Expression quantitative trait loci (eQTL) hotspot regions from whole genome analysis of Nellore steers

Aline S.M. Cesar[1*], Luciana C.A. Regitano[2], Dante P.D. Lanna[1], Mirele D. Poleti[1], James E. Koltes[3], James M. Reecy[4], Eric Fritz-Waters[4], Luke Kramer[4], Dorian Garrick[4], Maurício A. Mudadu[2], Rymer R. Tullio[2] and Luiz L. Coutinho[1]

[1] Department of Animal Science, University of São Paulo, Piracicaba, SP, Brazil, [2] Embrapa Southeast-Cattle Research Center, São Carlos, SP, Brazil, [3] Department of Animal Science, University of Arkansas, Fayetteville, AR 72701, USA, [4] Department of Animal Science, Iowa State University, Ames, IA 50011, USA.

*e-mail: alinecesar@usp.br

INTERNATIONAL PLANT & ANIMAL GENOME CONFERENCE
JANUARY 9-13, 2016
TOWN & COUNTRY HOTEL
SAN DIEGO, CA

INTRODUCTION

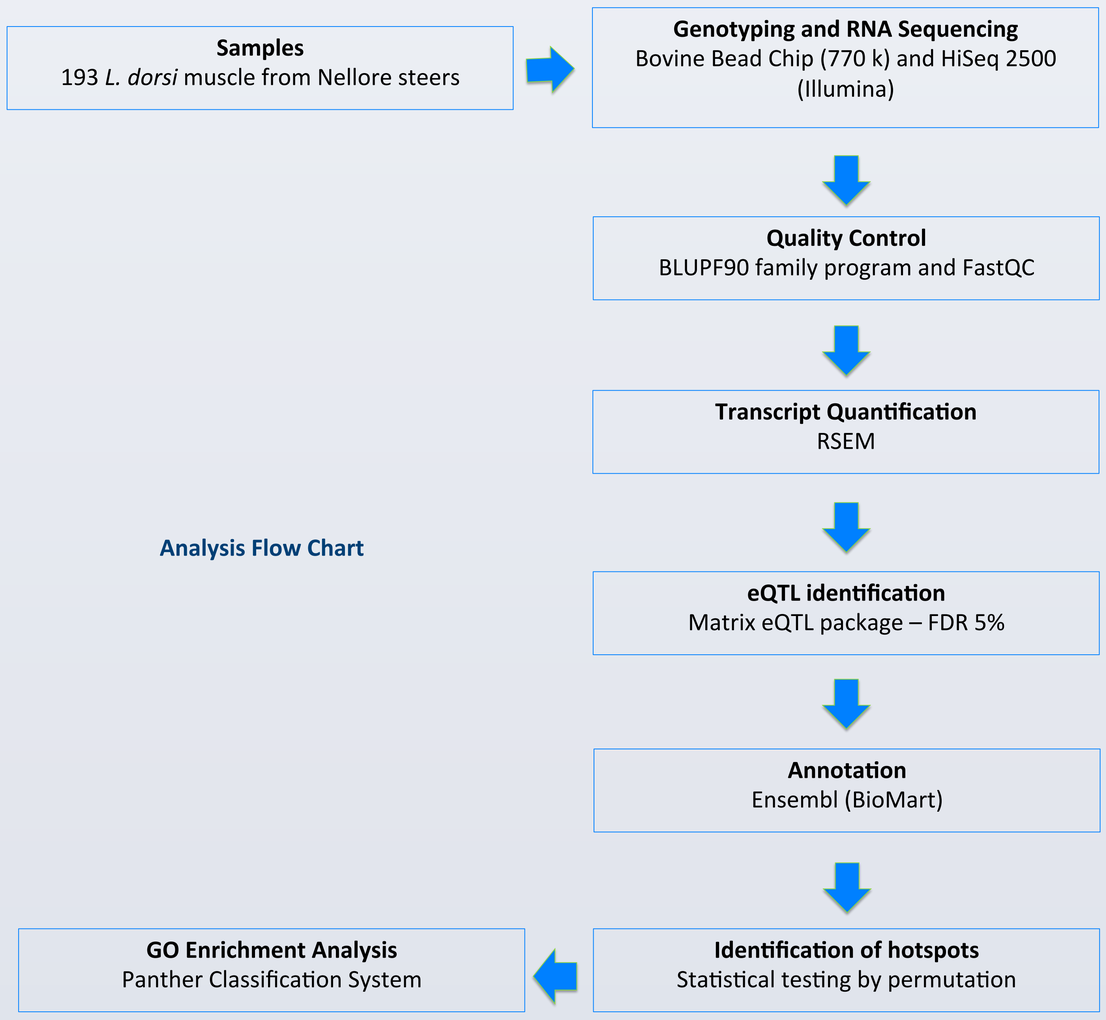
The use of panels of high-density SNPs and genome-wide association studies (GWAS) in animal breeding programs allows for the detection of significant QTLs for traits of economic importance with greater accuracy (Hayes et al., 2009). However, the majority of variants are found in non-coding regions, which makes it harder to understand the possible molecular mechanisms.

