

XII Simpósio Brasileiro de Melhoramento Animal Ribeirão Preto, SP – 12 e 13 de junho de 2017

Caracterização do desequilíbrio de ligação em uma população de bovinos da raça Nelore¹

Sabrina Kluska^{2*}, Rafael Lara Tonussi², Rafael Medeiros de Oliveira Silva⁷, Bianca Ferreira Olivieri², Fabieli Loise Braga Feitosa², Elisa Peripolli², Marcos Viniccius Antunes Lemos², Mariana Piatto Berton², Hermenegildo Lucas Justino Chiaia², Angelica Simone Cravo Pereira³, Raysildo Barbosa Lôbo⁴, Luiz Antônio Bezerra⁵, Cláudio de Ulhoa Magnabosco⁶, Ignácio Aguilar⁸, Fernando Di Croce⁹, Jason Osterstock⁹ e Fernando Baldi²

¹Fundação de Pesquisa de São Paulo (FAPESP, concessão 2013/25910-0);

²Departamento de Ciência Animal, Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, São Paulo, Brasil;

³Departamento de Nutrição e Produção Animal, Faculdade de Ciência Animal e Engenharia de Alimentos, Pirassununga, Brasil;

⁴Associação Nacional de Criadores e Pesquisadores (ANCP), Ribeirão Preto, Brasil;

⁵Departamento de Genética, Escola de Medicina de Ribeirão Preto, Ribeirão Preto, Brasil;

⁶Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Distrito Federal, Brasil;

⁷Departamento de Ciência Animal e laticínios, Universidade da Geórgia, Athens, Estados Unidos;

⁸Departamento de Melhoramento Animal, Instituto Nacional de Pesquisa e Agricultura, Las Brujas, Uruguay;

⁹Zoetis, 333 Portage St., KZO300-210SE, Kalamazoo, MI 49007, Estados Unidos

*Autor correspondente: sabrinakluska@gmail.com

Resumo: Tendo em vista a importância da estimação do desequilíbrio de ligação para a seleção genômica, o objetivo deste estudo foi estimar o desequilíbrio de ligação de uma população de bovinos da raça Nelore participantes do programa de melhoramento da ANCP. Foram utilizadas informações de 9.459 animais genotipados com um painel de alta densidade, totalizando 735.044 SNP's, antes do controle de qualidade. A estimação do desequilíbrio de ligação foi realizada através do programa SNP1101. Os valores de LD observados para os cromossomos autossômicos variaram de 0,18 a 0,25. Para marcadores distanciados até 1 Kb a média de r^2 foi de 0,53 e para marcadores distanciados entre 90 e 100 Kb 0,14. Para MAF a média variou de 0,23 a 0,25, considerando MAF mínimo de 5%. Os resultados obtidos neste estudo indicam que, a densidade de marcadores utilizados foi capaz de detectar altos níveis de LD. Adicionalmente, conclui-se que marcadores distanciados até 50 Kb ainda detectam consideráveis níveis de LD.

Palavras-chave: cromossomo, densidade de marcadores, genotipagem.

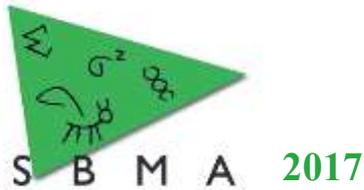
Characterization of linkage disequilibrium in a population of Nelore cattle

Abstract: Considering the importance of estimating linkage disequilibrium for genomic selection, the objective of this study was to estimate the linkage disequilibrium in a population of Nelore cattle participating in the ANCP breeding program. Information from 9,396 genotyped animals with a high density panel, totaling 735,044 SNP's before quality control were used. The estimation of linkage disequilibrium (LD) was performed using the SNP1101 program. The mean LD values observed for the autosomal chromosomes ranged from 0.18 to 0.25. For markers distanced lower than 1 Kb the r^2 mean was 0.53, and for markers distanced between 90 and 100 Kb was 0.14. For MAF, the mean ranged from 0.23 to 0.25, a minimum MAF of 0.05 was considered. The results obtained in this study indicated that the density of markers used was able to detect high levels of LD. Additionally, for markers distanced up to 50 Kb, considerable levels of LD was detected.

Keywords: chromosome, genotyping, marker density.

Introdução

O desequilíbrio de ligação (LD) é definido como a associação não aleatória entre alelos de diferentes locus gênicos, e é calculado entre a diferença das frequências observadas e esperadas dos háplotipos, considerando a segregação independente dos alelos (WEISS & CLARK, 2002). Na seleção genômica a estimação do desequilíbrio de ligação deve ser uma das primeiras análises a serem realizadas, a fim de investigar a informatividade dos marcadores utilizados. De acordo com Resende et al., (2008) a variância genética de uma característica de caráter quantitativo pode ser explicada pela presença dos marcadores em desequilíbrio de ligação com os QTL, sejam esses de pequenos, ou de grande efeito, uma vez que, somente os marcadores em desequilíbrio são utilizados na determinação dos fenótipos via genótipo. Diante disso, o objetivo do presente trabalho foi quantificar o desequilíbrio de ligação no genoma de uma população de bovinos da raça Nelore utilizando um painel de SNP's de alta densidade (Illumina High Density Bovine SNP BeadChip®).



