



## OPEN ACCESS

## EDITED BY

Nuno Carolino,  
Instituto Nacional Investigação Agrária  
e Veterinária (INIAV), Portugal

## REVIEWED BY

Zitong Li,  
Commonwealth Scientific and Industrial  
Research Organisation (CSIRO),  
Australia  
Tiago Bresolin,  
University of Illinois at Urbana-  
Champaign, United States

## \*CORRESPONDENCE

Gota Morota,  
morota@vt.edu  
Luiz Lehmann Coutinho,  
llcouthino@usp.br

## SPECIALTY SECTION

This article was submitted to Livestock  
Genomics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 19 May 2022

ACCEPTED 06 October 2022

PUBLISHED 21 October 2022

## CITATION

Novais FJd, Yu H, Cesar ASM, Momen M,  
Poleti MD, Petry B, Mourão GB,  
Regitano LCdA, Morota G and  
Coutinho LL (2022), Multi-omic data  
integration for the study of production,  
carcass, and meat quality traits in  
Nellore cattle.  
*Front. Genet.* 13:948240.  
doi: 10.3389/fgene.2022.948240

## COPYRIGHT

© 2022 Novais, Yu, Cesar, Momen,  
Poleti, Petry, Mourão, Regitano, Morota  
and Coutinho. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Multi-omic data integration for the study of production, carcass, and meat quality traits in Nellore cattle

Francisco José de Novais<sup>1</sup>, Haipeng Yu<sup>2</sup>,  
Aline Silva Mello Cesar<sup>3</sup>, Mehdi Momen<sup>2</sup>, Mirele Daiana Poleti<sup>4</sup>,  
Bruna Petry<sup>1</sup>, Gerson Barreto Mourão<sup>1</sup>,  
Luciana Correia de Almeida Regitano<sup>5</sup>, Gota Morota<sup>2\*</sup> and  
Luiz Lehmann Coutinho<sup>1\*</sup>

<sup>1</sup>Department of Animal Science, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba, Brazil, <sup>2</sup>Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States, <sup>3</sup>Department of Agri-Food Industry, Food and Nutrition, University of São Paulo, Piracicaba, Brazil, <sup>4</sup>Department of Veterinary Medicine, School of Animal Science and Food Engineering, University of São Paulo, Pirassununga, Brazil, <sup>5</sup>Embrapa Pecuária Sudeste, São Carlos, Brazil

Data integration using hierarchical analysis based on the central dogma or common pathway enrichment analysis may not reveal non-obvious relationships among omic data. Here, we applied factor analysis (FA) and Bayesian network (BN) modeling to integrate different omic data and complex traits by latent variables (production, carcass, and meat quality traits). A total of 14 latent variables were identified: five for phenotype, three for miRNA, four for protein, and two for mRNA data. Pearson correlation coefficients showed negative correlations between latent variables miRNA 1 (mirna1) and miRNA 2 (mirna2) (−0.47), ribeye area (REA) and protein 4 (prot4) (−0.33), REA and protein 2 (prot2) (−0.3), carcass and prot4 (−0.31), carcass and prot2 (−0.28), and backfat thickness (BFT) and miRNA 3 (mirna3) (−0.25). Positive correlations were observed among the four protein factors (0.45–0.83): between meat quality and fat content (0.71), fat content and carcass (0.74), fat content and REA (0.76), and REA and carcass (0.99). BN presented arcs from the carcass, meat quality, prot2, and prot4 latent variables to REA; from meat quality, REA, mirna2, and gene expression mRNA1 to fat content; from protein 1 (prot1) and mirna2 to protein 5 (prot5); and from prot5 and carcass to prot2. The relations of protein latent variables suggest new hypotheses about the impact of these proteins on REA. The network also showed relationships among miRNAs and nebulin proteins. REA seems to be the central node in the network, influencing carcass, prot2, prot4, mRNA1, and meat quality, suggesting that REA is a good indicator of meat quality. The connection among miRNA latent variables, BFT, and fat content relates to the influence of miRNAs on lipid metabolism. The relationship between mirna1 and prot5 composed of isoforms of nebulin needs further investigation. The FA identified latent variables, decreasing the dimensionality and complexity of the data. The BN was capable of generating interrelationships among latent variables from different types of data, allowing the integration of omics and

complex traits and identifying conditional independencies. Our framework based on FA and BN is capable of generating new hypotheses for molecular research, by integrating different types of data and exploring non-obvious relationships.

#### KEYWORDS

Bayesian network, factor analysis, meat quality, latent variables, omics data

## Introduction

Meat quality traits, which include meat tenderness, are an important aspect for consumers and are related to the customer's acceptability and beef repurchase (Rust et al., 2008). Meat quality traits are complex and influenced by diet, pre- and post-slaughter management, meat processing, storage methods, genetic factors, and genotype-by-environment interaction (Wheeler et al., 2005; Adzitey, 2011; Guerrero et al., 2013; Njisane and Muchenje, 2016). Most of the biological mechanisms involved in meat quality traits are not completely understood. In this context, systems biology has been proposed to elucidate the flux of molecular information, generating a holistic point of view for complex traits (Ideker et al., 2001). Data from the genome, transcriptome, proteome, microRNAome, and metabolome have been used independently to study the molecular architecture of complex traits and identify important genes, pathways, and networks that underlie economic livestock traits in the last decade (Tizioto et al., 2013; Carvalho et al., 2014; Cesar et al., 2014, 2016; Novais et al., 2019). However, studies using single omic data disregard the interactions among different levels of biomolecules, postulated by the central dogma of molecular biology (Ritchie M. D. et al., 2015).

Complex traits are regulated at different molecular levels, and considerable effort has been made to generate multi-level studies, integrating different omic data to understand the inherent biological meaning of livestock traits (Widmann et al., 2013; Suravajhala et al., 2016). However, omic data integration using a hierarchical analysis approach or considering just the common pathway enrichment may not reveal non-obvious relationships that exist among omic data (Misra et al., 2018). In this context, efforts to develop approaches to data omic integration have been proposed (Huang et al., 2017).

Factor analysis (FA) reduces the dimensionality of data, inferring latent (hidden) variables to explain dependencies among observed variables that share common variations (Meng et al., 2016). Furthermore, the Bayesian network (BN) has the potential to generate relationships among phenotypes and molecules by a graph-based model of joint multivariate probability distributions that represent conditional independence between variables (Rodin and Boerwinkle, 2005). Here, phenotypes of production, carcass, meat quality, and multi-omic data were fitted into the FA and BN framework

to explore the potential biological interrelationships to generate new hypotheses for complex traits in beef cattle.

## Materials and methods

### Animals and phenotypes

A total of 386 Nellore steers born between 2009 and 2011 at the Brazilian Agricultural Research Corporation (EMBRAPA/Brazil) were initially included in this study. The animals were raised in feedlots under identical diets, and environmental conditions, and slaughtered at age of 25 months. More details regarding animals, diet, and experimental design can be found in Cesar et al. (2014). The animals were handled and managed according to the Institutional Animal Care and Use Committee Guidelines from the Brazilian Agricultural Research Corporation—EMBRAPA approved by the president, Dr. Rui Machado.

Carcass ultrasound evaluations were performed by trained field technicians and followed the standards set by the Ultrasound Guidelines Council (UGC; [www.ultrasoundbeef.com](http://www.ultrasoundbeef.com)). An Aquila Pie Medical (Pie Medical Inc., Maastricht, Netherlands) equipped with a 172 mm-long linear transducer with a frequency of 3.5 MHz was used to measure the initial ribeye area (REAi) and initial backfat thickness (BFTi) obtaining sectional images of the longissimus dorsi (LD) muscle between the 12th and 13th ribs. The images were stored and measurements were obtained by ODT Eview R (Pie Medical Inc., Maastricht, Netherlands).

The details of carcass and meat quality trait evaluations were previously described by Nascimento et al. (2016). The visceral organs were removed during slaughter, and the heart, kidney, liver, and perirenal, pelvic, and inguinal fats were weighed. Carcasses were weighed and chilled for 24 h at 5°C. The carcass was weighted at 24 h, and the carcass depth was measured on the fifth rib from top to bottom, measuring the distance from the sternum to the middle of the spine where the marrowbone passes.

Steaks of 2.54 cm thick from the LD muscle between the 12th and 13th ribs were collected 24 h after slaughter. Steaks were vacuum packed and used to measure the shear force (SF; Kg), backfat thickness (BFT; mm), ribeye area (REA; cm<sup>2</sup>), myofibrillar fragmentation index (MFI), color parameters

( $L^*$  = lightness,  $a^*$  = redness, and  $b^*$  = yellowness), intramuscular fat (IMF; percentage), pH at 24 h, moisture, water holding capacity, and cook loss. Briefly, the final backfat thickness (BFTf) was measured using a ruler in millimeters (Filho, 2000). Color parameters  $L^*$ ,  $a^*$ , and  $b^*$  were measured after exposing the steaks to atmospheric oxygen for 30 min prior to analysis using a Hunter Lab colorimeter model MiniScan XE with Universal Software v. 4.10 (Hunter Associates Laboratory, Reston, VA), illuminant D65, and 10° standard observer. Additionally, muscle pH was measured at three locations across the steak using a Testo pH measuring instrument model 230 (Testo, Lenzkirch, Germany). The final ribeye area (REaf) was calculated as the area of LD muscle using a grid. Cooking losses were measured as the weight difference between the steaks before and after cooking. For IMF, approximately 100 g of muscle samples, previously lyophilized and ground, were obtained using an Ankom XT20 extractor as described in AOCS official procedure Am 5-04 (Horwitz, 2000). The myofibrillar fragmentation index was determined according to Hopkins et al. (2000). The SF values were obtained from 2.54 cm thick steaks after 24 h of aging at 2°C in a cold chamber using the texture analyzer TA-XT2i coupled to a Warner–Bratzler blade with 1.016 mm thickness.

