

**UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**IDENTIFICAÇÃO DE ASSINATURAS DE SELEÇÃO E
ANOTAÇÃO DO GENOMA EM BOVINOS DA RAÇA
NELORE**

**Adriana Luiza Somavilla
Zootecnista**

JABOTICABAL – SÃO PAULO - BRASIL
Julho de 2012

**UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**IDENTIFICAÇÃO DE ASSINATURAS DE SELEÇÃO E
ANOTAÇÃO DO GENOMA EM BOVINOS DA RAÇA
NELORE**

Adriana Luiza Somavilla

**Orientador: Prof. Dr. Maurício Mello de Alencar
Co-orientadora: Dr^a. Luciana Correia de Almeida Regitano**

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Genética e Melhoramento Animal.

**JABOTICABAL – SÃO PAULO - BRASIL
Julho de 2012**

S693i **Somavilla, Adriana Luiza**
Identificação de Assinaturas de Seleção e Anotação do Genoma
em Bovinos da Raça Nelore / Adriana Luiza Somavilla. – –
Jaboticabal, 2012
iv, 64 f. ; il. ; 28 cm

Dissertação (mestrado) - Universidade Estadual Paulista,
Faculdade de Ciências Agrárias e Veterinárias, 2012
Orientador: Mauricio Mello de Alencar
Co-orientadora: Luciana Correia de Almeida Regitano
Banca examinadora: Henrique Nunes de Oliveira, Luiz Lehmann
Coutinho
Bibliografia

1. Bos indicus. 2. Desequilíbrio de ligação. 3. Gado de corte. 4.
Genotipagem. 5. REHH. 6.SNPs. I. Título. II. Jaboticabal -
Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.2:636.082

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da
Informação - Serviço Técnico de Biblioteca e Documentação - UNESP, Campus de
Jaboticabal.

DADOS CURRICULARES DO AUTOR

Adriana Luiza Somavilla – nascida em 5 de maio de 1985, na cidade de Curitiba – PR, filha de Francisco Carlos Somavilla e Maria Heleno Somavilla. Iniciou em março de 2005 o curso de graduação em Zootecnia na Universidade Federal do Paraná, obtendo o título em novembro de 2010. Durante a graduação exerceu atividades de monitora e de iniciação científica em Melhoramento Genético Animal, sob orientação da Prof^a. Dr^a. Laila Talarico Dias. Em março de 2011, ingressou no Programa de Pós-graduação em Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus de Jaboticabal – SP, como bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), obtendo o grau de mestre em 30 de julho de 2012, sob orientação do Prof. Dr. Mauricio Mello de Alencar e co-orientação da Dra. Luciana Correia de Almeida Regitano.

*“Pelos caminhos que ando um dia vai
ser, só não sei quando.”*

Paulo Leminski

*“Pintou estrelas no muro, e teve o céu ao
alcance das mãos.”*

Helena Kolody

AGRADECIMENTOS

Aos meus pais, Francisco e Maria, por todo amor e dedicação incondicionais a mim e pelos sacrifícios que me permitiram chegar até aqui. Aos meus irmãos, Jefferson e Ramílio, pela amizade acima de tudo. Aos demais familiares que participaram e ajudaram de alguma forma em todas as dificuldades e que comemoram comigo cada conquista.

Aos professores da Universidade Federal do Paraná, Laila e Rodrigo, responsáveis pela minha escolha profissional. Agradeço o apoio, as conversas, a amizade e a confiança.

Ao meu orientador, Dr. Mauricio Mello de Alencar pela oportunidade e à Embrapa Pecuária Sudeste, pela concessão dos dados.

A minha Co-orientadora, Dra. Luciana Correia de Almeida Regitano, pelo desafio que foi a realização desse trabalho. Obrigada por confiá-lo a mim e por todos os ensinamentos transmitidos.

Ao Dr. Maurício Mudadu, por todo o suporte na condução do meu estudo, desde a análise dos dados, discussões e questionamentos.

Ao CNPq, pela concessão da bolsa de estudos.

Ao pessoal do Laboratório de Biotecnologia Animal da Embrapa: Flávia, Dri, Marina, Suelen, Andressa, Priscila, Polyana, Marcela, Alexandre, Wil e Giba, pela amizade, companheirismo, risos, conversas e apoio.

As meninas de Jaboticabal: Taciana, Raquel, Val, Marina, Camila e Bárbara, pela amizade e abrigo no início do Mestrado, essenciais na fase mais difícil da minha adaptação.

Ao meu namorado Rafael, que apareceu na minha vida e me ensinou, entre outras coisas, equilibrar vida pessoal e profissional.

Obrigada a todos que passaram pela minha vida durante este período e partiram deixando um pouco de si e levando um pouco de mim. De alguma forma todos fazem parte de mais essa etapa.

SUMÁRIO

	Página
LISTA DE ABREVIATURAS.....	ii
LISTA DE TABELAS	iii
LISTA DE FIGURAS	iv
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....	1
Objetivos	2
Revisão de Literatura	2
<i>Caracterização do Genoma</i>	2
<i>Assinaturas de Seleção</i>	6
<i>Anotação do Genoma</i>	8
CAPÍTULO 2 – DETECÇÃO DE ASSINATURAS DE SELEÇÃO RECENTE E IDENTIFICAÇÃO DE GENES CANDIDATOS POSICIONAIS EM BOVINOS DA RAÇA NELORE.....	10
Introdução	12
Material e Métodos.....	13
<i>Amostras</i>	13
<i>Genotipagem e Controle de Qualidade</i>	14
<i>Inferência da fase de ligação dos polimorfismos de nucleotídeo único (SNPs)</i>	15
<i>Visualização dos haplótipos e aplicação da metodologia da homozigose do haplótipo estendido (EHH) em regiões de genes candidatos posicionais</i>	16
<i>Aplicação da metodologia da homozigose do haplótipo estendido (EHH) em todo o genoma</i>	18
<i>Anotação do Genoma</i>	20
Resultados e Discussão.....	22
<i>Marcadores e haplótipos</i>	22
<i>Análises nas regiões dos genes candidatos posicionais</i>	27
<i>Aplicação do teste EHH no genoma</i>	33
<i>Anotação de genes posicionais nas regiões significativas</i>	41
Conclusões	46
REFERÊNCIAS.....	48

LISTA DE ABREVIATURAS

EHH – Homozigose do Haplótipo Estendido

EHW – Equilíbrio de Hardy-Weinberg

GWAS – Estudo de Associação Ampla do Genoma

LD – Desequilíbrio de Ligação

MAF – Frequência do Alelo Menos Frequente

N_e – Tamanho Efetivo Populacional

QTL – Locus de Característica Quantitativa

REHH - Homozigose Relativa do Haplótipo Estendido

SNP – Polimorfismo de Nucleotídeo Único

LISTA DE TABELAS

	Página
Tabela 1. Genes candidatos utilizados na aplicação da metodologia EHH em bovinos da raça Nelore.....	17
Tabela 2. Informação dos SNPs e blocos distribuídos ao longo do genoma de bovinos da raça Nelore.....	24
Tabela 3. Estatística da aplicação do teste EHH nas regiões dos genes candidatos no genoma de bovinos Nelore.....	30
Tabela 4. Estatística da aplicação do teste EHH no genoma de bovinos da raça Nelore.....	33
Tabela 5. Descrição dos QTLs próximos as regiões de assinaturas de seleção ($P<0,0001$) do genoma de bovinos da raça Nelore.....	37
Tabela 6. Genes relacionados com processos biológicos dos tecidos muscular e adiposo, crescimento, reprodução e sistema imune, próximos as regiões de assinaturas de seleção do genoma de bovinos da raça Nelore.....	42

LISTA DE FIGURAS

	Página
Figura 1. Distribuição do tamanho e número de SNPs formando os blocos no genoma de bovinos da raça Nelore.....	25
Figura 2. Distribuição do número de haplótipos em relação à frequência na população.....	26
Figura 3. EHH x distância para os genes SST, LEPR, FABP4 e GH1, mostrando o declínio da homozigosidade dos dois haplótipos mais frequentes, em função da distância do haplótipo núcleo, medida em cM.....	31
Figura 4. Visualização dos blocos e do desequilíbrio de ligação para os genes SST, LEPR, FABP4 e GH1.....	32
Figura 5. Distribuição dos valores de REHH em função da frequência dos haplótipos em bovinos da raça Nelore.....	34
Figura 6. Mapa das assinaturas de seleção no genoma de bovinos da raça Nelore em relação à posição nos cromossomos.....	36

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

Os resultados da bovinocultura de corte evidenciam a grande importância econômica que a atividade representa para o Brasil. Os dados de 2011 mostram que, com um rebanho de aproximadamente 180 milhões de cabeças, houve produção média de 7,45 milhões de toneladas de equivalente-carcaça, proveniente de pouco mais de 38 milhões de animais abatidos (ANUALPEC, 2012). Além disso, a exportação de mais de 1 milhão de toneladas foi, principalmente, de carne in natura (ABIEC, 2012). Portanto padrões sanitários e de qualidade, como cor, maciez, palatabilidade e suculência devem ser alcançados para atender essa demanda do mercado consumidor (MISSIO et al., 2010)

O rebanho nacional é composto predominantemente por animais de raças zebuínas (*Bos indicus*), originárias da Ásia (LENSTRA & BRADLEY, 1999). Esses animais são mais adaptados ao ambiente tropical, em função principalmente da tolerância ao calor e resistência a parasitas, além de apresentarem maior fertilidade quando comparados a indivíduos de raças taurinas, de origem européia (*Bos taurus*), fatores que contribuem para a maior produtividade desses animais nos climas tropicais (BRADLEY & CUNNINGHAM, 1999). Dentre elas, a raça Nelore pode ser considerada a de maior importância, já que é a mais utilizada nos sistemas extensivos brasileiros.

O aumento do consumo de carne bovina no Brasil, bem como das exportações, vem sendo acompanhado por uma maior exigência do mercado consumidor e, portanto, existe a necessidade não só de aumento na produtividade do rebanho nacional, mas também da qualidade dos produtos (COSTA et al., 2007). Sabe-se que os programas de melhoramento genético de raças zebuínas contemplam principalmente características de crescimento e, recentemente, maior atenção tem sido dada àquelas relacionadas à qualidade de carne e reprodução, a fim de melhorar a eficiência produtiva dos rebanhos e também o produto final.

A seleção artificial resulta em alterações nas frequências alélicas da população, que podem ser detectadas, por exemplo, por diferenças nos padrões de homozigose dos nucleotídeos dos indivíduos (INNAN & KIM, 2004). Esse efeito é localizado ao redor

de genes específicos e a identificação dessas alterações, ou assinaturas de seleção, no genoma da raça Nelore pode auxiliar na busca por regiões que controlam características quantitativas sob seleção, na identificação de genes associados às variações fenotípicas de interesse, bem como das vias metabólicas relacionadas a eles.

A disponibilidade de populações de bovinos com informações zootécnicas e os programas de melhoramento genético baseados na disseminação do material genético de reprodutores podem acelerar a descoberta de poligenes relacionados às características sob seleção (TELLAM et al., 2009). Assim, as informações obtidas a partir da análise genômica, quando incorporadas aos programas de melhoramento genético permitirão a seleção, para objetivos específicos, de reprodutores mais jovens com maior acurácia e, conseqüentemente, a diminuição do intervalo entre gerações e aumento do ganho genético na população.

OBJETIVOS

Identificar as regiões correspondentes a assinaturas de seleção em animais da raça Nelore, por meio da análise dos genótipos de polimorfismos de nucleotídeo único desses indivíduos, os genes e processos biológicos associados a essas regiões.

REVISÃO DE LITERATURA

Caracterização do Genoma

Estima-se que o genoma bovino tenha aproximadamente 2,87 Gb (Giga bases) e ao menos 22 mil genes, distribuídos em 29 pares de cromossomos autossômicos e 1

par de cromossomos sexuais (LIU et al., 2009). As características de interesse econômico são geralmente poligênicas e as regiões do genoma que contêm os genes relacionados a elas são denominadas QTLs, ou *loci* de características quantitativas (PAREEK et al., 2011). Essa caracterização é relevante na determinação dos delineamentos de estudos genômicos, como por exemplo, o número de marcadores necessários e indivíduos a serem genotipados.

Outro fator importante na identificação das bases genéticas de características quantitativas é a quantificação do desequilíbrio de ligação (LD) ao longo do genoma (McKAY et al., 2007). Segundo GODDARD et al. (2009) o LD é observado quando determinadas combinações de alelos de diferentes *loci* em um cromossomo ocorrem com maior frequência que aquelas esperadas em uma situação ao acaso. Portanto, para que seja possível identificar variações no genoma relacionadas a fenótipos de caracteres quantitativos, aplicam-se metodologias baseadas em marcadores genéticos, que devem estar fortemente ligados às mutações causais, ou seja, em LD com as mesmas (KHATKAR et al., 2006).

Recentemente esses estudos tornaram-se viáveis devido ao desenvolvimento de painéis de alta densidade de polimorfismos de nucleotídeo único (SNPs) e dos métodos de análise em larga escala (ANDERSSON, 2009), sendo que, para uma maior garantia da existência de LD em toda a população, os marcadores devem estar suficientemente próximos às mutações causais (DEKKERS, 2004). Por conta dessa proximidade, a taxa de recombinação é consideravelmente pequena, fazendo com que as frequências alélicas dos marcadores ligados sejam alteradas devido às mudanças nas frequências das mutações causais (FAY & WU, 2000), fenômeno conhecido como “efeito carona”.

Dentre as diversas metodologias que permitem estimar a extensão do LD ao longo do genoma, as mais utilizadas são baseadas no LD entre pares de genes. As primeiras medidas de LD entre os pares, estabelecidas por LEWONTIN (1964) foram: D (determinante de LD) e D' (valor relativo de desequilíbrio), sendo:

$$D_{ij} = x_{ij} - p_i q_j \text{ e } D' = \frac{|D|}{D_{max}}$$

em que D_{ij} é o LD entre os *loci* i e j ; x_{ij} é a frequência observada de gametas A_iB_j ; p_iq_j é a frequência esperada de gametas A_iB_j , assumindo que não haja nenhuma associação estatística entre os *loci* i e j ; D_{max} é o desvio máximo da frequência gamética real em relação ao equilíbrio, dadas as frequências gênicas.

Segundo os autores, a correção para a medida normalizada D' é importante para minimizar a influência que as frequências gênicas exercem sobre as estimativas, assim, é possível detectar as mudanças na intensidade do LD sem interferência das frequências iniciais dos alelos envolvidos nas análises.

HEDRICK (1987) comparou seis estimadores do LD entre pares de genes e também sugeriu que, para a comparação entre *loci*, sejam utilizadas medidas que não são dependentes das frequências alélicas ou do tamanho das amostras, a fim de evitar resultados viesados. Além disso, concluiu que o estimador D' garante maior confiança na determinação da associação estatística entre os *loci*.