Next Generation Sequencing (NGS) technologies have allowed whole transcriptome sequencing to occurred. When coupled with genotype information, it is possible to study quantitative variation in transcription regulation, e.g. expression quantitave loci (eQTL) (Michaelson et al., 2009). To date, the architecture of gene regulation and genetic basis of variation in gene expression has not been described for *Bos indicus breeds* and only to a limited extent in *Bos taurus* (Littlejohn et al . 2013).

OBJECTIVES

The aim of this study was to identify *cis* and *trans*-eQTL and discover eQTL hotspots (genetic marker directly or indirectly controlling a large amount of transcripts throughout the genome) across the bovine genome.

MATERIAL AND METHODS



Analysis Flow Chart

RESULTS

Herein, 925 *cis*-eQTLs (distance ≤ 1 Mb) and 10,685 *trans*-eQTLs were identified across autosomal chromosomes. Using the trans-eQTLs identified, twelve eQTL hotspots were identified by permutation test (p-value=0.05) on chr3 (8Mb and 9Mb), chr4 (108Mb), chr11 (11.5, 11.74 and 11.75 Mb), chr14 (73 Mb), chr16 (59 Mb), chr17 (55Mb) and chr28 (20Mb, 32.01 and 32.03 Mb) (Figure 1).

Enrichment analysis of gene ontology (GO) enrichment analysis was performed by Panther version 10.0 (Thomas et al., 2003) from the list of genes harbored on eQTL hotspot regions (Figure 2).

Potential candidate regulators were identified in close proximity to some hotspot regions, i.e. a 1 Mb for either side of the association location. Within these regions four transcription factors (TF) were identified, such as *EGR4* (chr 11: 145925-11148403) , *NR1I3* (chr3: 8290360-8295271, *RUNX1T1* (chr14: 74642806-74787356), *USF1* (chr3: 8455700-8461692). The circos plot (Figure 3) shows the association between three candidate genes within the hotspot regions and transcript expression level along with the average percentage of the variance explained by hotspot , R^2.

RESULTS

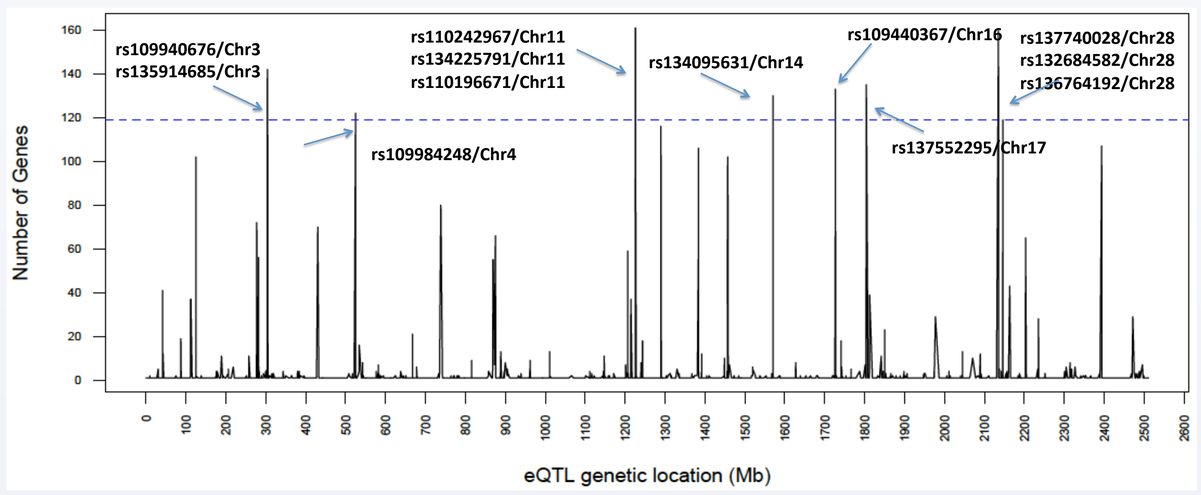


Figure 1. eQTL hotspot regions from whole genome analysis of L. dorsi muscle of Nellore steers.