## mRNA data processing and WGCNA

For total RNA extraction, a sample of 100 mg of the LD muscle was processed using the Trizol reagent (Life Technologies, Carlsbad, CA, United States), following the manufacturer's guidelines. After extraction, RNA integrity was verified using the Bioanalyzer 2100 (Agilent, Santa Clara, CA, United States), and the samples presenting RNA integrity numbers lower than 7.0 were removed from further analysis. A total of 2 µg of RNA from each sample was used for the cDNA library preparation, in accordance with the protocol described in the TruSeq RNA Sample Preparation kit v2 guide (Illumina, San Diego, CA, United States). The libraries were sequenced using the HiSeq2500 ultra-high-throughput sequencing system (Illumina, San Diego, CA, United States) with the TruSeq SBS kit v3-HS (200 cycles). All sequencing analyses were performed at the ESALQ Genomics Center (Piracicaba, São Paulo, Brazil).

The FastQC software v0.10.1 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was applied to check the quality of the sequencing data. Low-quality reads were filtered and adapter sequences were trimmed using Seqclean package version 1.4.13 (Neapolitan, 2004). The details of data acquisition were previously described by de Lima et al. (2020).

The read alignment was carried out against the bovine reference genome *Bos taurus* ARS-UCD1.2 (available at the Ensembl database [https://www.ncbi.nlm.nih.gov/assembly/GCF\\_002263795.1](https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1)) and read counts using STAR software (Spliced Transcripts Alignment to a Reference) version 2.7

(Dobin and Gingeras, 2015) with the Ensembl (release 95, January 2019) gene annotation file. Subsequently, genes with zero counts for all samples were removed. Next, the genes were filtered by the counts different from zero in at least 70% of the samples and counts per million (CPM) > 5 using the EdgeR Bioconductor package (Chen H.-J. et al., 2018). This was followed by normalizing counts using the DESeq2 Bioconductor package (Love et al., 2014), and a batch effect was identified using the limma R package (Ritchie M. E. et al., 2015).

Clustering analysis was performed on the mRNA dataset using the weighted gene co-expression network analysis (WGCNA) R package (Langfelder and Horvath, 2008). To measure the connectivity among genes, an adjacency matrix was generated by calculating the Pearson's correlation coefficients among all genes and raising it to a power  $\beta$  (soft threshold) of 6, which is chosen using a scale-free topology criterion ( $R^2 = 0.8$ ). Modules containing at least 30 genes were retained. Modules with hub genes that had a module membership (MM) > 0.95 and gene significance (GS) with a  $p$ -value < 0.001 were kept for further analysis. Enrichment analysis was performed using MetaCore software (MetaCore, 2021) to elucidate biological processes and pathways represented by the hub genes of modules.

## miRNA and data acquisition

Small RNA libraries were constructed from 1 µg of total RNA from each sample using the Illumina TruSeq small RNA Sample Prep Kit (Illumina Inc, San Diego, CA, United States), in accordance with the manufacturer's protocol. High Sensitivity DNA Chip and an Agilent 2100 Bioanalyzer (Agilent Technologies) was used to determine library quality and qPCR with the KAPA Library Quantification kit (KAPA Biosystems, Foster City, CA, United States) for quantification. Sequencing was performed using a Miseq Reagent Kit v3 for 150 cycles in an Illumina Miseq Sequencing System (Illumina Inc., San Diego, CA, United States). The Illumina CASAVA v1.8 was used to generate and de-multiplex the raw fastq sequences. The quality of Illumina deep sequencing data was determined using the FastQC program (version 0.9.5) (Andrews, 2010). Adapters and low-quality reads were trimmed using Cutadapt (version 1.2.1) (Martin, 2011). Filtered reads were then processed following the mirDeep2 analysis pipeline (Friedländer et al., 2012). Sequences were aligned to the *Bos taurus* ARS-UCD1.2 reference genome (available at the Ensembl database ([https://www.ncbi.nlm.nih.gov/assembly/GCF\\_002263795.1](https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1))). Only alignments with zero mismatches in the seed region (first 18 nucleotides of a read sequence) of a read mapped to the genome were retained. More details about data acquisition were provided by Kappeler et al. (2019).

Briefly, miRNAs with zero counts for all samples were removed. Next, the miRNAs were filtered by the counts that are different from zero in at least 70% of the samples and CPM > 5 using the EdgeR Bioconductor package (Chen Y.

et al., 2018). The miRNA counts were normalized using the DESeq2 Bioconductor package (Love et al., 2014), and the limma R package was used to identify a batch effect (Ritchie M. E. et al., 2015).

## Proteome and data acquisition

The details for data acquisition and processing are previously described in Poleti et al. (2018). Frozen muscles (500 µg) of 106 animals were ground on liquid nitrogen, then transferred to a microcentrifuge tube, and weighed to minimize protein degradation. The muscle was homogenized in 2.5 ml lysis buffer containing 8 M urea, 2 M thiourea, 1% DTT, 2% CHAPS, and 1% protease inhibitor cocktails (Sigma-Aldrich) in an ULTRA-TURRAX® IKA homogenizer on ice for 2 min. The extracts were vigorously shaken for 30 min on ice and centrifuged at 10,000 × g for 30 min at 4°C. The supernatants were collected, the total protein concentration was determined by the PlusOne 2-D Quant Kit (GE Healthcare), and then stored at -80°C for further analysis.

The protein extract was desalted with a 3-kDa cutoff Amicon® Ultra centrifugal filter (Millipore, Ireland), where the lysis buffer was exchanged using a solution of 50 mM ammonium bicarbonate and 2 M urea five times. The concentration of the retained protein solution was quantified using a Bradford Protein Assay Kit (BioRad). For protein digestion, 50 µg of proteins of each sample were denatured with 25 µL of 0.2% RapiGest SF (Waters Corporation, United States) at 80°C for 15 min, reduced with 2.5 µL of 100 mM dithiothreitol (DTT) (Sigma, United States) at 60°C for 30 min, and alkylated with 2.5 µL of 300 mM iodoacetamide (AA) (Sigma, United States) at room temperature in the dark for 30 min. Enzymatic digestion was performed with sequencing grade modified trypsin (Promega) at a 1:100 (w/w) enzyme: protein ratio at 37°C for 16 h. Digestion was stopped by the addition of 10 µL of 5% (V/V) trifluoroacetic acid and incubated at 37°C for 90 min to hydrolyze the RapiGest (Yu et al., 2003). The peptide mixture solution was then centrifuged at 18,000 × g for 30 min at 6°C. The supernatant was transferred to a new vial, dried down in a vacuum centrifuge, and stored at -20°C.

Qualitative and quantitative bidimensional nanoUPLC tandem nanoESI-HDMSE analyses were conducted using both 1-h reversed-phase gradient from 7% to 40% (v/v) acetonitrile (0.1% v/v formic acid) and 500 nL·min<sup>-1</sup> on a nanoACQUITY UPLC 2D Technology system (Gilar et al., 2005). A nanoACQUITY UPLC HSS T3 1.8 µm, 75 µm × 15 cm column (pH 3) was used in conjunction with a reverse-phase (RP) XBridge BEH130 C18 5 µm 300 µm × 50 mm nanoflow column (pH 10). The ion mobility cell was activated and filled with nitrogen gas, which operates at the cross-section resolving power of at least 40 Ω/ΔΩ (Lalli et al., 2013). The effective resolution has the conjoined ion mobility of >1.5 M FWHM

(Silva et al., 2014). The ionization of samples was performed using a NanoLockSpray ionization source (Waters, Manchester, United Kingdom) in the positive ion mode nanoESI (+). The mass spectrometer was calibrated with an MS/MS spectrum of [Glu1]-fibrinopeptide B human (Glu-Fib) solution (100 fmol·µL<sup>-1</sup>) delivered through the reference sprayer of the NanoLockSpray source. Data acquisition was performed using a Synapt G2-S HDMS mass spectrometer (Waters, Manchester, United Kingdom). A mass-charge value ranges from m/z 50 to 2000.

Mass spectrometry data were acquired with Waters MassLynx v.4.1 software and processed using Progenesis QI for Proteomics (QIP) 2.0 software (Nonlinear Dynamics, United Kingdom). Progenesis QIP software was used to run alignment, peak picking, ion drift time data collection, ion abundance measurements, normalization, quantification, peptide and protein identification, and statistical analysis. The processing parameters for Progenesis included the following: automatic tolerance for precursor and product ions based on peptide identification and normal distribution (Geromanos et al., 2009), one missed cleavage, carbamidomethylation of cysteine as a fixed modification, and oxidation of methionine as variable modification. For protein identification and quantification, the obtained raw data were searched against a Nellore transcriptome database built from RNA-sequencing data from LD muscle. Data quality assessment was performed accordingly (Souza et al., 2017), and proteins were selected based on the detection and identification in at least 80% of biological samples. The assembled data were compared to the NCBI's UniProt database (<https://www.uniprot.org/>) as functional analysis.

