**Novos polimorfismos e genes candidatos que influenciam a produção de proteína do leite na Raça Guzerá**Júlia Gazzoni Jardim^{1*}, Goutam Sahana², Celia Raquel Quirino¹, Maria Gabriela Campolina Diniz Peixoto³, Mogens Sando Lund²¹Universidade Estadual do Norte Fluminense Darcy Ribeiro, LRMGA, Campos dos Goytacazes, RJ, Brasil.²Aarhus University, Center for Quantitative Genetics and Genomics, Tjele, Dinamarca.³Centro Nacional de Pesquisa em Gado de Leite, EMBRAPA, Juiz de Fora, MG, Brasil.

*Autor correspondente: jugazzoni@hotmail.com

Resumo: O leite é fonte de proteína de alta qualidade, esse componente desempenha importantes funções nutricionais, fisiológicas, imunológicas e moleculares. Mapeamento de genes que promovam melhorias na qualidade e valor econômico do leite são desejáveis. Para tal, uma varredura genômica para componentes de leite e escore de células somáticas foi realizada em 900 vacas Guzerá genotipadas (54k) imputadas em 25 touros genotipados em alta densidade (770k) da mesma raça. Para análise de associação genômica ampla, um modelo linear misto foi utilizado e valores genéticos derregredidos foram considerados como fenótipos. Foram identificadas 4 regiões genômicas significativamente associadas à produção de proteína de leite. SNPs mais significativos foram observadas no cromossomo 11 - destacando os genes candidatos *CAMSAP1*, *UBAC1*, *NACC2*; variantes nos cromossomos 2, 5 e 20 também apresentaram significância. Em relação a produção de leite, gordura e escore de células somáticas não foram detectados sinais associativos significativos. A análise da associação identificou novos genes candidatos que influenciam a variação no rendimento da proteína do leite na raça Guzerá.

Palavras-chave: bos indicus, componentes do leite, gwas, imputação, produção de leite.

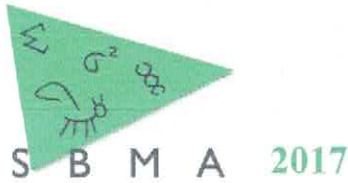
New polymorphisms and candidates genes that influence milk protein yield in Guzera Breed

Abstract: Milk is an important source of high-quality protein. The milk protein plays in role nutritional, physiological, immunological and molecular functions in Zebuine breeds. Mapping genes that promote improvements on milk quality and economic value in milk production are desirable. For this purpose, a genome-wide scan for milk components and somatic cell score were performed on 900 genotyped Guzera cows (54k) imputed to 25 high density genotyped bulls (770k) of the same breed. A linear mixed model was carried out to association analysis using deregressed breeding values as phenotypes. Were identified 4 genomic regions significantly associated with the milk protein yield. Most significant SNPs were observed on chromosome 11 - highlighting the candidate genes *CAMSAP1*, *UBAC1*, *NACC2*; variants on chromosomes 2, 5 and 20 also were presented significance. In relation to milk yield, fat yield and somatic cell score no significant associations signals were detected. Association analysis identified novel candidate genes influencing the variation in milk protein yield in Guzera breed.

Keywords: bos indicus, milk components, gwas, imputation, milk yield.

Introdução

O Guzerá (*Bos indicus*) é uma raça completa, pois atende às modernas tendências técnicas, produtivas e científicas, sendo adotada em diferentes sistemas de produção. Devido a versatilidade da raça, impressos na boa conformação corporal, termo tolerância, eficiência alimentar, fertilidade, produção de leite satisfatória e resistência a doenças, despertou o interesse no mapeamento fino para a melhoria das características produtivas e funcionais (Rosse et al., 2016). Estudos recentes de mapeamento de características quantitativas identificaram segmentos cromossômicos e polimorfismos de nucleotídeo único (SNPs) associados à resistência à mastite em diferentes populações bovinas (Kadri et al. 2015). Porém, este tipo de dado ainda não foi analisado na Raça Guzerá, apenas as associações relacionadas à produção de leite, gordura e proteína foram encontradas no cromossomo 10 e 21 (Rosse et al., 2016). Genes que codificam proteínas como a caseína, lactoglobulina e lactoferrina são utilizados como marcadores genéticos para produção, composição do leite e susceptibilidade à mastite em diversas raças zebuínas (Caroli et al.,



2009). Realizar estudos de associação entre polimorfismos e características quantitativas ou qualitativas do leite é uma estratégia poderosa, principalmente quando a seleção de marcadores moleculares é aplicada em características complexas. Portanto, nosso objetivo foi realizar um estudo de associação genômica ampla em bovinos Guzerá para características economicamente relevantes, especificamente os componentes do leite e mastite, utilizando marcadores de alta densidade imputados.

