

RESEARCH ARTICLE

Open Access



# Differential network analysis of bovine muscle reveals changes in gene coexpression patterns in response to changes in maternal nutrition

Lihe Liu<sup>1,2</sup>, Rocío Amorín<sup>2</sup>, Philippe Moriel<sup>3</sup>, Nicolás DiLorenzo<sup>4</sup>, Phillip A. Lancaster<sup>5</sup> and Francisco Peñagaricano<sup>1,2\*</sup> 

## Abstract

**Background:** Coexpression network analysis is a powerful tool to reveal transcriptional regulatory mechanisms, identify transcription factors, and discover gene functions. It can also be used to investigate changes in coexpression patterns in response to environmental insults or changes in experimental conditions. Maternal nutrition is considered a major intrauterine regulator of fetal developmental programming. The objective of this study was to investigate structural changes in gene coexpression networks in the muscle of bull beef calves gestated under diets with or without methionine supplementation. Both muscle transcriptome and methylome were evaluated using next generation sequencing.

**Results:** Maternal methionine supplementation significantly perturbed coexpression patterns in the offspring's muscle. Indeed, we found that neither the connection strength nor the connectivity pattern of six modules (subnetworks) detected in the control diet were preserved in the methionine-rich diet. Functional characterization revealed that some of the unpreserved modules are implicated in myogenesis, adipogenesis, fibrogenesis, canonical Wnt/ $\beta$ -catenin pathway, ribosome structure, rRNA binding and processing, mitochondrial activities, ATP synthesis and NAD(P) H oxidoreductases, among other functions. The bisulfite sequencing analysis showed that nearly 2% of all evaluated cytosines were differentially methylated between maternal diets. Interestingly, there were significant differences in the levels of gene body DNA methylation between preserved and unpreserved modules.

(Continued on next page)

\* Correspondence: [fpenagarican@wisc.edu](mailto:fpenagarican@wisc.edu)

<sup>1</sup>Department of Animal and Dairy Sciences, University of Wisconsin-Madison, 1675 Observatory Drive, Madison, WI 53706, USA

<sup>2</sup>Department of Animal Sciences, University of Florida, Gainesville, FL 32611, USA

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



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Conclusions:** Overall, our findings provide evidence that maternal nutrition can significantly alter gene coexpression patterns in the offspring, and some of these perturbations are mediated by changes in DNA methylation.

**Keywords:** Correlation network, Fetal programming, Maternal diet, Muscle transcriptome

## Background

Transcriptome analysis is an essential tool to uncover the molecular basis of phenotypic variation. The advent of RNA sequencing has dramatically improved the characterization and quantification of transcriptomes [1]. The most common use of RNA sequencing is the identification of differentially expressed genes, that is, genes that show differences in expression between conditions. However, genes and gene products do not usually work in isolation, but they are connected in complex networks. There is increasing interest in moving beyond differential expression and examine transcriptional profiles in the context of molecular networks [2]. It is well-accepted that genes that are controlled by the same set of transcription factors or are involved in the same biological processes tend to have similar expression profiles [3]. This principle is known as guilt-by-association and represents the basis for the reconstruction of gene networks using RNA sequencing data. These networks, commonly called gene coexpression networks, are undirected graphs where nodes correspond to genes and edges represent pairwise expression similarities.

Gene coexpression networks can be used for different purposes. One popular application consists in the characterization of the topology of the reconstructed network and examination of interesting nodes and coexpression structures. This single network analysis focuses on the mechanisms allowing the identification of transcription factors (hub genes), the functional annotation of unknown genes, i.e., the association of genes of unknown function with well-described biological processes, and the detection of transcriptional regulatory programs [4]. Another application consists of evaluating gene coexpression networks but across conditions. Here, the term condition is very broad and can refer to different tissues, different developmental stages, or even different treatments. This application, commonly known as differential network analysis, focuses on determining changes in the topology of the networks across conditions. For instance, it is possible to examine whether connections or subnetworks defined under normal conditions (control group) are reproducible and preserved in the testing group (treatment group) [5]. In this scenario, differences in the topology of these two networks would indicate that coexpression patterns were significantly perturbed by the treatment. Note that

expression similarities (coexpression) hint common regulatory mechanisms (coregulation), and hence, changes in the network might indicate that the treatment has disrupted coregulation mechanisms, functional links and biological processes. Undoubtedly, this shift in focus from differentially expressed genes to differentially connected genes provides more holistic insights about gene regulation.

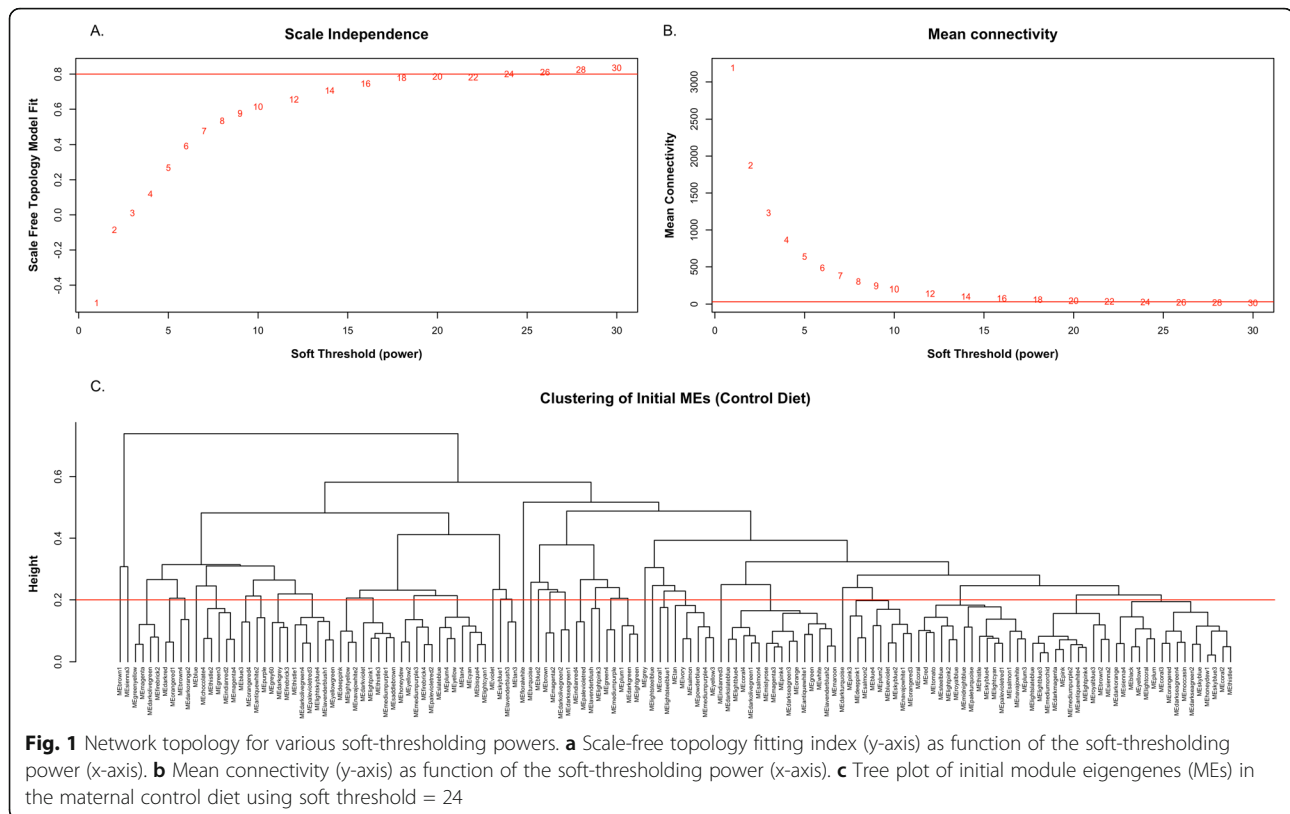
It is well-documented that different intrauterine insults can induce permanent changes to the structure, physiology, and metabolism of the offspring. This phenomenon has been termed fetal programming and may have lasting or lifelong consequences [6]. Maternal nutrition is considered a major intrauterine environmental factor and it is now known that maternal nutritional status during pregnancy can induce remarkable effects on fetal development [7]. There is growing evidence that maternal nutrition can alter epigenetic marks of the fetal genome, such as DNA methylation [8]. Indeed, this link between maternal nutrition and subsequent modification of fetal epigenome, including changes in gene expression, is one of the molecular mechanisms proposed to explain the phenomenon of fetal programming [9].

