.2. Integrity of RNA was determined using an Experion Automated Electrophoresis System (Bio-Rad Laboratories, Inc., Hercules, CA) to determine the ratio of bound, intact RNA to the 28 and 18S rRNA. Any samples with an RNA quality indicator (RQI) less than 7 were analyzed using an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). Samples with an RNA integrity number (RIN), or RQI, less than 7 were excluded. Selected RNA samples were shipped on dry ice to the Iowa State University DNA Facility (Ames, IA) for library construction and sequencing. A total of 90 samples were selected for whole transcriptome profiling to allow 30 samples to blocked on three sequencing lanes of an Illumina HiSeq 3000 (Illumina Inc., San Diego, CA) to allow for nearly equal representation of samples by pasture type and parity. Library construction was conducted using the QuantSeq 3' mRNA library prep kit (Lexicogen, Greenland, NH) to allow multiplexing of 30 samples on three different flow cell lanes. Samples were then sequenced to 150 bps by single-end sequencing to generate approximately 350 million reads per lane.

Data Analysis

Statistical analysis was conducted using the R software package QuasiSeq to conduct differential expression analysis and q-values to control the false discovery rate. A statistical model was fit to test the effect of pasture treatment, specifically the comparisons of the 3 pasture types (toxic, non-toxic, and toxic/non-toxic). In addition, the model accounted for differences in the two sire breeds (Hereford and Charolais), 3 levels of parity (indicated one month after the trial: 2, 3, and 4), and 2 pregnancy statuses (yes or no) as well as the effect of the sequencer lane (3 total). Following the statistical analysis to identify differentially expressed (DE) genes, a p-value histogram plot showed lower p-values than expected by random chance. The q-value package was used to calculate q-values for the resulting p-values. The q-value provides a risk

level for type 1 error (declaring genes DE when they are not) for a list of genes. Thus, a list of 100 genes with q< 0.05, by definition, is expected to have no more than 0.05*100 = 5 false discoveries (i.e. type 1 errors) and thus 95 genes are expected to be truly DE. The DE genes identified as significant were then entered into DAVID using the ENSEMBL gene IDs to identify statistically overrepresented (enriched) GO and KEGG pathways.

Results and Discussion

Results:

Data filtering and mapping statistics

Data was filtered after sequencing to remove samples based on data quality, including: poor RNA integrity/quality (RIN, n = 2); poor sequence quality (FASTQC statistics, n = 3); misidentification (unknown ID number, n = 4); low total number of read counts (n = 2). Thus, a total of 11 samples were excluded from analysis due to poor RNA quality, integrity, low read count numbers, or inconsistent sample identification. A total of 79 RNA samples were used for data analysis: P1 = 34 samples, P2 = 25 samples, and P3 = 20 samples. After data was mapped to the UMD3.1 Bos Taurus genome, a total of 743,040,014 reads were uniquely mapped to the genome at an average of 8,256,000 reads per sample (average of 71.8% uniquely mapped per sample). Genes were excluded from analysis based on the following criteria: 1) genes with no expression across all samples (0 read counts, n = 18,495); 2) genes with low gene expression (genes expressed in less than 20% of the samples and less than 2 read counts on average per sample). After data filtering was completed, 79 samples with 11,429 expressed genes remained with an average of 4,794,121 uniquely mapped, annotated genes, averaging 419 reads per gene.

Because Quant-Seq technology was used, this number is equivalent to fold coverage (419X) since one read is the same as transcript and only the 3' end of the gene is sequenced.

Differentially expressed genes as a function of fescue pasture type

A total of 527 differentially expressed (DE) genes ($q < 0.05$) were identified upon comparing the three pasture types used in this study. A total of 499 DE genes were identified as annotated, with 28 genes having no annotated biological function. Analysis of the functions of these genes are further summarized based on location, function and pathway in the ontology enrichment results below.