## Factor analysis

This section closely follows the work of Yu et al. (2020) and Momen, et al. (2021). The exploratory factor analysis (EFA) was applied to search the structure of underlying latent variables (factors) that drive the observed phenotypes and omic data. First, the caret R package (Kuhn, 2008) was used to check collinearity, and one of the features with correlation >0.9 was removed. Then, the Kaiser-Meyer-Olkin (KMO) test was applied to measure the sampling adequacy using the psych R package (Revelle, 2017) assessing the factor ability of the data (Cerny and Kaiser, 1977). The measure of sampling adequacy ranges between 0 and 1, and values closer to 1 are preferred. Here, KMO >0.7 was considered acceptable. The number of underlying latent variables  $q$  was determined using a parallel analysis (Horn, 1965) using the psych R package, as described in more detail in a previous work of our group (Momen et al., 2021). The EFA model is given as a function of latent factor scores.

$$Y = \Lambda F + \epsilon,$$

where  $Y$  is a  $p \times n$  matrix of  $p$  molecular features or phenotypes of  $n$  animals,  $\Lambda$  is the  $p \times q$  matrix of factor loading connecting the relation between features and latent common factors,  $F$  is the  $q \times n$  matrix of latent factor scores, and  $\varepsilon$  is the  $p \times n$  vector of unique effects that is not explained by  $q$  underlying common factors. The variance–covariance matrix of  $Y$  is

$$\Sigma = \Lambda\Phi\Lambda' + \Psi,$$

where  $\Sigma$  is the  $p \times p$  variance–covariance matrix of phenotypes,  $\Phi$  is the variance of factor scores, and  $\Psi$  is a  $p \times p$  diagonal matrix of unique variance. The elements of  $\Lambda$ ,  $\Phi$ , and  $\Psi$  are parameters of the model to be estimated from the data. With the assumption of  $F \sim N(0, I)$ ,  $\Lambda$  and  $\Psi$  were estimated by maximizing the log-likelihood of  $\mathcal{L}(\Lambda, \Psi|Y)$  using the R package *psych* (Revelle, 2017) along with a varimax rotation (Kaiser, 1958). A parallel analysis was performed to determine the number of underlying factors. A feature having loading  $> |0.55|$  was assigned to only one of the factors based on the factor loadings.

The Bayesian confirmatory factor analysis (BCFA) is an alternative to frequentist CFA generating an important role in the assessment of the reliability and validity of latent variables. We fitted BCFA to estimate the factor scores according to the phenotype-factor structure inferred from the earlier EFA step. BCFA was applied to concatenated data, including phenotypes, proteins, miRNA, and the hub genes of modules obtained from WGCNA. Briefly, the *blavaan* R package (Merkle and Rosseel, 2018) was used with three Markov Monte Carlo chains, each with 6,000 Gibbs samples after 6,000 burn-in. Then, the posterior means of the factor scores of latent variables were estimated and treated as the new phenotypes for further analysis.

## Bayesian network

In the Bayesian network (BN), a direct acyclic graph is generated, and each random variable is associated with a node, the edges represent conditional dependency between variables, whereas the absence of an edge implies that the variables are conditionally independent of other variables (Choi, 2015). The details of BN procedures can be found in more detail in Yu et al. (2019) and Momen et al. (2021). Briefly, the BN structure learning with the *bnlearn* R package (Scutari, 2010) was applied to study the probabilistic relationships among the omic and latent variables. The BN is given by

$$BN = (G, X_V),$$

where  $G$  represents a direct acyclic graph composed of nodes ( $V$ ) connected by edges ( $E$ ), describing the probabilistic relationships and the vector  $X_V = (X_1, \dots, X_k)$  where  $k$  is the random variable (Yu et al., 2019). The joint probability of distributions is therefore given by

$$P(X_V) = \prod_{v=1}^k P(X_v | P_a(X_v)),$$

where  $P_a(X_v)$  expresses a set of parent nodes of  $X_v$ . The score-based (hill climbing and tabu) and hybrid algorithms (max–min hill climbing and general 2-phase restricted maximization) were used to perform structure learning (Scutari, 2010). Candidate networks were compared based on the Bayesian information criterion (BIC) and Bayesian Gaussian equivalent score (BGe). The BIC score was calculated as a criterion for the selection of the candidate model, and BGe reflects the posterior probability of the networks. A larger BIC score is preferred since it is rescaled by  $-2$  in the *bnlearn* R package. In addition, 1,000 bootstrapping replicates were used to estimate the uncertainty of the edge's strength and the direction of the network. Edges showing presence in at least 80% (strength) among all the 1,000 models were kept in the BN through model averaging.

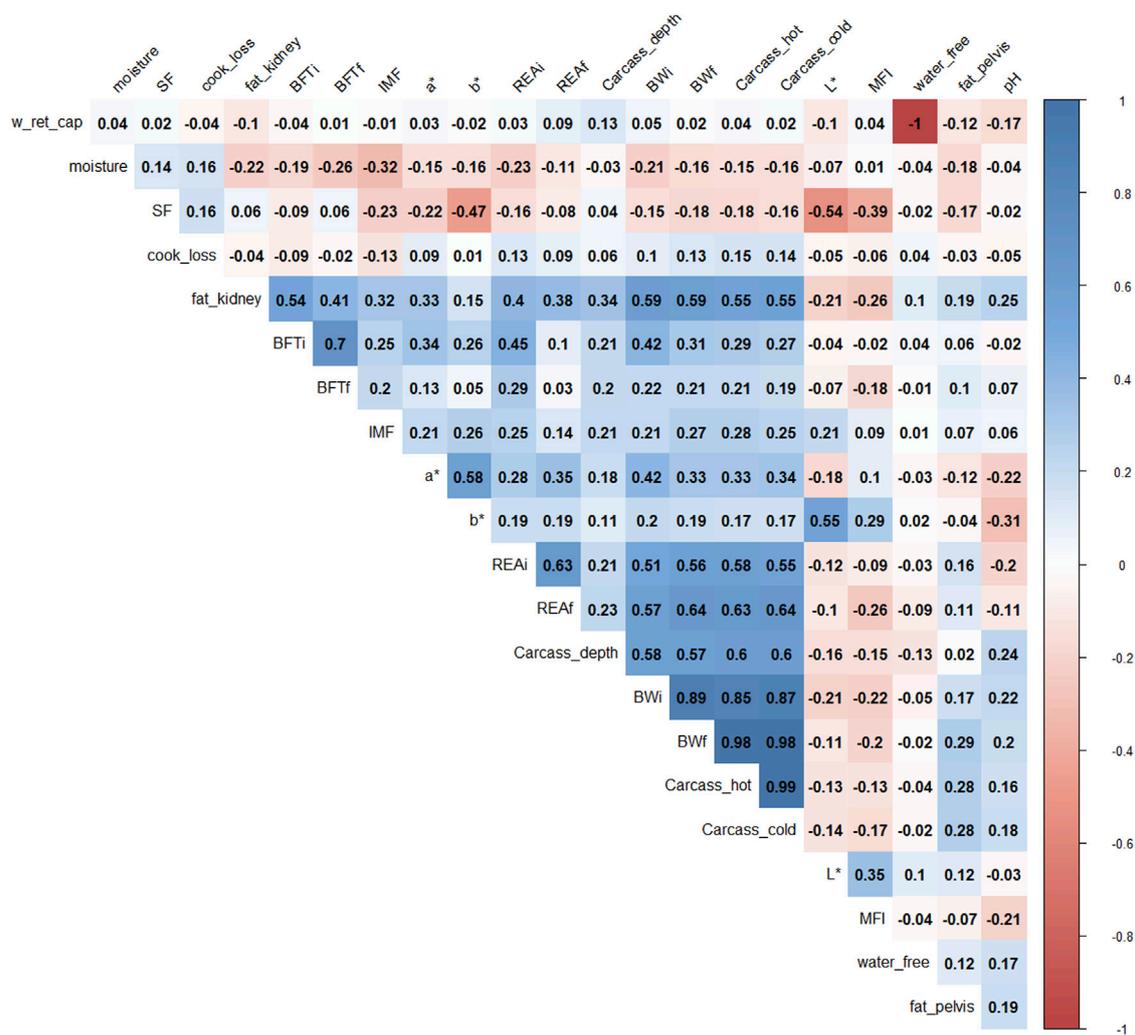
## Results

### Data preprocessing for analysis

In this study, we investigated the effective application of FA and BN framework to generate networks with biological meaning on for three different phenotypic categories: 1) production trait category included pre-feedlot body weight (BW<sub>i</sub>), post-feedlot body weight (BW<sub>f</sub>), initial backfat thickness (BFT<sub>i</sub>), and initial ribeye area (REAI); 2) carcass trait category included final backfat thickness (BFT<sub>f</sub>), final ribeye area (REAF), hot carcass weight (carcass\_hot), cold carcass weight (carcass\_cold), carcass depth (carcass\_depth), kidney fat content (fat\_kidney), and pelvis fat content (fat\_pelvis); and 3) meat quality category included the shear force at 24 h (SF), pH at 24 h (pH), meat moisture (moisture), free water (water\_free), water-holding capacity (w\_ret\_cap), cooking weight loss (cook\_loss), color parameters (L\*, a\*, and b\*), myofibrillar fragmentation index (MFI), and intramuscular fat (IMF) along with three different omic datasets: 1) mRNA sequencing, 2) miRNA sequencing, and 3) protein abundance.

Pearson's correlations (Figure 1) showed that BW<sub>f</sub>, carcass\_hot, and water\_free were highly correlated with carcass\_cold and water-holding capacity (correlation  $> 0.9$ ), therefore; they were removed for further analysis to avoid duplicate information. For example, the correlation between water\_free and w\_ret\_cap was  $-1$  because both traits represent oppositional and complementary information (Pearce et al., 2011).

For the RNA-Seq (mRNA) data, after the quality control and filtering procedure, 13,023 genes were included in WGCNA. The WGCNA method identified 20 modules, and two modules (mRNA1 and mRNA2) showed module membership (MM)  $> 0.95$  and gene significance  $p$ -value  $<$



**FIGURE 1**

Correlation plot of 22 phenotypes. The degree of shading and the value reported correspond to the correlations among the traits. BWi: pre-feedlot body weight; REAi: initial ribeye area by ultrasonography; REAf: final ribeye area on steak; BFTi: initial backfat thickness by ultrasonography; BFTf: final backfat thickness by carcass; fat\_pelvis: pelvis fat content at carcass; fat\_kidney: kidney fat content; carcass\_hot: hot carcass weight; carcass\_cold: cold carcass weight; carcass\_depth: carcass depth; pH: pH at 24 h; water\_free: free water; w\_ret\_cap: water holding capacity; moisture: meat moisture; SF: shear-force; MFI: miofibrillar fragmentation index; L\*, a\*, b\*: color parameters; and IMF: intramuscular fat.