The main objective of this study was to assess whether maternal nutrition in beef cattle can disrupt gene coexpression patterns in the offspring. Maternal nutritional treatments consisted of control or methionine-rich diets offered during periconceptional and early gestation periods. Both muscle transcriptome and methylome of bull beef calves were evaluated using next generation sequencing. Note that DNA methylation depends on the availability of methyl donors, such as methionine, and hence, we hypothesized that maternal methionine supplementation could alter the fetal epigenome, which in turn could induce significant changes in the topology of gene networks.

## Results

### RNA-sequencing analysis

The RNA-sequencing of the muscle transcriptome yielded about 50 million paired-end reads per sample. Roughly 87% of the reads were mapped to the ARS-UCD1.2 bovine genome assembly using the software Hisat2 (see Additional File 1). After removing highly abundant genes (such as myosins, tropomyosins, and troponins) and lowly expressed genes (genes with 5 or



less read counts in at least 9 biological replicates), a total of 12,786 genes were retained for the network analysis.

### Network construction and module identification

We first characterized the muscle transcriptome under normal conditions. As such, the inference of the gene coexpression network and subsequent module identification was performed using only samples derived from the maternal control diet. A total of 7034 genes with high across-sample expression variance were included in this network analysis. A soft-threshold equal to 24 was derived from the high scale-free fitting index ( $R^2 \geq 0.8$ ; Fig. 1a), resulting in a mean connectivity equal to 31.45 (Fig. 1b). A total of 147 preliminary modules were detected using a dynamic tree cut process, and after merging highly correlated modules (Pearson's correlation  $\geq 0.8$ ; Fig. 1c), a total of 14 modules, including the background set (grey module) were retained for subsequent analysis (Fig. 2a).

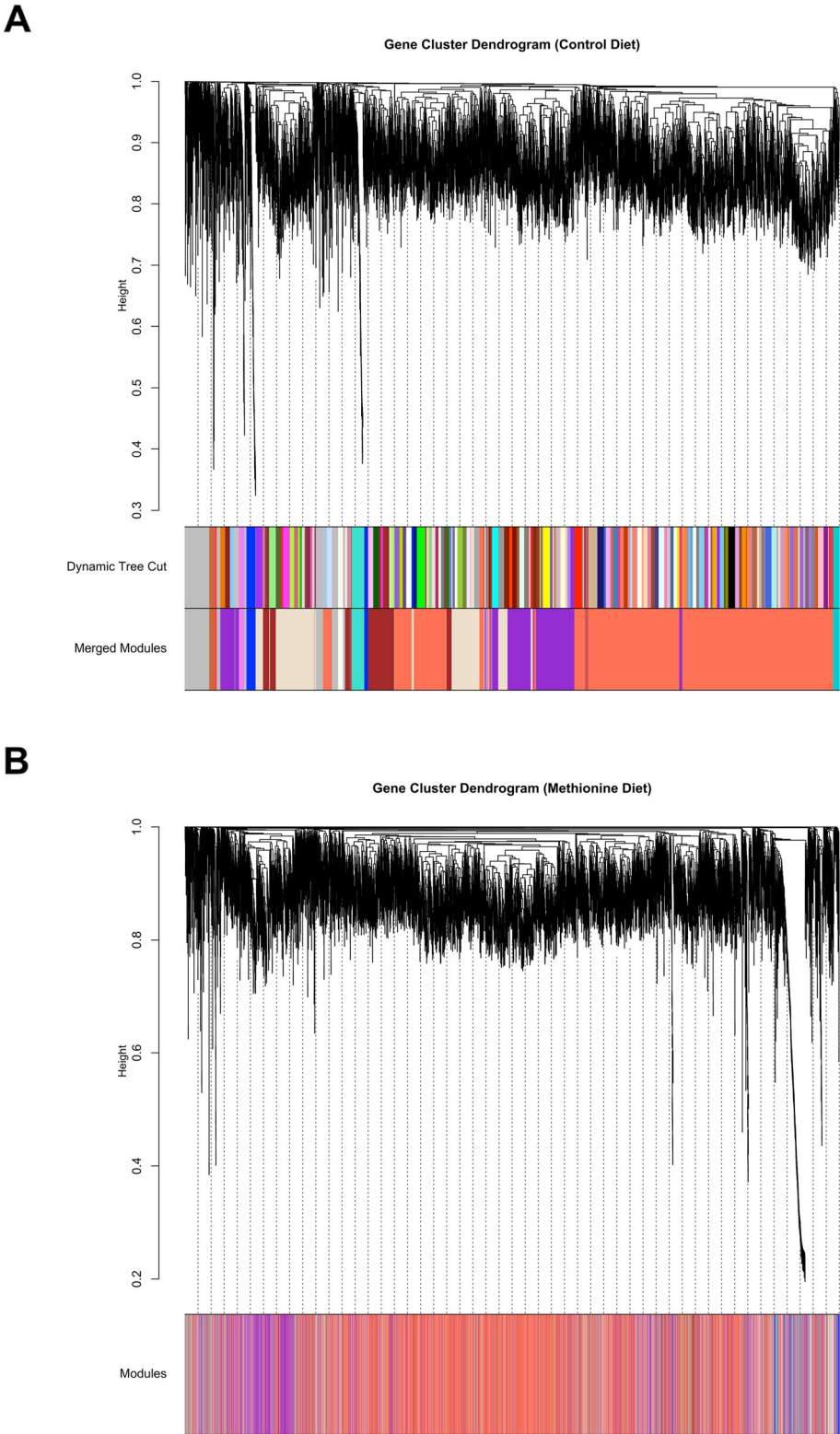
### Module preservation

After we characterized the muscle transcriptome in the control diet, we investigated the impact of the nutritional treatment (maternal methionine supplementation) on gene coexpression patterns. As such, we compared the structure of coexpression networks between control and methionine diets to identify changes in the topology (Fig. 2b). These changes were evaluated using a

