

## Identification of genomic regions related to lipid and cholesterol content in beef

## Identificación de las regiones genómicas relacionadas con el contenido de lípidos y colesterol en bovinos

F.M. Rezende<sup>1</sup>, G.A. Oliveira Junior<sup>2</sup>, M.E. Carvalho<sup>2</sup>, R.V. Ventura<sup>2,3</sup>, J.B.S. Ferraz<sup>2</sup>, and J.P. Eler<sup>2</sup>

<sup>1</sup>Federal University of Uberlandia, Uberlandia, MG, Brazil,

<sup>2</sup>Research Center of Animal Breeding, Biotechnology and Transgenesis of the University of Sao Paulo/FZEA Brazil,

<sup>3</sup>Beef Improvement Opportunities, Guelph, Ontario, Canada.

Email: frezende@ingeb.ufu.br

The aim of this study was to identify genomic regions that potentially have association with lipid and cholesterol content in Nellore cattle meat. Phenotypes of 615 Nellore bulls were obtained according to the methods described by Bligh and Dyer (1959) and Saldanha *et al.* (2004). Data of 658 genotyped Nellore bulls were used for association studies through the ssGWAS method. Those animals were genotyped with the Illumina Bovine HD® Beadchip (777 962 SNP) or Gene Seek GGPI (74,153 SNP). Based on another Nellore population genotyped for Illumina Bovine HD® Beadchip, genotypes were determined using the FImpute software. After quality control (MAF <2% and call rate <90%), 535 824 SNP in autosomal chromosomes were used in the association analyses. Single step analyses were performed using a pedigree composed by 4065 animals by the BLUPF90 program considering windows of 10 markers to estimate their effects. This procedure enables the identification of regions associated with lipid and cholesterol content by chromosome. Results of this research showed regions on chromosomes 5, 10, 12, 23 and 29 related to lipid content and on chromosomes 3, 10, 11, 12, 13, 17 and 18 associated with cholesterol deposition.

Key-words: *GWAS, Nellore, meat quality, SNP.*

### Introduction

The localization of genomic regions related to the expression of quantitative traits allows identifying genes and their pathways that provides a better comprehension of their genetic control. Lipid and cholesterol contents are becoming major concerns to consumers because of excessive consumption of high density calorie food has harmful effects on human health mainly on increasing cardiovascular diseases. Therefore, the aim of this study was to research on genomic association methods that allow the use of pedigree and genomic information for the identification of chromosomal regions that potentially have association with lipid and cholesterol content in Nellore beef.

### Materials and Methods

*Animals.* Nellore young bulls raised under pasture conditions until 18 months of age and, after, fed in feedlots, were slaughtered between 21 and 27 months of age in six different dates, what makes them standardized. All animals are progenies of bulls selected for production

El objetivo de este estudio fue identificar las regiones genómicas que tienen asociación potencial con el contenido de lípidos y colesterol de la carne de ganado vacuno Nellore. Se obtuvieron fenotipos de 615 toros Nellore, de acuerdo con los métodos descritos por Bligh y Dyer (1959) y Saldanha *et al.* (2004). Los datos de 658 toros Nellore genotipados se utilizaron para estudios de asociación a través del método ssGWAS. Estos animales se genotiparon con Illumina Bovine beadchip HD® (777 962 SNP) o GeneSeek GGPI (74 153 SNP). Sobre la base de otra población de Nellore genotipada por Illumina beadchip BovineHD®, los genotipos se determinaron con el software FImpute. Después del control de calidad (MAF <2% y call rate <90%), se utilizaron 535 824 SNP en cromosomas autosomales para el análisis de asociación. Los análisis de paso simple se realizaron utilizando un pedigrí compuesto por 4065 animales con el programa BLUPF90, considerando ventanas de 10 indicadores para estimar sus efectos. Este procedimiento posibilita la identificación de regiones asociadas al contenido de colesterol y lípidos por cromosoma. Los resultados de esta investigación demostraron regiones en los cromosomas 5, 10, 12, 23 y 29 relacionadas con el contenido de lípidos, y regiones en los cromosomas 3, 10, 11, 12, 13, 17 y 18 relacionadas a la deposición de colesterol.