0.001. The mRNA1 module was composed of seven hub genes and the mRNA2 module of four hub genes (Supplementary Table S1).

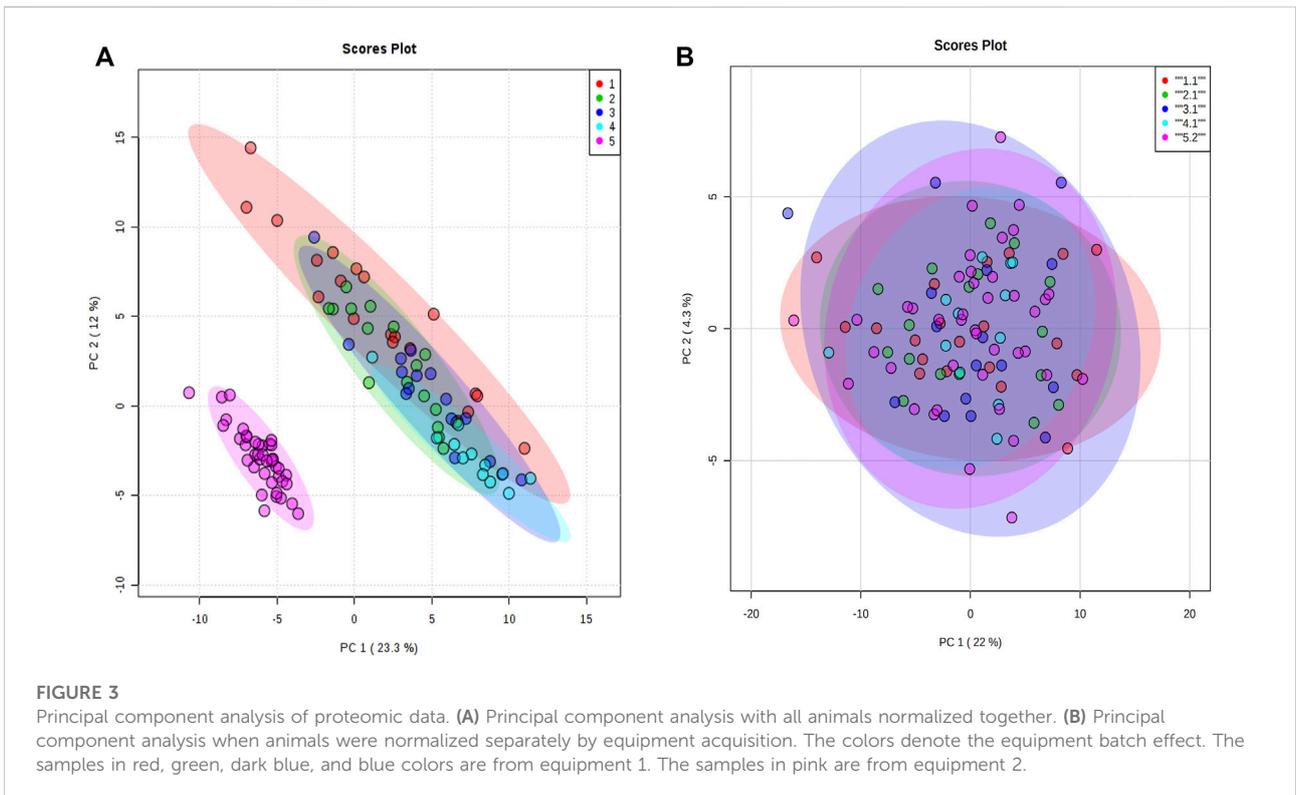
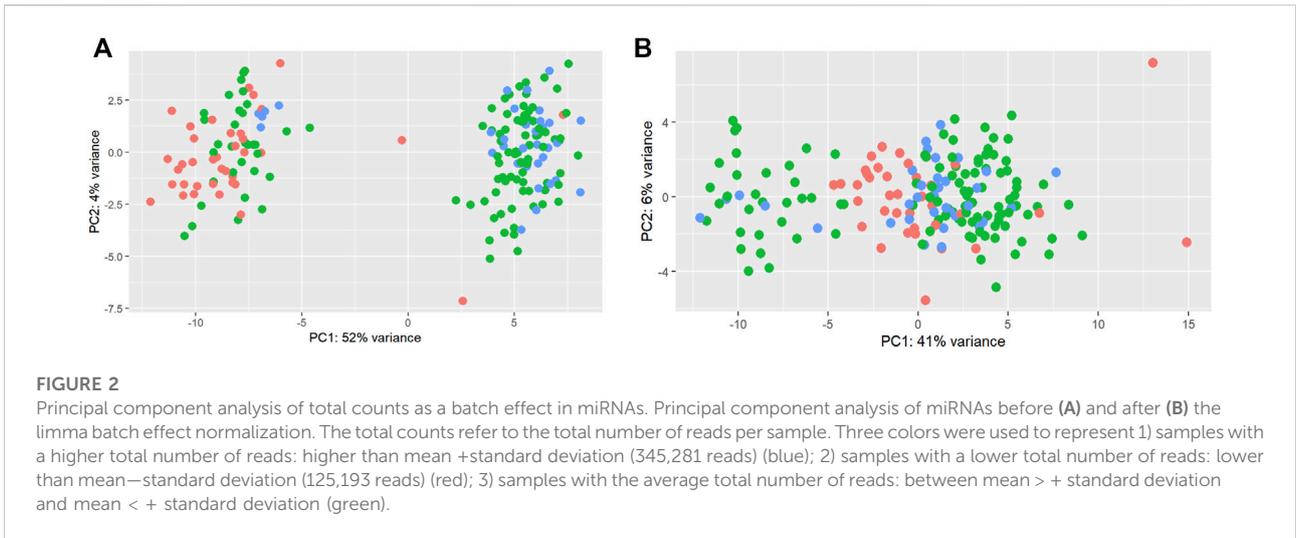
A total of 192 miRNAs were used for further analysis after the preprocessing steps. One animal was excluded as an outlier. After normalization, limma was used to identify a batch effect (Figure 2). Principal component analysis (PCA) revealed clusters based on the total counts of samples (Figure 2A). limma was applied to remove the batch effect in the miRNA data for further analysis (Figure 2B).

For proteomic data, 159 proteins from 106 animals were used in the analysis after the quality control steps. PCA was applied and a batch effect due to the equipment used was identified

(Figure 3A). The batch effect was accounted for by normalizing every data separately (Figure 3B).

### Exploratory and Bayesian confirmatory factor analysis

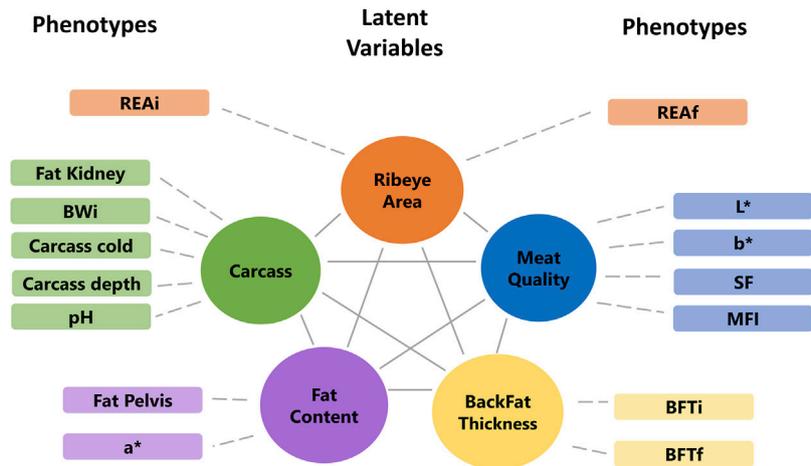
The factor analyses were performed using a subset of 102 animals that have phenotypes, miRNA, mRNA, and protein data. First, the phenotypes, miRNA, and protein data were used individually to fit an exploratory factor analysis (EFA). EFA can reduce data dimension without any prior assumptions about the observed data and latent factors structures. The parallel analysis suggested that phenotypes,



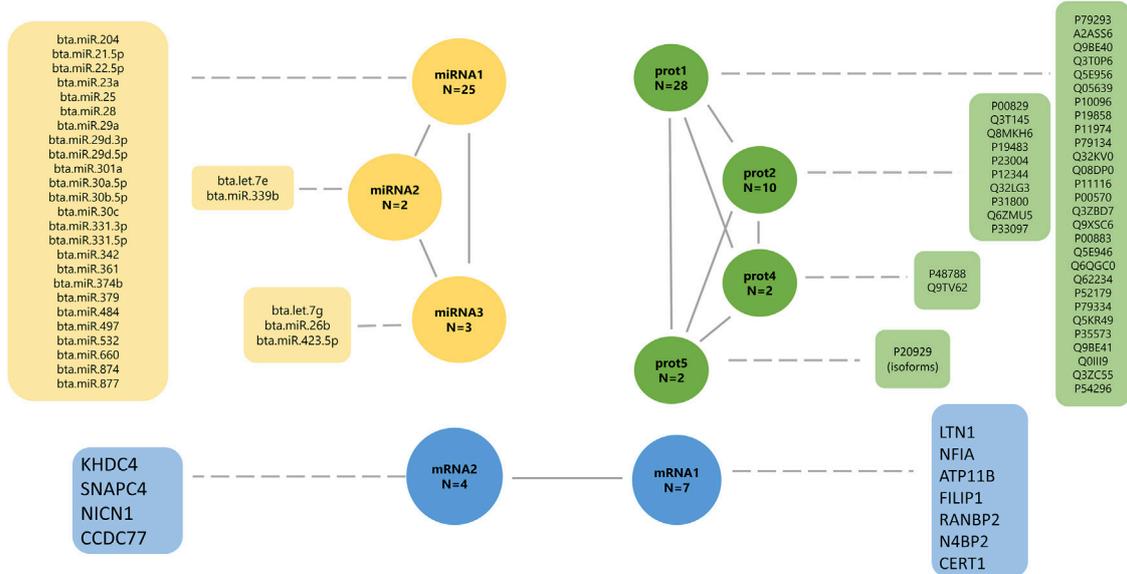
miRNA, and protein data were composed of five, ten, and eight latent variables, respectively. Each omic dataset was assigned to a factor according to the highest loading value ( $>|0.5|$ ), filtering some latent variables composed of a few features. The final underlying latent structures from EFA of the phenotype, miRNA, and protein data are shown in Figures 4, 5.