A exceção é o teste r^2 (HILL & ROBERTSON, 1968), que mesmo sendo dependente das frequências alélicas é também muito utilizado atualmente, já que o D' pode ser superestimado em amostras muito pequenas (WEISS & CLARK, 2002). O quadrado do coeficiente de correlação entre um par de *loci* é definido como:

$$r^2 = \frac{D^2}{fA_1fA_2fB_1fB_2}$$

em que $fA_1fA_2fB_1fB_2$ são, respectivamente, as frequências observadas dos alelos A_1 , A_2 , B_1 e B_2 , supondo que um par de *loci* A e B, com 2 alelos cada, esteja sendo testado.

Em uma população em equilíbrio, o valor esperado de r^2 pode ser definido, segundo WEISS & CLARK (2002) como:

$$E(r^2) = \frac{1}{4N_e c + 1}$$

sendo N_e o tamanho efetivo da população e c a taxa de recombinação, ou o comprimento do segmento cromossômico, dado em Morgans. A equação prediz um rápido declínio no LD conforme a distância genética aumenta e, esse declínio é mais acentuado em populações maiores. Além disso, os autores destacam uma propriedade importante dessa estimativa: multiplicando o valor encontrado de r^2 pelo tamanho da

amostra utilizada, obtém-se o valor de χ^2 que deve ser considerado para determinar se o teste é ou não significativo, ou seja, se os *loci* estão associados.

O LD do genoma bovino é considerado alto, inclusive em longos fragmentos de DNA e, essa caracterização pode auxiliar no entendimento de processos de recombinação e seleção, além da comparação entre populações e detecção de assinaturas de seleção e outros eventos evolucionários (KHATKAR et al., 2006). Observa-se que para indivíduos *Bos indicus*, O LD entre marcadores próximos parece ser substancialmente menor quando comparado com *Bos taurus*, possivelmente em consequência do menor tamanho efetivo da população (N_e) ancestral dos animais taurinos (THE BOVINE HAPMAP CONSORTIUM, 2009).

Além disso, os primeiros painéis de SNPs disponíveis para genotipagem foram desenvolvidos a partir do genoma de raças taurinas, tornando os marcadores mais informativos para animais do mesmo grupo genético que para zebuínos, que apresentam vários alelos com a frequência do alelo menos frequente (MAF) muito baixa (MATUKUMALLI et al., 2009). Em contrapartida, os painéis de alta densidade disponíveis atualmente já contemplam também raças zebuínas.

Outro fator que influencia o sucesso da genotipagem é a estrutura de genealogia disponível, ou seja, quanto maior o número de informações presentes na matriz de parentesco, mais acurados serão os genótipos e a fase de LD nos cromossomos (MCKAY et al., 2007), ou seja, é possível determinar, por probabilidade, as combinações de alelos transmitidos ao longo das gerações. A construção de mapas de LD de alta resolução pode proporcionar melhor identificação de regiões alvo, associadas às variações fenotípicas de interesse (MATUKUMALLI et al., 2009). Para o desenvolvimento desses mapas de LD, o uso de SNPs específicos para cada raça e o posicionamento físico desses marcadores nos cromossomos deve ser considerado, a fim de evitar maior viés no momento de estimar a associação dos mesmos com os QTLs de interesse (ARIAS et al., 2009).

A abordagem de QTLs é uma maneira de associar segmentos do genoma com características quantitativas de interesse econômico e, se for realizada uma varredura completa do genoma, permite estimação dos efeitos de todos, ou pelo menos grande

parte dos genes relacionados com os fenótipos (PAREEK et al., 2011). Segundo estes autores, já existem 4.682 QTLs detectados no genoma bovino, que afetam 376 características. Além disso, estudos indicam que a densidade de genes encontrada nas regiões contendo QTLs é maior em relação ao restante do genoma e, esse agrupamento parece ocorrer de acordo com a função ou influência em determinadas características (SALIH & ADELSON, 2009).

Em situações em que é possível genotipar um grande número de SNPs, como ocorre recentemente com os painéis de alta densidade contendo mais de 700 mil marcadores, é possível a formação de blocos de SNPs em LD, chamados de haplótipos, que são utilizados nos estudos de associação (VILLA-ANGULO et al., 2009). Os autores estudaram indivíduos *Bos taurus* e *Bos indicus* e concluíram que a abordagem utilizando somente SNPs individuais exige ao menos 28.700 marcadores informativos, em contrapartida, para a caracterização dos haplótipos, um mínimo de 574.000 marcadores é requerido.

Independente da abordagem utilizada, mapas de LD comparativos possibilitam detectar regiões específicas do genoma que tiveram as frequências gênicas alteradas pela seleção artificial (MCKAY et al., 2007). Essas regiões, denominadas assinaturas de seleção, são caracterizadas pelo aumento do LD e diminuição da variabilidade genética (WIENER et al., 2003), causada pela rápida fixação das mutações associadas às características selecionadas (KIM & NIELSEN, 2004).

Assinaturas de Seleção

A domesticação das espécies de interesse zootécnico está diretamente relacionada à seleção fenotípica de diferentes características, que diferiram ao longo do tempo, de acordo com a necessidade e objetivo dos criadores. Sabe-se que a seleção pode alterar a frequência dos alelos de diversos genes que controlam as características de interesse e, conseqüentemente, de regiões neutras ligadas a eles. Segundo LUIKART et al. (2003), os *loci* sob pressão de seleção artificial podem apresentar

comportamento diferenciado e revelar padrões que não atendem as propriedades típicas da distribuição, ou seja, são considerados “outliers” em relação às frequências esperadas.

A detecção de regiões do genoma que evidenciam a seleção por meio desses padrões de polimorfismos é extremamente afetada pelas frequências iniciais dos alelos na população antes da seleção (INNAN & KIM, 2004). Os autores verificaram que a probabilidade de detecção é consideravelmente maior quando essas frequências são baixas ($<0,2$), pois se torna mais fácil observar as diferenças nos padrões de variação.

Vários métodos têm sido desenvolvidos para permitir a observação de assinaturas de seleção, a maioria baseados nas diferenças entre as frequências alélicas (STELLA et al., 2010). Cada metodologia possui suas particularidades e pode ser aplicada em diferentes situações. As mais utilizadas são:

- Teste D (TAJIMA, 1989) e Teste H (FAY & WU, 2000) – Ambos são baseados nas diferenças de frequências alélicas de cada *loci* e foram desenvolvidos para sequências completas e não para análises amplas do genoma baseadas em SNPs (HAYES et al., 2008). Além disso, requerem o conhecimento da história demográfica da população para determinar os valores críticos dos testes (INNAN & KIM, 2004). O teste D não utiliza informações de alelos ancestrais e valores negativos revelam excesso de alelos raros, enquanto que valores negativos do teste H refletem a alta frequência de alelos derivados (PRZEWORSKI, 2002);
- Estatística F_{ST} – É definida como a correlação das frequências gênicas entre subpopulações e é utilizada para quantificar o nível de diferenciação entre as mesmas (WEIR & COCKERHAM, 1984).
- Escore de Integração dos Haplótipos – iHH – Abordagem eficiente para comparar o LD em torno de alelos selecionados e não selecionados (HAYES et al., 2008) que requer alta densidade de SNPs e a especificação completa do estado dos alelos ancestrais (THE BOVINE HAPMAP CONSORTIUM et al., 2009). Fornece uma medida de quão não usuais são os haplótipos ao redor de um determinado SNP (VOIGHT et al., 2006);

- Homozigose do Haplótipo Estendido – EHH (SABETI et al., 2002) – metodologia também conhecida por teste de haplótipo de longo alcance (“Long Range Haplotype” – LHR) que proporciona uma maneira de detectar seleções positivas recentes por meio da análise da estrutura dos haplótipos de indivíduos de um único grupo genético. O método consiste na relação entre as frequências dos alelos e a extensão do LD. Nesse caso, a mensuração do LD é realizada entre um alelo em um *locus* específico com múltiplos *loci*, a várias distâncias.

Anotação do Genoma

O sequenciamento em larga escala tem a capacidade de produzir enormes quantidades de dados brutos que precisam ser derivados em informações significativas e úteis biologicamente, como marcadores, genes, mRNA ou proteínas codificadas (ROUZÉ et al., 1999). Segundo os autores, o primeiro passo é a anotação do genoma, processo pelo qual sequências anônimas são documentadas. Para auxiliar o processo de anotação foi criado o “gene ontology”, que é uma representação formal do conhecimento biológico dos genes e seus produtos e contempla três abordagens: função molecular, processos biológicos (em quais processos as funções moleculares atuam) e componentes celulares (em qual parte das células as funções são desempenhadas) (THOMAS et al., 2007).

Pode-se classificar a identificação de genes no genoma ou a atribuição de funções aos transcritos observados como anotação unidimensional e, a determinação de interações entre os componentes, localização física dos genes e de mudanças na sequência do genoma resultantes da adaptação evolutiva, como multidimensional (REED et al., 2006). Outra possibilidade é a anotação de QTLs de características consideradas importantes para cada espécie nas regiões de interesse (POLINENI et al., 2006).

Dentre várias metodologias, a busca por sequências homólogas de genes conhecidos, por meio de banco de dados que dispõe de ferramentas computacionais

que mensuram a similaridade entre as mesmas, fornece os primeiros indicativos sobre a função da região em estudo (ALTSCHUL et al., 1990). De acordo com os autores, esses métodos atribuem escores de inserções, deleções e substituições e realizam o alinhamento com base no conjunto que indica os menores números de mutações.

Outra abordagem é a utilização de algoritmos sofisticados desenvolvidos para a predição de genes em eucariotos, capazes de detectar propriedades do DNA, como regiões com altas concentrações das bases nitrogenadas citosina e guanina (C+G), associadas a regiões codificadoras. Alguns preditores geram ainda um modelo de gene completo, definido como um conjunto de "exons" transcritos coordenadamente, e são mais confiáveis (STEIN, 2001).

Mesmo com o avanço tecnológico é preciso cuidado na realização da anotação, pois existem alguns problemas que podem ocorrer, como por exemplo, a baixa acurácia na inferência de sequências homólogas, diferentes funções molecular e celular de genes homólogos, ou ainda, a ausência de experimentos laboratoriais para comprovar as análises computacionais (BRENNER, 1999).

Há também a possibilidade da anotação ser realizada com base em SNPs significativamente associados com características de importância econômica, detectados nas análises de associação ampla do genoma (GWAS), com o uso de ferramentas como o SnpSites - ("Snp annotation tool") (JIANG et al., 2011). Essa plataforma foi desenvolvida para o uso em bovinos e compila informações de vários bancos de dados.

Existe ainda a ferramenta GBrowse ("Generic Genome Browser"), que pode ser usada para visualizar, de forma arbitrária, sequências de nucleotídeos ou proteínas, bem como SNPs, genes e QTLs nas regiões de interesse (STEIN et al., 2002).

A detecção de assinaturas de seleção delimita regiões do genoma que são, ou foram, funcionalmente importantes (BISWAS & AKEY, 2006) e pode também auxiliar na construção de mapas genéticos, ou mapas de seleção, nos quais são mapeados os genes que afetam as características produtivas, em função dos padrões de diversidade genética (WIENER et al., 2003).

CAPÍTULO 2 – DETECÇÃO DE ASSINATURAS DE SELEÇÃO RECENTE E IDENTIFICAÇÃO DE GENES CANDIDATOS POSICIONAIS EM BOVINOS DA RAÇA NELORE

Deteção de Assinaturas de Seleção Recente e Identificação de Genes Candidatos Posicionais em Bovinos da Raça Nelore

RESUMO - Um maneira de identificar regiões do genoma que controlam características de interesse econômico e que estejam sob seleção é por meio de genotipagem em plataformas de altas densidades de polimorfismos de único nucleotídeo (SNPs). Essas assinaturas de seleção podem ser identificadas principalmente pelas frequências alélicas, altas e não usuais, 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, por meio do software AlphaPhase. A reconstrução dos haplótipos e a detecção de assinaturas de seleção recentes foram realizadas por meio do software Sweep v1.1, que faz uso da aplicação da metodologia da homozigose do haplótipo estendido (EHH). A identificação dos genes foi realizada pela localização das regiões de assinaturas de seleção por meio da plataforma Gbrowse UMD 3.1 e a anotação funcional foi conduzida com o auxílio da ferramenta de classificação funcional de genes, disponível no programa DAVID. Foram detectadas mais de 2000 regiões de possíveis assinaturas de seleção recente e 545 genes posicionais. Vários genes foram relacionados com crescimento, metabolismo muscular e adiposo, características reprodutivas e sistema imunológico. Os resultados encontrados podem auxiliar o avanço das pesquisas na raça Nelore, como o estudo de associação ampla do genoma das regiões significativas com os fenótipos das características de interesse.

Palavras-Chave: *Bos indicus*, desequilíbrio de ligação, gado de corte, genotipagem, REHH, SNPs

IDENTIFICATION OF SELECTION SIGNATURES AND GENOME ANNOTATION IN NELLORE CATTLE

ABSTRACT - One way to identify genomic regions controlling complex traits under selection is by genotyping on high density platforms of single nucleotide polymorphisms (SNPs). Selection signatures can be identified mainly by allele frequencies, unusual high, of adjacent loci to favorable mutations. With the purpose to detect evidence of recent selection signatures in Nellore cattle genome, 796 individuals were genotyped with the Illumina BovineHD BeadChip, which has more than 777,000 SNPs and, after that, phase of SNPs were inferred by AlphaPhase 1.1. Haplotypes reconstruction and detection of recent selection signatures were realized using Sweep v1.1, applying Extended Haplotype Homozygosity methodology (EHH). Gene identification was performed by the location of selection signatures regions through the Gbrowse UMD 3.1 platform and the functional annotation was made using the tool of the functional classification of genes, available in the software DAVID. We detected more than 2,000 possible regions of recent selection signatures and 545 positional genes. Several genes were related to growth, muscle and fat metabolism, immune and reproductive characteristics. These results may help the progress of research in Nellore cattle, as the study of genome-wide association of significant regions with the phenotypes of interest.

Keywords: Beef cattle, *Bos indicus*, genotyping, linkage disequilibrium, REHH, SNPs

Introdução

A seleção artificial é capaz de promover a fixação de fenótipos desejáveis na população em um curto período de tempo, quando comparada à seleção natural (INNAN & KIM, 2008). Nesse contexto, a intensificação dos sistemas de produção animal foi viabilizada pela combinação de três fatores: melhoramento genético, nutrição e sanidade, necessários ao desenvolvimento dos indivíduos com genótipos superiores (BRADLEY & CUNNINGHAM, 1999).

O rebanho bovino brasileiro, composto principalmente por animais da raça Nelore (*Bos indicus*), vem sendo selecionado há décadas, por diversos programas de melhoramento genético, que focaram os esforços na seleção de características de crescimento e, mais recentemente, também em características reprodutivas, importantes para a eficiência econômica que no geral apresentam baixas herdabilidades. O ganho genético ao longo das gerações pode ser comprovado, por exemplo, pelos estudos de tendência genética de pesos, ganhos de peso, idade ao primeiro parto e perímetro escrotal (LAUREANO et al., 2011).