Palabras clave: *GWAS, Nellore, calidad de la carne, SNP*

### Introduccion

La localización de las regiones genómicas relacionadas con la expresión de rasgos cuantitativos permite la identificación de genes y sus trayectorias, las cuales proveen mejor comprensión de su control genético. Los contenidos de colesterol y lípidos se están convirtiendo en una mayor preocupación para los consumidores debido a que el consumo excesivo de alimentos con alta densidad de calorías tiene efectos negativos en la salud humana, principalmente en el aumento de enfermedades cardiovasculares. Por lo tanto, el objetivo de este estudio fue investigar los métodos genómicos de asociación que permitan el uso de pedigrí e información genómica para la identificación de regiones cromosómicas que tienen asociación potencial con el contenido de lípidos y colesterol de la carne de ganado vacuno Nellore.

### Materiales y Métodos

*Animales.* Se utilizaron toros jóvenes Nellore, criados en condiciones de pastoreo hasta los 18 meses de edad. Después de alimentarlos en el cebadero, se sacrificaron entre los 21 y 27 meses de edad, en seis

and reproduction traits.

*Phenotypic traits.* *Longissimus dorsi* sample was collected, individually identified and, after 7 days of ageing, kept frozen at  $-18^{\circ}\text{C}$  till the analyses. The determination of total lipids was based on methodology described by Bligh and Dyer (1959). Cholesterol extraction and quantification was made according to method described by Saldanha *et al.* (2004).

*Animal genotyping.* Animals were genotyped with Illumina Bovine Beadchip HD® (777962 SNP, N = 449 animals) or Illumina Gene Seek SNP Beadchips Bovine GGP-HDi (74153 SNP, N = 209 animals) technologies, according to the manufacturer protocols. After animal genotyping, the data from the lower density panel were imputed using the software FImpute, for panel Illumina Bovine HD®. The accuracy of imputation was determined by cross validation for each animal, and the concordance rate between the imputed and the real genotype was higher than 97.51%. The quality control of markers SNP was made, excluding those with unknown genomic position, placed in sexual chromosomes, with minor allele frequency (MAF) of 2%, markers that presented call rate lower than 90% and markers with heterozygous genotype excess. After quality control 658 animals samples and 535824 SNP remained to be analyzed.

*Association analysis.* Analyses were performed using a pedigree composed by 4,065 animals. Single step analyses were realized by BLUPf90 program considering windows of 10 SNP to estimate their effects, this procedure enables the identification of regions associated with lipid and cholesterol content along the chromosomes. The animal model considered as fixed effects slaughter group and analysis date, besides, as covariates, slaughter age and backfat and, as random effects, animal and residual effects. The variance components and genetic parameters were estimated using the Bayesian inference (Gianola and Fernando 1986), considering a linear animal model, and the GIBBS2F90 and ssGWAS computer programs were used (Misztal *et al.* 2002 and Aguilar *et al.* 2011). Chains with size of 100 000 interactions, with a burn in criteria of the first 1000 interactions were generated. For the estimation of parameters, the samples were stored every 100 cycles, forming samples with 990 data. The data convergence was verified with the graphic evaluation of the values sampled versus interaction the according to Heidelberger and Welch (1983) and Raftery and Lewis (1992) by package of analysis Bayesian Output Analysis (BOA) of software R 2.9.0 (The R Development Core Team, 2009).

*Genomic region prospecting.* The gene identification was realized with BioMart from Ensembl Genome Browser tool. It was possible to make the identification

fechas diferentes, lo cual se estandarizó. Todos los animales son descendientes de toros seleccionados por sus características para la producción y reproducción.

*Rasgos fenotípicos.* Se recolectó una muestra del *Longissimus dorsi*, identificada individualmente y, después de 7 días de envejecimiento, se congelaron a  $-18^{\circ}\text{C}$  hasta los análisis. La determinación de lípidos totales se realizó de acuerdo con la metodología descrita por Bligh y Dyer (1959). La extracción y cuantificación de colesterol se realizó según el método descrito por Saldanha *et al.* (2004).