The BCFA was used to estimate factor loadings and scores based on the structure obtained from the EFA analysis, assuming that these latent variables determine the observed phenotypes and molecular profile levels (Supplementary Tables S2, S3).

The five phenotype latent factors showed strong contributions to the observed phenotypes, with



**FIGURE 4**  
Final underlying latent structures of phenotypes generated by exploratory factor analysis. BWi: initial body weight on feedlot trial; REAi: initial ribeye area by ultrasonography; REAf: final ribeye area on steak; BFTi: initial backfat thickness by ultrasonography; BFTf: final backfat thickness by carcass; SF: shear-force; MFI: myofibrillar fragmentation index; and L\*, a\*, b\*: color parameters.



**FIGURE 5**  
Final underlying latent structures of miRNA (yellow), proteins (green), and mRNA (blue) generated by exploratory factor analysis. N denotes the total number of features in each latent variable.

standardized regression coefficients ranging from 0.989 to 0.986 for backfat thickness, -0.993 to 0.956 for meat quality, 0.654 to 1 for the carcass, 0.942 to 0.992 for fat content, and 0.973 to 0.991 for ribeye area. The seven latent variables for miRNA and protein also showed strong contributions to the molecular level profiles, with

standardized regression coefficients ranging from -0.999 to 0.999 for factor mirna1 (miRNA), -0.971 to 0.979 for factor mirna2 (miRNA), -0.914 to 0.989 for factor mirna3 (miRNA), 0.842 to 0.990 for factor prot1 (protein), 0.774 to 0.973 for factor prot2 (protein), 0.963 to 0.997 for factor prot4 (protein), and 0.976 to 0.990 for factor prot5 (protein).

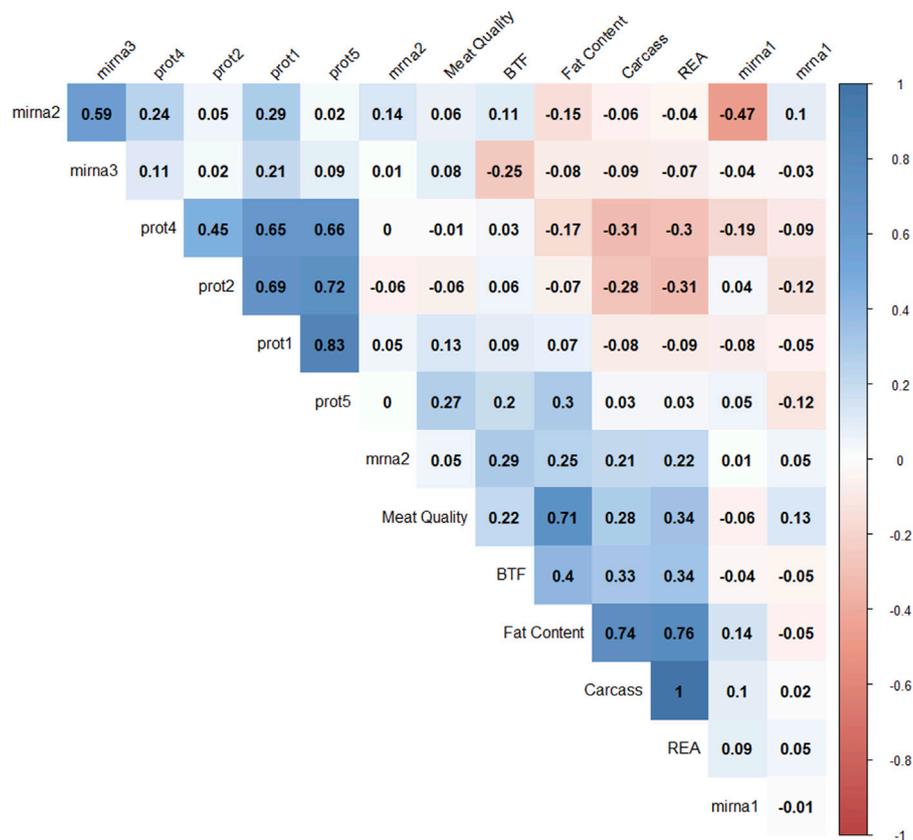


FIGURE 6

Correlation plot of 14 factor scores. The degree of shading and the value reported correspond to the correlation between each pair of latent variables.

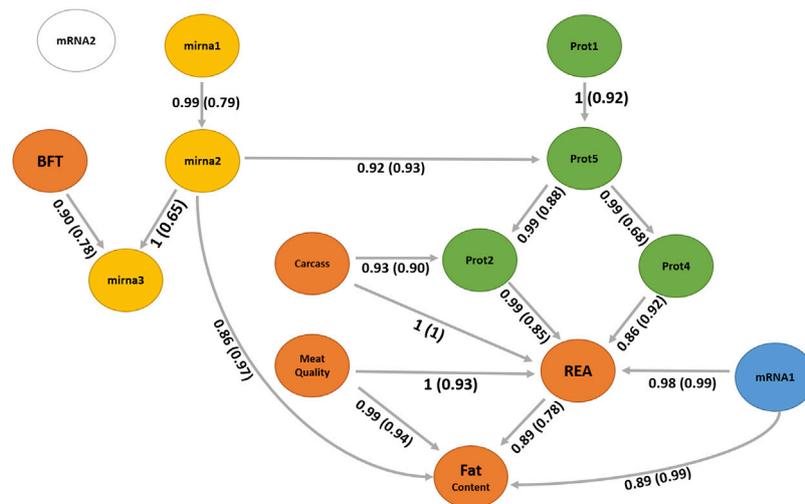
The latent factor backfat thickness (BFT) had a positive contribution to BFT<sub>i</sub> and BFT<sub>f</sub> (0.989 and 0.986, respectively; [Supplementary Table S2](#)), indicating that larger values for the latent factor can be interpreted as a greater thickness on the backfat content. The latent factor meat quality has a positive contribution to shear force (0.956; [Supplementary Table S2](#)), and a negative contribution to the colors b\*, L\*, and MFI (-0.993, -0.959, and -0.785, respectively) indicating that lower values on the latent factor can be interpreted as more tender meat. The latent factor carcass showed the largest positive contributions to traits describing carcass (e.g., weight to carcass cold, 1; weight to carcass depth, 0.990; weight to the kidney's fat content, 0.987; and pH of meat at 24 h, 0.863), suggesting that this latent factor is an overall representation of carcass. The latent factor ribeye area (REA) has a strong positive contribution to the REA<sub>f</sub> and REA<sub>i</sub> (0.991 and 0.973, respectively; [Supplementary Table S2](#)), indicating that larger values for the latent factor can be interpreted as a greater ribeye area.

The latent factor mirna1 has a positive contribution to 18 miRNAs (0.876–0.999; [Supplementary Table S2](#)), and a

negative contribution to seven miRNAs (-0.995 to -0.999, respectively). The mirna2 latent variable has a positive contribution to miRNA “bta.let.7e” (0.979; [Supplementary Table S2](#)), and a negative contribution to miRNA “bta.miR.339b” (-0.971, respectively; [Supplementary Table S2](#)). The latent factor mirna3 has a positive contribution of two miRNAs, “bta.let.7g” (0.889) and “bta.miR.26b” (0.987), and a negative contribution to miRNA “bta.miR.423.5p” (-0.914). The latent factors prot1, prot2, prot4, and prot5 have a positive contribution to all proteins, including 28 proteins (0.842–0.990), 10 proteins (0.774–0.973), two proteins (0.976–0.997), and two proteins (0.976–0.990), respectively.

## Correlation among latent variables

Pearson correlation coefficients were calculated to understand the relationships among latent variables ([Figure 6](#)). Negative correlations were observed between mirna1 and mirna2 (-0.47), REA and prot4 (-0.33), REA and prot2 (-0.3), carcass and prot4



**FIGURE 7**

Bayesian network between latent variables based on the score-based (hill climbing and tabu) algorithms. The quality of the structure was evaluated by bootstrap resampling and model averaging across 1,000 replications. Orange nodes: phenotype latent variables; yellow nodes: miRNA latent variables; green nodes: protein latent variables; blue node: gene expression of mRNA1 module (WGCNA); and white node: gene expression of mRNA2 module (WGCNA). The labels of the arcs correspond to the strength and direction (in parenthesis).

(−0.31), carcass and prot2 (−0.28), and BFT and mirna3 (−0.25). Positive correlations are observed between all protein factors; meat quality and fat content (0.71), fat content and carcass (0.74), fat content and REA (0.76), and mirna2 and mirna3 (0.59). The latent variables REA and carcass correlated at 0.996. These results suggest that protein levels might have a negative impact on carcass, REA, and fat content factors.

## Bayesian network

A BN was used to infer the interrelationships between latent variables. The BN algorithm learned with the most favorable network score in terms of BIC (1801.31) and BGe (1903.46) was the score-based hill climbing algorithm (Figure 7). The structure of BN was refined by model averaging with 1,000 networks from bootstrap resampling to reduce the impact of local optimal structures. The labels of the arcs measure the percentage of the uncertainty, corresponding to strength and direction (in parenthesis). The strength measures the frequency of the arc presented among all 1,000 networks from the bootstrapping replicates and the direction is the frequency of the direction shown conditionally in the presence of the arc.

We observed no difference in the structures between the two score-based algorithms used, the hill climbing and tabu. The two score-based algorithms produced a greater number of edges than the hybrid algorithms. The hill climbing algorithm produced 17 directed connections from the 14 latent variables.