Material e Métodos

Foram utilizadas informações de 9.459 animais genotipados da raça Nelore, participantes do programa de melhoramento Nelore Brasil, da Associação Nacional de Criadores e Pesquisadores (ANCP). Um total de 8.519 animais foram genotipados usando o painel de baixa densidade (Clarifide Nelore 2.0 - 22k), e posteriormente imputados a painéis com 54k e 777k, usando o software Fimpute, o qual apresenta alta eficiência para imputação de Genótipos. Para a imputação, uma população de referência com 940 animais genotipados com o painel Illumina High-Density Bovine BeadChip (777.962 SNP) foi utilizada. O controle de qualidade dos marcadores foi realizado em duas etapas. Primeiramente, um filtro de *call rate* (90%) foi aplicado nos arquivos de 777k e 22k separadamente. Em seguida, após a imputação dos genótipos, na segunda etapa, filtros para menores frequências de alelos (MAF) inferiores a 0,05 e valores de Equilíbrio de Hardy-Weinberg (HWE) inferiores a 0,01, foram aplicados às amostras. Além disso, foram excluídos marcadores com posição redundante e localizados em cromossomos não autossômicos. Após o controle de qualidade, permaneceram 463.124 SNPs e 9.396 animais na análise. A estimação do desequilíbrio de ligação (LD) médio, número de SNP's, menor frequência de alelos (MAF) e número de combinações de pares de SNP's em cada cromossomo foi realizadas através do *software* SNP1101 (Sargolzaei, 2014). O desequilíbrio de ligação entre dois SNP foi estimado utilizando o r^2 , o qual foi calculado como:

$$r^2 = \frac{(freq.AB * freq.ab - freq.Ab * freq.aB)^2}{(freq.A * freq.a * freq.b * freq.B)}$$

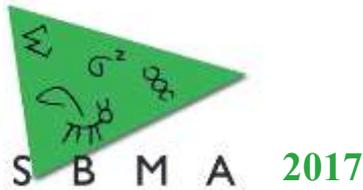
Em que $freq.A$, $freq.a$, $freq.B$ e $freq.b$ são as frequências dos alelos A, a, B e b, respectivamente, e $freq.AB$, $freq.ab$, $freq.aB$ and $freq.Ab$ são as frequências dos háplotipos Ab, ab, aB and Ab na população estudada. Se os dois locus são independentes a frequência do háplotipo AB ($freq.AB$) é calculada como o produto entre $freq.A$ e $freq.B$ e o r^2 é admitido como zero. Se a frequência do háplotipos é superior ou inferior ao valor esperado indica que estes dois loci em particular tendem a segregar em conjunto e estão em LD.

Resultados e Discussão

Nesta pesquisa 63% dos marcadores (463.124) atenderam os critérios de controle de qualidade e foram incluídos na análise. A média de LD no genoma estudado foi de 0,22, sendo o maior valor observado para o BTA5 (0,25) e o menor para o BTA27 (0,18) (Tabela 1). Para todos os cromossomos estudados o coeficiente de variação para r^2 foi alto (Tabela 1). Valores inferiores de r^2 foram observados por Espigolan et al., (2009) também para a raça Nelore. Os autores estimaram o desequilíbrio de ligação médio, variando de 0,003 (BTA29) a 0,21(BTA13) nos cromossomos autossômicos, com uma população de 795 animais genotipados. A estrutura das populações estudadas é um dos fatores que podem afetar a detecção do desequilíbrio de ligação, sendo esta a principal causa da variação entre estes trabalhos. Além disso, outro fator que interfere na identificação do LD é a densidade de marcadores utilizados.

Tabela 1. Resumo dos marcadores de SNP analisados e desequilíbrio de ligação médio (r^2) entre marcadores sintéticos adjacentes obtidos para cada cromossomo autossomo do genoma de uma população de bovinos da raça Nelore.

BTA	SNP (N)	Comparações (n)	Média $r^2 \pm SD$	CV(%) r^2	Média MAF $\pm SD$
1	30.842	684.993	0,23 \pm 0,28	122,30	0,24 \pm 0,13
2	24.990	534.818	0,24 \pm 0,29	120,90	0,24 \pm 0,13
3	22.363	477.224	0,23 \pm 0,28	122,14	0,25 \pm 0,13
4	20.846	436.293	0,21 \pm 0,27	127,93	0,25 \pm 0,13
5	20.509	407.113	0,25 \pm 0,30	119,10	0,25 \pm 0,13
6	24.915	580.004	0,23 \pm 0,28	118,71	0,25 \pm 0,13
7	21.004	464.469	0,23 \pm 0,28	120,80	0,25 \pm 0,13
8	22.506	496.161	0,23 \pm 0,28	119,05	0,26 \pm 0,13
9	21.147	479.995	0,23 \pm 0,28	121,98	0,25 \pm 0,13
10	19.423	429.148	0,22 \pm 0,28	129,16	0,23 \pm 0,13
11	18.916	399.333	0,23 \pm 0,29	123,39	0,24 \pm 0,13