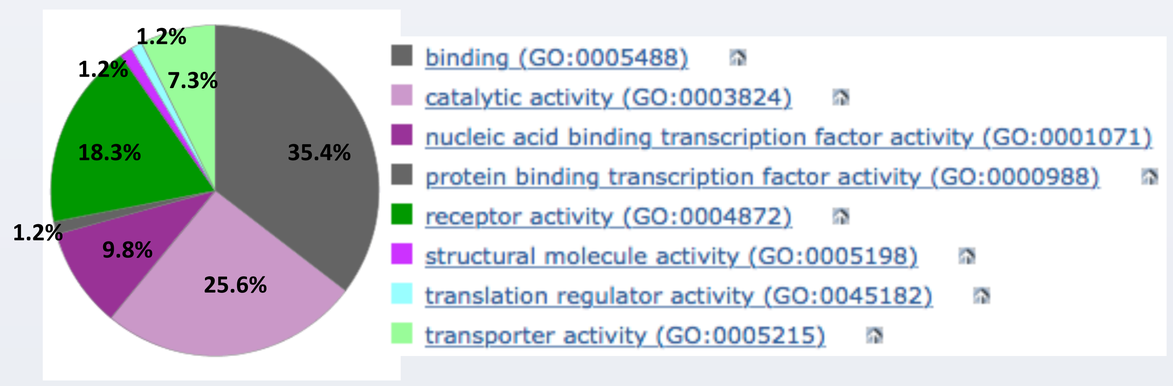


Figure 2. GO enrichment analysis from the list of genes harbored on eQTL hotspot regions from whole genome analysis of L. dorsi muscle of Nellore steers.

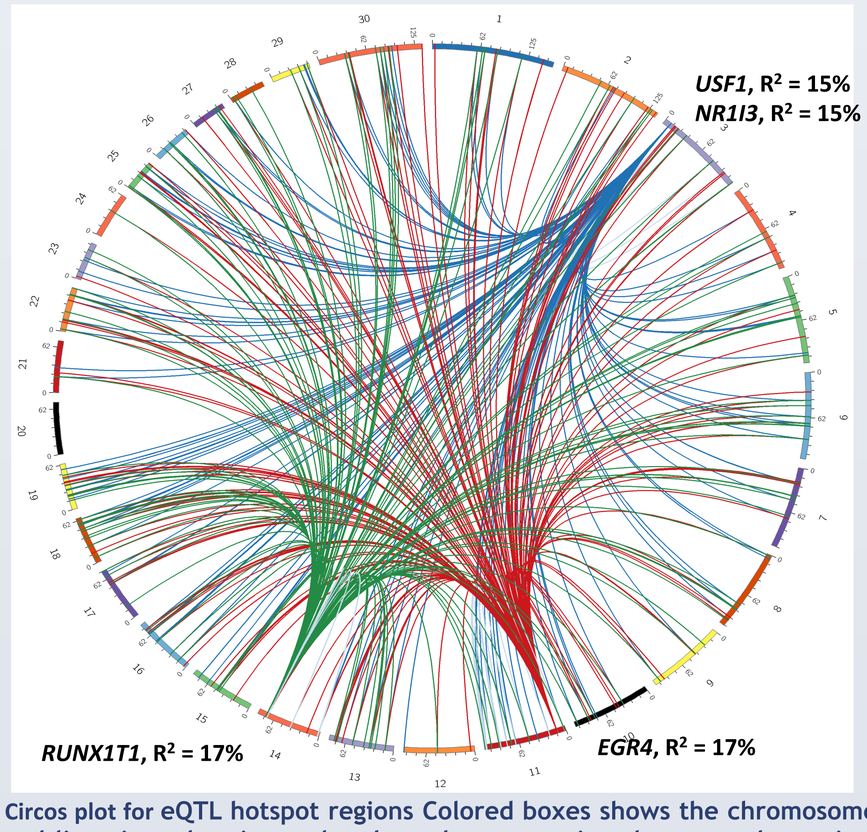


Figure 3. Circos plot for eQTL hotspot regions Colored boxes shows the chromosome number and the colored lines into the circos plot show the connections between the regions that harbor the TFs and the genes that showed the expression level associated. In blue are the connections from *USF1* and *NR1I3* TFs. In red are the connections from *EGR4* TF. In green are the connections from *RUNX1T1* TF. The percentage of genetic variance accounted for by a hotspot are show next to the candidate gene names (R^2).

CONCLUSIONS and FUTURE DIRECTIONS

Important candidate regulators involved to transcription mechanisms in bovine species, more specifically in *Bos indicus* are identified in this study. These findings provide new insight into the complex gene networks and genetic architecture that contributes to important economical traits in beef cattle. TFBS enrichment are being performed to better understanding of gene regulation process.

ACKNOWLEDGMENTS

This study is supported by FAPESP (processes number 2012/23638-8, 2014/11871-5 and 2014/22884-0) and EMBRAPA (Macroprograma 1, 01/2005).

My travel was partially supported by Agriculture and Food Research Initiative Competitive Grant no. 2016-67015-24579 (FAANG Workshop) from the USDA National Institute of Food and Agriculture.