Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science

Embrapa

Documentos 125

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science

Editors

Luiz Lehmann Coutinho
Luciana C. de A. Regitano
Gerson Barreto Mourão
Aline Silva Mello Cesar
Bárbara Silva Vignato
Mirele Daiana Poleti
Wellison Jarles da Silva Diniz

Editors

Luiz Lehmann Coutinho

Engenheiro Agrônomo, professor da ESALQ/USP, Piracicaba, SP.

llcoutinho@usp.br

Luciana C.de A. Regitano

Médica Veterinária, pesquisadora, Embrapa Pecuária Sudeste, São Carlos, SP.

luciana.regitano@.embrapa.br

Gerson Barreto Mourão

Zootecnista, professor da ESALQ/USP, Piracicaba, SP.

gbmourao@usp.br

Aline Silva Mello Cesar

Zootecnista, pós-doutoranda da ESALQ/USP, Piracicaba, SP.

alinecesar@usp.br

Bárbara Silva Vignato

Zootecnista, doutora pela FZEA/USP, Pirassununga, SP.

bsilvavignato@gmail.com

Mirele Daiana Poleti

Biomédica, pós-doutoranda da ESALQ/USP, Piracicaba, SP.

mdpoleti@gmail.com

Wellison Jarles da Silva Diniz

Zootecnista, doutorando da Universidade Federal de São Carlos, São Carlos, SP.

wjarles09@gmail.com

Foreword

The integration of national and international researchers is extremely important for the dissemination and perpetuation of knowledge in several areas of scientific research. The "Ist workshop on omics strategies applied to livestock science" aimed to promote discussion and experience exchange on innovative subjects and methods in livestock, addressing issues such as omics technologies for system biology, animal breeding, and genomic data integration for complex traits like meat quality and feed efficiency. This event was organized by the partnership between the College of Agriculture "Luiz de Queiroz" (ESALQ/USP) and the Brazilian Agricultural Research Corporation (EMBRAPA) Southeast Livestock Center, coordinated by Prof. Dr. Luiz Lehmann Coutinho, Dr. Luciana C. A. Regitano, Prof. Dr. Gerson Barreto Mourão and Dr. Adhemar Zerlotini Neto. The event took place in the city of Piracicaba/SP on April 24th - 26th, 2017. To attend these goals national and international speakers from partner institutions recognized in their areas of activity were invited, which enabled fruitful discussions on these themes, prompting new ideas for future research.

Organizing Committee

Coordinators

Luiz Lehmann Coutinho	Escola Superior de Agricultura Luiz de Queiroz
Luciana C.de A. Regitano	Empresa Brasileira de Pesquisa Agropecuária
Gerson Barreto Mourão	Escola Superior de Agricultura Luiz de Queiroz
Adhemar Zerlotini Neto	Empresa Brasileira de Pesquisa Agropecuária

Technical Team

Aline S. M. Cesar	Escola Superior de Agricultura Luiz de Queiroz
Andressa Oliveira de Lima	Universidade Federal de São Carlos
Bárbara Silva Vignato	Faculdade de Zootecnia e Engenharia de Alimentos
Carlos Eduardo Buss	Universidade Federal de São Carlos
Juliana Afonso	Universidade Federal de São Carlos
Juliana Petrini	Escola Superior de Agricultura Luiz de Queiroz
Marina Ibelli P. Rocha	Universidade Federal de São Carlos
Mirele D. Poleti	Escola Superior de Agricultura Luiz de Queiroz
Wellison J. S. Diniz	Universidade Federal de São Carlos

Collaborators

Amanda B. Alvarenga	Escola Superior de Agricultura Luiz de Queiroz
Bruna Petry	Escola Superior de Agricultura Luiz de Queiroz
Bruno G. N. Andrade	Empresa Brasileira de Pesquisa Agropecuária
Edyjoelson P. M. Luna	Universidade Federal Rural de Pernambuco
Fábio Pértille	Escola Superior de Agricultura Luiz de Queiroz
Gabriel Costa M. Moreira	Escola Superior de Agricultura Luiz de Queiroz
Marcela M. de Souza	Universidade Federal de São Carlos
Mayara Salvian	Escola Superior de Agricultura Luiz de Queiroz
Natália Morosini	Escola Superior de Agricultura Luiz de Queiroz
Nirlei Aparecida	Escola Superior de Agricultura Luiz de Queiroz
Octávio A. Costa Almeida	Escola Superior de Agricultura Luiz de Queiroz
Paula Suarez Henriques	Escola Superior de Agricultura Luiz de Queiroz
Polyana C. Tizioto	Escola Superior de Agricultura Luiz de Queiroz
Priscila S. N. de Oliveira	Empresa Brasileira de Pesquisa Agropecuária
Shilton Reino	Escola Superior de Agricultura Luiz de Queiroz

Institutional Partners

Universidade de São Paulo - Escola Superior de Agricultura Luiz de Queiroz – ESALQ/USP

Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA

Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP – (Projeto temático nº 2012/23638-8)

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) – *Programa de Apoio a Eventos no País (PAEP, proposta nº. 636152)*

Illumina®

Proceedings Table of Contents

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

11

00-01 — A hybrid algorithm to search for causal phenotypic networks in complex biological systems conditional to genetic confounders

Renan Mercuri Pinto[1,2*], Fernando Brito Lopes[2], Bruno Dourado Valente[3], Roseli Aparecida Leandro[1], Guilherme Jordão de Magalhães Rosa[2]

[1]*Department of Exact Sciences, Luiz de Queiroz College of Agriculture, University of São Paulo, campus Piracicaba, Av. Pádua Dias n.11, 13418-900, Piracicaba, SP, Brazil. *renanmp@usp.br*
[2]*Department of Animal Sciences, University of Wisconsin - Madison, Madison, WI 53706, USA*
[3]*Genus-PIC, Hendersonville, TN 37075*

