

IX Simpósio Brasileiro de Melhoramento Animal João Pessoa, PB – 20 a 22 de junho de 2012

Associação de SNPs com características de carcaça em uma população da raça Canchim¹

Fabiana Barichello Mokry², Andressa Oliveira de Lima³, Ismael Urbinati⁴, Roberto Augusto de Almeida Torres Júnior⁵, Roberto Hiroshi Higa⁶, Luciana Correa de Almeida Regitano⁷

¹Projeto financiado pela CAPES

²Pós-doutoranda do PPGGEv – UFSCar. Bolsista da CAPES. e-mail: fbmokry@gmail.com

³Mestranda do PPGGEv – UFSCar. e-mail: eng.agro.andressa@gmail.com

⁴Zootecnista. e-mail: ismaelurbinati@zootecnista.com.br

⁵Pesquisador da Embrapa Gado de Corte. e-mail: rtorres@cnpqc.embrapa.br

⁶Pesquisador da Embrapa Informática Agropecuária. e-mail: roberto@cnptia.embrapa.br

⁷Pesquisadora da Embrapa Pecuária Sudeste. e-mail: luciana@cnpse.embrapa.br

Resumo: Foram genotipados 400 animais Canchim que apresentavam fenótipos para área de olho de lombo (AOL) e espessura de gordura subcutânea (EGS). Após o controle de qualidade, foi realizado análise de associação dos SNPs com AOL e EGS por meio da metodologia de RandomForest. Os 400 indivíduos foram também reamostrados 10 vezes formando grupos de 200 indivíduos mais representativos e menos aparentados. Foi criado um grupo de SNPs contendo apenas os SNPs que foram selecionados em todas as análises, considerados os SNPs mais robustos totalizando 197 SNPs para AOL e 162 SNPs para EGS. Após o ajuste de uma regressão selecionou-se um conjunto de 20 SNPs para AOL com $R^2=0,59$ e um conjunto de 17 SNPs para EGS com $R^2=0,49$. Esses SNPs ainda precisam ser validados para então serem utilizados para seleção assistida por marcadores moleculares.

Palavras-chave: área de olho de lombo, bovino de corte, espessura de gordura subcutânea, randomforest

SNP association with carcass traits in a population of Canchim cattle

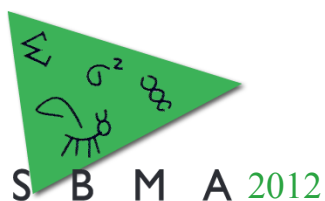
Abstract: A set of 400 Canchim cattle that were previously evaluated for ribeye area (AOL) and backfat thickness (EGS) were genotyped. After data quality control, SNP association analysis with AOL and EGS were carried out using RandomForest methodology. The 400 individuals were also resampled 10 times to build groups of 200 individuals that were more representative and less related. A SNP group was created only with the SNPs that were selected in all the analyses, with those being considered the most robust SNP yielding a total of 197 SNPs for AOL and 162 SNPs for EGS. After adjusting a regression model, one group of 20 SNPs for AOL with $R^2=0.59$, and one group of 17 SNPs for EGS with $R^2=0.49$ were selected. These SNP still need to be validated before they can be used for marker assisted selection.

Keywords: backfat thickness, beef cattle, random forest, ribeye area

Introdução

As características de área de olho de lombo (AOL) e espessura de gordura subcutânea (EGS) são importantes para o sistema de produção de carne, visto que AOL é relacionada à musculabilidade do animal e ao rendimento de cortes nobres, a EGS é relacionada com a precocidade do animal e importante para a proteção da carcaça durante o processo de resfriamento (Luchiari Filho, 2000). Tanto AOL quanto EGS são características avaliadas por ultra-som, o que demanda maior custo e tempo para a obtenção das medidas, sendo assim, o desenvolvimento de programa de seleção assistida por marcadores moleculares facilitará o processo seletivo e também aumentará o ganho genético esperado para as características consideradas.

O objetivo neste estudo foi de prospectar marcadores do tipo SNP (polimorfismo de base única) associados às características de AOL e EGS que possam ser utilizados em programas de seleção assistida por marcadores moleculares na raça Canchim.



IX Simpósio Brasileiro de Melhoramento Animal

João Pessoa, PB – 20 a 22 de junho de 2012

Material e Métodos

De uma população inicial de 987 animais que foram avaliados geneticamente, selecionou-se 400 animais extremos para AOL para serem genotipados no chip *Illumina Bovine HD*. Após a genotipagem, realizou-se o controle de qualidade por amostra e por SNPs. O controle de qualidade considerou o desvio máximo de heterozigosidade, “*call rate*” mínimo por amostra, desvio máximo do Equilíbrio de Hardy-Weinberg (HWE), frequência alélica mínima (MAF), frequência genotípica mínima e “*call rate*” mínimo por SNP. Para realizar as análises de associação foi utilizado o valor genético desregredido (VGD) (Garrick et al., 2009) como fenótipo. Como os animais genotipados foram selecionados pela estratégia de extremos para AOL, os VGD foram discriminados em maiores ou menores do que zero, tomando os valores 0 ou 1 para AOL. Para EGS, foi utilizado o valor genético desregredido como fenótipo contínuo.