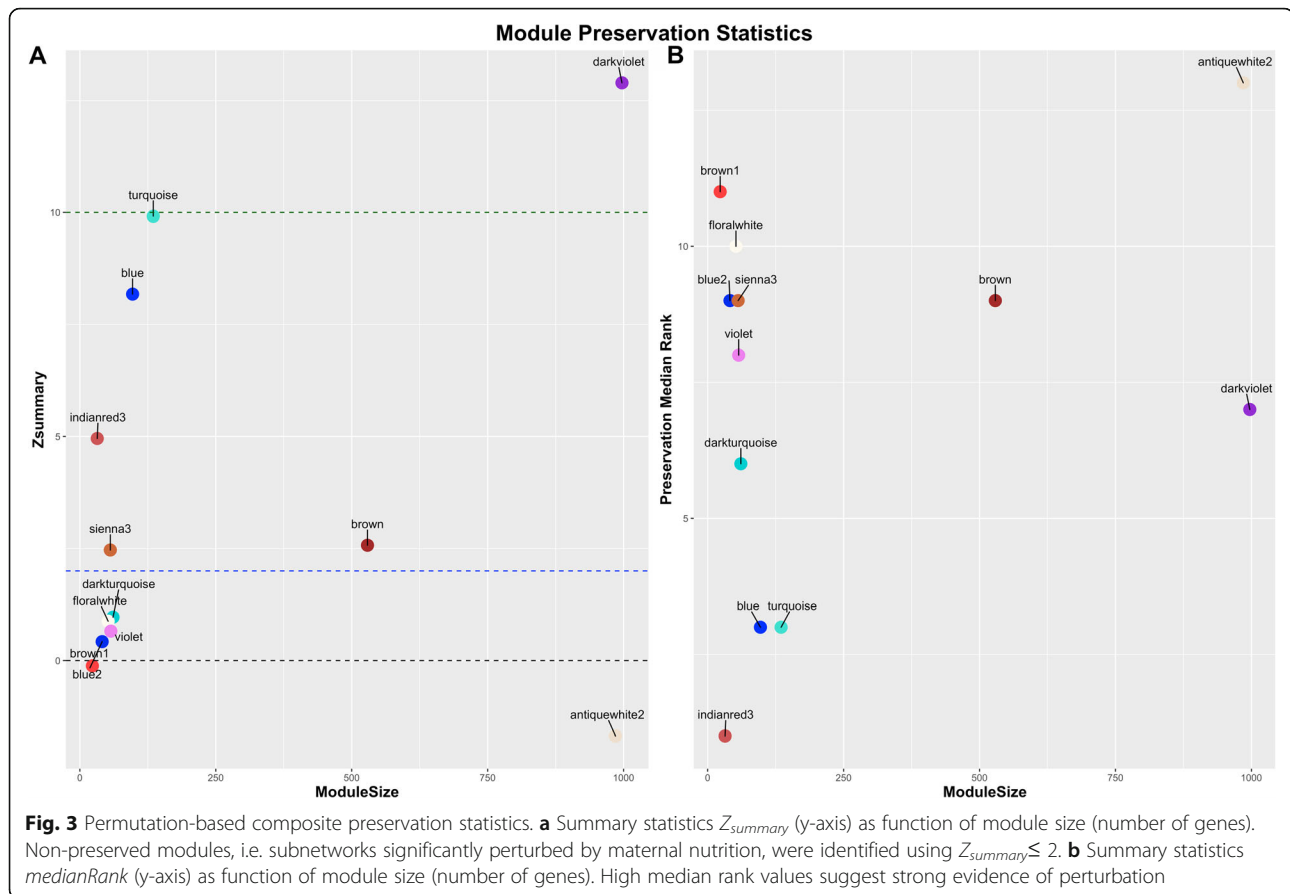
permutation test with 2000 iterations. Twelve module preservation statistics were calculated for each of the 14 modules previously identified in the control diet (see Additional File 2). A total of six modules were considered as unpreserved ( $Z_{summary} \leq 2$ ), five modules were considered as preserved with weak to moderate evidence of preservation ( $2 < Z_{summary} \leq 10$ ), and finally two modules showed strong evidence of preservation ( $Z_{summary} > 10$ ) (Fig. 3). Overall, the six unpreserved modules were considered as gene coexpression modules or subnetworks that were significantly perturbed by the maternal methionine diet.

### Module characterization

The six unpreserved modules were further investigated to reveal their functional roles and gain insights about the biological processes that were impacted by maternal methionine supplementation. This functional characterization was performed using a Fisher's exact test, a hypergeometric-based overrepresentation test commonly used to evaluate  $2 \times 2$  contingency tables. Six different biological databases were evaluated, including Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, InterPro, Medical Subject Headings (MeSH), and Molecular Signatures Database (MSigDB). Figure 4 shows the functional characterization for *antiquewhite2*, the most perturbed module. Interestingly, our analysis revealed that



**Fig. 2** Gene coexpression networks. **a** Gene cluster dendrogram in the maternal control diet. A total of 147 preliminary modules were detected using a dynamic tree cut process, and after merging highly correlated modules (correlation  $\geq 0.8$ ), a total of 14 modules were retained for subsequent analysis. **b** Gene cluster dendrogram in the maternal methionine diet. Changes in the structure of gene coexpression networks between maternal diets were evaluated using a permutation test



genes in this unpreserved module are closely related to (i) myogenesis, adipogenesis, and fibrogenesis, (ii) ribosome structure, (iii) rRNA binding and processing, (iv) mitochondrial activities, (v) ATP synthesis, and (vi) NAD(P) H oxidoreductases. Moreover, genes in the module *violet* are implicated in the regulation of canonical Wnt signaling pathway, a signal transduction pathway that is involved in different embryonic processes, such as cell fate specification, cell proliferation, and cell migration. Additional File 3 reports the full list of significant functional terms, including term ID, term name, total number of genes in the module, and Fisher's  $P$ -value.