In complex biological systems, such as livestock production, understanding causal networks underlying phenotypic traits is fundamental to the development of efficient management and breeding strategies. In this context, graphical models such as Bayesian networks (BN) have been successfully used in many areas to investigate causal relationships between variables and to estimate the magnitude of such effects. Inferring the structure of a causal network, however, is not a simple task, given a large number of potential networks to be compared, even for a modest set of variables. With high-dimensional data, for example in genetics and genomics applications in which a huge number of the variable is observed in each unit (animal or plant), things get even more complex. There are many structure learning algorithms for BN, which allow exploring the causal structure space compatible with the joint distribution function of the variables studied. In the area of quantitative genetics, Valente et al. (2010) proposed a constraint-based algorithm to search for recursive causal structures among phenotypes conditional to unobservable additive genetic effects, which act as confounders. However, their method is based on an infinitesimal model, and as such it does not take into account the possibility of major genes affecting the traits. Neto et al. (2008) suggested a score-based algorithm including information on quantitative trait loci (QTL) for each phenotype to help to determine causal directions in the phenotypic network. Here, we propose a hybrid method by combining these two approaches, to search for recursive causal structures among complex phenotypic traits with polygenic inheritance but also allowing the possibility of major gene effects. Briefly, a standard multiple-trait model is fitted using Bayesian methods considering major genes as covariates, in addition to unobservable additive genetic effects. Next, posterior samples of the residual covariance matrix are used as input for the Inductive Causation (IC) algorithm to search for plausible causal network structures. Finally, the Akaike information criterion (AIC) is used to compare each putative causal structure provided by the IC algorithm. Results of a simulated study considering a QTL mapping population showed that, in the presence of major genes, our method recovers the correct skeleton structure and causal direction with a higher rate than that of Valente et al. (2010).

Keywords: Bayesian networks, phenotypic causal network, graphical models.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

12

00-02   Genome-wide association study for feed conversion ratio in Nelore cattle

Samuel Wallace Boer dos Santos[1]*, Camila Urbano Braz[1], Daniel Gustavo Mansan Gordo[1], Gerardo Alves Fernandes Júnior[1], Lúcio Flávio Macedo Mota[1], Adriana Luize Bocchi[2], Josineudson Augusto II de Vasconcelos Silva[3,5], Émerson G. Moraes[4], Leonardo F. N. Souza[4], Lucia Galvão de Albuquerque[1,5]

[1]*São Paulo State University (Unesp), Animal Science Department, Jaboticabal, SP, 144884-900, Brazil. *samuel_wallace_eu@hotmail.com*
[2]*Goiás Federal University, Jataí, GO, Brazil*
[3]*São Paulo State University (Unesp), Animal Breeding and Nutrition Department, Botucatu, SP, 18618-681, Brazil*
[4]*Nelore Qualitas Breeding Program, Goiânia, GO, Brazil*
[5]*Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq*

The objective of this study was to perform a GWAS analysis to identify genomic regions associated with feed conversion ratio in Nellore cattle. The data set used provided by Nelore Qualitas breeding program. Phenotypic and genotypic information from 715 young bulls born between 2008 and 2013, were used. These animals participated in a feed efficiency performance test between 2010 and 2015. Contemporary groups (CG) were defined by the year of the feed efficiency test. We excluded the observations that presented measures with 3.5 standard deviations above or below the mean of the CG. The animals were genotyped using the high-density SNP panel (777 K) or had their genotypes imputed to it. After quality control of genotypes, 409.871 single nucleotide polymorphisms (SNP) located in autosomal chromosomes, with minor allele frequency (MAF) $\geq$ 0.02, Hardy-Weinberg Equilibrium p-value $\geq$ 10^{-5}, call rate $\geq$ 0.98 and $r^2 \leq$ 0.995, remained. In addition, samples with call rate $\leq$ 0.90 have also been excluded. The GWAS analysis was performed using the single-step methodology (ssGBLUP), using BLUPF90 family programs. A single trait model was considered: $y = Xb + Za + e$, where: y is the vector of phenotypic observations; X is the incidence matrix for fixed effects, b is the vector of fixed effects, including CG and, as covariate (linear effect), the age of the animal at the beginning of the test; Z is the incidence matrix that relates the animal to the phenotype; a is the vector of direct additive genetic effects; e is the vector of residual effects. The results of the GWAS analysis were presented based on the proportion of the additive genetic variance explained by windows of 100 adjacent SNPs. The MapViewer tool for bovine genome in NCBI was used for the identification of the genes. Three windows with major effects were identified explaining more than 1.8% of the additive genetic variance for feed conversion ratio. These regions are located on chromosomes 7, 9 and 13 with 37 genes, being 14 coding protein; 21 transcribing to non-coding RNAs (ncRNA), and two transcribing to transfer RNA. These genes are involved in fertility, lipid metabolism, feed efficiency and carcass traits, which may be related to feed conversion ratio. This study may contribute to a more detailed understanding of genetic mechanisms underlying feed conversion ratio in Nelore cattle.

Keywords: Association analysis, bovine, single-step, SNP.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

13

00-03 Population structure in genotype x environment interaction for days to calving in Nelore Cattle

Lúcio Flávio Macedo Mota[1*], Gerardo Alves Fernandes Júnior[1], Ana Cristina Herrera Rios[1], Roberto Carvalheiro[1,2], Lucia Galvão de Albuquerque[1,2]

[1]School of Agricultural and Veterinarian Sciences, São Paulo State University
*flaviommota.zoo@gmail.com
[2]CNPq Researcher