Gene Ontology (GO) and KEGG pathway enrichment analysis

All DE genes were analyzed for their involvement in specific gene expression pathways using gene set enrichment analysis using the DAVID software (Huang et al., 2009). The ENSEMBL gene identification numbers were used to identify gene function information and association with specific biological pathways (i.e. ontology information) within DAVID. Using Gene Ontology enrichment, the genes were summarized based on significant enrichment utilizing the following ontologies: biological process (BP: Table 1), cellular component (CC: Table 1) and molecular function (MF: Table 1). Only enrichment terms with a p-value < 0.10 corrected for multiple testing (Benjamini false discovery rate) were considered significant, including 1, 7, and 8 for BP, CC and MF respectively. The most significantly enriched GO terms were protein binding (BP; padj < 0.02), cytosol (CC; padj < 0.0053), and poly A RNA binding (MF; padj < $5.8E^{-10}$). The most significant KEGG pathway enrichment result was the Estrogen signaling pathway (padj < $8.3E^{-5}$). All significant KEGG pathway enrichment results (13 total)

are presented in Table 2. In addition, a total of 15 gene clusters were identified as significant with an enrichment score > 1.3 (equivalent to $p < 0.05$; Table 3).

Discussion:

Differentially expressed genes include those previously identified with toxic fescue exposure

Many of the differentially expressed genes discovered in this experiment have been linked, by previous research, to exposure to toxic fescue including heat shock protein genes (HSP70L and HSP90) (Bastin et al., 2014; Campbell et al., 2014), genes involved in prolactin and dopamine signaling pathways (Li et al., 2017), and genes involved in cell-mediated immune response pathways (Liao et al., 2015).

Relationship of enriched pathways with fescue toxicosis

Several other differentially expressed genes identified in this experiment were related to GO and KEGG pathways that have been proven to be affected by the presence of toxic fescue. While these genes may not have been targeted specifically by previous research, their presence in the enriched GO pathways of interest is significant because it indicates the far-reaching effects of fescue toxicosis in bodily function. The GO pathways affected by toxic fescue included, among others, cytosol (n = 54), nucleoplasm (n = 64), and nucleolus (n = 33), protein complex genes (n = 17), poly A RNA binding (n = 64), ATP binding (n = 67), and heat shock protein binding (n = 6), which is associated with heat stress in cattle (**Table 1**).

KEGG pathways affected by the presence of toxic fescue included genes related to hormone production such as estrogen (n = 15) and oxytocin (n = 13); protein processing in endoplasmic reticulum (n = 16); and immune response pathways, such as platelet activation and T cell response (n = 12 and n = 10 respectively). Estrogen and oxytocin are involved in weight

gain and loss in cattle, as well as fertility in beef cows. Several genes were found to be involved in pathways also indicated to affect reproductive health and function including oocyte meiosis (n = 10) and endometrial cancer (n = 7). It is also important to note that several genes identified are related to pathways involving nerve function, such as Huntington's disease (n = 15) and Alzheimer's disease (n = 16) (**Table 2**).

These genes and pathways are of particular interest to researchers of fescue toxicosis because of the hormone signaling that may be impacted by toxic fescue, especially reproductive hormones (estrogen and GnRH) and hormones related to vasodilation and vasoconstriction (oxytocin) as both are major factors affecting the health and productivity of beef cows.

Clusters of gene ontology terms enriched for genes DE in response to fescue toxin level have functions related to diverse processes

Clusters of various ontology terms have a host of pathways and biological processes consistent with previous knowledge of fescue toxicosis (**see Table 3**). Many of these gene ontologies can be formed into clusters that indicate terms for similar or related functions and processes. Cluster 1 includes genes that regulate chaperones, protein folding, and unfolded protein response. The primary function of protein chaperones is to assist in the folding and unfolding of proteins—many chaperones are heat shock proteins. Likewise, Clusters 4 (HSP20-like chaperone), 6 (HSP 70 family), and 10 (HSP 90, reserved site) include many of the heat shock proteins, which aggregate as proteins are denatured by abnormally high body temperatures often associated with fescue toxicosis. Genes found in Cluster 9, which involve toxin transport, chaperonins, TCP-1, and GroEL, are closely associated with chaperones and heat shock proteins. The aggregation of heat shock proteins affects their ability to assist in protein preparation for bodily processes. This is especially important to the biology of fescue toxicosis as heat stress

greatly impacts protein folding. Cluster 5 genes are related to mitochondrial function and oxidative phosphorylation, as well as nerve function and cardiac muscle contraction, and Cluster 15 genes deal mainly with nerve signaling and vasoconstriction: serotonergic synapse, glutamatergic synapse, dopaminergic synapse, long-term depression, cholinergic synapse, gap junction, long-term potentiation. As a group, these clusters reflect the fact that the presence of toxic fescue affects the expression of genes and pathways related to hormone, protein, and nerve function.

