

IX Simpósio Brasileiro de Melhoramento Animal

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

Caracterização de haplótipos e identificação de assinaturas de seleção em bovinos da raça Nelore¹

Adriana Luiza Somavilla², Roberto Hiroshi Higa³, Maurício de Alvarenga Mudadu⁴, Mauricio Mello de Alencar^{4,5}, Luciana Correia de Almeida Regitano^{4,5}, Rede BifeQuali⁶

¹Parte da dissertação de mestrado do primeiro autor

²Programa de Pós-Graduação em Genética e Melhoramento Animal – UNESP/Jaboticabal. Bolsista CNPq. e-mail: adri_somavilla@yahoo.com.br

³Pesquisador Embrapa Informática Agropecuária – Campinas/SP

⁴Pesquisador Embrapa Pecuária Sudeste – São Carlos/SP

⁵Pesquisador CNPq

⁶<http://www.macroprograma1.cnptia.embrapa.br/bifequali>

Resumo: A identificação de regiões do genoma que controlam características de interesse econômico e que estejam sob seleção é possível por conta da disponibilidade da genotipagem em plataformas de altas densidades de polimorfismos de único nucleotídeo (SNPs). A busca por assinaturas de seleção avalia, principalmente, altas e não usuais frequências alélicas dos loci adjacentes às mutações favoráveis. Com o objetivo de detectar evidências de assinaturas de seleção recente no genoma de bovinos da raça Nelore, foram genotipados 796 indivíduos com o chip *Illumina BovineHD BeadChip*, que possui mais de 777 mil SNPs e, posteriormente, inferida a fase de ligação dos SNPs e a reconstrução dos blocos de desequilíbrio de ligação (haplótipos). A detecção de assinaturas de seleção recentes foi realizada por meio da aplicação da metodologia EHH (homozigose do haplótipo estendido). 330.262 SNPs entraram na formação dos 68.054 haplótipos identificados, cobrindo 42% do genoma, sendo que a maior parte deles apresentou até 10 kb de comprimento e foram formados por 3 marcadores. Além disso, foram detectadas regiões de possíveis assinaturas de seleção com diferentes significâncias: 2.103 ($P < 0,01$), 269 ($P < 0,001$) e 31 ($P < 0,0001$).

Palavras-chave: *Bos indicus*, desequilíbrio de ligação, genotipagem, REHH, snps

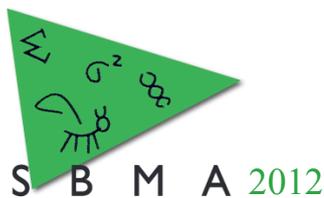
Haplotype block structure and detection of selection signatures in Nelore cattle

Abstract: Identification of genomic regions controlling complex traits under selection is possible due to the availability of genotyping on high density platforms of single nucleotide polymorphisms (SNPs). Genome scan for selection signatures evaluates, mainly, the fast increase in allele frequencies of adjacent loci to favorable mutations. With the purpose to detect evidence of selection signatures on Nelore cattle genome, 796 individuals were genotyped with the *Illumina BovineHD BeadChip*, which has more than 777,000 SNPs and, after that, with the fully phased SNPs, haplotypes were reconstructed. Detection of recent selection signatures was performed by applying EHH statistic (Extended Haplotype Homozygosity). 68,054 haplotypes were reconstructed with 330,262 SNPs, allowing 42% of genome coverage, being the prevalent blocks up to 10 kb long and formed by 3 markers. Furthermore, we detected regions of possible signatures of selection with different significances: 2,103 ($P < 0.01$), 269 ($P < 0.001$) and 31 ($P < 0.0001$).

Keywords: *Bos indicus*, genotyping, linkage disequilibrium, REHH, snps

Introdução

O rebanho bovino brasileiro, composto principalmente por animais da raça Nelore (*Bos indicus*), vem sendo selecionado há décadas, por diversos programas de melhoramento genético. Esses programas focaram os esforços na seleção de características de crescimento e reprodutivas, por serem importantes para a eficiência econômica e também por apresentarem herdabilidades que variam entre moderada a alta (Laureano et al., 2011). A identificação de regiões do genoma que controlam esses caracteres e que estejam sob seleção pela análise de marcadores em escala genômica, como é o caso dos *chips* de polimorfismos de único nucleotídeo (SNPs) *Illumina BovineHD BeadChip*. Assim, a varredura do



genoma por essas assinaturas de seleção avalia, principalmente, o rápido aumento nas frequências alélicas dos loci adjacentes às mutações favoráveis (Qanbari et al., 2010).