Recently, advances in genotyping technologies have made possible to detect population structure and explore genomic regions that provide great adaptability to cattle in different environments. Thus, the aim of the present study was to evaluate the effect of population structure in GxE interactions and the contribution of SNP markers to differentiation of populations according to environment levels. A total of 1,847 female genotypes with 416,555 SNP markers were used to evaluate population structure. The animals were raised on farms distributed in three regions of Brazil (Midwest, Southeast, and Northeast). Days to calving (DC) was estimated as the difference in days between the calving date and the first day of the breeding season. Environmental gradient (EG) was defined according to herd and year of birth and management group (from birth to yearling). EG levels were standardized to present a mean equal 0 and variance of 1, and with values ranging from -3 to +3 SD units. The population structure analysis was performed using the R package adegenet, applying discriminant principal component analysis (DPC) to evaluate the pattern of genetic variability across environments, using SNP markers. Heterozygosity was estimated using autosomal markers with plink v. 1.09, by the formula: $Het = (N-O)/N$, where N is the number of non-missing genotypes and O is the observed number of homozygous genotypes for a given individual. The genes that occur within 1 Mb for SNPs with higher effect were identified using Ensembl (www.ensembl.org/biomart/). DPC analysis for SNP markers provided the cluster of animals according to the environmental gradient. Animals from adjacent environmental levels clustered together with small variation within the population. Animals raised in environments more favorable showed higher heterozygosity (0.38 for EG5 and 0.34 for EG4) when compared with unfavorable (0.28 for EG1 and 0.26 for EG2). DPC analysis applied to animals raised in different environments, identified candidate genes with biological functions related to body fat, embryonic growth, growth traits, a decrease in levels of insulin, insulin-like growth factor I and leptin, muscle fiber type and decreased in female and male fertility, decreased oocyte number and spermatogenesis. In summary, DPC analysis indicated a difference in population structure with evidence for five clusters and genes with biological functions for reproductive traits, as well as observed differences in heterozygosity.

Keywords: Environment conditions, genetic diversity, genotype environment interaction.

Acknowledgements

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

14

00-04   Evidence of runs of homozygosity islands in Nellore Cattle[1]

Ana C. Herrera Rios[2,3*], Gerardo A. Fernandes Júnior[2], Lucio F.M. Mota[2], Tiago Bresolin[2], Danielly B.S. Silva[2], Camila Urbano Braz[2], Roberto Carvalheiro[2,4], Lucia G. Albuquerque[2,4]

[1]Financial Suport: São Paulo Research Foundation FAPESP (grant 2009/16118-5)
[2]FCAV-UNESP, Jaboticabal, SP, Brazil. * anacristinah@gmail.com
[3]Genetics Animal Improvement and Modeling Research Group, GaMMA, University of Antioquia,Colombia
[4]CNPq, Brasilia, DF, Brazil

Runs of homozygosity (ROH) islands describe genomic regions where more than 50% of the individuals of a population share ROH segments, which could, theoretically, be an indication of strong past selection in these populations. Thus, studies of ROH islands can provide insights into population structure, genomic properties and cultural habits, which can affect animal performance. This study aimed to detect ROH islands in the genome of Nellore cattle that may be involved in selection events. Information from 3,492 animals (2,007 females and 1,485 males) genotyped using the Illumina® BovineHD BeadChip was used. The autosomal and X chromosome markers with GenTrain $\geq$ 0.70 and Call Rate $\geq$ 0.95 and samples with Call Rate $\geq$ 0.90 were used in the analysis. The ROH segments were obtained using PLINK software, considering a sliding window of 50 consecutive SNPs and allowing 2 heterozygous per window and 5 missing calls. The minimum length of a ROH was set to be 1000 kb, the required minimum SNP density considered was 1 SNP per 50 kb, and the maximum gap between two consecutive SNPs was set to 500 kb. To identify the ROH islands, we used the command -homozyg-group of PLINK, obtaining pools of overlapping and potentially matching segments exceeding 50% of the whole population on autozygosity. Six ROH islands were identified on chromosomes 5, 7, 12, 21 and 26. In these regions, 43 genes were found, being 34 coding protein; 3 transcribing to non-coding RNAs (ncRNAs), 2 micro RNAs (miRNA); 2 ribosomal RNA (rRNA); 1 codingan open reading frame (ORF); and 1 coding for a transcription factor. A Gene Ontology enrichment analysis showed that these genes are involved in mammary gland cell proliferation, growth, spermatogenesis, skeletal muscle fiber development, immune system and apoptotic process. Some genes found in the ROH islands are involved in sexual development and maturity such as the PROB1 that was associated with puberty in *Bos indicus*. The genes *GLI1, BTRC, FGF8; MYO1A; SHMT2; STAT6 and SLC23A1* were previously associated with carcass and meat traits in cattle. Our results evidenced the presence of ROH islands across the genome of Nellore cattle, and some islands have been previously associated with economic importance traits.

Keywords: *Bos indicus*, Single Nucleotide Polymorphisms, runs of homozygosity.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

15

00-05 *NEDD4*: a putative candidate gene for ribeye area in Nellore steers

Bárbara Silva-Vignato[1]*, Luciana C. A. Regitano[2], Luiz L. Coutinho[3], Aline S. M. Cesar[3], Mirele D. Poleti[3], Júlio C. C. Balieiro[4*]

[1]*College of Animal Science and Food Engineering, University of São Paulo, Pirassununga, Brazil.*
**bsilvavignato@gmail.com*
[2]*EMBRAPA Pecuária Sudeste, São Carlos, Brazil*
[3]*College of Agriculture "Luiz de Queiroz", University of São Paulo, Piracicaba, Brazil;*
[4]**College of Veterinary Medicine and Animal Science, University of São Paulo, Pirassununga, Brazil*