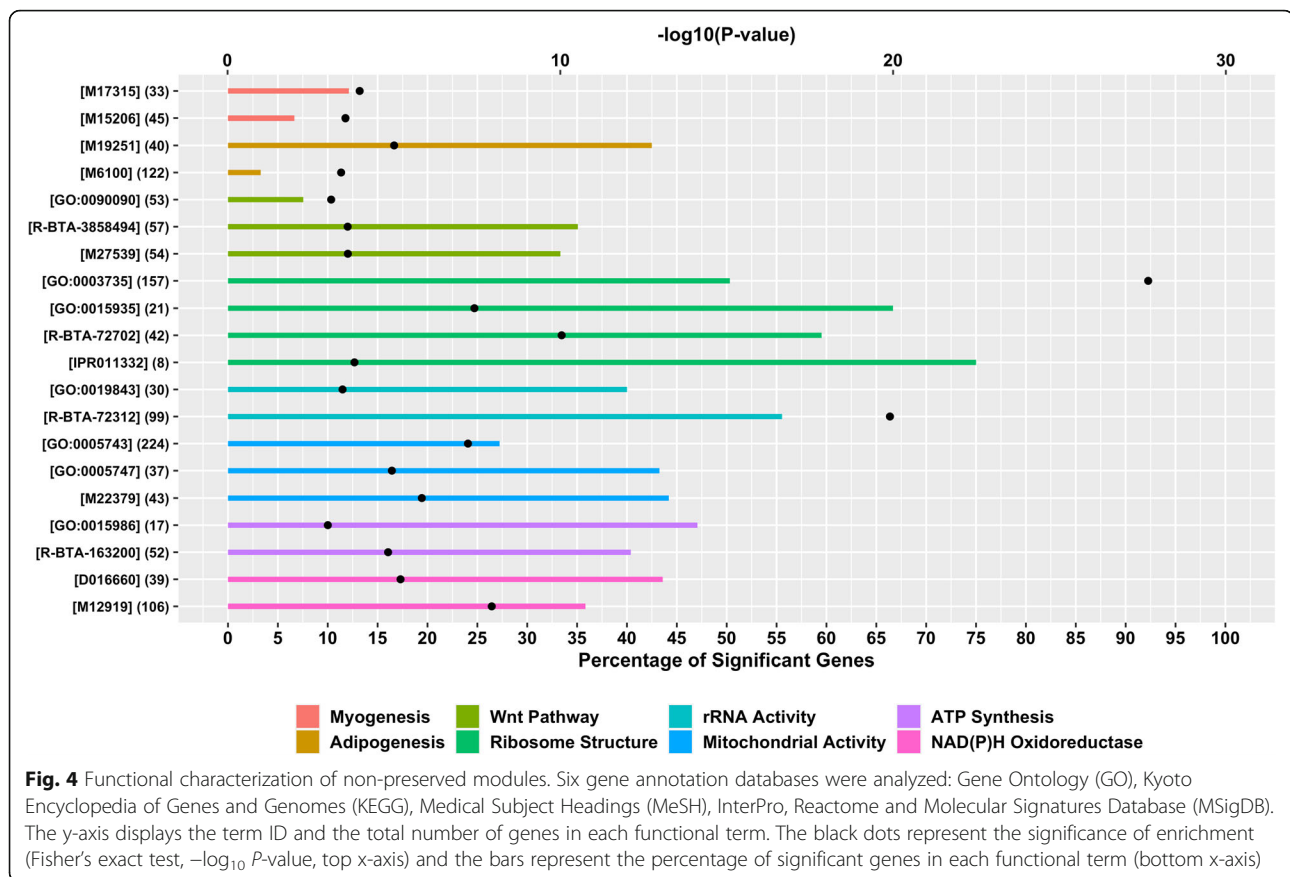
#### DNA methylation analysis

Whole-genome bisulfite sequencing produced roughly 350 M paired-end reads per sample. The software Bismark was used to map the reads to the ARS-UCD1.2 bovine genome assembly, yielding a 70% mapping rate (see Additional File 1). A total of 5,136,556 cytosines (CpG context) were evaluated (read coverage  $\geq 8$ ), and 101,094 were identified as differentially methylated between maternal diets (methylation change  $\geq 20\%$ ,  $q$ -value  $\leq 0.10$ ). Based on the ARS-UCD1.2 annotation file, cytosines were classified as (i) within a gene (gene body: exons and introns), (ii) within the regulatory region (5.5 kb

upstream the gene), or (iii) located in an intergenic region. As results, we targeted a total of 25,491 genes annotated in the cow genome that had at least one evaluated cytosine (either gene body or regulatory region), and 10,247 of the 25,491 had at least one differentially methylated cytosine. Of interest, a total of 6735 of the 7034 genes used in the network analysis had methylation data. Additional File 4 reports the DNA methylation results, including gene ID and number of cytosines per genomic region. Additional File 5 reports the full list of differentially methylated cytosines and the corresponding genomic regions.

#### DNA methylation and module preservation

We investigated if there were significant differences in DNA methylation between genes in preserved and unpreserved modules. For each gene, we calculated the methylation level as differentially methylated cytosines divided by all the cytosines evaluated. We calculated the methylation level either for the gene body or the regulatory region. Notably, the distribution of methylation level in the gene body was significantly different in genes located in unpreserved modules ( $n = 1146$ ) compared to genes in preserved modules ( $n = 5589$ ) (Kolmogorov-Smirnov test,  $P$ -value  $\leq 0.01$ , Fig. 5). We observed the



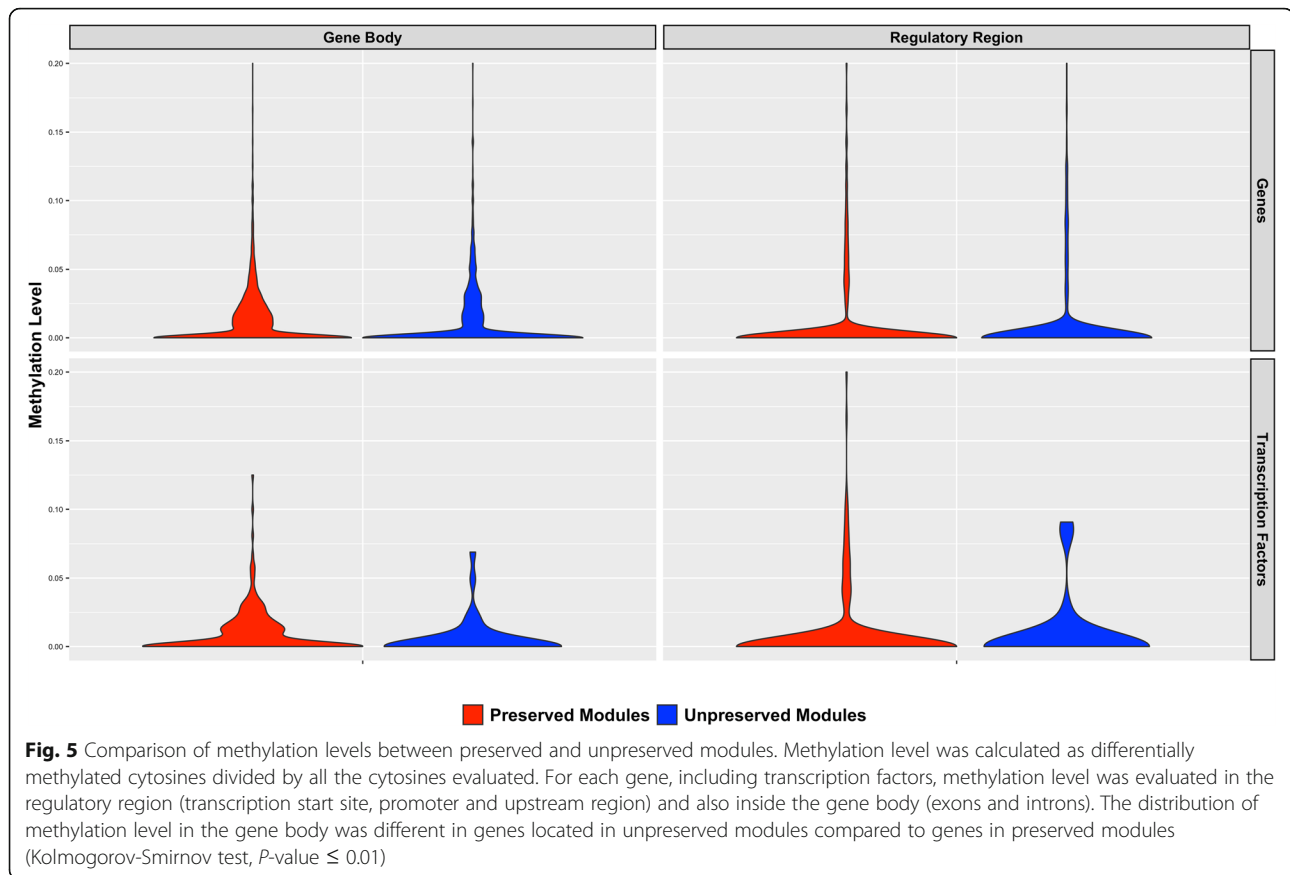
same result if only transcription factors are considered. On the other hand, there was no difference in the distribution of methylation level in the regulatory region between preserved or unpreserved modules, neither for all genes nor for only transcription factors.