Relationship of DE genes and pathways with those identified in previous gene expression studies

Many of the genes or pathways identified here have been identified in previous gene expression or genetic studies of cattle and toxic fescue. These studies have included research into specific hormone signaling pathways and hormone levels in blood serum as an indicator of the effect of toxic fescue on hormone regulation, especially GnRH and growth hormone (Li et al., 2017). Research conducted by Li et al. concluded that the pituitary transcriptome profiles, determined by targeted RT-PCR analysis, were altered by the consumption of endophyte-infected tall fescue. Many of these same genes were identified in this study which corroborate the conclusion that gene pathway expression is altered by the presence of tall fescue: GnRH gene expression pathways and changes in estrogen and prolactin gene expression pathways were also found through KEGG enrichment analysis. COX genes were also found to play some role in the regulation of altered bodily function in cows affected by toxic fescue (Li et al., 2017). COX genes regulate homeostasis within the body and have some function in immune response.

Conclusions

While a fairly low percentage of the DE genes in this experiment have been linked by previous research to toxic fescue, most of the DE genes that have not been previously implicated in playing a role in the pathways linked to fescue toxicosis. Genes identified in this study who were previously identified in SNP association studies include HSP90 and HSP70L. In addition, several pathways were identified that have been observed in previous studies, such as GnRH signaling. Significant pathways of interest from GO enrichment results include heat shock proteins and poly A RNA binding, especially as they are related to protein function. KEGG pathways discovered that have the potential to be of significant value to research on the effects of fescue toxicosis include estrogen, oxytocin, and T cell response pathways. Cluster enrichment results likewise implicate heat shock protein and nerve signaling pathway as being involved in the response to fescue toxicosis. The DE genes and expression pathways that have not yet been discussed may help to further our understanding of the impact of fescue toxicosis, such as those involved in immune responses. New findings that are consistent with the biology of fescue toxicosis include genes that regulate fatty liver syndrome (altered fat metabolism), T cell response (weight loss, stress), thermoregulation (heat stress, stress, vasoconstriction) and hormone regulation (weight loss, reduced reproductive performance, altered fat metabolism, etc.).While study of the expression of genes and gene pathways may tell us which biological processes are affected, it does not explain causatively how or why they are affected. Replication of these findings in an independent study would merit further investigation. The next step to further understand and mitigate the effects of the consumption of endophyte-infected tall fescue in beef cows would be to identify a gene expression pathway to disrupt or promote by means of pharmaceutical or gene editing procedures within cells to determine the cellular function of these

genes and pathways with or without ergovaline or ergovaline analogues. For example, after the identification of differentially expressed genes found in the T cell response pathway, attempting to confirm immune-response related cell signaling and function could determine if T cell disfunction plays a role in any fescue toxicosis related phenotypes.