Grande parte das características de importância econômica tem herança complexa, ou seja, são influenciadas por inúmeros *loci* e por fatores ambientais (ANDERSSON & GEORGES, 2004), o que dificulta a busca pelos genes que as controlam. Assim, a identificação de regiões do genoma que controlam caracteres que estejam sob seleção, por meio da análise dos padrões dos dados genômicos, sem a necessidade de informações fenotípicas, pode ser mais viável que a busca por esses genes (SIMIANER et al., 2010).

Além disso, a detecção dessas assinaturas de seleção permite a mineração de genes que tiveram papel importante nas mudanças genéticas recentes (INNAN & KIM, 2008). Isso torna possível um melhor entendimento da evolução e da biologia de fenótipos específicos, que pode ser utilizado no desenvolvimento de ferramentas para aumentar a eficiência de seleção (STELLA et al., 2010).

A disponibilidade da genotipagem dos indivíduos em plataformas de alta densidade de SNPs permite a varredura do genoma por assinaturas de seleção. Para

tal, é necessária a análise estatística dos dados, que avalia, principalmente, o rápido aumento nas frequências alélicas dos *loci* adjacentes às mutações favoráveis (QANBARI et al., 2010), pois, em função do LD, os mesmos tendem a ser transmitidos juntos, causando um efeito carona (FAY & WU, 2000).

Dentre as diversas metodologias desenvolvidas, a estatística EHH - Homozigose do Haplótipo Estendido – (SABETI et al., 2002) é a que se aplica melhor a um conjunto de dados composto por um único grupo genético e sem informações dos genótipos de alelos ancestrais (QANBARI et al., 2010).

Assim, os objetivos desse trabalho foram detectar evidências de assinaturas de seleção recente no genoma de bovinos da raça Nelore por meio da metodologia EHH, e identificar genes correspondentes a essas regiões.

Material e Métodos

Amostras

Foram utilizados 796 animais, sendo 762 novilhos machos, descendentes de 34 touros Nelore registrados, pertencentes às principais linhagens de uso comercial mais frequente no país, escolhidos com base nos coeficientes de parentesco de cerca de 600 touros com sêmen disponível, com o objetivo de minimizar o grau de parentesco e representar a variabilidade genética da raça no país. Além disso, não foram levados em consideração os valores genéticos para características de importância econômica. Os produtos foram gerados em três estações de monta, nos anos de 2007, 2008 e 2009, por meio de inseminação artificial em tempo fixo (IATF) e criados na Embrapa Pecuária Sudeste, localizada na cidade de São Carlos-SP, na Embrapa Gado de Corte, situada no município de Campo Grande-MS e em três propriedades particulares dos Estados de Mato Grosso do Sul e Mato Grosso. A média de idade dos touros foi de 13 anos, sendo variando entre 5 e 32 anos. Além disso, o número médio de filhos/touro foi de 21, com

variação entre 2 e 42. Em resumo, as etapas de coleta de material, extração de DNA e genotipagem ocorreram da seguinte maneira: para obtenção do DNA dos touros foram utilizadas as palhetas de sêmen congelado, provenientes das centrais de inseminação artificial e a extração do DNA foi realizada pelo método de desproteinação com solventes orgânicos (fenol, clorofórmio e álcool isoamílico). Para os novilhos, foram coletados 5 mL de sangue, por punção da veia jugular, em tubos para coleta a vácuo contendo EDTA potássico. As extrações de DNA foram realizadas a partir de leucócitos, por meio do método “salting out”. Após a extração, para avaliar a concentração, pureza e integridade do DNA, as amostras foram quantificadas em NanoDrop® e submetidas à eletroforese em gel de agarose 1% corado com GelRed.

Genotipagem e Controle de Qualidade

A genotipagem foi realizada no Laboratório de Genômica Funcional do “United States Department of Agriculture” (BFGL/USDA/EUA) financiada pelo projeto Rede Bife de Qualidade, Macroprograma 1 da Embrapa. Foi utilizado o chip “Illumina BovineHD BeadChip”, que possui mais de 777 mil SNPs em sua plataforma, com espaçamento médio de 3,4 kb. Após a genotipagem, um arquivo contendo informações dos indivíduos, SNPs e um escore correlacionado com a qualidade da genotipagem de cada SNP (“GC Score”) foi criado. Esse escore é uma qualidade métrica que indica a confiabilidade dos genótipos de cada marcador em cada amostra. Para que as análises posteriores, baseadas em relações de parentesco, não fossem prejudicadas, foi aplicado um teste de paternidade com base nos SNPs selecionados pelo algoritmo implementado por MUDADU et al. (2011), que identificou as incompatibilidades com a genealogia disponível. TEO (2008) já havia destacado a importância desse tipo de análise em estudos baseados em famílias.

Após as correções de paternidade, os dados foram submetidos a análises de controle e garantia de qualidade, em relação aos indivíduos e também aos *loci*, a fim de evitar, ao máximo, resultados espúrios (LAURIE et al., 2010). Para tal, foi utilizado o

pacote “snpMatrix” (CLAYTON & LEUNG, 2011), do programa estatístico R (R DEVELOPMENT CORE TEAM, 2012).

Foram selecionados os SNPs que apresentavam “GC Score” maior ou igual a 0,2. Esse limiar foi escolhido com base na quantidade de SNPs removidos e na diminuição do número de inconsistências mendelianas proporcionadas. Em seguida, os dados foram filtrados utilizando os seguintes parâmetros: indivíduos com “call rate” > 90% (porcentagem do genoma genotipado corretamente), SNPs com “call rate” > 95% (porcentagem de genotipagem de cada SNP em toda a população), excluindo resultados de baixa qualidade, o que contribui para o aumento da acurácia dos dados (OLIPHANT et al., 2002), pois erros de genotipagem podem interferir no estudo de haplótipos (GABRIEL et al., 2002). A MAF foi estabelecida em 1%. Foram excluídos ainda os SNPs localizados no cromossomo X, genoma mitocondrial e aqueles sem informação de posição no genoma ou considerados ambíguos.

Inferência da fase de ligação dos polimorfismos de nucleotídeo único (SNPs)

As análises posteriores dependiam da formação de haplótipos e, portanto, a fase de ligação dos SNPs deveria ser inferida, ou seja, determinar, por probabilidade, as combinações de alelos que são transmitidas ao longo das gerações. Como os animais eram provenientes de famílias de meio-irmãos, optou-se por utilizar um algoritmo desenvolvido para indivíduos relacionados e aparentemente não relacionados, disponível no programa AlphaPhase 1.1 (HICKEY et al., 2011). Resumidamente, o AlphaPhase 1.1 combina dois importantes métodos: “long-range phasing” (LRP), que infere a fase dos marcadores com base nos haplótipos compartilhados por indivíduos relacionados (chamados de possíveis pais) e “haplotype library imputation” (HLI), biblioteca de haplótipos criada a partir do próprio banco de dados, que aumenta o rendimento da inferência, corrige erros criados pela definição equivocada de possíveis pais e/ou erros de genotipagem. As etapas são executadas em cada cromossomo e, em regiões delimitadas dos arquivos, para aumentar a acurácia da inferência. Para tal,

devem ser definidos os tamanhos dos núcleos (cadeias consecutivas de SNPs) e das caudas (cadeias consecutivas de SNPs adjacentes a cada uma das extremidades dos núcleos).

Os parâmetros utilizados foram:

- Tamanho do núcleo: 100 SNPs;
- Tamanho das caudas para ambos os lados: 100 SNPs;
- Número de indivíduos analisados antes de inferir a fase: 10;
- Proporção máxima de discordância entre os indivíduos analisados: 10%;
- Proporção máxima de genótipos perdidos ou incorretos: 1%.

Visualização dos haplótipos e aplicação da metodologia da homozigose do haplótipo estendido (EHH) em regiões de genes candidatos posicionais

Para observar se a metodologia escolhida era eficaz, primeiramente observou-se os resultados em regiões que continham genes relacionados com crescimento, deposição de gordura, reprodução e sistema imune (Tabela 1). Os blocos foram analisados no programa Haploview v4.1 (BARRET et al., 2005) para estimar o LD entre os pares de SNPs e observar o padrão dos haplótipos formados e no programa Sweep v1.1 para identificar evidências de assinaturas de seleção. Para representar a região, o bloco que continha o SNP mais próximo de cada gene candidato foi escolhido. Foram então calculadas as estatísticas EHH e REHH dos dois haplótipos mais frequentes contidos em cada bloco e, para observar o declínio da homozigosidade nos genes, o EHH foi plotado contra a distância, em cM, do haplótipo núcleo.

Tabela 1. Genes candidatos utilizados na aplicação da metodologia EHH em bovinos da raça Nelore

Cromossomo	Gene	Nome	Característica	Referência
1	SST	Somatostatina	Crescimento e desenvolvimento	GAO et al. (2011)
2	MSTN/GDF8	Miostatina	Hipertrofia muscular	WIENER et al. (2009)
3	LEPR	Receptor da Leptina	Crescimento e deposição de gordura	GUO et al. (2008)
4	LEP	Leptina	Crescimento e deposição de gordura	GUO et al. (2008)
5	IGF1	Fator de crescimento semelhante a insulina	Crescimento e puberdade	ROGBERG-MUÑOZ et al. (2011); LIRON et al. (2012)
7	IL4	Interleucina 4	Aumento na secreção de cortisol e resposta inflamatória	WOODS & JUDD (2008)
14	FABP4	Proteína ligante de ácidos graxos 4	Deposição de gordura intramuscular e subcutânea	MICHAL et al. (2006)
17	IL2	Interleucina 2	Resistência à carrapato	PIPER et al. (2009)
19	GH1	Hormônio do crescimento 1	Crescimento	SILVEIRA et al. (2008)
20	GHR	Receptor do GH1	Deposição de gordura intramuscular e subcutânea	BAEZA et al. (2011)
21	IGF1R	Receptor do IGF1	Crescimento e puberdade	MICKE et al. (2011); FORTES et al. (2012)
23	BoLA-DRB3	Complexo Principal de Histocompatibilidade classe II	Resistência a carrapato	MARTINEZ et al. (2006)

Aplicação da metodologia da homozigose do haplótipo estendido (EHH) em todo o genoma

Essa abordagem, desenvolvida por SABETI et al. (2002) para a detecção de assinaturas de seleção, parece ser a mais utilizada para estudos com um único grupo genético. A metodologia proporciona uma maneira de detectar seleções positivas recentes por meio da análise da estrutura dos haplótipos dos indivíduos.

O teste foi aplicado por meio do programa Sweep 1.1 (SABETI et al., 2002) e, para tal, foram utilizados os arquivos de inferência das fases dos SNPs, gerados pelo AlphaPhase. Cada arquivo contendo 100 SNPs foi chamado de núcleo e, abrangia diversos blocos, definidos pelo algoritmo proposto por GABRIEL et al. (2002), que considera um par de SNPs em forte LD se a estatística D' estiver dentro de um intervalo de confiança de 95%, ou seja, aqueles com limite inferior maior que 0,7 e limite superior acima de 0,98. Foi especificado ainda que um bloco tivesse no mínimo 3 e, no máximo, 20 SNPs.

De acordo com SABETI et al. (2002), alelos comuns são geralmente antigos e têm LD de curto alcance. Já alelos raros podem ser antigos ou recentes e ter LD de longo ou curto alcance. Assim, o que caracteriza uma assinatura de seleção recente é o rápido aumento na frequência alélica, que ocorre num espaço de tempo curto o suficiente para que a recombinação não quebre o haplótipo no qual a mutação ocorreu. Uma assinatura de seleção positiva é então um haplótipo que apresenta LD de longo alcance e não usual, além de alta frequência, quando comparado aos outros haplótipos do mesmo *locus*.

O poder de detecção de assinaturas recentes é muito superior aos testes que levam em consideração somente as frequências alélicas, pois é baseado na associação entre um SNP com múltiplos *loci* a diversas distâncias. A homozigose do haplótipo estendido é definida como a probabilidade que dois cromossomos escolhidos aleatoriamente e que carreguem o haplótipo de interesse, sejam idênticos por descendência, o que permite a detecção da transmissão de um haplótipo sem

recombinação (SABETI et al., 2002). O EHH de um haplótipo núcleo testado (t) foi matematicamente calculado como:

$$EHH_t = \frac{\sum_{i=1}^s \binom{e_{ti}}{2}}{\binom{c_t}{2}}$$

onde c_t é o número de amostras de um haplótipo núcleo particular (em teste), e_{ti} é o número de amostras de cada haplótipo estendido presente no bloco e s é o número de haplótipos estendidos do bloco.

A metodologia foi inicialmente desenvolvida para dados do genoma humano e, na época, como não havia boas estimativas de taxas de recombinações locais, os autores elaboraram uma correção para a estimativa EHH. Sabe-se que em bovinos, as taxas de recombinação também não são constantes ao longo do genoma (KHATKAR et al., 2006), portanto, a Homozigose Relativa do Haplótipo Estendido (REHH), proposta por SABETI et al. (2002) foi utilizada para corrigir o EHH para as variações nas taxas de recombinação. Primeiramente foi calculada a homozigose relativa do cromossomo:

$$\overline{EHH} = \frac{\sum_{j=1, j \neq t}^n \left[\sum_{i=1}^s \binom{e_i}{2} \right]}{\sum_{i=1, i \neq t}^n \binom{c_i}{2}}$$

que é definido como o declínio do EHH em todos os outros haplótipos núcleos combinados e, n é número de diferentes haplótipos núcleos no cromossomo (em relação ao t). Em seguida, o REHH foi definido por:

$$EHH_t / \overline{EHH}$$

O teste foi aplicado a uma distância de 0,25 cM para ambos os lados do haplótipo núcleo, devido ao menor LD observado em relação a raças taurinas. Para determinar a significância dos valores de REHH, os haplótipos formados foram ordenados em 20 classes de frequência na população e, cada um foi comparado com os outros igualmente frequentes no cromossomo. Para obter uma distribuição próxima da normal, os valores de REHH foram transformados por escala logarítmica ($-\log_{10}$). Em seguida foi obtida a distribuição de probabilidade, calculados a média, os desvios-padrão e os P -valores, que também foram transformados para $-\log_{10}(P)$. Assim, aqueles haplótipos com P -valores extremos foram considerados significativos.

Anotação do Genoma

Para a anotação dos genes, foram consideradas somente as regiões com $-\log_{10}(P)$ superiores a 3 ($P < 0,001$). Além disso, estendeu-se a área de busca 250 kb para ambos os lados do haplótipo núcleo, distância menor que aquela utilizada por QANBARI et al. (2010) em bovinos Holstein, justamente pelo menor LD observado nos animais na raça Nelore. A ferramenta GBrowse versão 1.71, carregada com o genoma bovino versão UMD 3.1, foi utilizada para visualizar os genes nas regiões de interesse. Buscou-se também informações de QTLs no banco de dados CattleQTLdb (HU et al., 2010), para observar se as regiões identificadas como assinaturas de seleção estavam sobrepostas a QTLs previamente relacionados com características de importância econômica em bovinos de corte. Porém, para tal etapa, foram utilizadas somente aquelas regiões que apresentaram os P -valores mais extremos ($P < 0,0001$).