## Discussion

We integrated a multi-omic dataset with production, carcass, and meat quality traits and explored non-conventional relationships that led to new hypotheses in the meat quality field. Here, we applied EFA, BCFA, and BN to infer interrelationships among latent variables underlying complex traits and omic data. First, EFA and BCFA were used to reduce the dimensions of datasets by constructing latent variables and estimating their factor scores (de los Campos and Gianola, 2007). These latent variables represent more straightforward biological meanings than the original features measured in a population (Yu et al., 2019). Then, we applied a BN to understand the interrelationships among the latent variables (Neapolitan and others, 2004). We generated a network with 14 latent variables involving 17 directed connections. Moreover, this approach elucidated both direct and indirect relationships among latent variables. However, a precaution is essential to interpret the network as a causal relationship because causal statements require more assumptions (Pearl, 2009).

Yu et al. (2019) and Momen et al. (2021) applied a similar approach to obtain genetic insights on rice and wheat complex traits. Leal-Gutiérrez et al. (2018) studied the potential of using latent variables, obtained by structural equation analysis, on carcass and meat quality traits in beef cattle. They reduced the complexity of the data and reported biological mechanisms, such as postmortem proteolysis of structural proteins and cellular compartmentalization, cellular proliferation and

differentiation of adipocytes, and fat deposition. In recent work, Yu, et al. (2020) applied factor analysis to beef cattle behavior to better understand latent factors underlying temperament traits.

## Biological meaning of latent variables and their relationships

The latent variable for the carcass, mainly composed of carcass cold weight, carcass depth, and fat kidney content (Supplementary Table S2), can be interpreted as the overall representation of the carcass, a higher value indicates a larger and heavier carcass. Its direct and indirect relationships with the latent variable REA (Figure 7) suggest the positive impact of the carcass yield on the ribeye area. Aass (1996) reported a positive phenotypic correlation (0.26) between the carcass depth and ribeye area corroborating our findings. Dinkel and Busch (1973) also estimated a positive genetic correlation (0.32) between growth rate and carcass yield, impacting the ribeye area positively.

The carcass has a relationship with the latent variable prot2 that also impacts REA. The latent variable prot2 is composed of 10 proteins (Supplementary Tables S2, S4), including UQCRC2 related with proteolysis (GO:0006508); ATP5F1A and ATP5F1B related with ATP synthesis coupled proton transport (GO:0042776 and GO:0015986); TNNT1 and TRIM72 related with muscle contraction (GO:0006936), regulation of muscle contraction (GO:0006937), sarcomere organization (GO:0045214), and muscle organ development (GO:0007517); GOT1 and GOT2 related with aspartate biosynthetic and catabolic processes (GO:0006532, GO:0006533), cellular response to insulin stimulus (GO:0032869), fatty acid homeostasis (GO:0055089), glutamate catabolic process to aspartate (GO:0019550), glycerol biosynthetic process (GO:0006114), and oxaloacetate metabolic process (GO:0006107); and MDH1 and MDH2 related with the carbohydrate metabolic process (GO:0005975), malate metabolic process (GO:0006108), NADH metabolic process (GO:0006734), oxaloacetate metabolic process (GO:0006107), tricarboxylic acid cycle (GO:0006099), and aerobic respiration (GO:0009060).

The ATP synthase F (0) complex subunit B1 (ATP5F1) has been positively correlated with meat color parameter a\*, which impacts meat discoloration (Yu et al., 2017). Our findings show an indirect relationship between prot2 and the fat content latent variable that includes the parameter a\*. Although prot2 and meat quality are not directly connected (Figure 7), both impact REA. However, prot2 is mainly composed of enzymes involved with energy metabolism that have been reported as putative candidate proteins for meat tenderness. The aspartate aminotransferase (GOT1) has been considered a putative candidate protein usable for meat

tenderness prediction (Boudon et al., 2020). Rodrigues et al. (2017) reported that Nellore cattle have a higher abundance of malate dehydrogenase (MDH1) compared to Angus. This enzyme is important in gluconeogenesis, catalyzes the oxidation of malate to oxaloacetate, and is a relevant player in meat quality characteristics because this enzyme is involved in energy metabolism and affects how pH drops, changing the conversion of muscle to meat (Rodrigues et al., 2017). The ubiquinol-cytochrome C reductase core protein 2 (UQCRC2) gene, which is an important energy promoter for the development of cell functions was reported as up-regulated in a study analyzing gene expression on tough beef groups compared to the tender group in Nellore cattle (Muniz et al., 2021). The degradation of troponin T1 (TNNT1) proteins during post-mortem has been associated with meat tenderness (Zakrys-Waliwander et al., 2012; Contreras-Castillo et al., 2016; Wright et al., 2018).

The prot4 latent variable also shows a relationship with REA, which has two proteins (Supplementary Tables S2, S4) and includes the TNNI2 and MYH4. Troponin I, fast-twitch isoform (TNNI2) is a subunit of the troponin complex and plays a role in calcium regulation during muscle contraction and relaxation. The *TNNI2* gene was associated with pH, meat color value, and intramuscular fat content in pigs (Yang et al., 2010). The myosin heavy chains are relevant to muscle contraction velocity and power, MYH4 is one of the isoforms associated with IIB fibers types (Cho et al., 2016) and myotube hypertrophy in beef cattle (Bordbar et al., 2020). Our findings suggest new hypotheses of the impact of these proteins of prot2 could affect the REA and carcass traits.

The latent variable meat quality composed of shear force, myofibrillar fragmentation index, and the color parameters L\* and b\* can be interpreted as the overall representation of meat tenderness, and lower levels of this factor indicate more tender meat. It has a direct relationship with the latent variable REA and fat content. The relationship among tenderness, REA, and fat content has been discussed in the literature (Dinkel and Busch, 1973; Bonin et al., 2020). The mRNA1 latent variable has a relationship with REA and fat content. The mRNA1 factor is composed of the genes *LTN1*, *NFIA*, *ATP11B*, *FILIP1*, *RANBP2*, *N4BP2*, and *CERT1*. The nuclear factor IA gene (*NFIA*) has been studied indicating the potential to stimulate lipid accumulation in cattle (Chen H.-J. et al., 2018). According to the enrichment analysis (Supplementary Tables S5), these genes have been associated with an important cholesterol pathway called cholesterol and sphingolipid transport. Examples are the *RANBP2* gene which is associated with proteolysis and the *CERT1* gene which is related to intracellular cholesterol transport and sphingolipid metabolism. A further investigation is necessary to understand these relationships with REA or fat content.

The latent variable mirna3 is a child node of BFT and mirna2. The miRNAs are small RNA molecules that inhibit

translation or induce degradation of protein-coding mRNAs that contain complementary sequences to miRNAs. mirna3 is constituted by three miRNAs, namely, bta. let.7g, bta. miR.26b, and bta. miR.423.5p. bta. let.7g was found in studies related with lactation and infection in cattle (Ma et al., 2019; Rani et al., 2020). mirna2 is composed of two miRNAs, namely, bta. let.7e and bta. miR.339b. Gu et al. (2007) identified the expression of bta. let.7e on adipose tissue in cattle. bta. miR.339b was found in studies related to fatty acid metabolism and lactation (Do et al., 2017; Palombo et al., 2018; Poleti et al., 2018). mirna2 has a direct relationship with fat content (Figure 7). Further studies are necessary to better understand the functions of mirna2 and its association with fat metabolism in beef cattle.

The latent variable prot5 is composed of two isoforms of nebulin (NEBU) which are important structural components involved in meat aging (Koochmaraie et al., 1984; Ouali et al., 1995). Post-mortem degradation of nebulin has been associated with meat tenderness in cattle in which animals with a lower degradation have less tender meat (Anderson and Parrish, 1989; Wu et al., 2014). The prot5 latent variable is an important node that has relationships with prot2 and prot4 and has an indirect relationship with REA and fat content.

The generated network identified interomic relationships, bringing simplicity without losing complexity. This is one of the challenges found in studies of this nature. Often a methodology used ends up providing the interpretation of unfeasible results, which was not in our approach. Additional investigations are essential to understand the relationships of molecules and phenotypes on latent variables REA, prot2, prot4, prot5, mRNA1, carcass, mirna3, mirna2, and fat content. The network demonstrated a relationship between miRNAs and nebulin protein isoforms that will not be found in studies using single or multi-level omics. Finally, REA appears as a central node in the network, influenced by carcass, prot2, prot4, and meat quality, suggesting that REA is a good indicator phenotype for meat quality because it can be easily measured during slaughter or by ultrasonography.

## Conclusion

The FA identified latent variables, decreasing the dimensionality and complexity of data. The BN analysis was capable of identifying interrelationships among latent variables from different types of data, allowing the integration of different types of omic data and complex traits. The EFA, BCFA, and BN approaches can be used to generate new hypotheses on molecular research in the meat quality area, by integrating different types of data and exploring non-conventional relations.

## Data availability statement

The mRNA datasets supporting the conclusion of this article are available in the European Nucleotide Archive (ENA) repository (EMBL-EBI), under accession nos. PRJEB13188, PRJEB10898, PRJEB15314, and PRJEB19421 (<https://www.ebi.ac.uk/ena/browser/view/>). The miRNA dataset of this article is available in the European Nucleotide Archive (ENA) repository (EMBL-EBI), under accession no. PRJEB42280. The protein data presented in the study are publicly available. These data can be found at: <https://doi.org/10.1016/j.dib.2018.06.004>. Any other relevant data are available from the authors upon reasonable request.

## Ethics statement

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee Guidelines from EMBRAPA (CEUA 01/2013).

## Author contributions

FN performed the bioinformatics and data analysis and drafted the manuscript. LC and GM conceived and conducted this study and supervised FN in analytical and data analysis and revised the manuscript. LR and LC designed the experimental study. AC and MP conducted the sample collection and the omic extraction protocols. MM and HY contributed to the data analysis. BP contributed to the functional enrichment analysis. All the authors have read and approved the final manuscript.

