

X Simpósio Brasileiro de Melhoramento Animal

Uberaba, MG – 18 a 23 de agosto de 2013

### Desequilíbrio de ligação com um painel de alta densidade em genomas de bovinos da raça Guzerá<sup>1</sup>

Maria Gabriela Campolina Diniz Peixoto<sup>2</sup>, Daniel Jordan de Abreu Santos<sup>3</sup>, Maria Raquel Santos Carvalho<sup>4</sup>, Marco Antônio Machado<sup>2</sup>, Marcos Vinícius Gualberto Barbosa da Silva<sup>2</sup>, Humberto Tonhati<sup>5</sup>

<sup>1</sup>Trabalho financiado pela Fapemig, Fapesp e CNPq.

<sup>2</sup>Embrapa Gado de Leite, Juiz de Fora. e-mail: [gabriela.peixoto@embrapa.br](mailto:gabriela.peixoto@embrapa.br)

<sup>3</sup>Programa de Pós-Graduação em Zootecnia – FCAV - Unesp, Jaboticabal. Bolsista do CNPq.

<sup>4</sup>Instituto de Ciências Biológicas – UFMG, Belo Horizonte.

<sup>5</sup>Faculdade de Ciência Agrárias e Veterinárias – Unesp, Jaboticabal.

**Resumo:** O desequilíbrio de ligação (LD) de um modo geral refere-se à associação não aleatória de alelos de diferentes locos, que varia em função da amostra da população e da geração estudada. Sendo assim o objetivo deste estudo foi comparar as estimativas das estatísticas de desequilíbrio de ligação com base na distância física em amostra de bovinos da raça Guzerá. Neste estudo, foram utilizados vinte e cinco touros pertencentes ao Programa Nacional de Melhoramento de Guzerá para Leite. A genotipagem foi realizada utilizando *High Density Bovine SNP BeadChip* do *Illumina Infinium HD Assay*<sup>®</sup>. O desequilíbrio de ligação entre dois SNPs foi avaliado utilizando o  $r^2$  e o valor absoluto de  $D'$ . A média geral foi de 0,25 para  $r^2$  e 0,81 para  $|D'|$ . O  $|D'|$  tendeu a superestimar as estatísticas de DL para diferentes distâncias. O DL observado para os marcadores presentes nesse *chip*, principalmente para aqueles distanciados a menos de 60 Kb, podem ser adequados à estimação de valores genéticos genômicos para animais da raça Guzerá.

**Palavras-chave:** zebu, SNP, seleção genômica, genoma

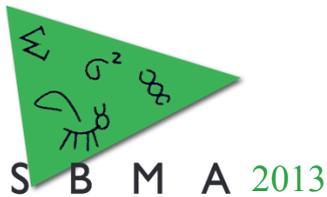
### Linkage disequilibrium with a high density panel in the bovine genomes of Guzerá breed

**Abstract:** Linkage disequilibrium (LD) generally refers to the non-random association of alleles of different loci, which varies according to the population sampled and generation studied. Therefore the objective of this study was to compare the estimates of the statistics for linkage disequilibrium based on the physical distance in bovines of Guzerá breed. In this study, we used twenty-five bulls belonging to the National Program for the Improvement of Guzerá Dairy Cattle. Genotyping was performed using High Density Bovine BeadChip SNP of Illumina's Infinium HD Assay<sup>®</sup>. Linkage disequilibrium between two SNP was evaluated using the absolute value of  $r^2$  and  $D'$ . The overall average was 0.25 and 0.81 for, respectively,  $r^2$  and  $|D'|$ . The  $|D'|$  tended to overestimate the LD statistics for different distances. The LD observed for markers present in this chip, especially for those distant less than 60 kb, may be suitable to estimation of genomic breeding values for Guzerá breed animals.

**Keywords:** zebu, SNP, genomic selection, genome

### Introdução

O desequilíbrio de ligação (DL) de um modo geral refere-se à associação não aleatória de alelos de diferentes locos. Este fenômeno está relacionado à ocorrência de recombinações históricas ocorridas antes mesmo do início da genotipagem, sendo propriedade de uma população e da geração em que é medida. A exploração deste conhecimento é ampla e mais recentemente utilizada nas análises de seleção genômica. Estas análises exploram o desequilíbrio de ligação entre os marcadores e genes de interesse econômico para identificar regiões genômicas com potencial, ou sob influência, de seleção, bem como no estabelecimento de mapas de desequilíbrio importantes para o estudo da diversidade genética dentro e entre raças (McKay et al., 2007). Os marcadores moleculares do tipo polimorfismo de nucleotídeo único (SNP) têm sido amplamente utilizados em estudos de DL (ZHANG et al., 2012). Sendo assim, o objetivo deste estudo foi comparar as estimativas das estatísticas do desequilíbrio de ligação em função da distância física em bovinos da raça Guzerá, usando marcadores SNP.