A etapa seguinte consistiu em analisar a lista de genes no programa DAVID ("Database for Annotation, Visualization, and Integrated Discovery") (DENNIS JR. et al., 2003). Os genes foram submetidos à classificação funcional disponível na plataforma, que os agrupa de acordo com a similaridade funcional. Essa etapa utiliza mais de 75 mil termos derivados de fontes de anotação funcional e, parte do princípio que genes com perfis de anotação semelhantes devem ser funcionalmente relacionados, o que facilita a

interpretação biológica dos mesmos (HUANG et al., 2007). Os autores explicam que esse agrupamento é realizado com base em uma medida de probabilidade corrigida de co-ocorrência entre dois conjuntos de dados categorizados, chamada de estatística “kappa”, que varia de 0 (co-ocorrência ao acaso) a 1 (co-ocorrência relacionada). Essa associação funcional entre dois genes é calculada com base nos perfis de anotação dos genes m e n , mais especificamente nas suas co-ocorrências observadas (O_{mn}) e as chances de co-ocorrência (A_{mn}):

$$K_{mn} = \frac{O_{mn} - A_{mn}}{1 - A_{mn}}$$

A análise foi conduzida utilizando os parâmetros padrões do programa, adequados para a maioria dos estudos (HUANG et al., 2007), definidos da seguinte forma:

- Sobreposição de termos similares: eram necessários, no mínimo, 4 termos sobrepostos entre um par de genes para qualificá-los para o cálculo de estatística “kappa”;
- Limiar de similaridade: o valor mínimo da estatística *kappa* considerado biologicamente significativo foi de 0,35;
- Membros do grupo inicial: um grupo inicial deveria ser formado por, no mínimo, 4 genes;
- Membros do grupo final: após a análise, os grupos deveriam ter, ao menos, 4 . Isso evita a formação de muitos grupos de tamanho pequeno;
- Limiar de multi-ligação: dois grupos que compartilhavam mais de 50% dos genes passavam a ser um único grupo.

A análise final consistiu na construção de redes associadas aos processos biológicos, funções moleculares e componentes celulares dos genes incluídos nos grupos com o auxílio do software GOrilla (“Gene Ontology enrichment analysis and visualization tool”) (EDEN et al., 2009). De acordo com os autores, essa ferramenta realiza o enriquecimento de termos super expressos do “Gene Ontology” a partir de uma lista de genes ordenados. A significância dos termos foi definida por meio de

distribuição de probabilidade hipergeométrica e, para isso, um escore denominado mHG é calculado. Esse escore pode ser definido como: a probabilidade mínima encontrada, com base em todas as configurações possíveis, dado o ranqueamento dos genes, corrigida para múltiplos testes. No presente trabalho, essa lista foi formada com base no agrupamento realizado pela ferramenta DAVID.

Resultados e Discussão

Marcadores e haplótipos

A Tabela 2 apresenta uma descrição resumida dos SNPs e blocos formados no conjunto de dados. Após o controle de qualidade, restaram 789 indivíduos e 560.565 SNPs cobrindo 2.522,54 Mpb (Mega bases) do genoma bovino (versão UMD 3.1), com espaçamento médio entre eles de 4,5 kb. A MAF média ($\pm$ erro-padrão) desse conjunto de marcadores foi de $0,26 \pm 0,0002$, estimativa próxima a observada recentemente para a raça Nelore, de 0,20 (THE BOVINE HAPMAP CONSORTIUM, 2009). Além disso, 66% dos SNPs apresentaram $MAF \geq 0,1$. Foram formados 68.054 blocos, com tamanho médio ($\pm$ desvio-padrão) de 15,69 kb ($\pm 17,94$ kb), abrangendo 1,07 Mpb (42%) dos 29 cromossomos estudados. O desvio-padrão reflete a alta variação no tamanho dos blocos, que apresenta amplitude de 0,65 a 491,27 kb. Dos SNPs utilizados, 330.262 (58,91%) participaram da formação dos blocos.

O teste do desvio do Equilíbrio de Hardy-Weinberg (EHW), que elimina SNPs acima de um desvio pré-determinado, não foi utilizado por diversas razões, sendo que a principal é o fato de que o teste EHH, que identifica as regiões de assinaturas de seleção depende da alta homozigose dos haplótipos e, consequentemente, dos SNPs. Além disso, de acordo com CHAN et al. (2008), de 13,6% a 23,6% dos SNPs desviam do EHW e isso pode ser consequência da ancestralidade compartilhada. Os filtros de “call rate” e MAF já proporcionam um conjunto de marcadores com genótipos

suficientemente acurados (TEO, 2008). Outra questão importante é a possibilidade que os desvios do EHW aconteçam em consequência das correlações genéticas entre indivíduos da mesma família e não por erros de genotipagem (FINNER et al., 2010). Nesse caso, SNPs em alta homozigose por conta da estrutura de pedigree e/ou seleção seriam excluídos das análises posteriores, comprometendo os resultados.

Em relação à caracterização dos blocos, mais de 50.000 tinham no máximo 20 kb de comprimento e, mais de 40.000 continham até quatro marcadores (Figura 1). Isso pode ser consequência da menor extensão do LD observado em indivíduos da raça Nelore, além do declínio mais intenso a curtas distâncias, quando comparados com animais *Bos taurus* (VILLA-ANGULO et al., 2009), resultando em blocos menores e formados por poucos marcadores.

Tabela 2. Estatística descritiva dos SNPs representados no “Illumina BovineHD BeadChip” e blocos distribuídos ao longo do genoma de bovinos da raça Nelore

Genomes da espécie <i>Artemia</i>											
Cromossomo	SNPs		Distância entre SNPs (Kb)	Nº Blocos	Comprimento médio blocos (Kb)		Cobertura blocos (Kb)	Máximo comprimento de bloco (Kb)	Cobertura a (%)	SNPs em blocos	SNPs em blocos/Total
	MAF>0,01										
1	36.565		4,33	4.176	15,58	65.082,24	177,97	0,41	20.820	0,57	
2	29.945		4,57	3.713	16,49	61.214,78	230,56	0,45	18.251	0,61	
3	27.248		4,45	3.388	15,76	53.378,45	251,52	0,44	16.513	0,61	
4	26.618		4,53	3.301	16,31	53.831,80	166,13	0,45	16.208	0,61	
5	25.016		4,84	2.987	17,98	53.699,22	395,44	0,44	15.197	0,61	
6	28.227		4,23	3.492	14,67	51.212,96	141,13	0,43	17.008	0,60	
7	25.487		4,42	3.033	15,87	48.128,71	303,66	0,43	14.944	0,59	
8	26.097		4,34	3.149	14,80	46.618,14	200,06	0,41	15.213	0,58	
9	24.699		4,28	2.997	15,14	45.366,80	150,17	0,43	14.618	0,59	
10	22.346		4,67	2.747	16,61	45.625,75	241,41	0,44	13.420	0,60	
11	24.379		4,40	3.019	16,96	51.208,60	260,11	0,48	15.324	0,63	
12	18.740		4,86	2.260	16,76	37.887,46	204,71	0,42	11.269	0,60	
13	17.329		4,86	2.032	17,87	36.312,63	245,46	0,43	10.072	0,58	
14	20.514		4,09	2.454	14,09	34.564,78	165,81	0,41	11.708	0,57	
15	19.445		4,38	2.422	15,50	37.535,50	281,49	0,44	11.587	0,60	
16	18.498		4,42	2.297	15,98	36.710,29	328,50	0,45	11.221	0,61	
17	16.914		4,44	2.060	15,34	31.610,69	162,47	0,42	9.877	0,58	
18	14.922		4,42	1.750	16,41	28.710,78	187,45	0,43	8.666	0,58	
19	13.478		4,75	1.623	16,45	26.692,06	177,48	0,42	7.646	0,57	
20	16.271		4,42	2.025	14,51	29.390,53	155,30	0,41	9.497	0,58	
21	16.679		4,29	2.069	15,83	32.745,16	424,92	0,46	10.193	0,61	
22	13.280		4,61	1.613	16,21	26.145,55	360,33	0,43	7.679	0,58	
23	11.889		4,41	1.377	14,76	20.320,72	149,44	0,39	6.296	0,53	
24	14.055		4,45	1.730	15,27	26.421,10	175,21	0,42	8.289	0,59	
25	8.965		4,77	1.065	15,90	16.928,95	187,33	0,39	4.985	0,56	
26	11.917		4,33	1.489	14,53	21.639,93	351,96	0,42	6.793	0,57	
27	9.929		4,57	1.197	13,39	16.028,08	491,27	0,35	5.210	0,52	
28	10.293		4,49	1.278	14,41	18.419,95	255,63	0,40	5.706	0,55	
29	10.820		4,75	1.311	15,57	20.417,81	239,96	0,40	6.052	0,56	
Total	560.565			68.054		1.073.849,40			330.262		
Média			4,50		15,69			0,42		0,58	

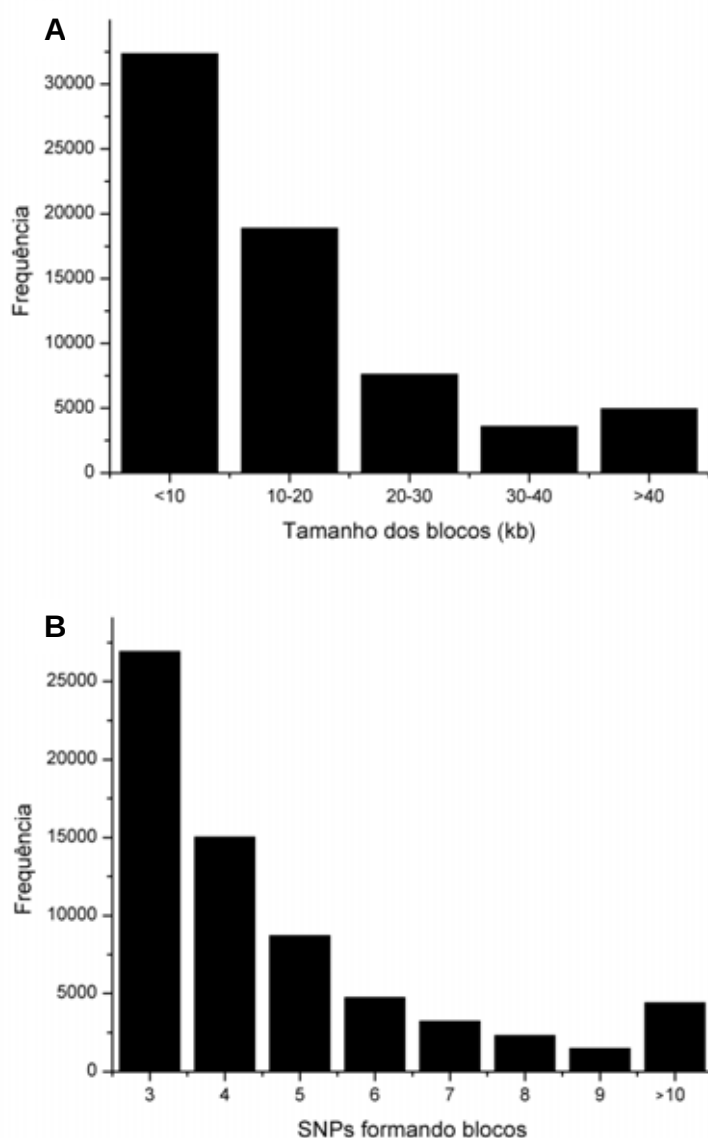


Figura 1. Distribuição do tamanho (A) e número de SNPs formando os blocos (B) no genoma de bovinos da raça Nelore.

Após análise dos blocos foram identificados 237.378 haplótipos. A Figura 2 representa a distribuição das classes de frequência desses haplótipos na população. É possível observar que há mais haplótipos frequentes na classe de 51 a 55% em relação às classes de 26 a 50%. Esse resultado pode ser explicado pelo número de haplótipos representados nos blocos. Aqueles com até quatro haplótipos são mais comuns e, para verificar como isso poderia influenciar a distribuição da Figura 2, as frequências

haplotípicas foram verificadas no conjunto de dados. Notou-se que, geralmente, com três haplótipos, existe um mais frequente (acima de 50%), um segundo mais frequente (em torno de 40%) e um terceiro com frequência menor que 10%. Quando são quatro haplótipos no bloco, existe um com frequência maior que 60% e três entre 10 e 20%. Dessa forma, ao juntar todas as informações em um único gráfico, fica visível o maior número de haplótipos frequentes em mais de 51% da população. Além disso, os haplótipos com baixa frequência (até 10%) são mais comuns, condizente com uma população que apresentou diminuição do N_e recentemente, devido a deriva genética (TAJIMA, 1989). De acordo com VILLA-ANGULO et al. (2009), estima-se que nas últimas cem gerações, o N_e da raça Nelore tenha passado de mais de 400 para menos de 100 animais. Corroborando essa hipótese, os resultados encontrados por FARIA et al. (2009) mostram que o N_e da raça no Brasil diminuiu de 85 para 68 animais em quatro gerações (de 1979 a 1998).

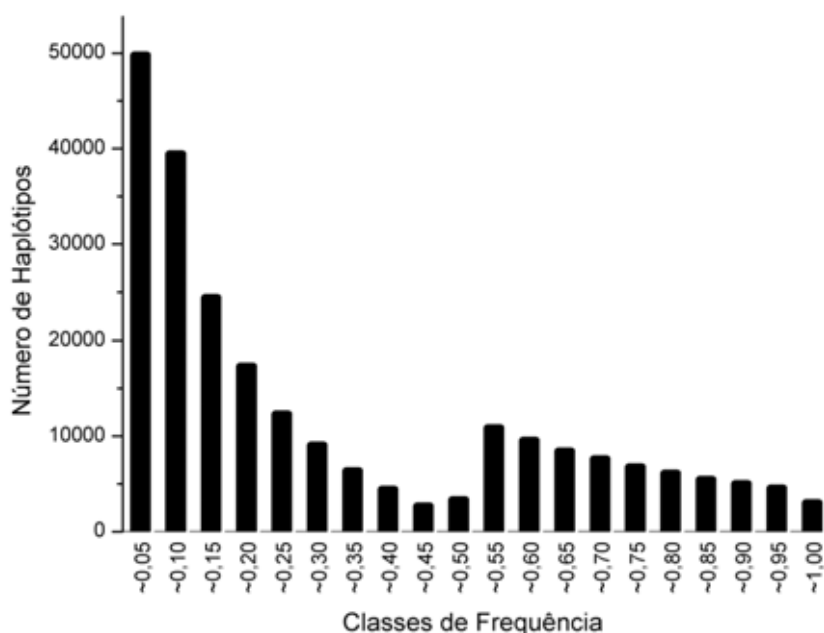


Figura 2. Distribuição do número de haplótipos em relação à frequência na população. Foram formadas 20 classes de frequência com intervalos de 5%.