#### DNA methylation and network properties

We also investigated the relationship between DNA methylation and three different gene network properties, namely differential coexpression score, module membership, and intramodular connectivity. Interestingly, for those genes located in preserved modules, we found a negative relationship between methylation level and intramodular connectivity, considering either the gene body (regression coefficient  $\beta = -0.23$ ,  $P$ -value = 0.009, Fig. 6) or the regulatory region (regression coefficient  $\beta = -0.10$ ,  $P$ -value = 0.013, Fig. 6). Contrary, there was no relationship ( $P$ -value > 0.05) between methylation level and intramodular connectivity for genes in unpreserved modules. Moreover, a significant negative relationship was found between methylation level in the regulatory region and module membership (regression coefficient  $\beta = -0.02$ ,  $P$ -value = 0.048), but again only for genes in preserved modules. Additional File 6 reports methylation levels and network properties for all the genes evaluated.

#### Discussion

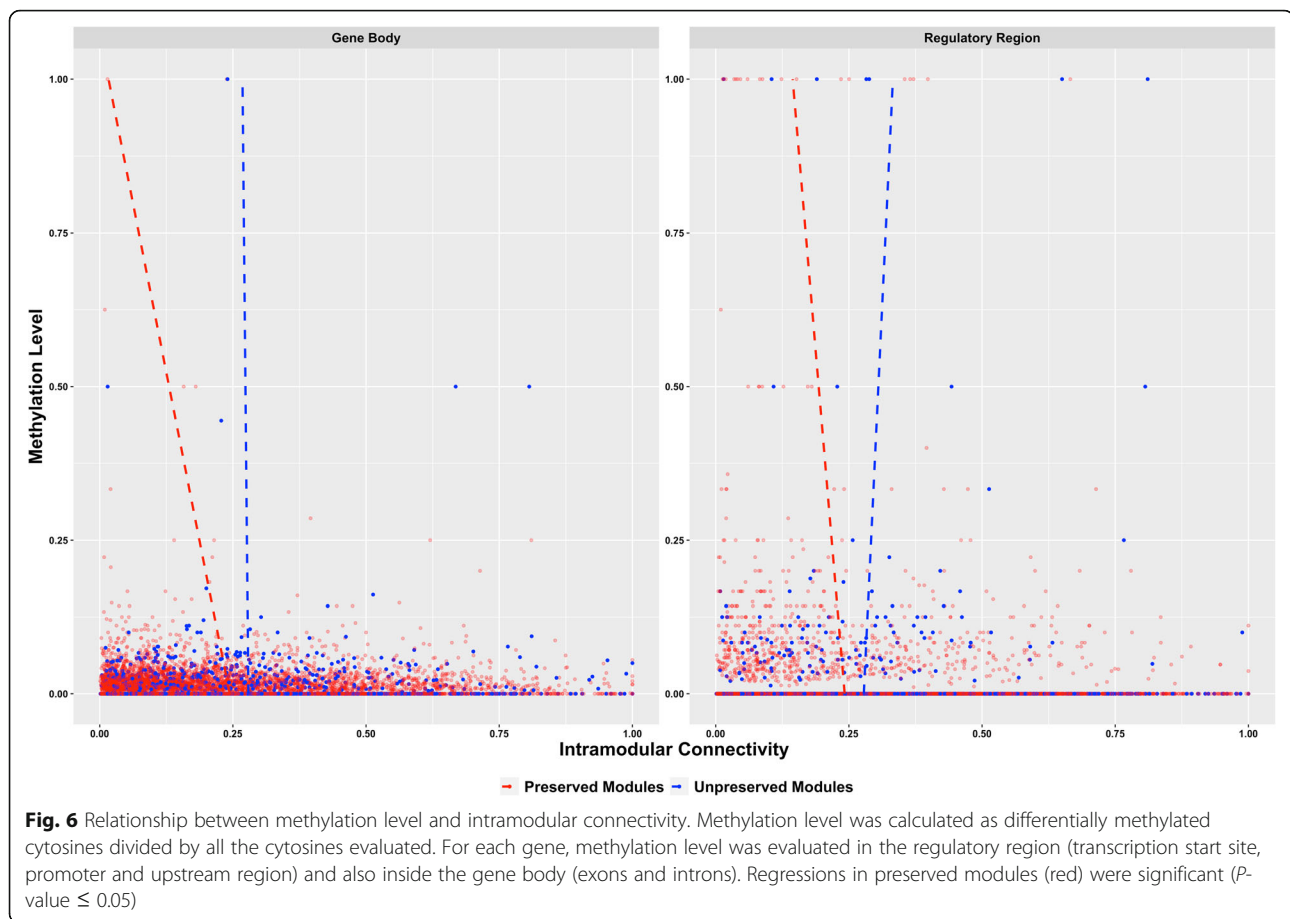
Global coexpression network analysis provides a powerful approach to uncover the molecular basis of phenotypic variation. Gene coexpression networks are typically used to infer and annotate gene functions, prioritize candidate regulatory genes, and reveal transcriptional regulatory mechanisms. Lately, there has been greater emphasis on the use of network analysis to elucidate the changes in gene expression patterns in response to changes in experimental conditions or environmental insults. The present study was specially designed to reveal structural changes in gene coexpression networks due to a maternal methionine-rich diet. Maternal nutrition represents a major intrauterine environmental insult that can induce permanent changes in the offspring. Here, we evaluated gene coexpression networks in the muscle of bull beef calves gestated under a control or methionine-rich diet, we functionally characterized the subnetworks altered by maternal methionine supplementation, and we investigated the link between network perturbation and DNA methylation. Our results provide evidence that maternal nutrition can significantly perturb gene coexpression patterns in the offspring, and some of these changes might be mediated by alterations in DNA methylation.



Maternal methionine supplementation significantly disturbed gene coexpression patterns in the offspring's muscle. In fact, following the methodology proposed by Langfelder and collaborators [5], we identified six modules or subnetworks that significantly changed between experimental conditions. This permutation-based method basically evaluates if the two most important aspects of module topology, namely density and connectivity, are preserved between a reference condition and a test condition. Here, we found that neither the connection strength nor the connectivity pattern of these six subnetworks detected in the control diet were preserved in the methionine-rich diet. There is growing evidence that certain intrauterine insults impact gene coexpression patterns, which in turn may alter fetal development programming. For instance, Deyssenroth et al. [10] reported that alterations in gene coexpression networks in human placenta are associated with abnormal fetal growth and development. Lombardo et al. [11] showed that maternal immune activation via infection during pregnancy disrupts fetal brain gene coexpression networks, and this disruption is associated with an increased risk for autism spectrum disorder. Recently, we reported that exposure to gossypol in utero and during lactation altered the development and gene expression

of the testicles, including a significant perturbation of coexpression patterns among spermatogenesis-related genes [12]. Overall, our findings provide further evidence that intrauterine insults, such as diet, not only can change gene expression but also alter coexpression patterns, which in turn suggests alterations in coexpression mechanisms.