*Genotipos animales.* Los animales se genotiparon con Illumina Bovine beadchip HD® (777 962 SNP, N = 449 animales) o Illumina GeneSeek SNP Beadchips Bovine GGP-HDi (74 153 SNP, N = 209 animales), según los protocolos del fabricante. Después de genotipar los animales, los datos del panel de menor densidad se analizaron utilizando en software FImpute, para el panel de Illumina BovineHD®. La precisión de atribución se determinó por validación cruzada para cada animal, y la tasa de concordancia entre el genotipo atribuido y el real fue mayor que 97.51%. El control de calidad de los indicadores SNP se realizó con frecuencia de alelos pequeños (MAF) de 2%, excluyendo a aquellos con posición genómica desconocida, colocados en cromosomas sexuales, indicadores que presentaron una tasa inferior a 90% y los indicadores con exceso de genotipo heterocigoto. Después del control de calidad de las muestras de 658 animales y 535 824 SNP quedaron por analizar.

*Análisis de asociación.* Los análisis se realizaron utilizando un pedigrí compuesto por 4 065 animales. Los análisis de paso simple se realizaron con el programa BLUPf90, considerando ventanas de 10 SNP para estimar sus efectos. Este procedimiento posibilita la identificación de regiones asociadas con el contenido de lípidos y colesterol en los cromosomas. El modelo animal considerado como efectos fijos fue un grupo sacrificado y la fechas de análisis. Además, como co-variantes, se utilizaron la edad al sacrificio y la grasa dorsal. Como efectos aleatorios se consideraron los efectos residuales y animales. Los componentes de la varianza y los parámetros genéticos se estimaron con la inferencia Bayesiana (Gianola y Fernando 1986), considerando un modelo lineal animal, y se utilizaron los programas computarizados GIBBS2F90 y ssGWAS (Misztal *et al.* 2002 and Aguilar *et al.* 2011). Se generaron cadenas de 100 000 interacciones, con un burn en el criterio de las primeras 1000 interacciones. Para los parámetros de interacción, las muestras se almacenaron cada 100 ciclos, formando muestras con 990 datos. La convergencia de los datos se verificó con la evaluación gráfica de los valores muestreados contra la interacción, según Heidelberger and Welch (1983) y Raftery y Lewis (1992), por el paquete de análisis Bayesian Output Analysis (BOA) del software R 2.9.0 (The R Development Core Team 2009).

*Prospección de la región genómica.* La identificación de genes se realizó con BioMart de la herramienta Ensembl Genome Browser. Fue posible realizar la

of gene association analyses for lipid and cholesterol content.

identificación de análisis de asociación de genes para el contenido de lípidos y colesterol.

**Results and Discussion**

**Resultados y Discusión**

The residual and the additive variance and heritability were 0.06, 0.34, 0.17 and 5.54, 31.98, 0.17 for lipid and cholesterol content, respectively. In table 1 were presented statistics description of analyzed data.

La varianza residual y de aditivos, y la heredabilidad fueron 0.06, 0.34, 0.17 y 5.54, 31.98, 0.17 para los contenidos de lípidos y colesterol, respectivamente. La tabla 1 presenta la descripción estadística de los datos analizados.

Figures 1 and 2 contain genomic regions and the percentage of variance explained by chromosome by windows of 10 SNP adjacent for lipid and cholesterol content, respectively.

Las figuras 1 y 2 contienen regiones genómicas y el porcentaje de varianza explicado a través de cromosomas y ventanas de 10 SNP adyacentes para el contenido de lípidos y colesterol, respectivamente.

Table 1. Description of phenotypic data.

Traits	N <sup>1</sup>	Mean	SD <sup>2</sup>	Min <sup>3</sup>	Max <sup>4</sup>
Lipid (g/100g of meat)	550	2.19	0.65	0.96	4.60
Cholesterol (mg/100g of meat)	615	56.42	8.26	28.76	83.95

<sup>1</sup>Number of animals; <sup>2</sup>Standart deviation; <sup>3</sup>Minimum; <sup>4</sup>Maximum.