## X Simpósio Brasileiro de Melhoramento Animal

Uberaba, MG – 18 a 23 de agosto de 2013

### Material e Métodos

Amostras de vinte e cinco touros pertencentes ao Programa de Melhoramento do Guzerá para Leite foram utilizadas neste estudo. A genotipagem foi realizada utilizando o *High Density Bovine SNP BeadChip* do *Illumina Infinium HD Assay*<sup>®</sup>. O *High Density Bovine SNP BeadChip* contém mais de 770 mil marcadores espaçados a uma distância média de 3.43 kb entre os marcadores. Foram utilizados apenas marcadores SNP dos cromossomos autossômicos e X. Para avaliar a qualidade da genotipagem foram utilizados os critérios de  $GC10 > 0,30$  (percentil 10 da distribuição do *Gencall*),  $call\ rate > 0,90$  (para amostra e SNP),  $MAF > 0,05$  e  $het\ excess > -0,7$  e  $het\ excess < 0,7$ . Após a edição, o conjunto de dados permaneceu com 25 animais, porém 448.316 marcadores autossômicos. O desequilíbrio de ligação entre dois SNP foi avaliado utilizando-se o  $r^2$  e o valor absoluto de  $D'$  (Hill, 1981). As medidas de desequilíbrio de ligação ( $r^2$  e  $|D'|$ ) foram calculadas para todos os pares de marcadores em cada cromossomo, utilizando o pacote *snpStats*, disponível no sistema computacional livre R.

### Resultados e Discussão

A média geral foi 0,25 para  $r^2$  e 0,81 para  $|D'|$ . Ao se analisar a média por cromossomo, o  $r^2$  variou de 0,20 a 0,28, e  $|D'|$  de 0,78 a 0,84. Silva et al. (2010), em análise de DL em 25 touros da raça Gir, usando painel de 54.000 SNP, obteve valor semelhante para a média de  $r^2$  (0,21), porém valor menor para média de  $|D'|$  (0,68). Por cromossomo, esses autores reportaram valores semelhantes de  $r^2$  (de 0,20 a 0,24) e valores ligeiramente mais baixos de  $|D'|$  (de 0,60 a 0,72). Os resultados encontrados neste trabalho também foram semelhantes aos encontrados para zebuínos por McKay et al. (2007). Os resultados, ligeiramente mais altos do que os da literatura citada, podem ser atribuídos ao baixo número de animais amostrados, o que pode ter colaborado para uma leve superestimação das medidas do desequilíbrio de ligação. Para analisar o declínio das medidas de LD de acordo com a distância física entre os marcadores, os pares de SNP foram classificados em intervalos de pares de base, e as médias do  $r^2$  e do  $|D'|$  foram estimadas por cromossomo e para todo o genoma (Tabela 1). Como esperado os valores de DL diminuíram com o aumento da distância física entre os marcadores. Verificou-se que a diminuição foi mais evidente para  $R^2$  do que para  $|D'|$ . Pela tabela 1, também foi possível observar maior diferença nas médias do  $r^2$  que nas médias do  $|D'|$  de acordo com a distância, indicando que a estatística  $r^2$  foi mais sensível para mensuração do LD da raça que o  $|D'|$ . Entretanto, o desvio padrão das estimativas de  $r^2$  foi mais alto que a do  $|D'|$  em distâncias menores. Isto, possivelmente, é devido ao fato da estatística  $|D'|$  estar superestimada, proporcionando valores próximos ao limiar superior (1), o que reduz a variância dos intervalos iniciais. Desse modo, conforme aumentou o intervalo, as estimativas diminuíram, possibilitando uma melhor estimativa do desvio padrão. Tendência oposta foi observada para o  $r^2$ , para o qual o desvio padrão diminuiu com o aumento do intervalo. Essa diminuição é esperada, uma vez que com o aumento do intervalo aumenta a frequência das estimativas iguais à zero, reduzindo a variação. Esses resultados indicam que o  $r^2$  é a medida mais adequada para analisar o DL que o  $|D'|$ , embora para intervalos maiores apresente algumas limitações. Níveis altos de  $r^2$  foram observados para distâncias físicas menores que 10 Kb e níveis moderados de  $r^2$  (0,20 – 0,30) foram observados para distâncias acima de 10 Kb e baixo de 70 Kb. Níveis baixos foram observados para distâncias maiores do que 70 Kb. Até 5 Kb, aproximadamente 50% dos marcadores apresentaram média de DL ( $r^2$ ) maior que 0,3. Esses valores foram mais baixos aos encontrados por Sargolzaei et al. (2008), que observaram mais de 60 % dos marcadores apresentaram média maior que 0,3, em uma distância entre 0 e 1 Kb, estudando 821 touros Holandês com 5564 SNP. Em uma distância de 90 Kb a 100 Kb, a média de DL foi 0,17. McKay et al. (2007) estimou o DL entre todos os pares de marcadores em uma distância de 100Kb e encontrou médias variando de 0,15 a 0,20 para seis raças taurinas e duas zebuínas. Os níveis de desequilíbrio de ligação entre os segmentos cromossômicos determinados por esses marcadores, e, conseqüentemente, entre os SNP e os QTL, seriam suficientes para utilização na estimativa de valores genéticos genômicos, principalmente em distâncias menores que 60 Kb ( $r^2$  moderado a alto). Para todos os cromossomos, não houve muita variação do  $r^2$  (0,21-0,26), indicando cobertura genômica suficiente por esses SNP. Foi observada tendência, embora pequena, do  $r^2$  diminuir com a diminuição do tamanho do cromossomo, ou seja, de diminuição da taxa de recombinação com o aumento do comprimento do cromossomo. Resultados semelhantes foram relatados por Bohmanova et al. (2010), ao estimar LD para animais da raça Holandesa.