## Funding

The experiments were financially supported by “Fundação de Amparo à Pesquisa do Estado de São Paulo,” Brazil (FAPESP grants 12/23638-8 and 19/04089-2). This work was conducted while FJN was visiting Virginia Polytechnic Institute and State University, supported by the Coordinating Agency for Advanced Training of Graduate Personnel (CAPES grant: 88887.367967/2019-00).

## Acknowledgments

LC, GM, and LR are recipients of CNPq productivity scholarships (304353/2019-1, 310714/2020-6, and 303754/2016-8, respectively).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.948240/full#supplementary-material>

## References

- Aass, L. (1996). Variation in carcass and meat quality traits and their relations to growth in dual purpose cattle. *Livest. Prod. Sci.* 46, 1–12. doi:10.1016/0301-6226(96)00005-X
- Adzitey, F. (2011). Effect of pre-slaughter animal handling on carcass and meat quality. *Int. Food Res. J.* 18.
- Anderson, T. J., and Parrish, F. C. (1989). Postmortem degradation of titin and nebulin of beef steaks varying in tenderness. *J. Food Sci.* 54, 748–749. doi:10.1111/j.1365-2621.1989.tb04695.x
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*.
- Bonin, M. de N., da Luz e Silva, S., Bünger, L., Ross, D., Feijó, G. L. D., da Costa Gomes, R., et al. (2020). Predicting the shear value and intramuscular fat in meat from Nellore cattle using Vis-NIR spectroscopy. *Meat Sci.* 163, 108077. doi:10.1016/j.meatsci.2020.108077
- Bordbar, F., Jensen, J., Du, M., Abied, A., Guo, W., Xu, L., et al. (2020). Identification and validation of a novel candidate gene regulating net meat weight in Simmental beef cattle based on imputed next-generation sequencing. *Cell Prolif.* 53, e12870. doi:10.1111/cpr.12870
- Boudon, S., Ounaissi, D., Viala, D., Monteils, V., Picard, B., and Cassar-Malek, I. (2020). Label free shotgun proteomics for the identification of protein biomarkers for beef tenderness in muscle and plasma of heifers. *J. Proteomics* 217, 103685. doi:10.1016/j.jpropt.2020.103685
- Carvalho, M. E., Gasparin, G., Poleti, M. D., Rosa, A. F., Balieiro, J. C. C., Labate, C. A., et al. (2014). Heat shock and structural proteins associated with meat tenderness in Nellore beef cattle, a *Bos indicus* breed. *Meat Sci.* 96, 1318–1324. doi:10.1016/j.meatsci.2013.11.014
- Cerny, B. A., and Kaiser, H. F. (1977). A study of A measure of sampling adequacy for factor-analytic correlation matrices. *Multivar. Behav. Res.* 12, 43–47. doi:10.1207/s15327906mbr1201\_3
- Cesar, A. S. M., Regitano, L. C. A., Poleti, M. D., Andrade, S. C. S., Tizioto, P. C., Oliveira, P. S. N., et al. (2016). Differences in the skeletal muscle transcriptome profile associated with extreme values of fatty acids content. *BMC Genomics* 17, 961. doi:10.1186/s12864-016-3306-x
- Cesar, A. S., Regitano, L. C., Mourão, G. B., Tullio, R. R., Lanna, D. P., Nassu, R. T., et al. (2014). Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genet.* 15, 39. doi:10.1186/1471-2156-15-39
- Chen, H.-J., Ihara, T., Yoshioka, H., Itoyama, E., Kitamura, S., Nagase, H., et al. (2018a). Expression levels of brown/beige adipocyte-related genes in fat depots of vitamin A-restricted fattening cattle. *J. Anim. Sci.* 96, 3884–3896. doi:10.1093/jas/sky240
- Chen, Y., Mccarthy, D., Ritchie, M., Robinson, M., and Smyth, G. (2018b). *edgeR: differential expression analysis of digital gene expression data User's Guide*, 1–110.
- Cho, E., Lee, K., Kim, J., Lee, S., Jeon, H., Lee, S., et al. (2016). Association of a single nucleotide polymorphism in the 5' upstream region of the porcine myosin heavy chain 4 gene with meat quality traits in pigs. *Anim. Sci. J.* 87, 330–335. doi:10.1111/asj.12442
- Choi, T. (2015). Bayesian networks with examples in R. *Biometrics* 71, 864–865. doi:10.1111/biom.12369
- Contreras-Castillo, C. J., Lomiwes, D., Wu, G., Frost, D., and Farouk, M. M. (2016). The effect of electrical stimulation on post mortem myofibrillar protein degradation and small heat shock protein kinetics in bull beef. *Meat Sci.* 113, 65–72. doi:10.1016/j.meatsci.2015.11.012
- de Lima, A. O., Koltes, J. E., Diniz, W. J. S., de Oliveira, P. S. N., Cesar, A. S. M., Tizioto, P. C., et al. (2020). Potential biomarkers for feed efficiency-related traits in nelore cattle identified by Co-expression network and integrative genomics analyses. *Front. Genet.* 11, 189. doi:10.3389/fgene.2020.00189
- de los Campos, G., and Gianola, D. (2007). Factor analysis models for structuring covariance matrices of additive genetic effects: A bayesian implementation. *Genet. Sel. Evol.* 39, 481–494. doi:10.1186/1297-9686-39-5-481
- Dinkel, C. A., and Busch, D. A. (1973). Genetic parameters among production, carcass composition and carcass quality traits of beef cattle. *J. Anim. Sci.* 36, 832–846. doi:10.2527/jas1973.365832x
- Do, D. N., Li, R., Dudemaine, P.-L., and Ibeagha-Awemu, E. M. (2017). MicroRNA roles in signalling during lactation: An insight from differential expression, time course and pathway analyses of deep sequence data. *Sci. Rep.* 7, 44605. doi:10.1038/srep44605
- Dobin, A., and Gingeras, T. R. (2015). Mapping RNA-seq reads with STAR. *Curr. Protoc. Bioinforma.* 51, 1–11. doi:10.1002/0471250953.bi1114s51
- Filho, L. (2000). *Pecuária da carne bovina*. first edit. São Paulo: Embrapa.
- Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40, 37–52. doi:10.1093/nar/gkr688
- Geromanos, S. J., Vissers, J. P. C., Silva, J. C., Dorschel, C. A., Li, G.-Z., Gorenstein, M. V., et al. (2009). The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* 9, 1683–1695. doi:10.1002/pmic.200800562
- Gilar, M., Olivova, P., Daly, A. E., and Gebler, J. C. (2005). Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J. Sep. Sci.* 28, 1694–1703. doi:10.1002/jssc.200500116
- Gu, Z., Eleswarapu, S., and Jiang, H. (2007). Identification and characterization of microRNAs from the bovine adipose tissue and mammary gland. *FEBS Lett.* 581, 981–988. doi:10.1016/j.febslet.2007.01.081
- Guerrero, A., Valero, M. V., Campo, M. M., and Sañudo, C. (2013). Some factors that affect ruminant meat quality: From the farm to the fork. Review. *Acta Sci. Anim. Sci.* 35. doi:10.4025/actascianimsci.v35i4.21756
- Hopkins, D., Littlefield, P., and Thompson, J. (2000). A research note on factors affecting the determination of myofibrillar fragmentation. *Meat Sci.* 56, 19–22. doi:10.1016/S0309-1740(00)00012-7
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* 30, 179–185. doi:10.1007/BF02289447
- Horwitz, W. (2000). *Official methods of analysis of AOAC International*. 17th ed. Gaithersburg, Md.
- Huang, S., Chaudhary, K., and Garmire, L. X. (2017). More is better: Recent progress in multi-omics data integration methods. *Front. Genet.* 8, 84–12. doi:10.3389/fgene.2017.00084
- Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: Systems biology. *Annu. Rev. Genomics Hum. Genet.* 2, 343–372. doi:10.1146/annurev.genom.2.1.343
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200. doi:10.1007/BF02289233