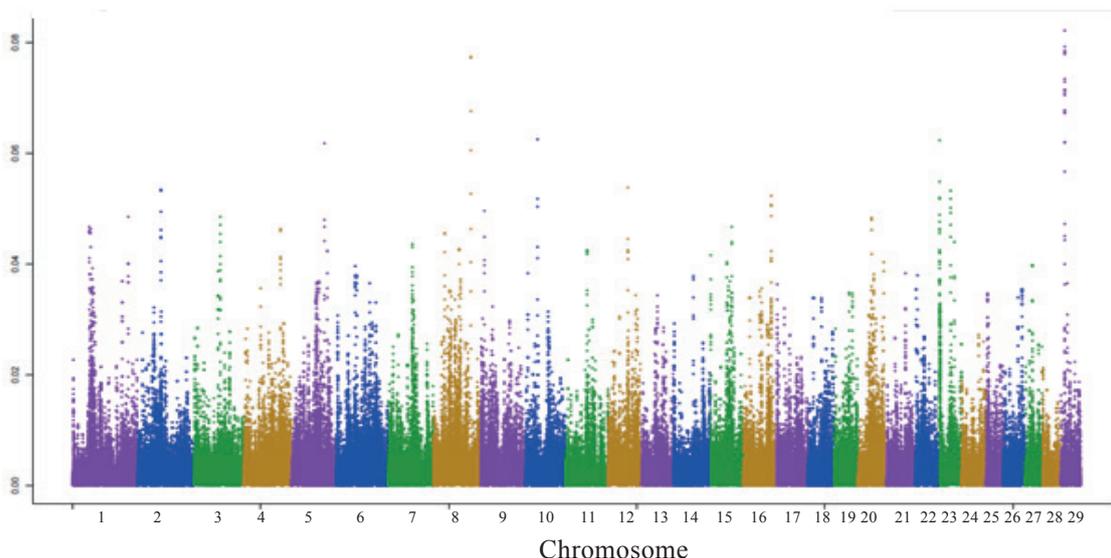


Figure 1. Manhattan Plot of genomic regions associated with lipid content

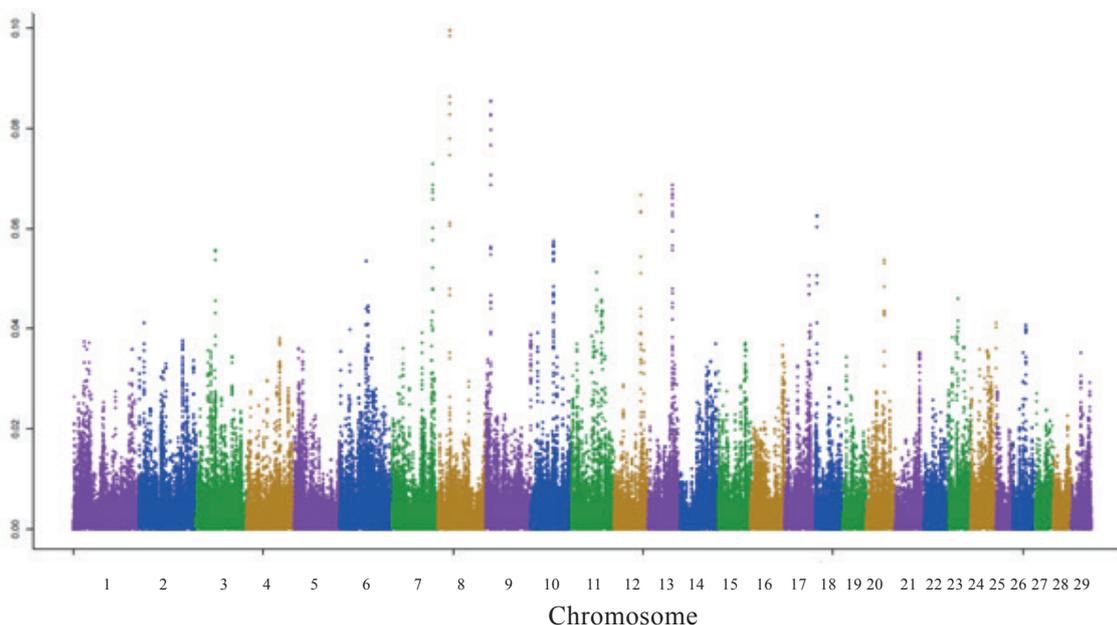


Figure 2. Manhattan Plot of genomic regions associated with cholesterol content

A total of 7 and 10 associated genomic regions, distributed in 5 and 7 different chromosomes were obtained for lipid and cholesterol content, respectively. In these regions, 7 and 9 genes were identified for each trait in this order (tables 2 and 3).

Un total de 7 y 10 regiones genómicas asociadas, distribuidas en 5 y 7 cromosomas diferentes, se obtuvieron para el contenido de lípidos y colesterol, respectivamente. En estas regiones, 7 y 9 genes se identificaron para cada rasgo en este orden (tablas 2 and 3).

Table 2. Chromosome, Ensembl\_ID and name of gene identified as associated with lipid content.

Chromosome	Ensembl_ID	Gene
5	ENSBTAG00000019603	LDHB
10	ENSBTAG00000025642	RYR3
12	ENSBTAG00000014227	NDFIP2
23	ENSBTAG00000043990	KHDRBS2
23	ENSBTAG00000005674	NEU1
29	ENSBTAG00000019995	C29H11orf73
29	ENSBTAG00000007847	EED

Table 3. Chromosome, Ensembl\_ID and name of gene identified as associated with cholesterol content.

Chromosome	Ensembl_ID	Gene
3	ENSBTAG00000017655	PALMD
10	ENSBTAG00000011936	ATP8B4
11	ENSBTAG00000004567	C1D
12	ENSBTAG00000047181	--
13	ENSBTAG00000001015	TM9SF4
13	ENSBTAG00000018334	PLAGL2
13	ENSBTAG00000016231	POFUT1
13	ENSBTAG00000003215	KIF3B
17	ENSBTAG00000044119	KSR2
18	ENSBTAG00000009805	BCNT

With ssGWAS method, using a high density panel, it was possible to identify regions related to lipid and cholesterol content in Nellore beef. Posteriorly, those genes and their pathways will be researched to evaluate their then importance for meat quality traits.