Material e Métodos

Um total de 25 touros Guzerá foram genotipados utilizando o Illumina® BovineHD BeadChip (marcador de alta densidade - HD) e 964 fêmeas Guzerá tiveram o genoma sequenciado com o Illumina® BovineSNP50 BeadChip (marcador de baixa densidade - LD). Os critérios aplicados para o controle de qualidade foram excluir indivíduos das análises se houvesse mais de 10% de genótipos ausentes, seus SNPs tivessem uma taxa de erro mendeliana acima de 2%, se *call rate* fosse menor que 90%, sua *MAF* menor que 3% ou se *HWE* tivessem um valor de *P* menor que 10^{-6} . Um total de 435718 SNPs, 900 fêmeas e 25 touros passaram pelo controle de qualidade e então selecionados para análise de associação. A imputação de genótipos de LD (54k) para HD (770k) foi realizada usando o software BEAGLE v.3.3. As análises estatísticas foram realizadas utilizando o pacote do software Rdmu (<http://dmu.agrsci.dk>). Os valores genéticos derregredidos foram considerados como fenótipos no modelo linear misto empregado. A correção de Bonferroni foi aplicada ($6.94e-07$) para o controle da taxa de erro e o nível de significância adotado foi 5%. SNPs com maior valor de *P* para cada região QTL foram nomeados SNPs candidatos e sua posição definida de acordo com a montagem do genoma de *Bos taurus* UMD3.1. As informações dos SNP foram obtidas do NCBI (<http://www.ncbi.nlm.nih.gov/snp/>). Gráficos do tipo *Manhattan* foram criados com o pacote Manhattanly v. 0.2.0 e qqman no software R.

Resultados e Discussão

Este estudo revelou novas regiões genômicas e genes candidatos significativamente associadas a produção de proteína do leite na raça Guzerá. Embora nenhuma associação significativa foi revelada para a produção de leite, produção de gordura e mastite (Figura 1). Para produção de proteína, a região mais significativa foi observada no cromossomo 11, destacando os genes candidatos *CAMSAP1*, *UBAC1*, *NACC2* - localizados próximos ao gene *PAEP* que tem efeito no conteúdo de β -lactoglobulina no leite. Estes genes também estão envolvidos na resposta inflamatória do metabolismo da glândula mamária e ácidos graxos monoinsaturados do leite. Assim como variantes do gene *PDE3A* no cromossomo 2, sobrepôs-se com QTLs anteriormente mapeados para produção de leite. SNPs no cromossomo 5 e 20 também apresentaram significância a 1%, porém estão em regiões intergênicas (Tabela 1). Todos os genes candidatos apresentam um transcrito para proteína do leite. Fontanesi et al. (2014) mapearam importantes genes candidatos nos cromossomos 2, 11 e 20 associados à produção de proteína do leite. Assim como Kadri et al. (2015), que reportaram QTLs para produção de leite e resistência a mastite no cromossomo 20.

A raça Guzerá é conhecida pelo conteúdo de gordura no leite (4.57% em média), polimorfismos no gene *DGATI* têm sido amplamente associados com a produção e composição de gordura no leite em raças taurinas, porém a frequência alélica desse gene é muito baixa em raças zebuínas (Rosse et al., 2014). Além disso, muitos marcadores em homozigose foram eliminados da análise de associação após o controle de qualidade, explicando o baixo número de regiões significativas para as características estudadas. É importante lembrar que o Programa de Melhoramento do Guzerá Leiteiro é recente, iniciou-se em 1994 com um pequeno número de fundadores ($n = 47$) que contribuíram com 50% da população atual. Porém, com a estratégia de seleção de núcleos MOET aplicado pelos criadores, a intensidade de seleção e os coeficientes de endogamia na população são baixos (Peixoto et al., 2010). Provavelmente, os respectivos alelos dessas características não estão fixos na população, dificultando o mapeamento, especialmente em pequenas populações, como a deste estudo.