- Kappeler, B. I. G., Regitano, L. C. A., Poleti, M. D., Cesar, A. S. M., Moreira, G. C. M., Gasparin, G., et al. (2019). miRNAs differentially expressed in skeletal muscle of animals with divergent estimated breeding values for beef tenderness. *BMC Mol. Biol.* 20, 1. doi:10.1186/s12867-018-0118-3
- Koohmaraie, M., Kennick, W. H., Elgasim, E. A., and Anglemier, A. F. (1984). Effects of postmortem storage on muscle protein degradation: Analysis by SDS-polyacrylamide gel electrophoresis. *J. Food Sci.* 49, 292–293. doi:10.1111/j.1365-2621.1984.tb13732.x
- Kuhn, M. (2008). Building predictive models in R using the caret package. *J. Stat. Softw.* 28. doi:10.18637/jss.v028.i05
- Lalli, P. M., Corilo, Y. E., Fasciotti, M., Riccio, M. F., de Sa, G. F., Daroda, R. J., et al. (2013). Baseline resolution of isomers by traveling wave ion mobility mass spectrometry: Investigating the effects of polarizable drift gases and ionic charge distribution. *J. Mass Spectrom.* 48, 989–997. doi:10.1002/jms.3245
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinform.* 9, 559. doi:10.1186/1471-2105-9-559
- Leal-Gutiérrez, J. D., Rezende, F. M., Elzo, M. A., Johnson, D., Peñagaricano, F., and Mateescu, R. G. (2018). Structural equation modeling and whole-genome scans uncover chromosome regions and enriched pathways for carcass and meat quality in beef. *Front. Genet.* 9, 532. doi:10.3389/fgene.2018.00532
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8
- Ma, S., Tong, C., Ibeagha-Awemu, E. M., and Zhao, X. (2019). Identification and characterization of differentially expressed exosomal microRNAs in bovine milk infected with *Staphylococcus aureus*. *BMC Genomics* 20, 934. doi:10.1186/s12864-019-6338-1
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10. doi:10.14806/ej.17.1.200
- Meng, C., Zeleznik, O. A., Thallinger, G. G., Kuster, B., Gholami, A. M., and Culhane, A. C. (2016). Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief. Bioinform.* 17, 628–641. doi:10.1093/bib/bbv108
- Merkle, E. C., and Rosseel, Y. (2018). Blavaan: Bayesian structural equation models via parameter expansion. *J. Stat. Softw.* 85. doi:10.18637/jss.v085.i04
- MetaCore (2021). MetaCore login | clarivate. Available at: <https://portal.genego.com/> (Accessed June 8, 2021).
- Misra, B. B., Langefeld, C., Olivier, M., and Cox, L. A. (2018). Integrated omics: Tools, advances and future approaches. *J. Mol. Endocrinol.* 62, R21–R45. doi:10.1530/jme-18-0055
- Momen, M., Bhatta, M., Hussain, W., Yu, H., and Morota, G. (2021). Modeling multiple phenotypes in wheat using data-driven genomic exploratory factor analysis and Bayesian network learning. *Plant Direct* 5, e00304. doi:10.1002/pld3.304
- Muniz, M. M. M., Fonseca, L. F. S., dos Santos Silva, D. B., de Oliveira, H. R., Baldi, F., Chardulo, A. L., et al. (2021). Identification of novel mRNA isoforms associated with meat tenderness using RNA sequencing data in beef cattle. *Meat Sci.* 173, 108378. doi:10.1016/j.meatsci.2020.108378
- Nascimento, M. L., Souza, A. R. D. L., Chaves, A. S., Cesar, A. S. M., Tullio, R. R., Medeiros, S. R., et al. (2016). Feed efficiency indexes and their relationships with carcass, non-carcass and meat quality traits in Nellore steers. *Meat Sci.* 116, 78–85. doi:10.1016/j.meatsci.2016.01.012
- Neapolitan, R. E. (2004). *Learning bayesian networks*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Njisane, Y. Z., and Muchenje, V. (2016). Farm to abattoir conditions, animal factors and their subsequent effects on cattle behavioural responses and beef quality — a review. *Asian-Australas. J. Anim. Sci.* 30, 755–764. doi:10.5713/ajas.16.0037
- Novais, F. J., Pires, P. R. L., Alexandre, P. A., Dromms, R. A., Iglesias, A. H., Ferraz, J. B. S., et al. (2019). Identification of a metabolomic signature associated with feed efficiency in beef cattle. *BMC Genomics* 20, 8. doi:10.1186/s12864-018-5406-2
- Ouali, A., Demeyer, D., and Smulders, F. (1995). "Editorial," in *Tissue proteinases and regulation of protein degradation as related to meat quality* (Nijmegen: ECCEAMST), V–VI.
- Palombo, V., Milanese, M., Sgorlon, S., Capomaccio, S., Mele, M., Nicolazzi, E., et al. (2018). Genome-wide association study of milk fatty acid composition in Italian Simmental and Italian Holstein cows using single nucleotide polymorphism arrays. *J. Dairy Sci.* 101, 11004–11019. doi:10.3168/jds.2018-14413
- Pearce, K. L., Rosenvold, K., Andersen, H. J., and Hopkins, D. L. (2011). Water distribution and mobility in meat during the conversion of muscle to meat and ageing and the impacts on fresh meat quality attributes — a review. *Meat Sci.* 89, 111–124. doi:10.1016/j.meatsci.2011.04.007
- Pearl, J. (2009). *Causality*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511803161
- Poleti, M. D., Regitano, L. C. A., Souza, G. H. M. F., Cesar, A. S. M., Simas, R. C., Silva-Vignato, B., et al. (2018). Longissimus dorsi muscle label-free quantitative proteomic reveals biological mechanisms associated with intramuscular fat deposition. *J. Proteomics* 179, 30–41. doi:10.1016/j.jprot.2018.02.028
- Rani, P., Onteru, S. K., and Singh, D. (2020). Genome-wide profiling and analysis of microRNA expression in buffalo milk exosomes. *Food Biosci.* 38, 100769. doi:10.1016/j.fbio.2020.100769
- Revelle, W. R. (2017). *psych: Procedures for personality and psychological research*.
- Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., and Kim, D. (2015a). Methods of integrating data to uncover genotype–phenotype interactions. *Nat. Rev. Genet.* 16, 85–97. doi:10.1038/nrg3868
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015b). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Rodin, A. S., and Boerwinkle, E. (2005). Mining genetic epidemiology data with Bayesian networks I: Bayesian networks and example application (plasma apoE levels). *Bioinformatics* 21, 3273–3278. doi:10.1093/bioinformatics/bti505
- Rodrigues, R. T. de S., Chizzotti, M. L., Vital, C. E., Baracat-Pereira, M. C., Barros, E., Busato, K. C., et al. (2017). Differences in beef quality between angus (*Bos taurus taurus*) and Nellore (*Bos taurus indicus*) cattle through a proteomic and phosphoproteomic approach. *PLoS One* 12, e0170294. doi:10.1371/journal.pone.0170294
- Rust, S. R., Price, D. M., Subbiah, J., Kranzler, G., Hilton, G. G., Vanoverbeke, D. L., et al. (2008). Predicting beef tenderness using near-infrared spectroscopy. *J. Anim. Sci.* 86, 211–219. doi:10.2527/jas.2007-0084
- Scutari, M. (2010). Learning bayesian networks with thebnlearnRPackage. *J. Stat. Softw.* 35, 1–22. doi:10.18637/jss.v035.i03
- Silva, W. M., Carvalho, R. D., Soares, S. C., Bastos, I. F., Folador, E. L., Souza, G. H., et al. (2014). Label-free proteomic analysis to confirm the predicted proteome of *Corynebacterium pseudotuberculosis* under nitrosative stress mediated by nitric oxide. *BMC Genomics* 15, 1065. doi:10.1186/1471-2164-15-1065
- Souza, G. H. M. F., Guest, P. C., and Martins-de-Souza, D. (2017). LC-MSE, multiplex MS/MS, ion mobility, and label-free quantitation in clinical proteomics. *Methods Mol. Biol.* 1546, 57–73. doi:10.1007/978-1-4939-6730-8\_4
- Suravajhala, P., Kogelman, L. J. A., and Kadarmideen, H. N. (2016). Multi-omic data integration and analysis using systems genomics approaches: Methods and applications in animal production, health and welfare. *Genet. Sel. Evol.* 48, 38. doi:10.1186/s12711-016-0217-x
- Tizioto, P. C., Decker, J. E., Taylor, J. F., Schnabel, R. D., Mudadu, M. A., Silva, F. L., et al. (2013). Genome scan for meat quality traits in Nellore beef cattle. *Physiol. Genomics* 45, 1012–1020. doi:10.1152/physiolgenomics.00066.2013
- Wheeler, T. L., Cundiff, L. V., Shackelford, S. D., and Koohmaraie, M. (2005). Characterization of biological types of cattle (Cycle VII): Carcass, yield, and longissimus palatability traits. *J. Anim. Sci.* 83, 196–207. doi:10.2527/2005.831196x
- Widmann, P., Reverter, A., Fortes, M. R. S., Weikard, R., Suhre, K., Hammon, H., et al. (2013). A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics* 14, 798. doi:10.1186/1471-2164-14-798
- Wright, S. A., Ramos, P., Johnson, D. D., Scheffler, J. M., Elzo, M. A., Mateescu, R. G., et al. (2018). Brahman genetics influence muscle fiber properties, protein degradation, and tenderness in an Angus-Brahman multibreed herd. *Meat Sci.* 135, 84–93. doi:10.1016/j.meatsci.2017.09.006
- Wu, G., Farouk, M. M., Clerens, S., and Rosenvold, K. (2014). Effect of beef ultimate pH and large structural protein changes with aging on meat tenderness. *Meat Sci.* 98, 637–645. doi:10.1016/j.meatsci.2014.06.010
- Yang, H., Xu, Z. Y., Lei, M. G., Li, F. E., Deng, C. Y., Xiong, Y. Z., et al. (2010). Association of 3 polymorphisms in porcine troponin I genes (TNNI1 and TNNI2) with meat quality traits. *J. Appl. Genet.* 51, 51–57. doi:10.1007/BF03195710
- Yu, H., Campbell, M. T., Zhang, Q., Walia, H., and Morota, G. (2019). Genomic Bayesian confirmatory factor analysis and Bayesian network to characterize a wide spectrum of rice phenotypes. *G3* 9, 1975–1986. doi:10.1534/g3.119.400154
- Yu, H., Morota, G., Celestino, E. F., Dahlen, C. R., Wagner, S. A., Riley, D. G., et al. (2020). Deciphering cattle temperament measures derived from a four-platform standing scale using genetic factor Analytic modeling. *Front. Genet.* 11, 599. doi:10.3389/fgene.2020.00599
- Yu, Q., Wu, W., Tian, X., Hou, M., Dai, R., and Li, X. (2017). Unraveling proteome changes of Holstein beef M. semitendinosus and its relationship to meat discoloration during post-mortem storage analyzed by label-free mass spectrometry. *J. Proteomics* 154, 85–93. doi:10.1016/j.jprot.2016.12.012
- Yu, Y.-Q., Gilar, M., Lee, P. J., Bouvier, E. S. P., and Gebler, J. C. (2003). Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Anal. Chem.* 75, 6023–6028. doi:10.1021/ac0346196
- Zakrys-Waliwander, P. I., O'Sullivan, M. G., O'Neill, E. E., and Kerry, J. P. (2012). The effects of high oxygen modified atmosphere packaging on protein oxidation of bovine M. longissimus dorsi muscle during chilled storage. *Food Chem.* x, 131, 527–532. doi:10.1016/j.foodchem.2011.09.017