Understanding muscle growth and development is one of the principal goals of animal science, given the role of these processes in beef yield and quality. The ribeye area (REA) is a carcass trait that has been used to predict the amount of lean meat in the animal and is important as a direct parameter of muscular development. This study aimed to identify differentially expressed genes (DEG), and metabolic pathways in Nellore steers with extreme genomic estimated breeding values (GEBV) for REA. For that, the GEBV of 385 animals were obtained by the GenSel program. After that, twelve animals with extreme GEBV values were selected, and divided into two groups of six animals each, High and Low. The total RNA was extracted from the *Longissimus dorsi* muscle of these selected animals. Next generation sequencing technology (HiSeq2500 from Illumina platform) was used, and the quality of raw reads was analyzed using the FastQC software. The alignment and counting were performed using TopHat2 and HTSeq software, respectively. The differential expression analysis was carried out by DESeq2 software from R. It was identified 101 DEG (FDR$<$0.1), 72 down-regulated and 29 up-regulated in the Low REA group. Among the down-regulated genes, *NEDD4* (neural precursor cell expressed, developmentally down-regulated 4, E3 ubiquitin protein ligase), enriched in the endocytosis pathway by DAVID v6.7 (nominal p-value$<$0.1), encodes a ubiquitin protein ligase belonging to the Nedd4 family. This gene is required for cell surface expression of the IGF-1 (insulin-like growth factor, type 1) receptor and insulin receptor, and is a positive regulator of IGF-1 and insulin signaling. The IGFs are responsible for fetal and postnatal growth and are strongly related to muscle differentiation. Studies with knockout mice for the *NEDD4* gene showed that loss of *NEDD4* reduced IGF-1 and insulin signaling, delayed embryonic development, and reduced growth and body weight. These results indicate that *NEDD4* can be a putative candidate gene for REA in Nellore cattle, but more research is needed to better understand its complete role in muscular growth.

Keywords: Bovine, differential expression analysis, muscle growth, RNA-Seq.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

16

00-06   DNA methylation profiles in red blood cells of adult hens correlate to their previous rearing conditions

Fábio Pértille[1,2*], Margrethe Brantsæter[3], Janicke Nordgreen[3], Luiz Lehmann Coutinho[2], Andrew M. Janczak[3], Per Jensen[1], Carlos Guerrero-Bosagna[1]

[1]*Avian Behavioral Genomics and Physiology Group, IFM Biology, Linköping University, Linköping, Sweden, 58183.*pertille@usp.br
[2]*Animal Biotechnology Laboratory, Animal Science and Pastures Department, University of São Paulo (USP)/ Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, São Paulo, Brazil*
[3]*Animal Welfare Research Group, Department of Production Animal Clinical Science, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway*

Stressful conditions are common in the environment where production animals are raised. Stress in animals is usually determined by the levels of stress-related hormones. A big challenge, however, is in determining the history of the exposure of an organism to stress because the release of stress hormones can show a recent but not a sustained exposure to stressful conditions. Epigenetic tools are useful to investigate long-term stress and detect environmental exposures that affect gene regulation during the lifetime of organisms. It is expected that if animals are constantly subjected to stress and systemic hormonal changes, this exposure will imprint the epigenome of many cells types, including blood cells. Epigenomic effect of stress in red blood cells have been reported in monkeys and humans. Chickens, however, provide a unique model to study stress effects in red blood cells (RBCs), which are nucleated in birds. Moreover, chickens are the most consumed meat source in the world and are therefore subjected to a variety of conditions in their production environment. The fact that chickens contain nucleated red blood cells allows for a straightforward measuring of the epigenome in a cell type of easy access, and in live animals. The present study investigates in chickens whether two different rearing conditions can be identified by looking at DNA methylation patterns in their RBCs later in life. The conditions tested are rearing in a system of open aviaries versus in cages. These rearing conditions are likely to differ regarding the amount of stress to which birds are exposed, as suggested by observations showing long-term differences in fearfulness and cognitive functions. We found 115 differentially methylated regions in RBCs (P < 0.0005) between experimental groups. Network analyses of genes associated with these regions showed connections with important biological pathways, mainly related to immune system and signal transduction in opioid signaling. The objective of the present study is to generate a proof-of-concept for future detection of long-term stress in production animals, using epigenetic measurements in cell types of easy accessibility in live animals.

Keywords: Stress, epigenetics, chicken, DNA methylation, animal welfare.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

17

00-07 Bayesian methods and different π for genomic selection on residual feed intake using small training population

Amanda Botelho Alvarenga[1*], Juliana Petrini[1], Aline S. M. Cesar[1], Gregori A. Rovadoski[1], Luiz L. Coutinho[1], Gleidson G. P. de Carvalho[2], Luís F. Pinto[2], Gerson B. Mourão[1]

[1]ESALQ-USP, Piracicaba, SP, Brazil. *abalvarenga@usp.br, gbmourao@usp.br
[2]UFBA, Salvador, BA, Brazil

Genomic analyses presented a different scenario to the traditional statistic, where the number of observations is much smaller than the number of covariates. Some methodologies are able to work with these high-parametrization regression models, such as the Bayesian approach. In this study, we evaluated some Bayesian methods considering different probabilities of the genetic marker has zero effect on the phenotype (π) for genomic prediction. Residual feed intake (RFI, = -0.000055, SD= 0.153) from 392 Santa Inês sheep were used. The RFI is a deviation of a standardized feed intake (SFI) and expected feed intake (EFI). An SFI is the daily-observed feed intake mean, being the deviation of dry matter feed provided and leftovers. To calculate the EFI, average daily gain and metabolized mid-point weight were used to model SFI. Muscle samples were used for DNA extraction, and genotyping was performed using the Illumina Ovine SNP50 BeadChip (Illumina Inc., San Diego, CA, 54,241 single nucleotide polymorphisms, SNP). After the quality control for genotyping data, 43,779 SNP and 395 animals were maintained for the analyses. Analyses were performed using the following regression models: BayesA, BayesB with π equal to 0.990, 0.995, 0.999 and 0.490, BayesC with π equal to 0.490, and BayesCπ. For all the methods a similar number of iterations, burnin and thin was adopted. However, BayesCπ and BayesB with $\pi=0.999$ methods did not converge. For the remained methods, the genetic additive variance, residual variance, and heritability converged. The posterior distribution of correlations and regression between the real and estimated breeding value presented an adequate behavior considering all the methods, except the methods that not converged. In addition to this evaluation, it was calculated the accuracy average for each method initially selected by convergence and posterior distribution. The accuracy was estimated in a cross-validation procedure. In this, eight groups (seven groups with 50 animals and one with 43 animals) were formed. On each cycle, seven groups were considered as a training population, and one was the validation group. The accuracy was estimated using the Pearson correlation between the real (phenotype adjusted for environmental factors) and the estimated breeding value. Among the methods, higher accuracies were obtained for BayesB $\pi=0.99$ and BayesC $\pi=0.490$, with mean accuracy around of 0.079 and a standard deviation of 0.1434 and 0.1285, respectively. Therefore, based on these partial results, the best method was BayesB with $\pi=0.99$. This method optimizes the association analysis for detection of accurate SNP effect. Moreover, for genomic selection analysis, it economically impacts due to the marker density, being lower expensive low-density panel than high density.