Análises nas regiões dos genes candidatos posicionais

Inicialmente, regiões que continham genes relacionados ao sistema imunológico, crescimento e deposição de gordura foram avaliadas. Observou-se o LD entre os marcadores bem como o padrão de frequência e homozigose dos haplótipos, para detectar possíveis evidências de seleção.

Na Tabela 3 são apresentados os resultados da aplicação do teste EHH nos haplótipos observados em cada gene candidato e a informação do SNP mais próximo. Partiu-se do pressuposto que a medida de REHH é um indicativo do LD ao redor do haplótipo em estudo e, valores superiores a 1 significam alta homozigosidade daquele haplótipo em relação aos outros de igual frequência (SABETI et al., 2002).

A análise dos resultados sugere que os genes relacionados ao sistema imune não apresentam indicativos de assinatura de seleção recente. Indivíduos da raça Nelore são mais bem adaptados as condições ambientais em que estão inseridos e são mais resistentes a patógenos, quando comparados aos animais *Bos taurus* (BRADLEY & CUNNINGHAM, 1999). Provavelmente esses genes estejam sob pressão de seleção natural, que não causa aumento rápido das frequências alélicas como a seleção artificial (FAY & WU, 2000) e por isso não é detectável com a metodologia utilizada.

O gene da miostatina (MSTN) é membro da super família dos fatores de crescimento, sendo que sua inativação pode resultar em um expressivo desenvolvimento muscular (WIENER et al., 2009). McPHERRON & LEE (1997) já haviam relatado que mutações não silenciosas nesse gene causam hipertrofia muscular e, como consequência, o fenótipo chamado de dupla musculatura, que pode ser observado em várias raças européias, com destaque para Belgian-Blue e Piemontês. Em indivíduos da raça Nelore, GRISOLIA et al. (2009) encontraram 37 polimorfismos no gene MSTN, todos diferentes daqueles observados nas raças que apresentam dupla musculatura. Os autores argumentam que a identificação desses polimorfismos pode indicar uma alta taxa de mutação nesse gene. Dessa forma, a metodologia aplicada, baseada principalmente na diminuição da variabilidade genética pode não ter sido capaz de identificar evidências de assinatura de seleção nesse gene.

Em relação aos genes do fator de crescimento semelhante a insulina (IGF1) e de seu receptor (IGF1R), esperava-se algum indicativo de seleção, pois eles podem estar associados a características de crescimento (ROGBERG-MUÑOZ et al., 2011; MICKE et al., 2011) e com a idade à puberdade (FORTES et al., 2012). GE et al. (2001) observaram, por meio da utilização da técnica de PCR-RFLP, um polimorfismo no gene IGF1 em bovinos da raça Angus que dava origem a dois genótipos. Os autores concluíram que o genótipo BB conferia maior ganho de peso quando comparado ao AA. Em contrapartida, em uma população de indivíduos Nelore, CURI et al. (2005) verificaram que o genótipo BB estava fixado. No presente estudo, a metodologia foi capaz de detectar assinatura de seleção somente no gene do hormônio do crescimento 1 (GH1), que é o principal regulador do metabolismo e do crescimento (BREIER, 1999) e integra o eixo somatotrópico, juntamente com o gene IGF1, proteínas e receptores associados (CURI et al., 2006).

O mesmo pode ter ocorrido com os genes da leptina (LEP) e seu receptor (LEPR), pois somente o segundo apresentou evidências de assinatura de seleção. A leptina está envolvida com mudanças no apetite, metabolismo e partição de nutrientes e age através de receptores (TARTAGLIA, 1997). Já o gene LEPR está associado ao balanço energético, características de deposição de gordura e crescimento em bovinos (GUO et al., 2008).

Outros dois genes apresentaram evidências de assinaturas de seleção. O gene da somatostatina (SST), hormônio protéico que apresenta regulação negativa sobre o GH1 (MARTIN & MILLARD, 1986), já foi relacionado a QTLs para características como marmoreio, área de olho-de-lombo e espessura de gordura subcutânea em animais da raça Angus (MORSCI et al., 2006), além de apresentar SNPs associados com altura e comprimento em *Bos taurus* (GAO et al., 2011).

A proteína ligante de ácidos graxos 4 (FABP4) parece exercer papel importante no metabolismo de lipídios e na homeostase dos adipócitos em bovinos (HOASHI et al., 2008) e, juntamente com a lipase, media o processo de transporte de lipídios durante a lipólise (SHEN et al., 1999). Dois polimorfismos no gene FABP4 foram associados com marmoreio e deposição de gordura subcutânea em indivíduos *Bos taurus* (MICHAL et

al., 2006). Já os estudos com a raça Nelore indicaram a fixação dos alelos previamente descritos e nenhuma associação com as características relacionadas a deposição de gordura (AYRES et al., 2010; CURI et al., 2011). Anteriormente, BARENDSE et al. (2009), avaliando indivíduos *Bos taurus* e *Bos indicus*, concluíram que possíveis resultados não significativos de associação entre polimorfismos do gene FABP4 e marmoreio poderia ser explicada pelo pequeno efeito que o gene apresenta sobre os fenótipos de interesse e, por esse motivo, as análises deveriam contemplar um maior número de indivíduos. Como a metodologia utilizada no presente trabalho não leva em consideração os polimorfismos no gene, e sim a homozigose e a frequência do haplótipo em relação aos outros da mesma região do genoma, foi possível observar indícios de assinatura de seleção margeando esse gene.

Tabela 3. Estatística da aplicação do teste EHH nas regiões dos genes candidatos no genoma de bovinos Nelore

Gene Candidato	Cromossomo	Posição do gene candidato (bp)	SNP mais próximo ao gene (posição em bp)	Localização do bloco (bp)	Frequência haplotípica (%) ^a	EHH ^b	REHH ^b	P-valores REHH ^b
SST	1	80250205-	BovineHD0100023090 (80272087)	80272087-	H1: 67	0,13	12,69	0,001
		80251648			H2: 25	0,01	0,07	0,09
MSTN	2	6213566-	BovineHD0200001825 (6225434)	6225434-	H1: 50	0,17 ^c	4,33	1,00
		6220196			H2: 47	-	-	0,89
LEPR	3	80071689-	BovineHD0300023059 (80107536)	80107536-	H1: 56	0,04	0,23	-
		80147000			H2: 37	0,16	7,24	0,96
LEP	4	93249874-	BovineHD0400026037 (93250208)	93245659-	H1: 69	0,02	3,21	0,02
		93266624			H2: 18	0,04	0,14	0,28
IGF1	5	66532879-	BovineHD0500018608 (66536500)	66529557-	H1: 63	0,02	0,31	0,63
		66604699			H2: 18	0,11	1,67	0,12
IL4	7	22993178-	BovineHD0700006311 (23004080)	23004080-	H1: 63	0,07	0,25	0,38
		23001067			H2: 18	0,07	0,96	0,87
FABP4	14	46833665-	BovineHD1400013253 (46823262)	46821487-	H1: 86	0,07	0,10	0,73
		46838053			H2: 14	0,07	0,76	0,90
IL2	17	35618115-	BovineHD1700009839 (35619076)	35637592	H1: 86	0,05	1,04	0,20
		35622861			H2: 14	0,08	0,81	0,29
GH1	19	48768618-	BovineHD1900013561 (48772161)	48765934-	H1: 77	0,06	1,24	0,63
		48772014			H2: 26	0,05	0,89	0,82
GHR	20	31890736-	BovineHD2000009184 (31891078)	3189529-	H1: 86	0,07	0,74	0,65
		32064200			H2: 14	0,07	3,16	0,03
IGF1R	21	7967724-	BovineHD2100021163 (7962990)	7951484-	H1: 94	0,05	1,35	0,14
		8268932			H2: 14	0,05	0,32	0,86
BoLA-DRB3	23	25472161-	BovineHD2300007077 (25475120)	25475120-	H1: 94	0,09	0,49	0,87
		25476885			H2: 14	0,07	1,67	0,22
					H1: 71	0,09	1,12	0,77
					H2: 13	0,10	0,55	0,97
					H1: 71	0,09	1,10	0,01
					H2: 13	0,09	2,78	0,15
					H1: 71	0,09	1,21	0,69
					H2: 28	0,07	1,08	0,75
					H1: 60	0,11	0,83	0,20
					H2: 33	0,09	0,92	0,15
					H1: 60	0,11	1,27	0,63
					H2: 33	0,09	1,59	0,76
					H1: 78	0,08	0,62	0,29
					H2: 12	0,08	2,43	0,28
					H1: 78	0,03	0,37	0,74
					H2: 12	0,03	0,32	0,80

^aTestes realizados nos 2 haplótipos mais frequentes de cada bloco (H1 e H2); ^bOs valores da homozigose do haplótipo estendido (EHH), homozigose relativa do haplótipo estendido (REHH) e P-valores foram calculados para ambos os lados do haplótipo núcleo; ^csó foi possível entender o haplótipo para a esquerda

Pode-se observar na Figura 3 que para os genes SST, LEPR, FABP4 e GH1 a homoziguidade não se estende além de 0,2 cM, evidenciando que o LD na raça Nelore é menor que o observado em bovinos da raça Holstein por QANBARI et al. (2010). BOLORMAA et al. (2011) também haviam relatado que a extensão do LD em indivíduos *Bos indicus* é menor quando comparada a animais *Bos taurus* ou cruzados.

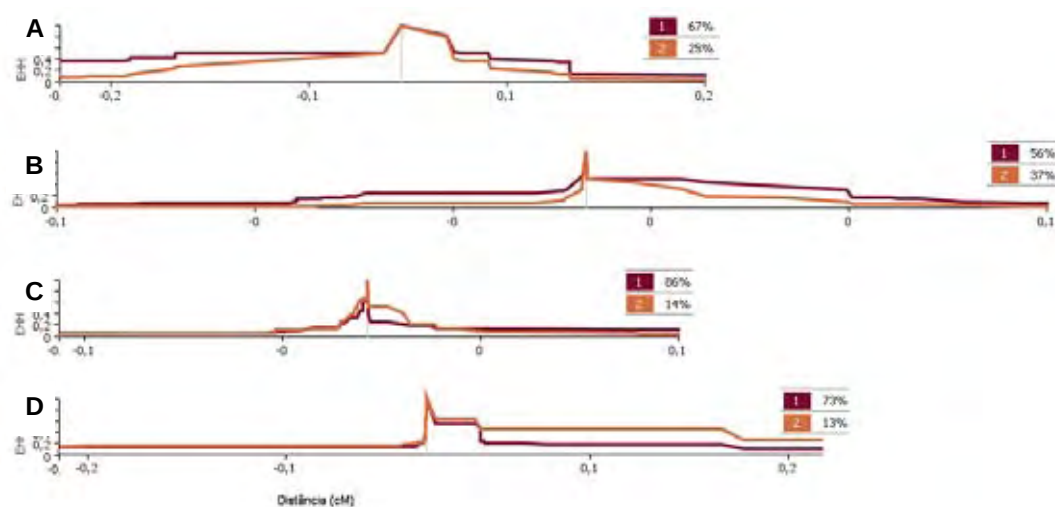


Figura 3. EHH x distância para os genes SST (A), LEPR (B), FABP4 (C) e GH1 (D), mostrando o declínio da homoziguidade dos dois haplótipos mais frequentes, em função da distância do haplótipo núcleo, medida em cM. As legendas representam a frequência de cada haplótipo para cada gene.

A Figura 4 apresenta a visualização dos blocos, bem como o LD entre os marcadores, quantificado por meio da estatística r^2 , que varia de 0 (ausência de LD) a 1 (completo LD). O quadrado do coeficiente de correlação entre dois marcadores (r^2) é corrigido para as respectivas frequências alélicas e é indicativo da quantidade de informação de um *locus* que é explicada pelo outro (ARDLIE et al., 2002). O comprimento dos blocos variou de 1 a 38 kb, o que sugere a pequena extensão do LD já observada anteriormente.

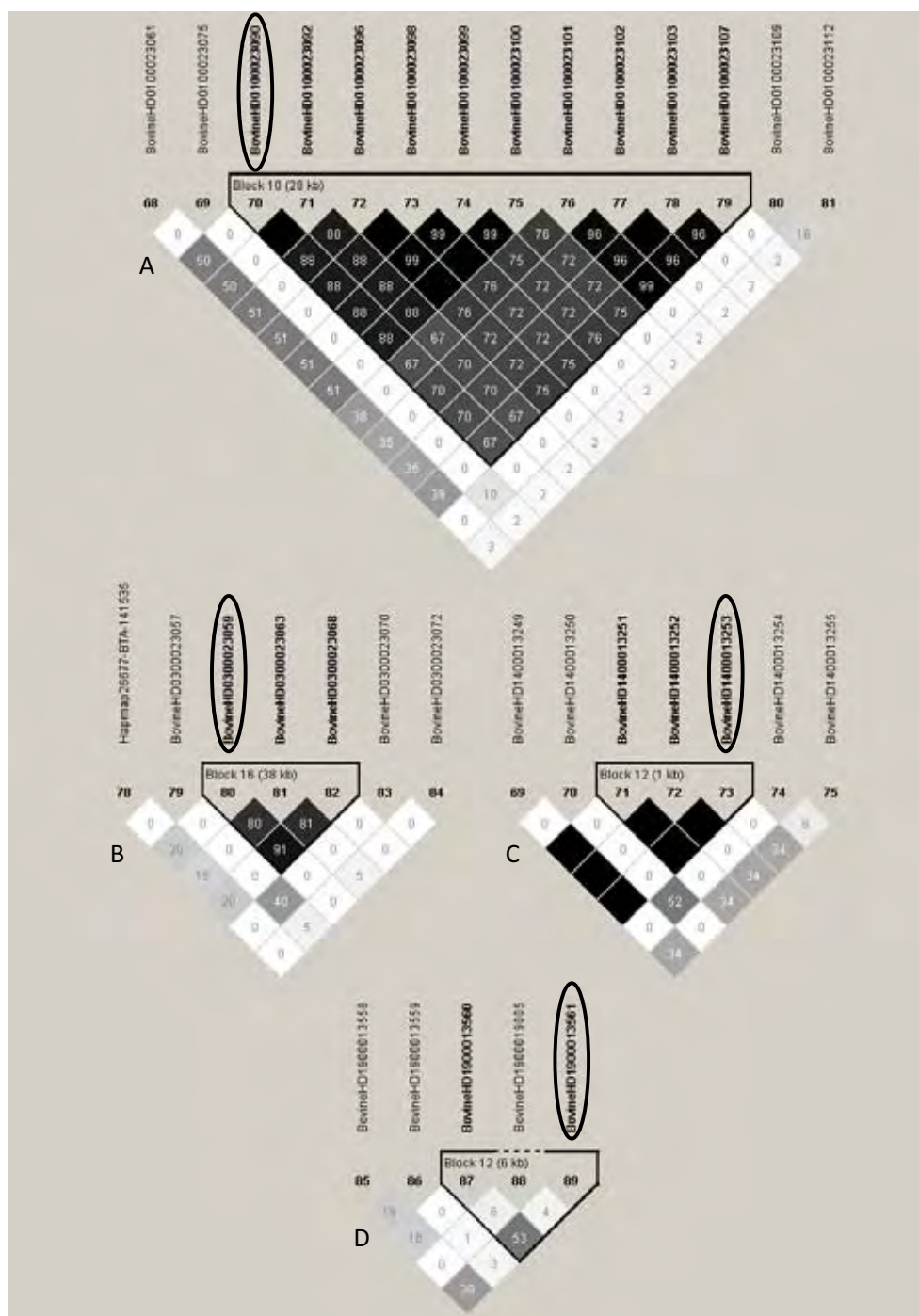


Figura 4. Visualização dos blocos e do desequilíbrio de ligação (LD) para os genes SST (A), LEPR (B), FABP4 (C) e GH1 (D). O LD foi mensurado pela estatística r^2 (em %). Cor branca representa $r^2 = 0$ (ausência de LD), preta $r^2 = 1$ (ou 100% - completo LD) e cores intermediárias representam valores intermediários. Os SNPs circulos são aqueles mais próximos ao gene candidato.