Table 1. Gene Ontology enrichment results

Category	Term	Count	%	PValue	List Total	Pop Hits	Pop Total	Fold Enrichment	Benjamini
Biological_Process(BP)									
GOTERM_BP_DIRECT	GO:0006457~protein folding	14	2.81	1.33E-05	403	107	13824	4.49	0.0198
Cellular_Component(CC)									
GOTERM_CC_DIRECT	GO:0005829~cytosol	54	10.82	1.41E-05	438	1025	15461	1.86	0.0053
GOTERM_CC_DIRECT	GO:0005654~nucleoplasm	64	12.83	2.00E-05	438	1310	15461	1.72	0.0038
GOTERM_CC_DIRECT	GO:0043234~protein complex	17	3.41	5.67E-05	438	181	15461	3.32	0.0071
GOTERM_CC_DIRECT	GO:0005730~nucleolus	33	6.61	5.17E-04	438	605	15461	1.93	0.0478
GOTERM_CC_DIRECT	GO:0002199~zona pellucida receptor complex	4	0.80	0.0011	438	8	15461	17.65	0.0821
GOTERM_CC_DIRECT	GO:0005925~focal adhesion	20	4.01	0.0018	438	318	15461	2.22	0.1082
GOTERM_CC_DIRECT	GO:0070062~extracellular exosome	84	16.83	0.0019	438	2162	15461	1.37	0.0978
Molecular_Function(MF)									
GOTERM_MF_DIRECT	GO:0044822~poly(A) RNA binding	64	12.83	1.35E-12	377	840	13095	2.65	5.85E-10
GOTERM_MF_DIRECT	GO:0005524~ATP binding	67	13.43	3.36E-07	377	1224	13095	1.90	7.30E-05
GOTERM_MF_DIRECT	GO:0051082~unfolded protein binding	11	2.20	1.25E-06	377	50	13095	7.64	1.80E-04
GOTERM_MF_DIRECT	GO:0000166~nucleotide binding	24	4.81	6.82E-06	377	282	13095	2.96	7.39E-04
GOTERM_MF_DIRECT	GO:0031072~heat shock protein binding	5	1.00	4.27E-05	377	8	13095	21.71	0.0037
GOTERM_MF_DIRECT	GO:0051087~chaperone binding	6	1.20	6.28E-04	377	25	13095	8.34	0.0444
GOTERM_MF_DIRECT	GO:0004129~cytochrome-c oxidase activity	6	1.20	0.0011	377	28	13095	7.44	0.0650
GOTERM_MF_DIRECT	GO:0004674~protein serine/threonine kinase activity	16	3.21	0.0012	377	213	13095	2.61	0.0654

Table 2. KEGG pathway enrichment results

Category	Term	Count	%	PValue	List Total	Pop Hits	Pop Total	Fold Enrichment	Benjamini
KEGG_PATHWAY	bta04915:Estrogen signaling pathway	15	3.01	3.77E-07	209	98	7583	5.55	8.29E-05
KEGG_PATHWAY	bta04141:Protein processing in endoplasmic reticulum	16	3.21	5.74E-05	209	169	7583	3.44	0.0063
KEGG_PATHWAY	bta05010:Alzheimer's disease	16	3.21	1.04E-04	209	178	7583	3.26	0.0076
KEGG_PATHWAY	bta04932:Non-alcoholic fatty liver disease (NAFLD)	14	2.81	4.31E-04	209	160	7583	3.17	0.0234
KEGG_PATHWAY	bta04912:GnRH signaling pathway	10	2.00	5.46E-04	209	86	7583	4.22	0.0237
KEGG_PATHWAY	bta04022:cGMP-PKG signaling pathway	14	2.81	6.49E-04	209	167	7583	3.04	0.0235
KEGG_PATHWAY	bta04611:Platelet activation	12	2.40	6.97E-04	209	127	7583	3.43	0.0217
KEGG_PATHWAY	bta04921:Oxytocin signaling pathway	13	2.61	9.77E-04	209	153	7583	3.08	0.0265
KEGG_PATHWAY	bta05016:Huntington's disease	15	3.01	0.0011	209	199	7583	2.73	0.0266
KEGG_PATHWAY	bta04810:Regulation of actin cytoskeleton	15	3.01	0.0021	209	213	7583	2.56	0.0453
KEGG_PATHWAY	bta05213:Endometrial cancer	7	1.40	0.0025	209	51	7583	4.98	0.0498
KEGG_PATHWAY	bta04660:T cell receptor signaling pathway	10	2.00	0.0028	209	108	7583	3.36	0.0495
KEGG_PATHWAY	bta04114:Oocyte meiosis	10	2.00	0.0035	209	112	7583	3.24	0.0582