XII Simpósio Brasileiro de Melhoramento Animal
Ribeirão Preto, SP – 12 e 13 de junho de 2017

12	15.719	342.997	0,22 ±0,27	127,43	0,24±0,13
13	14.542	290.500	0,21 ±0,27	126,57	0,25±0,13
14	17.770	412.753	0,24±0,27	114,70	0,26±0,13
15	15.301	327.057	0,20 ±0,26	128,10	0,24±0,13
16	14.830	321.580	0,23 ±0,28	121,39	0,24±0,13
17	13.896	303.872	0,22 ±0,27	124,14	0,24±0,13
18	11.808	261.965	0,22 ±0,28	126,14	0,24±0,13
19	10.878	231.717	0,20±0,27	133,04	0,24±0,13
20	13.164	283.451	0,22±0,27	124,70	0,25±0,13
21	12.944	282.440	0,22±0,27	122,27	0,24±0,13
22	10.954	242.147	0,20 ±0,26	129,59	0,24±0,13
23	9.831	232.201	0,19 ±0,25	132,16	0,25±0,13
24	11.401	245.787	0,22 ±0,27	123,74	0,25±0,13
25	7.757	164.882	0,20 ±0,27	131,94	0,24±0,13
26	10.023	229.450	0,20 ±0,26	131,62	0,24±0,13
27	8.336	178.951	0,18 ±0,25	137,86	0,25±0,13
28	7.647	160.276	0,21 ±0,26	123,82	0,25±0,13
29	8.862	187.224	0,19 ±0,26	134,98	0,23±0,13

BTA: Cromossomo; SNP: Polimorfismo de nucleotídeo único; r^2 : desequilíbrio de ligação; MAF: menor frequência de alelos.

Os resultados para MAF obtidos no presente trabalho (Tabela 1) são similares aos descritos por McKay et al., (2007) para a mesma frequência, em bovinos da raça Nelore. O valor de r^2 médio obtido no presente estudo para SNP's distanciados até 1 Kb foi de 0,53. Marcadores distanciados até 50 kb apresentaram r^2 de 0,20, o que de acordo com Calus et al., (2008) resulta em uma acurácia de 0,82 na seleção genômica. Valores de r^2 acima de 0,20 foram encontrados por Espigolan et al., (2013) para distancias menores entre SNP'S (20 a 30 kb). O alto valor de r^2 para os SNP's mais distantes indica que menores densidades de marcadores podem ser usadas com eficiência na seleção genômica, sem reduzir a acurácia, devido à informatividade dos marcadores.

Conclusão

A média de desequilíbrio de ligação por cromossomo observada neste estudo foi alta, indicando que a seleção genômica poderia ser aplicada nesta população, uma vez que os marcadores utilizados foram informativos. Adicionalmente, conclui-se que densidades de marcadores mais baixas também podem ser utilizadas com esse intuito, uma vez que conseguem captar elevados níveis de LD nesta população.

Literatura citada

- CALUS, M. P.L., MEUWISSEN, T. H. E.; DE ROOS, A. P. W., VEERKAMP, R. F. Accuracy of genomic selection using different methods to define haplotypes. **Genetics** 178:553–561, 2008.
- ESPIGOLAN, R.; BALDI, F.; BOLIGON, A.A.; SOUZA, F.R.P.; GORDO, D.G.M.; TONUSSI, R.L.; CARDOSO, D.F.; OLIVEIRA, H. N.; TONHATI, H.; SARGOLZAEI, M.; SCHENKEL, F.S.; CARVALHEIRO, R.; FERRO, J.A.; ALBUQUERQUE, L.G. Study of whole genome linkage disequilibrium in Nelore cattle. **BMC Genomic**, 14:305, 2013.
- MCKAY, D.S.; SCHNABEL, R.D.; MURDOCH, B.M.; MATUKUMALLI, L.K.; AERTS, J.; COPPIETERS, W.; CREWS, D.; DIAS NETO, E.; GILL, C.A.; GAO, C.; MANNEN, H.; STOTHARD,P.; WANG, Z.; VAN TASSEL, C.P.; WILLIAMS, J.L.; TAYLOR, J.F.; MOORE, S.S. Whole genome linkage disequilibrium maps in cattle. **BMC Genomic**, 8:74, 2007.
- RESENDE, M.D.V de.; LOPES, P.S.; SILVA, R.L da.; PIRES, I.E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa Florestal Brasileira**, Colombo, n.56, p.63-77, jan/jun, 2008.
- SARGOLZAEI, M. 2014. **SNP1101 User's Guide. Version 1.0.0**. Mehdi Sargolzaei, Semex Alliance, Ontario, Canada.
- WEISS, K.M.; CLARK, A.G. Linkage disequilibrium and the mapping of complex human traits. **Trends Genetic**, v.18, p.19–24, 2002.