Keywords: Santa Inês breed ovine, feed efficiency, regression model, accuracy, cross-validation, Pearson correlation

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

18

00-08 Putative genes involved in muscle functioning are differentially expressed in Nelore steers divergent for sodium and potassium concentration

Juliana Afonso[1]*, Laura A. Lins[2], Erica G. Cassaro[3], Polyana C. Tizioto[4], Priscila S.N. de Oliveira[3], Wellison J. da S. Diniz[1], Andressa O. de Lima[1], Marcela M. de Souza[1], Marina I.P. Rocha[1], Juliana V. da Silva[1], Carlos E. Buss[1], Caio F. Gromboni[4], Gerson B. Mourão[5], Ana R. A. Nogueira[2], Luiz L. Coutinho[5], Luciana C. de A. Regitano[3]

[1]Department of Genetics and Evolution, Federal University of São Carlos, São Carlos, SP, Brazil. *juafonsobio@gmail.com
[2]Zootecnical department, São Paulo State University – UNESP, Jaboticabal, SP, Brazil.
[3]Embrapa Southeast-Cattle Research Center, São Carlos, SP, Brazil
[4]Federal Institute of Education, Bahia Science and Technology, Valença, BA, Brazil
[5]Department of Animal Science, University of São Paulo, Piracicaba, SP, Brazil

Sodium and potassium play a major role in muscle functioning acting together in the Sodium-Potassium pump. Finding genes that are involved in sodium and potassium homeostasis can lead to a better understanding regarding the unknown shared roles of these minerals in the muscle physiology. In this study, we detected differentially expressed genes (DEGs) among extreme genomic estimated breeding values for sodium and potassium concentration in muscle (n=6 per group) of Nelore steers. Gene expression was measured with RNA-Seq in *Longissimus dorsi* muscle and the Tuxedo suite modified pipeline was adopted to identify the DEGs. We identified 73 DEGs for sodium and 79 DEGs for potassium, being 35 of them shared between the minerals (four upregulated and 31 downregulated in the positive group for both minerals). Among the downregulated we found seven genes (*CD44, COL11A1, TNC, THBS4, COMP, COL11A2,* and *ITGA10*) that are part of the Extracellular matrix (ECM) receptor interaction and the Focal Adhesion pathways. The ECM-receptor pathway is responsible for signaling molecules that act as receptors for interaction between cells and the extracellular matrix. The Focal adhesion pathway is liable for cell adhesions and, in conjunction with the ECM-receptor pathway, act on migration, differentiation, proliferation, apoptosis and other cellular activities. Products of both pathways together can link membrane receptors to the actin cytoskeleton using focal adhesion components, controlling thus the shape of the cell and acting as a sensor of an ECM-receptor condition affecting the cell behavior, being part of muscle growth. Our shared DEGs were previously associated with muscle and tendons healing process, but their expression were not related to sodium and potassium concentration. The other 24 downregulated shared DEGs have also functions related to signaling and extracellular matrix interaction. Our results provide inferences that the pathways discussed and other shared DEGs probably affect or are being affected by the maintenance of the ideal sodium and potassium concentration in muscle cells that are involved in functions such as muscle growth. We theorize that sodium and potassium concentration can act as messengers for the receptors in the extracellular matrix affecting muscle growth and healing.

Keywords: Cattle, gene regulation, RNA-Seq, minerals.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

19

00-09 | **Co-expression network analysis identifies genes associated with iron content in bovine muscle**

Wellison Jarles da Silva Diniz[1*], Aline Silva Mello Cesar[2], Ludwig Geistlinger[3], Polyana Cristine Tizioto[2], Priscila Silva Neubern de Oliveira[3], Juliana Afonso[1], Marina Ibelli Pereira Rocha[1], Andressa Oliveira de Lima[1], Carlos Eduardo Buss[1], Luiz Lehmann Coutinho[2], Luciana Correia de Almeida Regitano[3]

*Department of Genetic and Evolution, Federal University of São Carlos, São Carlos, São Paulo, Brazil.
*wjarles09@gmail.com
Department of Animal Science, University of São Paulo/ESALQ, Piracicaba, São Paulo, Brazil
Embrapa Pecuária Sudeste, São Carlos, São Paulo, Brazil*

Iron is an essential cofactor required for several biological functions. The iron content in the body is known to be influenced by several factors including genetic composition. However, there is limited knowledge about potentially functional genes and metabolic pathways related to the iron content in cattle. To identify candidate gene networks associated with muscle iron content, we performed weighted gene co-expression network analysis (WGCNA) in a Nelore population of 60 steers. Gene expression was measured with RNA-Seq in *Longissimus dorsi* muscle samples and genomic estimated breeding values for iron content (GEBV-IC) were used as a quantitative trait. After quality control of raw reads and normalization, blockwise modules function was carried out using the 800 highest connected genes, and the signed network was constructed using a $\beta = 14$ ($R^2 = 0.89$). From the co-expression modules detected by WGCNA, we found six modules to be significantly correlated with GEBV-IC. Response to the hormone, mRNA metabolic process, and mitochondrial organization were biological process found significantly enriched for the genes presented in the associated modules. We identified as hub *Pin1* and *MRPL* genes which are iron responsive and play a major role in metabolism, cell division, and growth. Genes involved with muscle structure development such as *ACTA1*, *SRF*, and *GSK3A* were identified. The transcription factor *SRF* stimulates both cell proliferation and differentiation, and it is a master regulator of the actin cytoskeleton. The detected co-expression modules in this study provide evidence for substantial interplay between genes in processes influencing iron content. Further studies will be required to understand the underlying regulatory mechanisms of the observed co-expression networks.