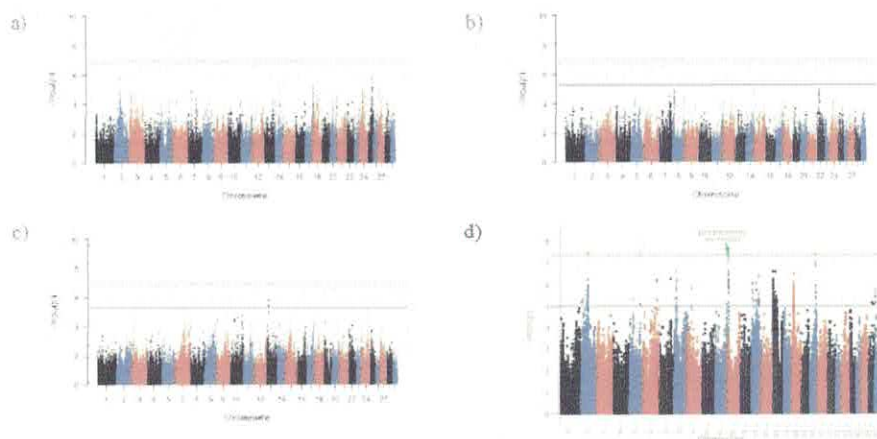


Figura 1. *Manhattan plot* dos valores $-\log_{10}(P)$ para associação genômica da produção de leite (a), gordura (b), escore de células somáticas (c) e proteína (d). A linha vermelha representa a correção de Bonferroni ($6.94e^{-07}$) e a linha azul representa o limite da taxa de falsos positivos ($5.30e^{-05}$).

Tabela 1. Regiões genômicas com SNPs mais significativos associados com a produção de proteína do leite

BTA	SNP	Posição	MAF	P-value	Genes Candidatos
2	rs135505973	61426777	0,16	4.14e-08	Região Intergênica
5	rs136054412	89730211	0,24	3.45e-08	PDE3A
11	rs132688470	103492400	0,14	2.83e-08	CAMSAP1
11	rs133139096	103536624	0,21	2.99e-08	UBAC1
11	rs109608280	103596127	0,05	4.12e-08	NACC2
20	rs135246329	50628539	0,45	1.10e-07	Região Intergênica

Conclusão

Não foi possível identificar regiões cromossômicas associadas a produção de leite, gordura e escore de células somáticas. Foram identificados genes candidatos no BTA 5 e 11 que contribuem para a produção de proteína de leite na raça Guzerá. As informações dessas variantes podem ser incluídas no modelo para aumentar a acurácia de predição dos valores genéticos na seleção genômica.

Literatura citada

- CAROLI, A. M.; ERHARDT, G. J. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. **Journal of dairy science**, v.92(11), p. 5335-5352, 2009.
- FONTANESI, L.; CALÒ, D. G.; GALIMBERTI, G.; NEGRINI, R.; MARINO, R.; NARDONE, A. A candidate gene association study for nine economically important traits in Italian Holstein cattle. **Animal genetics**, v.45 (4), p. 576-580, 2014.
- KADRI, N. K.; SAHANA, G. Genetic dissection of milk yield traits and mastitis resistance quantitative trait loci on chromosome 20 in dairy cattle. **Journal of dairy science**, v.98, p. 9015-9025, 2015.
- PEIXOTO, M. G. C. D.; POGGIAN, C. F.; VERNEQUE, R. S.; EGITO, A. A., CARVALHO, M. R. S. Genetic basis and inbreeding in the Brazilian Guzerat (*Bos indicus*) subpopulation selected for milk production. **Livestock Science**, v.131 (2), p. 168-174, 2010.
- ROSSE, I. C.; ASSIS, J. G.; OLIVEIRA, F. S.; LEITE, L. R.; ARAUJO, F.; MACHADO, M. A. Whole genome sequencing of Guzerá cattle reveals genetic variants in candidate genes for production, disease resistance, and heat tolerance. **Mammalian Genome**, p. 1-15, 2016.
- ROSSE, I.C.; STEINBERG, R.S.; COIMBRA, R. S.; PEIXOTO, M.G.C.D; VERNEQUE, R.S. Novel SNPs and INDEL polymorphisms in the 3' UTR of DGAT1 gene: in silico analyses and a possible association. **Molecular Biology Reports**, v. 42, p. 1, 2014.