Table 3. Cluster Annotation of GO and KEGG pathway analysis

Annotation Cluster 1	**Enrichment Score: 6.622999118572811**
Category	**Term**
UP_KEYWORDS	Chaperone
GOTERM_MF_DIRECT	GO:0051082~unfolded protein binding
GOTERM_BP_DIRECT	GO:0006457~protein folding
Annotation Cluster 2	**Enrichment Score: 4.363773614935281**
Category	**Term**
GOTERM_MF_DIRECT	GO:0000166~nucleotide binding
INTERPRO	IPR012677:Nucleotide-binding, alpha-beta plait
INTERPRO	IPR000504:RNA recognition motif domain
SMART	SM00360:RRM
Annotation Cluster 3	**Enrichment Score: 4.351574255130137**
Category	**Term**
GOTERM_MF_DIRECT	GO:0005524~ATP binding
UP_KEYWORDS	ATP-binding
UP_KEYWORDS	Nucleotide-binding
Annotation Cluster 4	**Enrichment Score: 2.6826417808753886**
Category	**Term**
INTERPRO	IPR008978:HSP20-like chaperone
UP_SEQ_FEATURE	domain:CS
INTERPRO	IPR007052:CS-like domain
Annotation Cluster 5	**Enrichment Score: 2.1240359265670072**
Category	**Term**
KEGG_PATHWAY	bta05010:Alzheimer's disease
KEGG_PATHWAY	bta04932:Non-alcoholic fatty liver disease (NAFLD)
UP_KEYWORDS	Mitochondrion inner membrane
GOTERM_MF_DIRECT	GO:0004129~cytochrome-c oxidase activity
KEGG_PATHWAY	bta05016:Huntington's disease
INTERPRO	IPR003177:Cytochrome c oxidase, subunit VIIa
GOTERM_BP_DIRECT	GO:1902600~hydrogen ion transmembrane transport
KEGG_PATHWAY	bta05012:Parkinson's disease
GOTERM_CC_DIRECT	GO:0005746~mitochondrial respiratory chain
KEGG_PATHWAY	bta00190:Oxidative phosphorylation
GOTERM_CC_DIRECT	GO:0005751~mitochondrial respiratory chain complex IV
KEGG_PATHWAY	bta04260:Cardiac muscle contraction
GOTERM_CC_DIRECT	GO:0045277~respiratory chain complex IV
UP_KEYWORDS	Transit peptide
UP_SEQ_FEATURE	transit peptide:Mitochondrion
Annotation Cluster 6	**Enrichment Score: 2.0574508398446696**
Category	**Term**
INTERPRO	IPR013126:Heat shock protein 70 family
INTERPRO	IPR018181:Heat shock protein 70, conserved site
KEGG_PATHWAY	bta04612:Antigen processing and presentation

Table 3. Continued

Annotation Cluster 7	**Enrichment Score: 2.0235801357783747**
Category	**Term**
GOTERM_MF_DIRECT	GO:0004674~protein serine/threonine kinase activity
INTERPRO	IPR011009:Protein kinase-like domain
UP_KEYWORDS	Serine/threonine-protein kinase
INTERPRO	IPR008271:Serine/threonine-protein kinase, active site
INTERPRO	IPR017441:Protein kinase, ATP binding site
INTERPRO	IPR000719:Protein kinase, catalytic domain
SMART	SM00220:S_TKc
UP_KEYWORDS	Kinase
GOTERM_BP_DIRECT	GO:0018105~peptidyl-serine phosphorylation
UP_KEYWORDS	Transferase
Annotation Cluster 8	**Enrichment Score: 1.9965939090299136**
Category	**Term**
KEGG_PATHWAY	bta04611:Platelet activation
KEGG_PATHWAY	bta04810:Regulation of actin cytoskeleton
KEGG_PATHWAY	bta05205:Proteoglycans in cancer
KEGG_PATHWAY	bta04510:Focal adhesion
KEGG_PATHWAY	bta04670:Leukocyte transendothelial migration
Annotation Cluster 9	**Enrichment Score: 1.983885393279854**
Category	**Term**
GOTERM_BP_DIRECT	GO:0007339~binding of sperm to zona pellucida
GOTERM_CC_DIRECT	GO:0002199~zona pellucida receptor complex
GOTERM_BP_DIRECT	GO:0032212~positive regulation of telomere maintenance via telomerase
GOTERM_BP_DIRECT	GO:1901998~toxin transport
INTERPRO	IPR027409:GroEL-like apical domain
INTERPRO	IPR027413:GroEL-like equatorial domain
INTERPRO	IPR002423:Chaperonin Cpn60/TCP-1
GOTERM_BP_DIRECT	GO:1904871~positive regulation of protein localization to Cajal body
GOTERM_CC_DIRECT	GO:0044297~cell body
INTERPRO	IPR002194:Chaperonin TCP-1, conserved site
GOTERM_CC_DIRECT	GO:0005832~chaperonin-containing T-complex
INTERPRO	IPR027410:TCP-1-like chaperonin intermediate domain
INTERPRO	IPR017998:Chaperone tailless complex polypeptide 1 (TCP-1)
GOTERM_BP_DIRECT	GO:1904874~positive regulation of telomerase RNA localization to Cajal body
GOTERM_CC_DIRECT	GO:0005874~microtubule
GOTERM_BP_DIRECT	GO:0050821~protein stabilization
Annotation Cluster 10	**Enrichment Score: 1.8932545334239885**
Category	**Term**
INTERPRO	IPR019805:Heat shock protein Hsp90, conserved site
PIR_SUPERFAMILY	PIRSF002583:heat shock protein, HSP90/HTPG types
INTERPRO	IPR001404:Heat shock protein Hsp90
INTERPRO	IPR020575:Heat shock protein Hsp90, N-terminal
INTERPRO	IPR003594:Histidine kinase-like ATPase, ATP-binding domain
SMART	SM00387:HATPase_c
INTERPRO	IPR020568:Ribosomal protein S5 domain 2-type fold
Annotation Cluster 11	**Enrichment Score: 1.8000214708066433**
Category	**Term**
GOTERM_BP_DIRECT	GO:0032212~positive regulation of telomere maintenance via telomerase
GOTERM_BP_DIRECT	GO:0051973~positive regulation of telomerase activity
GOTERM_BP_DIRECT	GO:1904355~positive regulation of telomere capping