Keywords: Gene expression, network analysis, systems biology.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

20

00-10	Analysis of functional polymorphisms associated with feed efficiency in Nelore cattle

Gabriela Ribeiro[1*], Pâmela Almeida Alexandre[1], Aline Silva Mello Cesar[2], Heidge Fukumasu[1]

[1]*Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo*
*gabriela.ribeiro.br@gmail.com, fukumasu@usp.br
[2]Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo*

Due to population growth and the consequent increase in beef demand, studies regarding feed efficiency (FE) in beef cattle are important, since selection for efficient animals improve productivity. However, evaluating FE is a slow and onerous activity, which opens space for other approaches that could select superior animals in less time and with lower cost. One possible strategy is the identification of molecular markers that allow the earlier identification of high FE animals. Considering that, the aim of this study is to identify functional genomic variations associated with the FE by analyzing the sequence of transcripts expressed in the liver of Nellore bulls. For this purpose, samples from 8 high feed efficient animals and 8 low feed efficient animals were collected based on the measure of residual feed intake. The mRNA from samples were extracted and sequenced using Illumina HiSeq 2500 equipment (100bp, paired-end) and the reads generated aligned to the bovine reference genome (UMD3.1) using TopHat. As quality control, PCR duplicates were excluded, and identification of single nucleotide polymorphisms (SNPs) in all samples was performed using GATK. A total of 268,459 SNPs were identified after filtering for Quality $\leq$ 30, Depth $<$ 4, QualByDepth $<$ 3, FisherStrand $<$ 30 and RMSMappingQuality $>$ 35. Then, using PLINK, 174,079 SNPs were excluded due to MAF (minor allele frequency) $<$ 0.40 and a genotypic association test was performed. Significant SNPs ($P<0.05$) were analyzed by VEP online tool, which predicts functional consequences of each polymorphism. Finally, 4 significant polymorphisms were identified as deleterious by SIFT, which is a program in VEP that predicts whether protein function is affected by an amino acid substitution. It calculates normalized probabilities that up to 0.05 indicate a deleterious SNP and above that the polymorphism is considered to be tolerated. It is still necessary to validate those results in a larger population so that it is possible to confirm the association with the FE and thus obtain interesting markers that will be important to select more efficient animals.

Keywords: RNAseq, residual feed intake, single nucleotide polymorphism.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

21

00-11 Absence of CpG observed/expected ratio depletion in *Haemonchus contortus* evidences a non-methylated genome

Simone Cristina Méo Niciura[*], Patricia Tholon

*Embrapa Pecuária Sudeste, São Carlos, SP, Brazil. *simone.niciura@embrapa.br*

Haemonchus contortus is a blood-sucking helminth, which parasites the abomasum of small ruminant and leads to production losses, anemia, and death. DNA methylation is an epigenetic event that regulates gene expression, cell and tissue differentiation and also cell cycle transition in parasites. Due to the absence of DNA methylation in the genome of *Caenorhabditis elegans*, it was assumed that this epigenetic mark was absent in all nematode species. However, studies confirmed DNA methylation in *Trichinella spiralis* and nine other nematodes. Thus, the objective of this study was to *in silico* investigate whether DNA methylation is present in the *H. contortus* genome. DNA methylation occurs in cytosine in the CpG context, and methylated cytosines are prone to deamination to thymine. Then, in methylated genomes there is a depletion in the ratio between observed and expected CpG (CpGo/e) content in methylated regions of DNA over time compared to unmethylated regions, leading to a bimodal CpGo/e distribution. To assess this, a total of 23,100 expressed sequenced tags (EST) of *H. contortus* were retrieved from NCBI-dbEST public database and filtered by size ($\geq$ 500 bp), resulting in 10,940 EST subjected to CpGo/e analysis with CpG Island Promoter Detection (http://doua.prabi.fr/software/cpgprod). Normality and Kolmogorov-Smirnov tests performed in SAS (SAS Inst., Inc., Cary, NC) fitted the data in a normal curve with mean 0.91 ($\pm$0.18), which indicates a unimodal CpGo/e distribution. This absence of CpGo/e depletion in *H. contortus* suggests that its genome lacks DNA methylation.

Keywords: Epigenetics, DNA methylation, gastrointestinal nematode, barber's pole worm.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

22

00-12 Identification of a pleiotropic locus for beef quality and feed efficiency in cattle using bi-trait genome-wide association analysis

Carlos Eduardo Buss[1]*, Juliana Petrini[2], Gerson Barreto Mourão[2], Ludwig Geistlinger[3], Jason B. Wolf[4], Wellison Jarles da Silva Diniz[1], Marina Ibelli P. Rocha[1], Juliana Afonso[1], Andressa Oliveira de Lima[1], Polyana Cristine Tizioto[2], Rymer Ramiz Tullio[3], Bruno Gabriel Andrade[3], Luiz Lehmann Coutinho[2], Luciana Correia de Almeida Regitano[3]

[1]Department of Genetic and Evolution, Federal University of São Carlos, São Carlos, SP, BR.
*cadu_buss@hotmail.com
[2]Department of Animal Science, University of São Paulo/ESALQ, Av. Pádua Dias, Piracicaba, SP, BR
[3]Embrapa Southeast Cattle, Rodovia Washington Luiz, Km 234, s/n, Fazenda Canchim, São Carlos, SP, BR
[4]Department of Biology & Biochemistry, University of Bath, Milner Centre for Evolution Bath, BA2 7AY, UK