The functional characterization revealed that some of the unpreserved modules are directly implicated in myogenesis, adipogenesis, and fibrogenesis. Notably, it is well-documented that maternal nutrition alters fetal skeletal muscle development by interfering with these three important processes [13]. For instance, Zhu and collaborators have shown that nutrient deficiency in ruminants from early to mid-gestation negatively impacts myogenesis, reducing muscle fiber number and also muscle mass [14, 15]. In addition, Tong and collaborators have reported that maternal overnutrition enhances adipogenesis in fetal skeletal muscle [16, 17]. Similarly, Du et al. [13] reported that maternal undernutrition with supplementation of ruminal bypass protein from day 60 to day 180 of gestation significantly affect adipogenesis, changing marbling scores of steer progeny. Moreover, pigs with reduced birth weight due to malnutrition in utero have greater content of collagen in their skeletal



muscle [18]. Note that myogenesis, adipogenesis and fibrogenesis are vital aspects of muscle physiology, directly impacting lean muscle mass, marbling and also collagen content, and hence, any disturbances in these processes may have long-term consequences, impacting muscle growth and meat quality.

We also found unpreserved subnetworks related to regulation of canonical Wnt/ $\beta$ -catenin pathway. By acting through autocrine and/or paracrine mechanisms, the Wnt family of secreted glycoproteins affects different aspects of cell physiology, such as cell proliferation, cell differentiation or maintenance of precursor cells [19, 20]. Interestingly, in skeletal muscle,  $\beta$ -catenin regulates the expression of two transcription factor, namely *PAX3* and *GLI*, which are essential for skeletal myogenesis [21, 22]. Indeed, some studies have shown that blocking the  $\beta$ -catenin pathway reduces the total number of myocytes [23, 24]. Shang and colleagues showed that the upregulation of Wnt signaling promotes myogenesis, whereas downregulation of this pathway promotes adipogenesis [25]. Moreover, we identified terms related to basic cell structures/functions, including ribosome structure, rRNA binding and processing, mitochondrial activities, ATP synthesis and NAD(P) H oxidoreductases. Notably,

previous studies have shown that these cell activities can be affected by maternal nutrient restrictions. For instance, Peñagaricano et al. [26] reported that maternal diets significantly impact functional terms closely related to ribosome in fetal muscle. Moreover, Mayeur et al. [27] found that maternal undernutrition induces placental mitochondrial abnormalities and reduced ATP level in mice offspring. Similarly, Zhu et al. [14] found that maternal nutrient restriction can induce downregulation of key enzymes involved in mitochondrial function in offspring's muscle.

The exact mechanisms by which maternal diet can affect gene coexpression patterns in the offspring are not yet known. Here, we examined the hypothesis that changes in DNA methylation cause changes in the topology of gene networks. Nearly 2% of all the evaluated cytosines in a CpG context were found to be differentially methylated between maternal diets. Notably, we found significant differences in the level of gene body methylation between genes in preserved modules versus genes in unpreserved modules. The same trend was observed when only transcription factors were considered. Despite the function of DNA methylation in regulatory regions is well-known, the role of DNA methylation



within the gene is not yet well understood. Some studies have suggested that DNA methylation in gene body might be involved in the regulation of alternative splicing [28, 29]. Of special interest, Saha et al. [30] reported that the regulation of alternative splicing is coordinated across functionally related genes. Therefore, changes in DNA methylation might cause changes in isoform expression, which in turn can alter gene coexpression patterns. Moreover, changes in DNA methylation also altered subnetwork properties. Indeed, in preserved modules, there was a clear negative relationship between methylation level and intramodular connectivity, i.e., more methylation, less connectivity, that is, less gene activity. Notably, this functional relationship completely disappeared in unpreserved modules. Overall, our findings suggest that maternal methionine supplementation may induce changes in the offspring epigenome, such as changes in DNA methylation, which in turn alter coexpression patterns and gene network properties.

## Conclusions

Our study has shown that maternal nutrition levels during preconception and early pregnancy can significantly impact gene coexpression patterns in the offspring. Some of the perturbed gene functions are directly implicated in the development of skeletal muscle, such as myogenesis, adipogenesis, and Wnt/ $\beta$ -catenin pathway. Notably, some of the changes in gene coexpression patterns are associated with changes in DNA methylation. To the best of our knowledge, this is the first study that investigates the link between maternal nutrition, DNA methylation and gene coexpression networks. Our findings suggest that maternal nutrition perturbs gene coexpression patterns, and these alterations are in part mediated by changes in the epigenome.

## Methods

### Ethics statement

All the animal procedures used in this study were approved by the Institutional Animal Care and Use Committee (IACUC #2014408583) of the University of Florida. All experiments were performed in accordance with relevant guidelines and regulations.

### Animals and experimental design

Brangus-Angus crossbred beef cows from the University of Florida Range Cattle Research and Education Center (Ona, Florida, US) were assigned to one of two nutritional treatments from days -30 to +90 relative to the beginning of the breeding season. These treatments consisted of a control diet based on limpograss hay (*Hemarthria altissima*) supplemented with molasses and urea (22% crude protein, 1.7 kg per head per day) and a methionine-rich diet equal to the control diet but

supplemented with 10 g per head per day of MetaSmart Liquid (Adisseo, Alpharetta, GA) providing 3.7 g per head per day of rumen-protected methionine. Longissimus dorsi muscle samples were collected from 20 bull calves, 10 per maternal diet, at 1 month of age. After sampling, the bull calves were released. Maternal diets did not affect birth or weaning weight, but altered post weaning calf growth performance, calves derived from the methionine-rich diet had greater average daily gain (0.985 kg vs 0.810,  $P$ -value = 0.043) and also greater gain-to-feed ratio (0.191 vs 0.159,  $P$ -value = 0.025) post weaning.

### RNA extraction, library preparation and sequencing

Total RNA was extracted using Qiagen RNeasy Mini kit. RNA yield and quality were evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). RNA-sequencing libraries were prepared from 50 ng RNA samples using a poly(A) capture method and then sequenced using Illumina's HiSeq 3000 at the University of Florida. A total of 19 muscle samples from 19 bull calves derived from 9 control and 10 methionine-rich maternal diets were successfully processed and sequenced, and hence used for subsequent RNA-sequencing analyses. RNA-sequencing data can be accessed by NCBI GEO with the accession number GSE116974.

### RNA-seq quality control and mapping

The quality of the sequencing reads was evaluated using the software FastQC (v0.11.7, Babraham Bioinformatics, UK). Adaptor removal and trimming were conducted with Trim Galore (version 0.4.4, Babraham Bioinformatics, UK) using the following parameters: --paired, --clip\_R1 10, --clip\_R2 10, --three\_prime\_clip\_R1 10, --three\_prime\_clip\_R2 10, and --length 20. The resulting paired-end sequencing reads were mapped to the latest bovine reference genome (ARS-UCD1.2) using the software Hisat2 (v2.1.0) [31].

### Read counting, processing and normalization

Number of reads that mapped to each annotated gene in the bovine GTF file was obtained using the python script *htseq-count* (v0.6.1p1) using the option *intersection-nonempty* [32]. Both highly abundant genes ( $n = 25$  genes, such as myosins, tropomyosins, and troponins) and lowly expressed genes (read counts  $\leq 5$  in at least 9 biological replicates) were removed from the raw expression data and not included in subsequent analyses. After data processing, read counts were normalized using the trimmed mean of M-values (TMM) normalization method available in the *R* package edgeR (v3.14) [33].