Aplicação do teste EHH no genoma

Para os 68.054 blocos formados, foram realizados 193.944 testes EHH. Somente aqueles haplótipos com frequência e homozigosidade maiores que 30% foram incluídos nas etapas posteriores, pois valores inferiores poderiam ser indicativos de eventos de recombinação (SABETI et al, 2002). A Tabela 4 mostra a descrição do número de testes e das regiões significativas por cromossomo. O teste indicou mais de 2.000 regiões que podem estar sob seleção recente, porém apenas as significativas ao menos a 0,1% de probabilidade ($P < 0,001$) foram utilizadas para a anotação do genoma ($n = 269$).

Tabela 4. Estatística da aplicação do teste EHH no genoma de bovinos da raça Nelore

Cromossomo	Testes nos haplótipos (n)	$P < 0,01$ (n)	$P < 0,001$ (n)	$P < 0,0001$ (n)
1	10741	113	22	4
2	10510	105	8	1
3	9762	107	17	1
4	9417	107	10	0
5	8429	90	12	2
6	9644	125	14	3
7	8729	91	12	0
8	9483	92	15	2
9	8989	91	15	3
10	8060	86	11	0
11	8262	96	14	0
12	6349	66	8	2
13	6202	65	9	4
14	7108	78	11	0
15	6656	65	8	0
16	6631	85	10	1
17	5939	67	1	0
18	4615	36	2	0
19	4526	50	6	1
20	5746	76	15	2
21	5724	54	4	0
22	4405	41	2	0
23	4264	33	2	0
24	5162	69	17	0
25	3164	37	5	2
26	4286	38	5	1
27	3667	43	4	0
28	3774	48	5	2
29	3700	49	5	0
Total	193944	2103	269	31

QANBARI et al. (2010) utilizaram a metodologia EHH em indivíduos da raça Holstein genotipados com painel de 50 mil SNPs e identificaram 702 regiões indicativas de assinaturas de seleção ($P < 0,05$). No presente trabalho, a utilização da plataforma com mais de 770 mil SNPs pode ter influenciado o número de blocos formados ao longo do genoma e, conseqüentemente, da quantidade de regiões significativas.

A Figura 5 apresenta a distribuição da estatística REHH em função da frequência dos haplótipos, com distinção entre os P -valores. Pode-se observar que o gráfico tem padrão semelhante a distribuição dos haplótipos em função das frequências (Figura 2). A homozigosidade relativa dos haplótipos decresce até o limite de cerca de 50%. A partir dessa classe de frequência a homozigosidade relativa deve ser superior para ser considerada significativa. A alta homozigose presente nos haplótipos de alta frequência na população pode ser consequência da diminuição da variabilidade genética aditiva (ALLENDORF, 1986), consequência da seleção artificial baseada na disseminação de material genético de poucos reprodutores (VISSCHER et al, 2008).

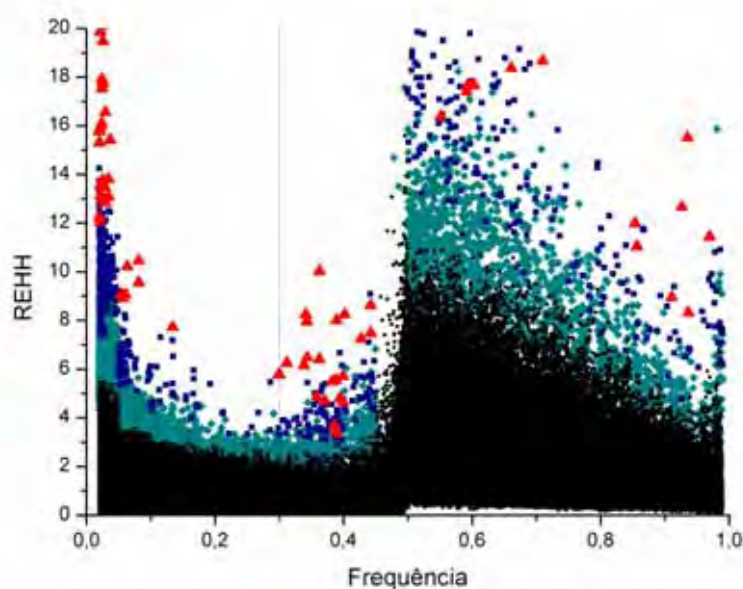


Figura 5. Distribuição dos valores de REHH em função da frequência dos haplótipos em bovinos da raça Nelore. São representados os haplótipos significativos a 0,001% (triângulos vermelhos), 0,01% (quadrados azuis) e 0,1% (losangos verdes). Somente foram considerados nas análises aqueles com frequência maior que 30% (linha vertical tracejada).

Para melhor visualização da distribuição das assinaturas de seleção ao longo dos cromossomos, foram plotados os P -valores transformados ($-\text{Log}_{10}$) da estatística REHH contra a posição do haplótipo no cromossomo (Figura 6). É possível observar que a distribuição não é uniforme ao longo do genoma, porém assinaturas de seleção foram identificadas em todos os cromossomos, resultado consistente com a hipótese de que as características sob seleção são complexas e controladas por diferentes regiões do genoma.

Os cromossomos 1, 3, 8, 9, 20 e 24 apresentaram, ao menos, 15 regiões significativas cada um ($P < 0,001$). QTLs relacionados com características reprodutivas já haviam sido identificados nos cromossomos 1, 3, 9, 20 e 24 em bovinos da raça Holstein (SAHANA et al., 2010), sendo que 15 regiões identificadas no presente trabalho estão sobrepostas a eles. Além disso, BOLORMAA et al. (2011) encontraram SNPs associados ao consumo alimentar residual, medida relacionada à eficiência alimentar, nos cromossomos 3, 5, 7 e 8 em uma população composta por indivíduos *Bos taurus* e *Bos indicus*.

Algumas regiões significativas ($P < 0,0001$) estavam ainda sobrepostas a diversos QTLs (Tabela 5) relacionados a características reprodutivas, crescimento, eficiência, qualidade de carne, perfil de ácidos graxos e imunidade. Embora nenhum dos estudos tenha sido realizado em animais da raça Nelore, a identificação de QTLs sobrepostos às assinaturas de seleção pode ser utilizada como informação inicial para a exploração de regiões do genoma que controlam características de importância econômica específicas. Existe ainda a possibilidade de estudar regiões que podem ter sido alteradas em função da resposta correlacionada, como por exemplo, QTLs relacionados ao perfil de ácidos graxos, já que esse fenótipo específico não é contemplado nos programas de melhoramento genético.

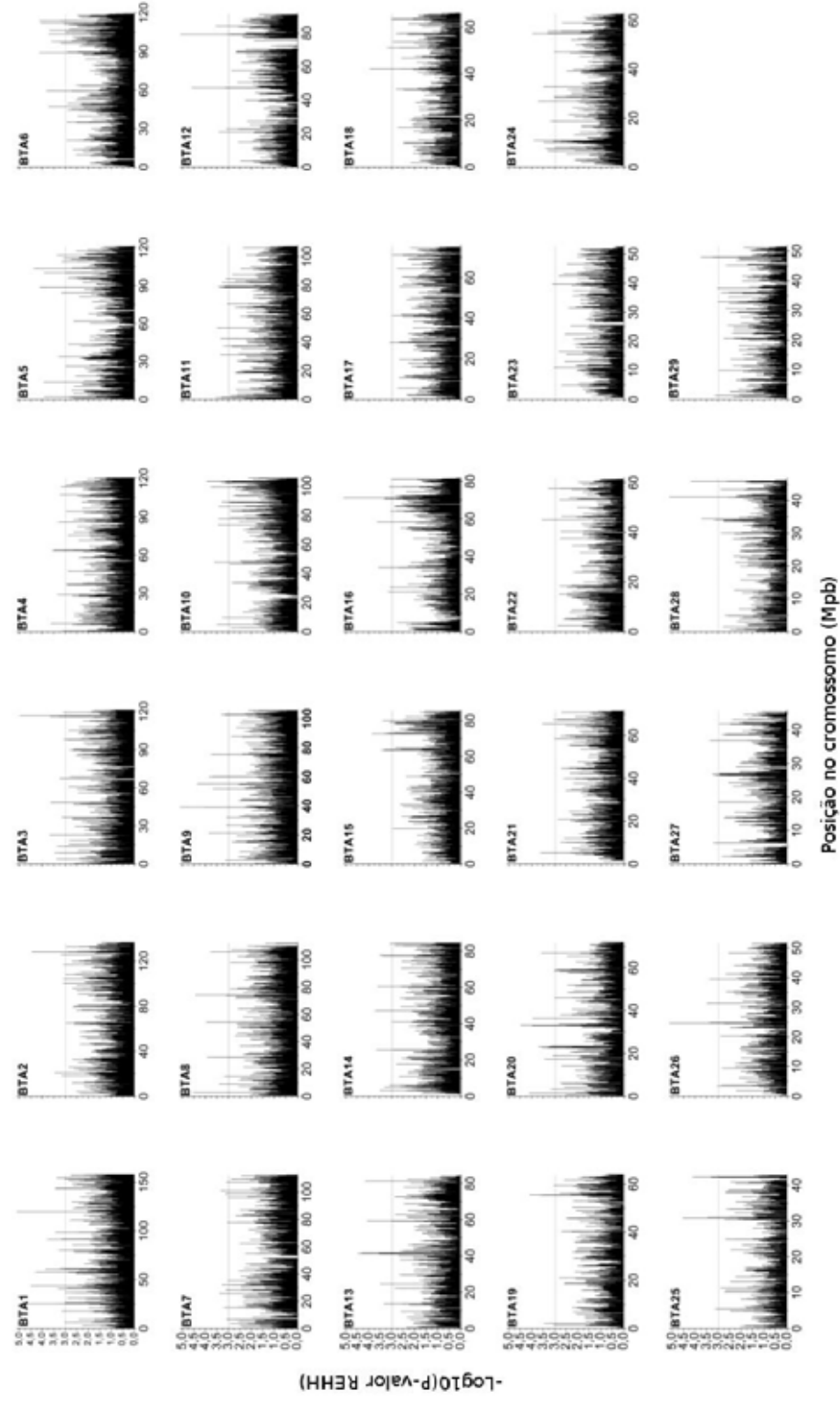


Figura 6. Mapa das assinaturas de seleção no genoma de bovinos da raça Nelore em relação à posição nos cromossomos. Foram considerados somente os haplótipos com frequência maior que 30% na população e as linhas horizontais indicam $P < 0,001$.

Tabela 5. Descrição dos QTLs próximos as regiões de assinaturas de seleção ($P < 0,0001$) do genoma de bovinos da raça Nelore

Cromossomo	Posição do bloco	Posição do QTL (bp)	Posição do QTL (cM)	Pico (cM)	Característica	População	Referência
1	25321721-25327725	11439999-31095406	15,42-38,08	16,00	Facilidade de parto	Angus	MCCLURE et al. (2010)
1	25321721-25327725	11439999-31095406	15,42-38,08	18,00	Peso ao nascer	Angus	MCCLURE et al. (2010)
1	25321721-25327725	11439999-31095406	15,42-38,08	19,00	Escore de marmoreio	Angus	MCCLURE et al. (2010)
1	25321721-25327725	11439999-31095406	15,42-38,08	29,00	Altura	Angus	MCCLURE et al. (2010)
1	57948469-57976904	50444332-61780206	51,01-62,45	54,00	Peso adulto	Angus	MCCLURE et al. (2010)
1	57948469-57976904	50901321-66472973	52,16-67,67	53,00	Rendimento de carcaça	Angus <i>Bos indicus</i> x <i>Bos taurus</i>	CASAS et al. (2003)
1	57948469-57976904	50901321-66472973	52,16-67,67	50,00	Porcentagem de gordura	<i>Bos indicus</i> x <i>Bos taurus</i>	CASAS et al. (2003)
1	119571093-119578005	106924804-119586413	99,71-111,3	104,00	Peso à desmama	Angus	MCCLURE et al. (2010)
2	128497489-128525783	127803326-128014817	nd	115,70	Peso da carcaça	Angus Hereford, Polish Red, Red Angus	MACNEIL & GROSZ (2002)
5	88769424-88798755	86569875-97454424	90,84-103,1	93,00	Peso à desmama	Angus	MCCLURE et al. (2010)
5	88769424-88798755	86569875-97454424	90,84-103,1	95,00	Peso adulto	Angus	MCCLURE et al. (2010)
5	88769424-88798755	86569875-97454424	90,84-103,1	101,00	Peso ao sobreano	Angus	MCCLURE et al. (2010)
5	103678422-103733253	100132930-108852142	105,6-116,9	106,00	Escore de marmoreio	Angus	MCCLURE et al. (2010)
5	103678422-103733253	100132930-108852142	105,6-116,9	111,00	Altura	Angus	MCCLURE et al. (2010)
5	103678422-103733253	986323396-115069577	125,0-135,6	127,00	Perímetro escrotal	Angus	MCCLURE et al. (2010)
8	74180453-74185747	60496303-80091301	66,03-83,98	71,00	Peso ao nascer	Angus	MCCLURE et al. (2010)
9	55094702-55096974	42253154-66606362	46,00-76,00	67,00	Escore de marmoreio	Angus <i>Bos indicus</i> x <i>Bos taurus</i>	CASAS et al. (2003)