Table 3. Continued

Annotation Cluster 12	**Enrichment Score: 1.6045807598834931**
Category	**Term**
GOTERM_CC_DIRECT	GO:0022625~cytosolic large ribosomal subunit
UP_KEYWORDS	Ribonucleoprotein
KEGG_PATHWAY	bta03010:Ribosome
UP_KEYWORDS	Ribosomal protein
GOTERM_BP_DIRECT	GO:0006412~translation
GOTERM_MF_DIRECT	GO:0003735~structural constituent of ribosome
Annotation Cluster 13	**Enrichment Score: 1.5176768694085245**
Category	**Term**
INTERPRO	IPR000253:Forkhead-associated (FHA) domain
INTERPRO	IPR008984:SMAD/FHA domain
SMART	SM00240:FHA
Annotation Cluster 14	**Enrichment Score: 1.5151341003401342**
Category	**Term**
INTERPRO	IPR001440:Tetratricopeptide TPR-1
UP_KEYWORDS	TPR repeat
SMART	SM00028:TPR
INTERPRO	IPR013026:Tetratricopeptide repeat-containing domain
INTERPRO	IPR019734:Tetratricopeptide repeat
UP_SEQ_FEATURE	repeat:TPR 3
INTERPRO	IPR011990:Tetratricopeptide-like helical
UP_SEQ_FEATURE	repeat:TPR 2
UP_SEQ_FEATURE	repeat:TPR 1
Annotation Cluster 15	**Enrichment Score: 1.3143883936088534**
Category	**Term**
KEGG_PATHWAY	bta04912:GnRH signaling pathway
KEGG_PATHWAY	bta04022:cGMP-PKG signaling pathway
KEGG_PATHWAY	bta04611:Platelet activation
KEGG_PATHWAY	bta04921:Oxytocin signaling pathway
KEGG_PATHWAY	bta04114:Oocyte meiosis
KEGG_PATHWAY	bta04720:Long-term potentiation
KEGG_PATHWAY	bta04270:Vascular smooth muscle contraction
KEGG_PATHWAY	bta05205:Proteoglycans in cancer
KEGG_PATHWAY	bta04750:Inflammatory mediator regulation of TRP channels
KEGG_PATHWAY	bta04725:Cholinergic synapse
KEGG_PATHWAY	bta04540:Gap junction
KEGG_PATHWAY	bta04922:Glucagon signaling pathway
KEGG_PATHWAY	bta04070:Phosphatidylinositol signaling system
KEGG_PATHWAY	bta04925:Aldosterone synthesis and secretion
KEGG_PATHWAY	bta04062:Chemokine signaling pathway
KEGG_PATHWAY	bta04730:Long-term depression
KEGG_PATHWAY	bta04924:Renin secretion
KEGG_PATHWAY	bta04971:Gastric acid secretion
KEGG_PATHWAY	bta04728:Dopaminergic synapse
GOTERM_BP_DIRECT	GO:0051209~release of sequestered calcium ion into cytosol
KEGG_PATHWAY	bta04724:Glutamatergic synapse
KEGG_PATHWAY	bta04020:Calcium signaling pathway
KEGG_PATHWAY	bta04972:Pancreatic secretion
KEGG_PATHWAY	bta04918:Thyroid hormone synthesis
KEGG_PATHWAY	bta04723:Retrograde endocannabinoid signaling
KEGG_PATHWAY	bta04261:Adrenergic signaling in cardiomyocytes
KEGG_PATHWAY	bta04726:Serotonergic synapse
KEGG_PATHWAY	bta04970:Salivary secretion
KEGG_PATHWAY	bta04713:Circadian entrainment
KEGG_PATHWAY	bta04916:Melanogenesis