### Gene coexpression network construction

Genes with high expression variance across samples (top 50%) were used for network modeling. The *R* package WGCNA (v1.69) was used for network construction [34, 35]. First, an unsigned adjacency matrix was constructed based on pairwise Pearson correlation coefficients using the function *adjacency*. This adjacency matrix was then remodeled as Topological Overlap Matrix (TOM) using the function *TOMsimilarity*. The TOM-based dissimilarity matrix, simply defined as  $dissTOM = (1 - TOM)$ , was used as pairwise distance matrix for hierarchical clustering. Note that genes in the same cluster share strong interconnections, and they might define modules or subnetworks. Module detection was performed by cutting the branches of the clustering dendrogram using the function *cutreeDynamic*. Genes that could not be assigned to any module were considered as background genes (grey module) and were not included in subsequent module preservation analyses. For each module, the variable module eigengene (ME) was calculated as the first principal component of the module expression data. Modules detected in the dynamic cutting process were merged if their eigengene values were highly correlated (Pearson correlation  $\geq 0.8$ ). The module membership was calculated using the function *signedKME*, where the expression profile of each gene was correlated with the module eigengene in order to quantify how connected a gene was to a given module.

### Module preservation

The function *modulePreservation* implemented in the *R* package WGCNA was used to evaluate the preservation of each module (subnetwork) across conditions, i.e. between maternal control and maternal methionine-rich diets. A total of 12 different module preservation statistics were investigated using a permutation test with 2000 resamples. These 12 different preservation statistics were combined in two composite preservation significance scores, namely  $Z_{density}$  and  $Z_{connectivity}$ . Finally,  $Z_{summary}$  was computed as the mean of  $Z_{density}$  and  $Z_{connectivity}$ , representing the general preservation status of a given module across the two experimental conditions [5]. In addition, an alternative module-size-independent statistic *medianRank* was also calculated to assess module preservation. This rank-based statistic is defined as the mean of observed median ranks for individual connectivity and density preservation statistics in each module, and thus modules with low median rank values are considered as preserved between conditions/treatments. Finally, following Langfelder and colleagues, an unpreserved module was defined as a module with

$Z_{summary} \leq 2$  and *medianRank* greater than half of the total number of the modules detected [5].

### Module characterization

The functional characterization of the unpreserved modules was performed using a Fisher's exact test, a test of proportions based on the cumulative hypergeometric distribution. Genes within each unpreserved module were scrutinized using six different databases, including *Gene Ontology* (GO) [36], *KEGG* [37], *Interpro* [38], *Reactome* [39], *Medical Subject Headings* (MeSH) [40] and *Molecular Signatures Database* (MSigDB) [41]. This over-representation analysis, i.e., evaluate whether a given biological pathway, molecular function or functional term is enriched or over-represented with genes in the unpreserved module, was performed using the *R* package *EnrichKit* (<https://github.com/liulihe954/EnrichKit>).

### DNA extraction, library preparation and sequencing

Total DNA was extracted from muscle samples for whole-genome bisulfite sequencing analysis. Extraction, library construction, bisulfite treatment and sequencing were performed by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). Libraries were sequenced with Illumina's HiSeq 3000 using 150-bp paired-end reads. A total of 16 muscle samples from 16 bull calves derived from 7 control and 9 methionine-rich maternal diets were successfully processed and sequenced, and hence used for subsequent bisulfite-sequencing analyses. Whole-genome bisulfite sequencing data can be accessed by NCBI GEO with the accession number GSE117194.

### Bisulfite-seq quality control and mapping

The quality of the sequencing reads was evaluated using the software FastQC (v0.11.7, Babraham Bioinformatics, UK). Adaptor removal and trimming was performed when needed using the software Trim Galore (v0.4.4, Babraham Bioinformatics, UK). After quality control and processing, the resulting paired-end sequencing reads were aligned to ARS-UCD1.2 bovine reference genome using the software Bismark (v0.17.0, Babraham Bioinformatics, UK) [42]. The tool *deduplicate\_bismark* was used to remove duplicate read alignments. Methylation calls were performed using Bismark *methylation extractor* (v0.17.0, Babraham Bioinformatics) using the following parameters: `--paired-end`, `--comprehensive`, `--bedGraph`, and `--cytosine_report` [42].

### Differentially methylated cytosines and genes

Differential methylation between maternal diets was analyzed using a logistic regression implemented in the *R* package *MethylKit* (v1.0.0) [43]. Only cytosines with read coverage equal or greater than 8 in a CpG context were

evaluated. Differentially methylated cytosines were defined as those having methylation percentage changes between treatments greater than 20% and  $q$ -values  $\leq 0.10$ . The software *Rgmatch* was used to match the cytosines to different gene features, such as transcription start site, exons, introns, and upstream regions [44]. In addition, we defined methylation level as the ratio of differentially methylated cytosines to all the cytosines evaluated in a given region. For each gene, we calculated the methylation level for the regulatory region (transcription start site, promoter and upstream region) and also inside the gene body (exons and introns).

### Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-020-07068-x>.

**Additional file 1.** Mapping Stats.xlsx: Summary of sequencing read alignments to the reference genome for both RNA-Sequencing and Whole-Genome Bisulfite-Sequencing.

**Additional file 2.** Module Preservation Stats.xlsx: Summary of composite and individual module preservation statistics.

**Additional file 3.** Overrepresentation Summary.xlsx: List of significant functional terms.

**Additional file 4.** Gene CpG Count by Region.xlsx: Count of cytosines (CpG context) evaluated in each of the six genomic regions per gene.

**Additional file 5.** Associations of Differentially Methylated CpGs.xlsx: List of differentially methylated cytosines.

**Additional file 6.** Gene Methylation level and Network Properties Measurements.xlsx: Summary of statistics at gene level including differential coexpression score, connectivity, module eigengenes, count of differentially methylated cytosines and methylation levels.

### Abbreviations

GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; ME: Module eigengene; MeSH: Medical Subject Headings (MeSH); MSigDB: Molecular Signatures Database; RNA-seq: RNA-sequencing; TMM: Trimmed mean of M-values; TOM: Topological Overlap Matrix

### Acknowledgements

Not Applicable.

### Authors' contributions

FP, LL, PM, ND, and PL conceived and designed the study. PM, ND, and PL performed the field trial and collected the muscle samples. LL and RA analyzed the sequencing data. LL, RA and FP contributed to the interpretation of the results. LL wrote the first draft of the manuscript. All authors have read and approved this manuscript.

### Funding

This research was supported by the Early Career Scientist Seed Fund Program from the University of Florida, Institute of Food and Agricultural Sciences. The funding body did not contribute to the design of the study or collection, analysis and interpretation of data and writing the manuscript.

### Availability of data and materials

Bisulfite-Seq (GSE117194) and RNA-Seq (GSE116974) data are available on NCBI GEO with accession number (SuperSeries) GSE117195. The ARS-UCD1.2 bovine reference genome is available on Ensembl: [ftp://ftp.ensembl.org/pub/release-101/fasta/bos\\_taurus/dna/](ftp://ftp.ensembl.org/pub/release-101/fasta/bos_taurus/dna/).

### Ethics approval and consent to participate

The University of Florida Institutional Animal Care and Use Committee approved the procedures used in this study (IACUC #2014408583).

### Consent for publication

Not applicable.

### Competing interests

The corresponding author, Francisco Peñagaricano, is a member of the editorial board for the BMC Genomics journal.

### Author details

<sup>1</sup>Department of Animal and Dairy Sciences, University of Wisconsin-Madison, 1675 Observatory Drive, Madison, WI 53706, USA. <sup>2</sup>Department of Animal Sciences, University of Florida, Gainesville, FL 32611, USA. <sup>3</sup>Range Cattle Research and Education Center, University of Florida, Ona, FL 33865, USA. <sup>4</sup>North Florida Research and Education Center, University of Florida, Marianna, FL 32351, USA. <sup>5</sup>Department of Clinical Sciences, Kansas State University, Manhattan, KS 66506, USA.