Com base nisso, o objetivo desse trabalho foi detectar evidências de assinaturas de seleção recente no genoma de bovinos da raça Nelore.

Material e Métodos

Foram genotipados 796 indivíduos da raça Nelore (764 novilhos machos e 32 touros) com o chip *Illumina BovineHD BeadChip*, que possui mais de 777 mil SNPs. Os dados foram submetidos a um controle de qualidade para eliminar marcadores e amostras com erros de genotipagem e SNPs nos quais o alelo menos frequente apresentasse menos de 1% de frequência (MAF<0,01) na população, ou seja, considerados não polimórficos e, conseqüentemente, não informativos.

A inferência da fase de ligação dos SNPs, para posterior formação dos blocos de desequilíbrio de ligação (haplótipos) foi feita por meio do software AlphaPhase (Hickey et al., 2011). A detecção de assinaturas de seleção recentes foi realizada por meio do programa Sweep 1.1 (Sabeti et al., 2002), que aplica a metodologia EHH (homozigose do haplótipo estendido), definida como a probabilidade que 2 cromossomos escolhidos aleatoriamente e que carreguem o haplótipo de interesse sejam idênticos por descendência para todo o intervalo da região do core a uma distância “x”, analisando a estrutura dos haplótipos de indivíduos de um único grupo genético amostrados aleatoriamente em uma população. Como em bovinos, assim como outras espécies, as taxas de recombinação não são constantes ao longo do genoma, os autores propuseram uma correção, a homozigose relativa do haplótipo estendido (REHH), definida como:

$$EHH_t / \overline{EHH}$$

sendo que,

$$EHH_t = \frac{\sum_{i=1}^s \binom{e_{ti}}{2}}{\binom{c_t}{2}} \quad \text{e} \quad \overline{EHH} = \frac{\sum_{j=1, j \neq t}^n \left[\sum_{i=1}^s \binom{e_{ij}}{2} \right]}{\sum_{i=1, i \neq t}^n \binom{c_i}{2}}$$

onde c_t é o número de amostras de um haplótipo núcleo particular (em teste), e_{ti} é o número de amostras de um haplótipo estendido particular, s é o número de haplótipos estendidos únicos e n é número de diferentes haplótipos núcleos (em relação ao haplótipo teste - t).

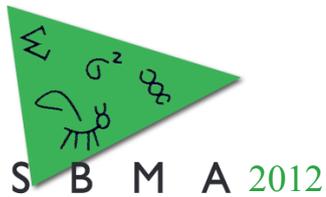
Para determinar a significância dos valores de REHH, os haplótipos formados foram ordenados de acordo com a sua frequência na população e cada um foi comparado com os outros igualmente frequentes. Para obter uma distribuição normal dos valores de P , utilizou-se a transformação $-\text{Log}_{10}$.

Resultados e Discussão

Após o controle de qualidade, restaram 560.565 SNPs com espaçamento médio de 4,5 kb e MAF média de 0,26, sendo que 58% (330.262) entraram na formação dos 68.054 haplótipos, cobrindo 42% do genoma. Somente os blocos com frequência e homozigosidade maiores que 30% foram incluídos nas etapas posteriores, pois valores inferiores são indicativos de eventos de recombinação (Sabeti et al., 2002). A Figura 1 apresenta a caracterização dos haplótipos formados, sendo que a maior parte deles apresentou até 10 kb de comprimento e foi formada por 3 SNPs.

Qanbari et al. (2010) caracterizaram os blocos no genoma de bovinos da raça Holstein (*Bos taurus*) e também observaram haplótipos formados por 3 marcadores, porém, com tamanhos variando de 100 a 150 Kb. Essa diferença entre Holstein e Nelore pode ser explicada pelo menor desequilíbrio de ligação no genoma dos indivíduos *Bos indicus* (Villa-Angulo et al., 2009), resultando em maiores taxas de recombinação e menores blocos conservados.