Acknowledgments

The author would like to acknowledge the members of the honors thesis committee for their help and support, including Drs. Charles Rosenkrans, Ph.D.; Jeremy G. Powell, Ph.D., D.V.M.; Rick Rorie, Ph.D. of the University of Arkansas and James E. Koltes, Ph.D. (Iowa State University), who designed/ oversaw the experiment and provided mentorship for this project. Special thanks also to Steven C. Ricke, Ph.D. for his support and understanding, and Laura Meyer, M.S., and Sarah Chewning for their support, advice, and assistance.

References

USDA (2012). 2012 Census of Agriculture. *USDA Department of Agriculture*, 2018. https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_1_State_Level/Arkansas/st05_1_012_013.pdf

NCBA (2018). Arkansas Pasture and Forages. *National Cattlemen's Beef Association*, 2018. https://www.uaex.edu/farm-ranch/animals-forages/pastures/ . Web 2018.

Leuchtmann A, Bacon CW, Schardl CL, White JF Jr, and Tadych M (2017) Nomenclatural realignment of Neotyphodium species with genus Epichloë, *Mycologia*, 106:2, 202-215, DOI: 10.3852/13-251

Clancy, S. (2008) RNA transcription by RNA polymerase: prokaryotes vs eukaryotes. *Nature Education* 1(1):125. https://www.nature.com/scitable/topicpage/rna-transcription-by-rna polymerase-prokaryotes-vs-961

Li Q, Hegge R, Bridges PJ, and Matthews JC (2017). Pituitary genomic expression profiles of steers are altered by grazing of high vs. low endophyte-infected tall fescue forages. *PLoS ONE, 12*(9), e0184612. http://doi.org/10.1371/journal.pone.0184612

NIH. "How Genes Work". *National Institute of Health*, 2018. https://ghr.nlm.nih.gov/primer/howgeneswork/protein

Illumina. "RNA Sequencing". *Illumina Sequencing, Inc*. 2018. https://www.illumina.com/techniques/sequencing/rna-sequencing.html

Pratt et al. (2015). Effects of grazing tall fescue containing ergot alkaloids on bull sperm cryopreservation. Journal of the American College of Radiology, 2017(6), 181:24-29. DOI: https://doi.org/10.1016/j.anireprosci.2017.03.011

Kurkurba KR, & Montgomery SB (2015). RNA Sequencing and Analysis. Cold Spring Harbor Protocols, 2015(11), 951–969. http://doi.org/10.1101/pdb.top084970

NHGRI (2018). Biological Pathways. *National Human Genome Research Institute* 2018. https://www.genome.gov/27530687/biological-pathways-fact-sheet/

Moisa SJ, Shike DW, Shoup L, and Loor JJ (2015). "Maternal Plane of Nutrition During Late-Gestation and Weaning Age Alter Steer Calf Longissimus Muscle Adipogenic MircoRNA and Target Gene Expression". *AOCS* 2015. Web 2017.

Cock PJA, Fields CJ, Goto N, Heuer ML, and Rice PM (2009) The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research* 38: 1767 1771.