Received: 1 June 2020 Accepted: 14 September 2020

Published online: 02 October 2020

### References

- Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet*. 2011;12(2):87–98.
- Han JD. Understanding biological functions through molecular networks. *Cell Res*. 2008;18(2):224–37.
- Allocco DJ, Kohane IS, Butte AJ. Quantifying the relationship between co-expression, co-regulation and gene function. *BMC bioinformatics*. 2004;5(1):18.
- van Dam S, Vosa U, van der Graaf A, Franke L, de Magalhaes JP. Gene co-expression analysis for functional classification and gene-disease predictions. *Brief Bioinform*. 2018;19(4):575–92.
- Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? *PLoS Comput Biol*. 2011;7(1):e1001057.
- Godfrey KM, Barker DJ. Fetal programming and adult health. *Public Health Nutr*. 2001;4(2b):611–24.
- Wu G, Bazer FW, Cudd TA, Meininger CJ, Spencer TE. Maternal nutrition and fetal development. *J Nutr*. 2004;134(9):2169–72.
- Burdge GC, Lillycrop KA. Nutrition, epigenetics, and developmental plasticity: implications for understanding human disease. *Annu Rev Nutr*. 2010;30:315–39.
- Chmurzynska A. Fetal programming: link between early nutrition, DNA methylation, and complex diseases. *Nutr Rev*. 2010;68(2):87–98.
- Deysenroth MA, Peng SE, Hao K, Lambertini L, Marsit CJ, Chen J. Whole-transcriptome analysis delineates the human placenta gene network and its associations with fetal growth. *BMC Genomics*. 2017;18.
- Lombardo MV, Moon HM, Su J, Palmer TD, Courchesne E, Pramparo T. Maternal immune activation dysregulation of the fetal brain transcriptome and relevance to the pathophysiology of autism spectrum disorder. *Mol Psychiatry*. 2018;23(4):1001–13.
- Louvandini H, Correa PS, Amorin R, Liu LH, Ieda EH, Jimenez CR, Tsai SM, McManus CM, Peñagaricano F. Gestational and lactational exposure to gossypol alters the testis transcriptome. *BMC Genomics*. 2020;21(1):59.
- Du M, Tong J, Zhao J, Underwood KR, Zhu M, Ford SP, Nathanielsz PW. Fetal programming of skeletal muscle development in ruminant animals. *J Anim Sci*. 2010;88:E51–60.
- Zhu MJ, Ford SP, Means WJ, Hess BW, Nathanielsz PW, Du M. Maternal nutrient restriction affects properties of skeletal muscle in offspring. *J Physiol-London*. 2006;575(1):241–50.
- Zhu MJ, Ford SP, Nathanielsz PW, Du M. Effect of maternal nutrient restriction in sheep on the development of fetal skeletal muscle. *Biol Reprod*. 2004;71(6):1968–73.
- Tong J, Zhu MJ, Underwood KR, Hess BW, Ford SP, Du M. AMP-activated protein kinase and adipogenesis in sheep fetal skeletal muscle and 3T3-L1 cells. *J Anim Sci*. 2008;86(6):1296–305.
- Tong JF, Yan X, Zhu MJ, Ford SP, Nathanielsz PW, Du M. Maternal obesity downregulates myogenesis and  $\beta$ -catenin signaling in fetal skeletal muscle. *American Journal of Physiology-Endocrinology and Metabolism*. 2009;296(4):E917–E24.
- Karunaratne JF, Ashton CJ, Stickland NC. Fetal programming of fat and collagen in porcine skeletal muscles. *J Anat*. 2005;207(6):763–8.
- Novakofski J. Adipogenesis: usefulness of in vitro and in vivo experimental models. *J Anim Sci*. 2004;82(3):905–15.

20. Johnson ML, Rajamannan N. Diseases of Wnt signaling. *Rev Endocr Metab Dis.* 2006;7(1–2):41–9.
21. Capdevila J, Tabin C, Johnson RL. Control of dorsoventral somite patterning by Wnt-1 and  $\beta$ -catenin. *Dev Biol.* 1998;193(2):182–94.
22. Borycki AG, Brown AMC, Emerson CP. Shh and Wnt signaling pathways converge to control Gli gene activation in avian somites. *Development.* 2000;127(10):2075–87.
23. Pan WJ, Jia YY, Wang JY, Tao DL, Gan XQ, Tsiokas L, Jing NH, Wu DQ, Li L. Beta-catenin regulates myogenesis by relieving I-mfa-mediated suppression of myogenic regulatory factors in P19 cells. *P Natl Acad Sci USA.* 2005;102(48):17378–83.
24. Yamanouchi K, Hosoyama T, Murakami Y, Nishihara M. Myogenic and adipogenic properties of goat skeletal muscle stem cells. *J Reprod Develop.* 2007;53(1):51–8.
25. Shang Y, Zhang C, Wang S, Xiong F, Zhao C, Peng F, Feng S, Yu M, Li M, Zhang Y. Activated  $\beta$ -catenin induces myogenesis and inhibits adipogenesis in BM-derived mesenchymal stromal cells. *Cytotherapy.* 2007;9(7):667–81.
26. Peñagaricano F, Wang X, Rosa GJM, Radunz AE, Khatib H. Maternal nutrition induces gene expression changes in fetal muscle and adipose tissues in sheep. *BMC Genomics.* 2014;15:1034.
27. Mayeur S, Lancel S, Theys N, Lukaszewski MA, Duban-Deweer S, Bastide B, Hachani J, Cecchelli R, Breton C, Gabory A, et al. Maternal calorie restriction modulates placental mitochondrial biogenesis and bioenergetic efficiency: putative involvement in fetoplacental growth defects in rats. *Am J Physiol – Endocrinol Metab.* 2013;304(1):E14–22.
28. Maunakea AK, Chepelev I, Cui KR, Zhao KJ. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* 2013;23(11):1256–69.
29. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature.* 2011;479(7371):74–U99.
30. Saha A, Kim Y, Gewirtz ADH, Jo B, Gao C, McDowell IC, Engelhardt BE, Battle A, Consortium G. Co-expression networks reveal the tissue-specific regulation of transcription and splicing. *Genome Res.* 2017;27(11):1843–58.
31. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–60.
32. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31(2):166–9.
33. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139–40.
34. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
35. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol.* 2005;4:Article17.
36. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. *Gene Ontology Consortium Nat Genet.* 2000;25(1):25–9.
37. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
38. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang HY, El-Gebali S, Fraser MI, et al. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 2019;47(D1):D351–D60.
39. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2020;48(D1):D498–503.
40. Nelson SJ, Schopen M, Savage AG, Schulman JL, Arluk N. The MeSH translation maintenance system: structure, interface design, and implementation. *Stud Health Technol Inform.* 2004;107(Pt 1):67–9.
41. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 2011;27(12):1739–40.
42. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. *Bioinformatics.* 2011;27(11):1571–2.
43. Akalin A, Korkmakkon M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012;13(10):R87.
44. Furio-Tari P, Conesa A, Tarazona S. RGMATCH: matching genomic regions to proximal genes in omics data integration. *BMC Bioinformatics.* 2016;17(Suppl 15):427.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