Cromossomo	Posição do bloco	Posição do QTL (bp)	Posição do QTL (cM)	Pico (cM)	Característica	População	Referência
9	55094702-55096974	37971718-57619653	42,48-57,08	43,00	Perímetro escrotal	Angus	MCCLURE et al. (2010)
12	47380621-47387967	31257437-54373225	43,00-64,00	60,00	Rendimento de carcaça	<i>Bos indicus</i> x <i>Bos taurus</i>	CASAS et al. (2003)
12	47380621-47387967	44623413-54244072	46,30-63,84	48,00	Área de olho-de-lombo	Angus	MCCLURE et al. (2010)
12	47380621-47387967	44623413-54244072	46,30-63,84	50,00	Peso ao nascer	Angus	MCCLURE et al. (2010)
12	47380621-47387967	44623413-54244072	46,30-63,84	50,00	Peso adulto	Angus	MCCLURE et al. (2010)
12	79273190-79274408	54026339-90212864	101,9-108,9	108,00	Peso da carcaça	Angus	MCCLURE et al. (2010)
12	79273190-79274408	54026339-90212864	101,9-108,9	108,00	Peso ao sobreano	Angus	MCCLURE et al. (2010)
13	41745142-41749885	39345372-49635934	51,71-61,84	nd	Teor de ácido graxo araquidônico	Charolais x Holstein	GUTIÉRREZ-GIL et al. (2010)
13	59335960-59356186	58977694-71827292	53,06-80,98	83,00	Teor de ácido graxo palmítico	Jersey x Limousin	MORRIS et al. (2010)
13	59335960-59356186	39501717-71509042	51,90-80,60	56,80	Taxa de conversão alimentar	<i>Bos taurus</i>	SHERMAN et al. (2009)
13	81451727-81456114	79688485-84240350	98,4-102,4	102,40	Ingestão de matéria seca	<i>Bos taurus</i>	SHERMAN et al. (2009)
13	81451727-81456114	60488154-84240350	62,80-73,63	67,00	Espessura de gordura subcutânea na 12ª costela	Angus	MCCLURE et al. (2010)
13	81451727-81456114	77335394-84240350	62,80-91,37	72,00	Peso à desmama devido ao leite materno	Angus	MCCLURE et al. (2010)
16	71301753-71311771	63178802-72861974	57,00-73,00	62,00	Porcentagem de gordura visceral	<i>Bos indicus</i> x <i>Bos taurus</i>	CASAS et al. (2003)
16	71301753-71311771	70631045-81724687	84,00-93,50	84,00	Espessura de gordura subcutânea na 12ª costela	Angus	MCCLURE et al. (2010)
16	71301753-71311771	70631045-77462145	84,00-94,45	85,00	Área de olho-de-lombo	Angus	MCCLURE et al. (2010)
16	71301753-71311771	70631045-81724687	84,00-93,50	85,00	Altura ao sobreano	Angus	MCCLURE et al. (2010)
19	55428844-55433197	52826649-55704624	78,00-86,00	82,00	Ganho de peso diário	Angus e Brahman	TAYLOR et al. (1998)

Cromossomo	Posição do bloco	Posição do QTL (bp)	Posição do QTL (cM)	Pico (cM)	Característica	População	Referência
19	55428844-55433197	45248309-55709981	69,83-86,01	71,00	Peso ao nascer	Angus	MCCLURE et al. (2010)
19	55428844-55433197	38422165-55709981	59,20-86,01	76,00	Peso adulto	Angus	MCCLURE et al. (2010)
19	55428844-55433197	45248309-55709981	69,83-86,01	77,00	Área de olho-de-lombo	Angus	MCCLURE et al. (2010)
19	55428844-55433197	45248309-55709981	69,83-86,01	79,00	Escore de marmoreio	Angus	MCCLURE et al. (2010)
19	55428844-55433197	45248309-55709981	69,83-86,01	79,00	Peso à desmama	Angus	MCCLURE et al. (2010)
19	55428844-55433197	45248309-55709981	69,83-86,01	80,00	Perímetro escrotal	Angus	MCCLURE et al. (2010)
19	55428844-55433197	45248309-55709981	69,83-86,01	80,00	Altura ao sobreano	Angus	MCCLURE et al. (2010)
19	55428844-55433197	57868775-52655231-	90,04-77,68-	87,60	pH muscular	Charolais x Holstein	GUTIÉRREZ-GIL et al. (2010)
19	55428844-55433197	58089400-52904864-	95,03-78,30-	86,36	Espessura de gordura subcutânea	<i>Bos taurus</i>	LI et al. (2004)
19	55428844-55433197	60055349-44268013-	98,70-68,00-	98,70	Consumo alimentar residual	<i>Bos taurus</i>	SHERMAN et al. (2009)
19	55428844-55433197	62102274-22206990-	108,00-31,86-	76,00	Teor de ácido graxo mirístico	Jersey x Limousin	MORRIS et al. (2010)
20	33170072-33186600	35800031-22206990-	31,86-49,72	nd	Volume de células vermelhas inicial (PCV _i) menos volume mínimo (PVC _M)	N'Dama x Boran	HANOTTE et al. (2003)
20	33170072-33186600	22206990-35800031	31,86-49,72	nd	Variação do PCV	N'Dama x Boran	HANOTTE et al. (2003)
20	33170072-33186600	22206990-35800031	31,86-49,72	nd	Porcentagem de declínio no PCV até 150 dias após o desafio	N'Dama x Boran	HANOTTE et al. (2003)
20	33170072-33186600	22206990-35800031	31,86-49,72	nd	Porcentagem de declínio no peso corporal até 150 dias após o desafio	N'Dama x Boran	HANOTTE et al. (2003)
20	33170072-33186600	22206990-35800031	31,86-49,72	nd	Número médio de parasitas (Log _n)	N'Dama x Boran	HANOTTE et al. (2003)
20	33170072-33186600	30964880-35800031	42,00-49,72	48,00	Peso adulto	Angus	MCCLURE et al. (2010)
25	30787340-30857744	23716582-31366660	22,63-46,43	32,00	Facilidade de parto	Angus	MCCLURE et al. (2010)

Cromossomo	Posição do bloco	Posição do QTL (bp)	Posição do QTL (cM)	Pico (cM)	Característica	População	Referência
25	30787340-30857744	23716582-31366660	31,59-46,43	38,00	Peso à desmama	Angus	MCCLURE et al. (2010)
25	30787340-30857744	23716582-31366660	31,59-46,43	40,00	Peso à desmama devido ao leite materno	Angus	MCCLURE et al. (2010)
25	30787340-30857744	23716582-31366660	31,59-46,43	42,00	Peso ao sobreano	Angus	MCCLURE et al. (2010)
25	30787340-30857744	23716582-34013168	31,59-53,37	50,50	Escore de maciez	Charolais x Holstein	GUTIÉRREZ-GIL et al. (2010)
25	42364359-42401717	31366660-42775844	46,43-68,41	66,00	Altura	Angus	MCCLURE et al. (2010)
25	42364359-42401717	38690885-42775844	61,66-68,41	68,00	Espessura de gordura subcutânea na 12ª costela	Angus	MCCLURE et al. (2010)
26	24644707-24667660	20334217-25229124	26,95-29,62	29,00	Escore de marmoreio	Angus	MCCLURE et al. (2010)
26	24644707-24667660	20334217-25229124	26,95-29,62	29,00	Peso à desmama	Angus	MCCLURE et al. (2010)
26	24644707-24667660	10723763-25449045	15,45-31,65	nd	Porcentagem de declínio no PCV até 150 dias após o desafio	N'Dama x Boran	HANOTTE et al. (2003)
26	24644707-24667660	10723763-25449045	15,45-31,65	nd	Porcentagem de declínio no PCV até 100 dias após o desafio	N'Dama x Boran	HANOTTE et al. (2003)
26	24644707-24667660	10723763-25449045	15,45-31,65	nd	Peso final dimensionado pelo peso inicial	N'Dama x Boran	HANOTTE et al. (2003)
26	24644707-24667660	10723763-25449045	15,45-31,65	nd	Peso corporal médio	N'Dama x Boran	HANOTTE et al. (2003)
26	24667660	10723763-25449045	15,45-31,65	nd	Porcentagem de declínio no peso corporal até 150 dias após o desafio	N'Dama x Boran	HANOTTE et al. (2003)
26	24667660	10723763-25449045	15,45-31,65	29,00	Teor de ácido graxo margárico	Jersey x Limousin	MORRIS et al. (2010)
28	41016612-41018711	34072385-42592570	37,00-47,84	59,00	Teor de ácido graxo palmítico	Wagyu x Limousin	ALEXANDER et al. (2007)
28	41016612-41018711	35304150-45759633	49,39-61,65	58,00	Facilidade de parto	Angus	MCCLURE et al. (2010)
28	41016612-41018711	35304150-45759633	49,39-61,65	61,00	Peso à desmama devido ao leite materno	Angus	MCCLURE et al. (2010)

nd (não disponível)

Anotação de genes posicionais nas regiões significativas

Para identificação dos genes, foi realizado o alinhamento das regiões de assinaturas de seleção por meio da plataforma “Gbrowse”. Como os resultados de tamanho de haplótipos sugerem que o LD na raça Nelore pode ser menor quando comparado à raça Holstein, cada região foi estendida somente 250 kb para cada lado e, os genes presentes nesse intervalo foram anotados. Ao final, foram identificados 545 genes.

A anotação funcional foi conduzida com o auxílio da ferramenta de classificação funcional de genes, disponível no programa DAVID (DENNIS JR et al., 2003). A análise resultou em 287 genes distribuídos em 20 grupos. Os outros 258 genes foram considerados órfãos e não participam de nenhum grupo. Os grupos foram ordenados em função de sua importância biológica no conjunto de genes disponível, mensurada pelo escore de enriquecimento, que é a média geométrica do escore EASE (“Expression Analysis Systematic Explorer”) de todos os termos anotados (HUANG et al., 2007). Por sua vez, o escore EASE é uma modificação do teste Exato de Fisher, aplicado após a remoção de um gene de cada vez de um determinado grupo (HOSACK et al., 2003). De acordo com os autores, ele penaliza a significância dos termos comuns a poucos genes e favorece categorias mais robustas. Sendo assim, os genes do grupo 1 compartilhavam um maior número de termos que aqueles do grupo 2 e assim sucessivamente.

Semelhante aos resultados observados por QANBARI et al. (2011), o presente estudo identificou vários genes envolvidos com processos biológicos referentes ao crescimento, metabolismo muscular e adiposo e características reprodutivas, além daqueles relacionados ao sistema imunológico (Tabela 6).

Tabela 6. Genes relacionados com processos biológicos dos tecidos muscular e adiposo, crescimento, reprodução e sistema imune, próximos as regiões de assinaturas de seleção do genoma de bovinos da raça Nelore

Gene	Cromossomo	Posição (pb)	Processos biológicos
TFF3	1	144067290-144071071	Resposta imune
COL8A1	1	43541936-43717620	Desenvolvimento muscular
GAP43	1	60843662-60957254	Crescimento
BCL6	1	80188691-80202541	Resposta imune humoral
XRCC5	2	104656345-104744654	Resposta imune
CCL20	2	116759408-116762490	Resposta imune humoral
PAFR	2	125808444-125809472	Resposta imune humoral
TTN	2	18054928-18330824	Desenvolvimento do músculo estriado
PLEKHA3	2	18382395-18410532	Ligante de lipídeos
GBX2	3	116266172-116268268	Resposta imune humoral
SH2D2A	3	14075266-14085163	Resposta imune
CRABP2	3	14173256-14178240	Ligante de lipídeos
F11R	3	8483558-8508123	Resposta imune humoral
CD244	3	8794793-8831787	Resposta imune humoral
RARRES2	4	113564406-113567492	Resposta a patógenos e estresse
C1R	5	103737161-103747961	Metabolismo de lipídeos
C1S	5	103768065-103779250	Ativação do complemento
TPI1	5	103942402-103945759	Ativação do complemento
PDGFB	5	111180167-111200490	Biossíntese e metabolismo de lipídeos e ácidos graxos
CACNA1I	5	111506218-111615642	Resposta imune
CMAS	5	88564697-88580838	Crescimento celular
LDHB	5	88962685-88981219	Biossíntese e metabolismos de lipídeos e ácidos graxos
PPIC	7	32114445-32128247	Metabolismo de reserva energética
TNFSF8	8	105841516-105870631	Resposta imune humoral
MSRA	8	8635550-9019576	Resposta imune
T	9	102666696-102676022	Metabolismo de aminoácidos
WISP3	9	38856662-38871640	Desenvolvimento embrionário
GALC	10	100733081-100789765	Ligante de fatores de crescimento
RPS6KA5	10	103508678-103703817	Catabolismo e metabolismo de lipídeos e glicolipídeos
PLCB2	10	36023882-36044688	Resposta imune
CCNB2	10	51247755-51272733	Catabolismo e metabolismo de lipídeos
			Crescimento celular

Gene	Cromossomo	Posição (pb)	Processos biológicos
CPLX2	10	5215474-5217493	Crescimento celular
MNAT1	10	73153517-73363532	Resposta imune e reparo de DNA
FNTB	10	77421464-77502097	Biossíntese e metabolismos de lipoproteína
ALDH6A1	10	85821883-85839587	Metabolismo de aminoácidos
FSHR	11	31110744-31305197	Desenvolvimento das gônadas femininas
SPTBN1	11	37030009-37241384	Metabolismo de aminoácidos
TMSB10	11	49933204-49934214	Desenvolvimento espermático
CUGBP2	13	13036990-13248003	Desenvolvimento muscular
XRN2	13	40962751-41025285	Produção de espermatozoides
SPO11	13	59349056-59363436	Regulação da prófase I da meiose
NOV	14	47005557-47013935	Ligante de fatores de crescimento
PAX6	15	63356618-63384294	Resposta imune
PECR	16	104589994-104638006	Biossíntese de aminoácidos
PPP1R12B	16	71061936-71090170	Regulação da contração muscular
CEACAM8	18	51138353-51156406	Resposta imune
LIPE	18	51216019-51227395	Metabolismo de lipídeos, ácidos graxos e esteróides
PAFAH1B3	18	51308584-51311417	Catabolismo e metabolismo de lipídeos
DEDD2	18	51380803-51395994	Produção de espermatozoides
IL6ST	20	23212633-23247944	Ligante de fatores de crescimento e resposta imune
C6	20	33328558-33405557	Ativação do complemento
ANKH	20	58376939-58546813	Resposta imune
SRD5A1	20	66681685-66718713	Diferenciação sexual
CRABP1	21	103612392-103617658	Ligante de lipídeos
NEURL	26	24330861-24403159	Desenvolvimento e motilidade espermática
SLK	26	24786019-24844846	Resposta imune
POLB	27	36877980-36902799	Resposta imune
FADD	29	48027926-48030331	Resposta imune

Alguns genes identificados como atuantes do sistema imune também desempenham funções em processos envolvendo o desenvolvimento e viabilidade de folículos e oócitos, além da manutenção da gestação em mamíferos. O gene RPS6KA5 (ribosomal protein S6 kinase alpha-5) foi associado à reprogramação na fase de blastocisto de embriões bovinos provenientes da técnica de transferência nuclear, ou clonagem (SMITH et al., 2005). Já o gene PTAFR (platelet-activating factor receptor) está envolvido com a regulação do ciclo celular, mais especificamente com as alterações no ciclo das células da granulosa, reduzindo o número de células nas fases de replicação de DNA e divisão (S, G₂ e M) e aumentando aquelas nas fases de manutenção ou crescimento (G₀/G₁) (VIERGUTZ & LOHRKE, 2007).