Liao SF, Boling JA, and Matthews JC (2015). Gene expression profiling indicates an increased capacity for proline, serine, and ATP synthesis and mitochondrial mass by the liver of steers grazing high vs. low endophyte-infected tall fescue. *Journal of Animal Science*, Volume 93, Issue 12, 1 December 2015, Pages 5659–5671, https://doi.org/10.2527/jas.2015-9193

Science (2014). The Evolution of Sequencing Technology. *Science*, 2014. Accessed 1 March 2018. http://www.sciencemag.org/posters/evolution-sequencing-technology

Li H, et al. and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009 Aug 15; 25(16): 2078–2079. Published online 2009 Jun 8. DOI: 10.1093/bioinformatics/btp352 PMCID: PMC2723002 PMID: 19505943

Andrews S. FASTQC software. 2010. Downloaded from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Joshi NA, Fass JN. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at https://github.com/najoshi/sickle

Stowe HM, Calcatera SM, Dimmick MA, Andrae JG, Duckett SK, et al. (2014) The Bull Sperm MicroRNAome and the Effect of Fescue Toxicosis on Sperm MicroRNA Expression. *PLoS ONE* 9(12), 2014. doi:10.1371/journal.pone.0113163. Print 2017.

Coufal-Majewski S, Stanford K, McAllister T, Blakely B, McKinnon J, et al. (2016) "Impacts of Cereal Ergot in Food Animal Production". *Front Vet Sci* 3(15), 2016. DOI:10.3389/fvets.2016.00015. Web 2017.

Roberts, Craig (2000). Tall Fescue Toxicosis. *University of Missouri Extension*, April 2000. Web 2017.

Campbell BT, Kojima CJ, Cooper TA, Bastin BC, Wojakiewicz L, Kallenbach RL, Schrick FN, and Waller JC (2013) A Single Nucleotide Polymorphism in the Dopamine Receptor D2 Gene May Be Informative for Resistance to Fescue Toxicosis in Angus-Based Cattle. *Animal Biotechnology*, 25:1, 1-12, DOI: 10.1080/10495398.2013.796960

Strickland JR, Looper ML, Matthews JC, Rosenkrans Jr. CF, Flythe MD, and Brown KR (2011). Board Invited Review: St. Anthony's Fire in livestock: Causes, mechanisms, and potential solutions. *J. Anim. Sci* 89(1603-1626), 2011. DOI:10.2527/jas.2010-3478. Print 2017.

NBCA. "Beef Industry Statistics". *National Cattlemen's Beef Association*, 2017. http://beefusa.org/beefindustrystatistics.aspx Web 2018.

Aleksey V Zimin, et al. (2009). A whole-genome assembly of the domestic cow, Bos Taurus. *Genome Biology* 2009 10:R42 https://doi.org/10.1186/gb-2009-10-4-r42

Cannell IG, Kong YW, and Bushell M (2008). How Do MicroRNAs Regulate Gene Expression?. *Biochem Soc. Trans*. 2008 Dec: 36(1224-31). DOI:10.1042/BST0361224. Web 2017

Strickland JR, Spiers DE, Fletcher LR, and Oliver JW (2009). Physiological Basis of Fescue Toxicosis. *American Society of Agronomy* 53, 203-227.

Allen VG & Segarra E (2001). Anti-Quality Components in Forage: Overview, Significance, and Economic Impact. *J Range Manag* 54: 409-412.

Zerbino DR, et al. (2018). ENSEMBL 2018. PubMed PMID: 29155950. DOI:10.1093/nar/gkx1098

Huang da W., Sherman BT, and Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 4(1):44-57. DOI: 10.1038/nprot.2008.211.

Rattray RM, Perumbakkam S, Smith F, Craig AM (2010). Microbiomic Comparison of the Intestine of the Earthworm *Eisenia fetida* Fed Ergovaline. *Current Microbiology*, 60(3), 229-35. http://0-dx.doi.org.library.uark.edu/10.1007/s00284-009-9530-8

Thermo Fisher Scientific (2018). The Basics: RNA Isolation. *Thermo Fisher Scientific*, 2018. https://www.thermofisher.com/us/en/home/references/ambion-tech-support/rna-isolation/general articles/the-basics-rna-isolation.html

Anders S, Pyl PT, and Huber W (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. http://doi.org/10.1093/bioinformatics/btu638