The molecular mechanisms of pleiotropy indicate multiple functions of a single gene product and many consequences of a single molecular function related to animal breeding. However, single-trait GWAS are not per se suitable for systematically identifying such loci. In this study, we used a bi-trait GWAS for detecting pleiotropic effects between twenty beef quality and feed efficiency traits. Based on Illumina High-Density SNP-chip data for phenotyped population of 387 Nelore steers. Covariance components were estimated using the restricted maximum likelihood method under a mixed animal model, considering a genomic relationship matrix. Highlighted, we found one locus significantly associated with the sum of omega-3 (N-3) and relative growth rate (RGR). This locus resides in the gene encoding TBC1 Domain Family Member 5 (*TBC1D5*). In agreement with our results, recent studies show that *TBC1D5* regulates the activation of the Ras-related Rab-7, a protein that plays a pivotal role in the regulation of the autolysosome-mediated lipid degradation in fat cells. The detected SNP in *TBC1D5* explains up to 0.45% of the variance observed for N-3 and 3.77% for RGR. The locus identified in this study show pleiotropic effects suggesting a role in lipolysis in both phenotypes provides evidence for the genetic interrelation between omega-3 and relative growth rate and indicates candidate gene for future investigation of causal mutations.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

23

00-13 Potential regulatory elements on *PCDH7* gene affecting residual feed intake in Nelore cattle

Andressa O. de Lima[1*], Polyana C. Tizioto[2], Priscila S.N. de Oliveira[3], Ludwig Geistlinger[3], Marina I.P. Rocha[1], Wellison J. da S. Diniz[1], Juliana Afonso[1], Carlos E. Buss[1], Juliana Petrini[2], Gerson B. Mourão[2], Luiz L Coutinho[2], Luciana C. de A. Regitano[3]

[1]*Departament of Genetic and Evolution, Federal University of São Carlos, São carlos, São Paulo, Brazil*
[2]*Department of Animal Science, University of São Paulo/ESALQ, Piracicaba, São Paulo, Brazil*
[3]*Embrapa Southeast Livestock, São Carlos, São Paulo, Brazil*

Residual feed intake (RFI) is a measure of feed efficiency and can improve the profitability of cattle herds and potentially reduce methane emission, but it has late and costly measurements. Identify the causes of gene expression variation, like regulatory polymorphisms, can be helpful to understand the regulatory elements that affect residual feed intake and can be useful in animal breeding programs. Recent studies performed by our research group in a Nelore cattle population, such as Genome-wide association (GWAS), Association weight matrix (AWM) and RNA-Seq analysis of liver tissue revealed that the *PCDH7* gene plays a role in RFI. To identify regulatory elements and polymorphisms for this gene, we analyzed the promoter region of the *PCDH7* (*chr4:30716201bp–30722201bp*) described in Ensembl database for human (*GRCh38.p10*) at UCSC genome browser. This region is conserved among several species and has a binding affinity with the transcription factor *E2F1,* which is involved in regulation of fat cell proliferation and differentiation. We converted the human promoter region coordinates to the correspondent on the bovine genome *UMD 3.1* (*chr6:51530833bp-51537647bp*) using the lift genome annotation tool from UCSC genome browser. In this region, we identified nine SNPs of the 20 genome sires sequenced by Illumina Hiseq2500®. Briefly, after the sequencing, we performed BWA-MEM for alignment and GATK for variant calling. We annotated the SNPs using the Variant ensembl prediction (VEP), five of which are located in an upstream region of the *PCDH7*. We found one SNP located in a transcription factor binding site (TFBS) for *CCAAT/enhancer-binding protein beta (CEBPB)* and one SNP located in TFBS for *Neurofibromin (NF1)*. These Transcription factors (TFs) are related to regulation of brown fat cell differentiation and skeletal muscle tissue development, respectively. Our findings indicate putative regulatory elements in the *PCDH7* gene that could have a role in RFI variation. Nevertheless, more studies considering variants in regulatory regions in this gene will be performed to understand its effect on feed efficiency.

Keywords: Residual feed intake, promoter, transcription factor, SNPs.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

24

00-14   Genome-wide association study for milk fat using principal component analysis in Holstein cows

Mayara Salvian[1*], Juliana Petrini[1], Eula R. Carrara[1]; Gregori A. Rovadoscki[1], Amanda B. Alvarenga[1], Luiz L. Coutinho[1], Paulo F. Machado[1], George R. Wiggans[2], Gerson B. Mourão[1]

[1]Department of Animal Science, University of São Paulo, Piracicaba, Brazil. *mayara.salvian@usp.br
[2]Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD, USA

Milk fat is composed mainly of triglycerides, which contain different structure, such as saturated, monounsaturated and polyunsaturated fatty acids. Some of these fatty acids are indicated as precursors of low-density blood cholesterol, responsible for cardiovascular diseases. To improve human health, it is interesting to increase the proportion of long-chain fatty acids, such as mono and polyunsaturated, in the fat milk composition. The fatty acids synthesized by the mammary glands have a significant genetic variation and, therefore, has the potential of alteration by selection. Thus, changes directed to these traits may be useful for the animal breeding of the nutritional quality of milk for human consumption. Genome-wide association studies (GWAS) have been effective in the identification of genetic variants associated with complex traits. This methodology analyzed traits separately, in a univariate manner. Meanwhile, a multivariate analysis could be better when a single variant is related to multiple traits, pleiotropic SNPs effects. Principal components analysis is a multivariate method of analysis, whose objective was to create an index from a group of variables. Milk samples were collected monthly on the test day milk (May 2012 until December 2014) from 4,781 Holstein cows. DNA was extracted from samples of the hair roots for the genotyping performed by the Illumina BovineLD BeadChip panel (Illumina Inc., San Diego, CA) providing 6,909 SNPs. Posteriorly, the panel was imputed to 60,671 SNPs. After the quality control remained to analyses 57,369 SNPs. The principal components analysis was carried out using the PROC PRINCOMP SAS® procedure to group the fat traits (fat percentage and fatty acids - palmitic, stearic, oleic, total saturated and unsaturated). The GWAS analysis was performed by BLUPF90 software. Therefore, the aim of this study was to quantify SNP effects and to identify major effects SNP for milk fat using principal components analysis. The first principal component explained 77% of the total variance of the system across the 6 traits. The SNP Hapmap41556-BTA-48495 (chromosome 2, position 102772529) presented a significant association with an effect equal to 5.13. Nevertheless, the SNP found in the current study were not reported in dairy cattle studies yet.