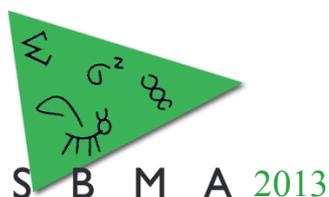


Tabela 1. Desequilíbrio de ligação ( $r^2$  and  $|D'|$ ) entre pares (N) de SNP\* intervalados a diferentes distâncias para todos os cromossomos autossômicos.

Distância (kb)	N	$r^2$ Média±SD	$r^2$ Mediana	$ D' $ Média±SD	$ D' $ Mediana	% $r^2 > 0,3$ **	% $ D'  > 0,8$ **
0 – 1	58908	0,56±0,40	0,55	0,96±0,15	1,00	62	93
1 – 2	106545	0,51±0,40	0,44	0,95±0,18	1,00	58	91
2 – 3	110692	0,50±0,39	0,35	0,94±0,19	1,00	54	90
3 – 4	123518	0,44±0,38	0,31	0,93±0,20	1,00	51	88
4 – 5	113827	0,42±0,37	0,28	0,92±0,21	1,00	49	87
5 – 10	530850	0,38±0,36	0,23	0,90±0,24	1,00	44	83
10 – 20	989455	0,32±0,33	0,18	0,87±0,27	1,00	38	78
20 – 30	943887	0,28±0,31	0,14	0,84±0,29	1,00	32	74
30 – 40	920628	0,25±0,29	0,12	0,82±0,30	1,00	29	70
40 – 50	905601	0,23±0,27	0,11	0,80±0,31	1,00	26	68
50 – 60	894763	0,21±0,26	0,10	0,79±0,32	1,00	24	66
60 – 70	884148	0,20±0,25	0,09	0,78±0,32	1,00	22	64
70 – 80	878477	0,18±0,24	0,08	0,77±0,33	1,00	20	63
80 – 90	874729	0,17±0,23	0,08	0,76±0,33	1,00	19	62
90 – 100	871026	0,17±0,22	0,07	0,75±0,33	1,00	18	61

\*SNP: polimorfismo de nucleotídeo único; \*\*Porcentagem de pares de SNP com  $r^2 > 0,3$  ou  $|D'| > 0,8$

### Conclusões

O nível de desequilíbrio de ligação observado para os marcadores SNP presente no *High Density Bovine SNP BeadChip*, principalmente os SNP intervalados a distâncias menores que 60 Kb, mostrou-se suficiente para auxílio à estimação de valores genéticos genômicos para animais da raça Guzerá.

### Literatura citada

- MCKAY, S.D., SCHNABEL, R.D., MURDOCH, B.M. et al. Whole genome linkage disequilibrium maps in cattle. **BMC Genet.**, v.8, p.74, 2007.
- SILVA, C.R., NEVES, H.H.R., QUEIROZ, S.A. et al. Extent of linkage disequilibrium in Brazilian Gyr dairy cattle based on genotypes of AI sires for dense SNP markers. In: Proceedings of the 9th World Congress on Genetics Applied to Livestock Production: Leipzig, Germany, 2010.
- HILL, W. G. Estimation of effective population size from data on linkage disequilibrium. **Genet. Res.**, v.38, p.209-216, 1981.
- SARGOLZAEI, M., SCHENKEL, F.S., JANSEN, G.B., SCHAEFFER L.R.. Extent of linkage disequilibrium in Holstein cattle in North America. **J. Dairy Sci.**, v.91, p.2106-2117, 2008.
- BOHMANOVA, J., SARGOLZAEI, M., SCHENKEL, F.. Characteristics of linkage disequilibrium in North American Holsteins. **BMC Genomics**, v.11, p.421, 2010.
- ZHANG, H., WANG, Z., WANG, S., LI, H.. Progress of genome wide association study in domesticated animals. **J. Anim. Sci. Biotech.** 2012. Disponível em: <<http://www.biomedcentral.com/content/pdf/2049-1891-3-26.pdf>> Acessado em: 24 de junho de 2013.