Após a aplicação do teste EHH, para ambos os lados de cada haplótipo, foram detectadas regiões de possíveis assinaturas de seleção, de acordo com a frequência e homozigosidade do haplótipo estendido, com diferentes significâncias, sendo 2.103 para $P < 0,01$, 269 para $P < 0,001$ e 31 para



$P < 0,0001$. Essas regiões podem ser usadas para entender o papel biológico nos quais os genes nelas contidos estão envolvidos (Stella et al., 2010).

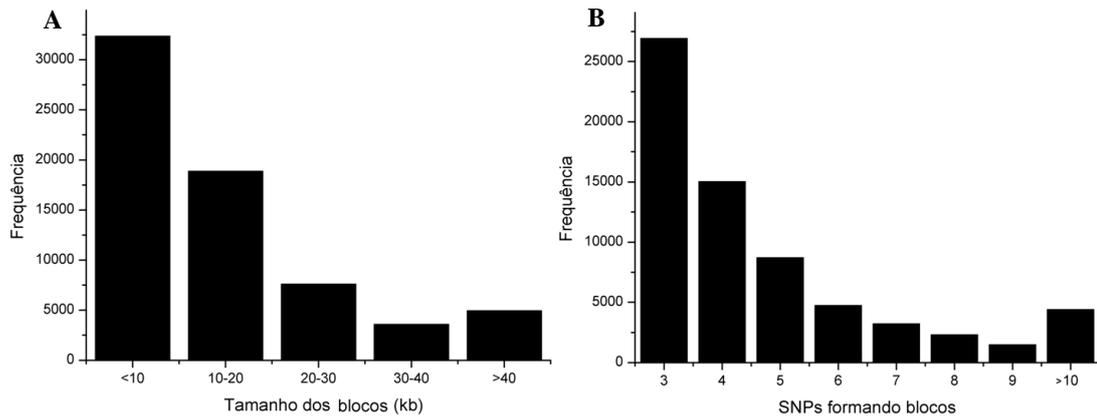


Figura 1 Distribuição do tamanho dos haplótipos (A) e número de SNPs formando os blocos (B) no genoma de bovinos da raça Nelore.

Conclusões

Os haplótipos formados no genoma de bovinos Nelore apresentam tamanho relativamente pequeno, provavelmente pelo menor desequilíbrio de ligação existente nesse grupo genético. Além disso, foram identificadas mais de 2.000 possíveis regiões de assinaturas ao longo do genoma, que podem ser analisadas detalhadamente para anotação de quais genes elas contêm e quais funções eles desempenham no organismo dos animais.

Agradecimentos

À Embrapa Pecuária Sudeste pela concessão dos dados e ao CNPq pelo apoio financeiro.

Literatura citada

- LAUREANO, M.M.M.; BOLIGON, A.A.; COSTA, R.B. et al. Estimativas de herdabilidade e tendências genéticas para características de crescimento e reprodutivas em bovinos da raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.63, n.1, p.143-152, 2011.
- QANBARI, S.; PIMENTEL, E.C.G.; TETENS, J. et al. A genome-wide scan for signatures of recent selection in Holstein cattle. **Animal Genetics**, v.41, p.377-389, 2010.
- HICKEY, J.M.; KINGHORN, B.P.; TIER, B. et al. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. **Genetics Selection Evolution**, v.43, n.12, doi:10.1186/1297-9686-43-12, 2011.
- SABETI, P.C.; REICH, D.E.; HIGGINS, J.M. et al. Detecting recent positive selection in the human genome from haplotype structure. **Nature**, v.419, p.832-837, 2002.
- VILLA-ANGULO, R.; MATUKUMALLI, L.K.; GILL, C.A. et al. High-resolution haplotype block structure in the cattle genome. **BMC Genetics**, v.10, n.19, doi:10.1186/1471-2156-10-19, 2009.
- STELLA, A.; AJMONE-MARSAN, P.; LAZZARI, B. et al. Identification of selection signatures in cattle breeds selected for dairy production. **Genetics**, v.185, p.1451-1461, 2010.