A análise de associação deu-se pela metodologia de RandomForest (Breiman, 2001), que é uma combinação de árvores de classificação e regressão (CART) (Breiman et al., 1984) com *bootstrap* agregado (chamado de *bagging*) (Breiman, 1996). CART é uma ferramenta efetiva para a construção de classificadores, mas tende a ser dado-dependente. *Bagging* é um processo de amostragem de dados com reposição, e o classificador aumenta usando esta reamostragem. Após muitas iterações, os resultados de todas as árvores obtidas são combinados para criar um classificador menos variável com menor erro de predição quando comparado com o classificador original. No processo de *bagging*, a redução da variância é limitada pela correlação entre as árvores: quando a correlação é diminuída ou minimizada, o potencial para redução é aumentado.

Assim, selecionou-se inicialmente 1% dos SNPs mais discriminantes por cromossomo, e após, foi realizada nova análise por RandomForest (Breiman, 2001) selecionando-se apenas 1% dos SNPs entre os previamente selecionados. Esta análise foi realizada com os 400 indivíduos da população. No entanto, como os 400 indivíduos não eram balanceados entre famílias, optou-se por re-amostrar 200 indivíduos de forma que estes 200 indivíduos fossem os mais representativos e menos aparentados dentro da amostra inicial de 400 indivíduos. Esse processo repetiu-se por 10 vezes. Cada grupo de 200 indivíduos foi então submetido à análise de associação dos SNPs pela metodologia de RandomForest (Breiman, 2001), como descrito para a amostra total.

Resultados e Discussão

Após a seleção dos SNPs nas análises com os 400 indivíduos e com as 10 subamostras de 200 indivíduos, foi realizada uma busca pelos SNPs mais robustos, ou seja, SNPs que foram associados às características (AOL e EGS) em todas as análises realizadas pelo RandomForest (Breiman, 2001). Esta busca resultou na seleção de 197 SNPs para AOL e 162 SNPs para EGS.

Foi então, ajustada uma análise de regressão seguindo um processo “stepwise” de seleção de variáveis de interesse (SNPs) para AOL e EGS. Assim obtendo-se um conjunto de 20 SNPs para AOL com $R^2=0,59$ e um conjunto de 17 SNPs para EGS com $R^2=0,49$.

Dentro do conjunto de 20 SNPs para AOL foi possível verificar pelo NCBI BioSystems database (Geer et al., 2010) que 13 desses SNPs encontram-se dentro de 13 diferentes genes, e para EGS 11 desses SNPs encontram-se dentro de 11 diferentes genes. Ainda é preciso verificar a função desses genes encontrados, e também verificar as regiões para QTLs descritos para as características de interesse.

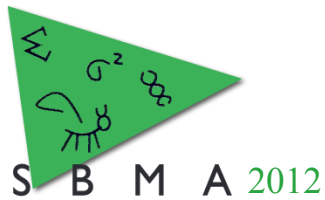
Conclusões

Esses resultados ainda precisam ser melhor analisados em conjunto com informações de blocos de desequilíbrio de ligação, anotação gênica, rotas metabólicas e afins para que então possam ser validados na população e utilizados pelos programas de melhoramento para seleção assistida por marcadores moleculares.

Literatura citada

BREIMAN, L. Random Forests. **Machine Learning**, v.45, p.5-32, 2001.

BREIMAN, L.; FRIEDMAN, J.; OLSHEN, R.; STONE, C. **Classification and Regression Trees**. New York: Chapman & Hall; 1984.



IX Simpósio Brasileiro de Melhoramento Animal
João Pessoa, PB – 20 a 22 de junho de 2012

BREIMAN, L. Bagging Predictors. **Machine Learning**, v.24, p.123-140, 1996.

GARRICK, J.; TAYLOR, F.; FERNANDO, R.L. Deregressing estimated breeding values and weighting information for genomic regression analyses. **Genetics Selection Evolution**, 41:55, 2009.

GEER, L.Y.; MARCHLER-BAUER, A.; GEER, R.C.; et al. The NCBI BioSystems database. **Nucleic Acids Research**, v.38(Database issue), 2010.

LUCHIARI FILHO, A. **Pecuária da carne bovina**. 1ed. São Paulo, 2000. 134p.