Con el método ssGWAS, utilizando paneles de alta densidad, fue posible identificar las regiones relacionadas con el contenido de lípidos y colesterol en la carne del ganado vacuno Nellore. Posteriormente, estos genes y sus trayectorias se investigarán para evaluar su importancia para los rasgos de calidad de la carne.

### References

- Aguilar, I., Misztal, I., Legarra, A. & Tsuruta, S. 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. *Journal of Animal Breeding Genetics*. 128: 422.
- Bligh, E.G. & Dyer, W.J. 1959. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology*. 37: 911-917.
- Ensembl Genome Browser tool. Available on: [www.ensembl.org](http://www.ensembl.org)
- Gianola, D. & Fernando, R.L. 1986. Bayesian methods in animal breeding theory. *Journal of Animal Science*. 63:217.
- Heidelberger, P. & Welch, P. 1983. Simulation run length control in the presence of an initial transient. *Operations Research*. 31:1109.
- Misztal, I., Tsuruta, S., Strabehl, T., Auvray, B., Drueta, T. & Lee, D.H. 2002. *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*. Montpellier: INRA. 28: 21.
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available on: <<http://www.R-project.org>>
- Raftery, A. E. & Lewis, S.M. 1992. Comment on "The Gibb sampler and Markov chain Monte Carlo". *Statistical Science*. 7: 493

Saldanha, T., Mazalli, M.R., & Bragagnolo, N. (2004) Avaliação comparativa entre dois métodos para determinação do colesterol em carnes e leite. *Ciência e Tecnologia de Alimentos*. 24(1), 109-113

**Received: July 5, 2016**