Keywords: BLUP, dairy cattle, fatty acids, multivariate analysis, SNP.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

25

00-15 *LDHB* gene has allele-specific expression in liver of Nelore cattle extremes for feed efficiency

Marina Ibelli Pereira Rocha[1*], Marcela Maria de Souza[1], Adhemar Zerlotini Neto[2], Polyana Cristine Tizioto[3], Priscila Silva Neubern de Oliveira[4], Andressa Oliveira de Lima[1], Juliana Afonso[1], Wellison Jarles da Silva Diniz[1], Carlos Eduardo Buss[1], Luiz Lehmann Coutinho[3], Luciana Correia de Almeida Regitano[4], Simone Cristina Méo Niciura[4]

[1]*Department of Genetics and Evolution, Federal University of São Carlos, Rodovia Washington Luiz, Km 235, s/n - Jardim Guanabara, São Carlos, SP, BR. *Grant number FAPESP 2015/17802-8. *marinaiprocha@gmail.com*
[2]*Embrapa Informática Agropecuária, Av. André Tosello, 209, Campus Unicamp, Barão Geraldo, Campinas, SP, BR*
[3]*Department of Animal Science, University of São Paulo/ESALQ, Av. Pádua Dias, Piracicaba, SP, BR*
[4]*Embrapa Pecuária Sudeste, Rodovia Washington Luiz, Km 234, s/n, Fazenda Canchim, São Carlos, SP, BR*

Feed efficiency is a multi-factorial trait of a large economic importance for cattle. Although previous studies reported gene expression differences associated with this trait the contribution of allele-specific expression (ASE) remains largely unknown. In this study, we analyzed ASE in liver samples of 30 Nelore cattle steers in two genetically divergent groups for residual feed intake (RFI) searching for genes with ASE linked to feed efficiency. Based on genotype data obtained using the Illumina BovineHD BeadChip, we computed the frequency of reads from RNA-seq data mapped to each allele of heterozygous individuals and applied a binomial test to identify loci of ASE. We detected significant differences in expression among alleles for seven SNPs (single nucleotide polymorphisms) that were tested significantly in >90% of the samples. Amid them, we selected the *LDHB* gene that has been previously associated with feed efficiency in chickens. The *LDHB* gene carries out functions in carbohydrate, carboxylic acid and oxidation-reduction metabolic processes important for glycolysis. These biological processes were associated with feed efficiency and cell energy balance in previous studies. Since the SNP found on *LDHB* is located at 3'UTR, we studied whether a putative microRNA binding site near or at the SNP site could exist and account for the regulation of gene expression. We found that the Bta-miR-139 binds at the SNP site of *LDHB*. This miRNA was also identified by the group in a previous study with expressed microRNA in the liver. Thus, we theorized that the Bta-miR-139 miRNA affects the expression of the *LDHB* gene and may contribute to ASE. Therefore, these results complement our understanding of the ASE profile of the *LDHB* gene and contribute to explain more accurately the differences in gene expression identified in Nelore steers genetically divergent for RFI.

Keywords: Cattle, residual feed intake, differential expression, microRNA.

Proceedings of the 1st Workshop on Omics Strategies Applied to Livestock Science.
April 24th -26th, 2017. Piracicaba, São Paulo.

26

00-16   **Fine mapping of genomic windows associated with carcass traits in a purebred Brazilian chicken population reveals potential causative mutations**

Gabriel C.M. Moreira[1*], Clarissa Boschiero[1], James M. Reecy[2], Aline S.M. Cesar[1], Thaís F. Godoy[1], Maurício E. Cantão[3], Mônica C. Ledur[3], Adriana M.G. Ibelli[3], Dorian Garrick[2], Luiz Lehmann Coutinho[1#]

[1]*University of São Paulo (USP) / Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, São Paulo, Brazil. #llcoutinho@usp.br*
[2]*Iowa State University (ISU), Department of Animal Science, Ames, Iowa, United States*
[3]*Embrapa Suínos e Aves, Concórdia, Santa Catarina, Brazil*

Carcass quality is a critical issue for sales of chicken meat worldwide. The discovery of genomic regions associated with carcass traits is a first step in the identification of putative candidate genes and putative candidate mutations responsible for this traits. The aim of this study was to perform fine mapping in genomic windows associated with carcass traits to identify possible causative mutations annotated in candidate genes, in a purebred broiler (TT) Brazilian chicken population. A total of 24 unique 1-Mb windows were associated previously with carcass traits on GGA1-6, 10, 12, 13, 14, 17, 20, 22, 25, 26 and 28. Imputation to sequence was performed by FImpute software (default parameters) using the parental whole-resequencing data as the reference population (14 sires and 666.488 SNPs) and genotyping information from offsprings as target population (1,394 offsprings and 27.495 SNPs). All the animals from offsprings were previously were genotyped with the 600K SNP genotyping array (Affymetrix®). Our imputation accuracy was around 99%, and after the removal of markers with MAF$\leq$2%, we kept 637,702 informative SNPs in our genomic windows. After imputation, the unique set of variants was annotated using Variant Effect Predictor (VEP) tool. Most of the SNPs were annotated in potential neutral regions (intronic and intergenic regions) comprising 71% of the total. A total of 285,707 SNPs were annotated as potentially functional (missense variations, stop gained, stop lost, start lost, stop retained, splicing, mature miRNA variant, up/downstream from the gene and 3'- or 5'-untranslated region (UTR)). We identified 68 high-impact mutations annotated in 62 different genes and two of those, related to skeletal muscle regulation: TNNT2 and TPM3. All these mutations and genes will be deep studied to select possible causative mutations for carcass traits regulation in Brazilian chicken lines.

Keywords: Sequence imputation, genomic windows, SNP discovery, high impact variant.

Embrapa

Pecuária Sudeste