O gene POLB (polymerase DNA directed beta) foi expresso após a maturação *in vitro* de oócitos de bovinos (DALBIES–TRAN & MERMILLOD, 2003). MURDOCH & VAN KIRK (2001) já haviam observado que essa regulação é mediada por estradiol e confere viabilidade aos oócitos em ovinos. Estudando animais *Bos taurus*, DORJI et al. (2012) observaram que o gene PAX6 (paired box 6) era um dos que apresentavam maior expressão em oócitos de fêmeas adultas, em relação a outras que não haviam atingido a puberdade, e também é um potencial candidato associado à viabilidade dos oócitos e ao desenvolvimento embrionário precoce. Já o gene SLK (STE20-like kinase) é parte de uma rede envolvida no processo final de maturação dos oócitos após a fase dependente do hormônio luteinizante (LH) (ASSIDI et al., 2010).

Os transcritos do gene FADD (Fas TNFRSF6-associated via death domain), considerados pró-apoptóticos, apresentam níveis decrescentes nas células da granulosa de folículos dominantes em crescimento (MIHM et al., 2006), enquanto que nos folículos subordinados, os níveis crescentes causam diminuição da secreção de estradiol, ativação da apoptose, inibição da proliferação celular e consequente artresia folicular (EVANS et al., 2004; YANG et al., 2011). Em contrapartida, os transcritos anti-apoptóticos do gene TNFSF8 foram encontrados em blastocistos bovinos tratados com um importante regulador da função embrionária, o CSF2 (Colony-stimulating factor 2), e estão relacionados com a inibição da apoptose e garantem a pré implantação do embrião (LOUREIRO et al., 2011).

O gene PDGFB (platelet-derived growth factor beta polypeptide) e seu receptor são expressos no corpo lúteo de ratos e inibem a secreção de esteróides estimulada por LH (TAYLOR, 2000). Podem estar relacionados também com a migração e proliferação das células da teca e na microvascularização do parênquima luteal, após a fase dependente de LH (SLEER et al., 2007). A expressão no endométrio pode estar relacionada com a participação no recrutamento de macrófagos, importantes na gestação para a regulação imune, remodelamento do tecido e angiogênese e, no período pós parto, para a apoptose (OLIVEIRA et al., 2010).

Além disso, dos genes relacionados com a idade à puberdade em bovinos encontrados por FORTES et al. (2010), oito também foram identificados no presente estudo (ADAMTS3, ALDH7A1, FSHR, FYN, MSRA, RTL1, SPATA6 e TRIO), além de outros membros das mesmas famílias identificadas pelos autores. Considerando os ganhos genéticos que vêm sendo observados em características reprodutivas, principalmente idade ao primeiro parto, em bovinos da raça Nelore no Brasil, é possível considerar que o aumento da homozigose ao redor desses genes seja consequência dos programas de seleção.

Foram identificados ainda 18 genes envolvidos no transporte e/ou ligação de íons cálcio (C1R, C1S, CACNA1C, CACNA1E, CACNA1I, CALM1, CDH19, CDH2, CHGA, CLSTN3, DLK1, FKBP7, FKBP9, FYN, MAN1A1, PLEK, RCN1, SNTB1) e alguns deles já foram relacionados com características de importância econômica em bovinos de corte. Por exemplo, o gene CACNA1C (calcium channel, voltage-dependent, L type, R 1C subunit) foi correlacionado significativamente com maciez e suculência da carne em indivíduos da raça Charolês (BERNARD et al., 2007). O gene CALM1 (calmodulin 1) foi mais expresso em células adiposas, após a indução da adipogênese (TAN et al., 2006) e também no músculo *Psoas major* (filé mignon) quando comparado ao *Flexor digitorum* (músculo) (MORENO-SÁNCHEZ et al., 2010). SHACKELFORD et al. (1995) já haviam observado que, de 10 músculos estudados em animais cruzados (*Bos indicus* x *Bos taurus*), o *Psoas major* apresentava maior maciez e, a mesma decaia com o aumento da proporção de *Bos indicus*. Desta forma, pode-se sugerir a

investigação desses genes e de suas vias metabólicas para futuros estudos de qualidade de carne na raça Nelore.

Para visualizar a interação dos processos biológicos, funções moleculares e componentes celulares relacionados com os 287 genes que formaram os 20 grupos, a lista em ordem de importância, disponibilizada pela ferramenta DAVID (DENNIS JR et al., 2003) foi processada na plataforma GOrilla (EDEN et al., 2009). Os processos biológicos foram classificados no nível de maior hierarquia em metabolismo, resposta a estímulos, processos celulares, regulação biológica, morte celular e desenvolvimento. Dentro de cada classificação, inúmeras ramificações foram observadas, corroborando a hipótese de que as características sob seleção artificial são, não só controladas por muitos genes e regiões do genoma, como também envolvem diversos processos biológicos no organismo animal, incluindo a regulação de processos metabólicos, biossíntese, regulação da expressão gênica e apoptose. O mesmo padrão pôde ser observado em relação às funções moleculares e aos componentes celulares determinados pelos genes analisados.

Conclusões

Os haplótipos formados por polimorfismos de nucleotídeo único (SNPs) no genoma de bovinos Nelore apresentam tamanho relativamente pequeno, provavelmente pela menor extensão do desequilíbrio de ligação existente nesse grupo genético. A aplicação da metodologia da homozigose relativa do haplótipo estendido (REHH) identificou mais de 2.000 possíveis regiões de assinaturas de seleção, distribuídas de forma não uniforme ao longo do genoma, sugerindo que as características sob pressão de seleção são poligênicas.

Os padrões de assinaturas de seleção observados nos genes testados previamente evidenciam os efeitos que os programas de melhoramento genético têm causado no genoma de bovinos da raça Nelore, uma vez que alguns genes

relacionados com as características selecionadas ao longo dos anos apresentam padrões de homozigose e frequência compatíveis com assinaturas de seleção recente.

Os resultados encontrados no presente trabalho podem auxiliar o avanço das pesquisas na raça Nelore em diferentes aspectos, já que a população estudada foi escolhida de modo a representar o rebanho brasileiro. Primeiramente, a identificação das regiões que estão sob pressão de seleção artificial pode direcionar estudos para regiões ou grupos de genes específicos. Além disso, os polimorfismos e haplótipos observados podem ser utilizados em estudos de associação e validação em diferentes sub-populações da raça Nelore, com o intuito de viabilizar, futuramente, a seleção genômica de indivíduos superiores para características específicas, diminuindo a necessidade de genotipar todos os indivíduos em painéis de larga escala.

REFERÊNCIAS

ABIEC – Associação Brasileira das Indústrias de Carne. **Exportações de Carne Bovina do Brasil**. 2012.

ALEXANDER, L. J.; MACNEIL, M. D.; GEARY, T. W.; SNELLING, W. M.; RULE, D. C.; SCANGA, J. A. Quantitative trait loci with additive effects on palatability and fatty acid composition of meat in a Wagyu–Limousin F₂ population. **Animal Genetics**, v.38, n.5, p.506–513, 2007.

ALLENDORF, F.W. Genetic drift and the loss of alleles versus heterozygosity. **Zoo Biology**, v.5, p.181-190, 1986.

ALTSCHU, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Molecular Biology**, v.215, p.403-410, 1990.

ANDERSSON, L. Genome-wide association analysis in domestic animals: a powerful approach for genetic dissection of trait loci. **Genetica**, v.136, p.341-349, 2009.

ANDERSSON, L.; GEORGES, M. Domestic-animal genomics: Deciphering the genetics of complex traits. **Nature Reviews Genetics**, v.5, p.202-212, 2004.

ARDLIE, K.G.; KRUGLYAK, L.; SEIELSTAD, M. Patterns of linkage disequilibrium in the human genome. **Nature Reviews Genetics**, v.3, p.299-310, 2002.

ARIAS, J.A.; KEEHAN, M.; FISHER, P.; COPPIETERS, W.; SPELMAN, R. A high density linkage map of the bovine genome. **BMC Genetics**, v.10, n.18, 2009. Disponível em: <<http://www.biomedcentral.com/1471-2156/10/18>>. Acesso em: 06 jun. 2011.

ASSIDI, M.; DIELEMAN, S.J.; SIRARD, M.A. Cumulus cell gene expression following the LH surge in bovine preovulatory follicles: potential early markers of oocyte competence. **Reproduction**, v.140, p.835-852, 2010.

AYRES, D.R.; SOUZA, F.R.P.; MERCADANTE, M.E.Z.; FONSECA, L.F.S.; TONHATI, H.; CYRILLO, J.N.S.G.; BONILHA, S.F.M.; ALBUQUERQUE, L.G. Evaluation of TFAM and FABP4 gene polymorphisms in three lines of Nellore cattle selected for growth. **Genetics and Molecular Research**, v.9, n.4, p.2050-2059, 2010.

BAEZA, M.C.; CORVA, P.M.; SORIA, L.A.; RINCON, G.; MEDRANO, J.F.; PAVAN, E.; VILLARREAL, E.L.; SCHOR, A.; MELUCCI, L.; MEZZADRA, C.; MIQUEL, M.C. Genetic markers of body composition and carcass quality in grazing Brangus steers. **Genetics and Molecular Research**, v.10, n.4, p.3146-3156, 2011.

BARENDSE, W.; BUNCH, R.J.; THOMAS, M.B.; HARRISON, B.E. A splice site single nucleotide polymorphism of the fatty acid binding protein 4 gene appears to be associated with intramuscular fat deposition in longissimus muscle in Australian cattle. **Animal Genetics**, v.40, p.770-773, 2009.

BARRETT, J.C.; FRY, B.; MALLER, J.; DALY, M.J. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v.21, n.2, p.263-265, 2005.

BERNARD, C.; CASSAR-MALEK, I.; LE CUNFF, M.; DUBROEUCQ, H.; RENAND, G.; HOCQUETTE, J.F. New indicators of beef sensory quality revealed by expression of specif genes. **Journal of Agricultural and Food Chemistry**, v.55, p.5229-5237, 2007.

BISWAS, S.; AKEY, J.M. Genomic insights into positive selection. **Trends in Genetics**, v.22, n.8, p. 437-446, 2006.

BOLORMAA, S.; HAYS, B.J.; SAVIN, K.; HAWKEN, R.; BARENDSE, W.; ARTHUR, P.F.; HERD, R.M.; GODDARD, M.E. Genome-wide association studies for feedlot and growth traits in cattle. **Journal of Animal Science**, v.89, p.1684-1697, 2011.

BRADLEY, D.G.; CUNNINGHAM, E.P. Genetic aspects of domestication. In: R. FRIES, R.; RUVINSKY, A. (ed) **Genetics of cattle**, p.15-32, New York, USA: Ed CABI Publishing, 1999.

BREIER, B.H. Regulation of protein and energy metabolism by the somatotropic axis. **Domestic Animal Endocrinology**, v.17, p.209-218, 1999.

BRENNER, S.E. Errors in genome annotation. **Trends in Genetics**, v.15, n.4, p.132-133, 1999.

CASAS, E.; KEELE, J.W.; SHACKELFORD, S.D.; KOOHMARAIE, M.; STONE, R.T. Identification of quantitative trait loci for growth and carcass composition in cattle. **Animal Genetics**, v.35, n.1, p.2-6, 2003.

CHAN, E.K.F.; HAWKEN, R.; REVERTER, A. The combined effect of SNP-marker and phenotype attributes in genome-wide association studies. **Animal Genetics**, v.40, n.2, p.149-156, 2008.

CLAYTON, D.; LEUNG, H.T. **SnpMatrix: The snp.matrix and X.snp.matrix classes**. R package version 1.16.5, 2011. Disponível em <http://www-gene.cimr.cam.ac.uk/clayton/software/>>. Acesso em: 26 jan. 2011.

COSTA, D.; ABREU, J.B.R.; MOURÃO, R.C.; SILVA, J.C.G.; RODRIGUES, V.C.; SOUSA, J.C.D.; MARQUES, R.A.F.S. Características de carcaça de novilhos inteiros Nelore e F1 Nelore x Holandês. **Ciência Animal Brasileira**, v. 8, n. 4, p. 687-696, 2007.

CURI, R.A.; OLIVEIRA, H.N.; SILVEIRA, A.C.; LOPES, C.R. Association between IGF-I, IGF-IR and GHRH gene polymorphisms and growth and carcass traits in beef cattle. **Livestock Production Science**, v.94, p.159-167, 2005.

CURI, R.A.; PALMIERI, D.A.; SUGISAWA, L.; OLIVEIRA, H.N.; SILVEIRA, A.C.; LOPES, C.R. Growth and carcass traits associated with GH1/Alu I and POU1F1/Hinf I gene polymorphisms in Zebu and crossbred beef cattle. **Genetics and Molecular Biology**, v.29, n.1, p.56-61, 2006.

CURI, R.A.; CHARDULO, L.A.L.; ARRIGONI, M.D.B.; SILVEIRA, A.C.; OLIVEIRA, H.N. Association between LEP, DGAT1 and FABP4 gene polymorphisms and carcass and

meat traits in Nelore and crossbred beef cattle. **Livestock Science**, v.135, p.244-250, 2011.

DALBIÈS-TRAN, R.; MERMILLOD, P. Use of heterologous complementary DNA array screening to analyze bovine oocyte transcriptome and its evolution during in vitro maturation. **Biology of Reproduction**, v.68, p.252-261, 2003.

DEKKERS, J.C.M. Commercial application of marker and gene assisted selection in livestock: Strategies and lessons. **Journal of Animal Science**, v.82, p.313-328, 2004.

DENNIS JR, G.; SHERMAN, B.T.; HOSACK, D.A.; YANG, J.; GAO, W.; LANE, H.C.; LEMPICKI, R.A. DAVID: Database for Annotation, Visualization, and Integrated Discovery. **Genome Biology**, v.4, n.9, 2003. Disponível em: <<http://genomebiology.com/2003/4/9/r60>>. Acesso em: 23 mar. 2012.

DORJI; OHKUBO, Y.; MIYOSHI, K.; YOSHIDA, M. Gene expression differences in oocytes derived from adult and prepubertal Japanese Black cattle during in vitro maturation. **Reproduction in Domestic Animals**, v.47, p.392-402, 2012.

EDEN, E.; NAVON, R.; STEINFELD, I.; LIPSON, D.; YAKHINI, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. **BMC Bioinformatics**, v.10, n.48, 2009. Disponível em: <<http://www.biomedcentral.com/1471-2105/10/48/>>. Acesso em: 06 jun. 2012.

EVANS, A.C.O.; IRELAND, J.L.H.; WINN, M.E.; LONERGAN, P.; SMITH, G.W.; COUSSENS, P.M.; IRELAND, J.J. Identification of genes involved in apoptosis and dominant follicle development during follicular waves in cattle. **Biology of Reproduction**, v.70, p.1475-1484, 2004.

FARIA, F.J.C.; FILHO, A.E.V.; MADALENA, F.E.; JOSAHKIAN, L.A. Pedigree analysis in the Brazilian Zebu breeds. **Journal of Animal Breeding and Genetics**, v.126, p.148-153, 2009.

FAY, J.C.; WU, C. Hitchhiking under positive Darwinian selection. **Genetics**, v.155, p.1405-1413, 2000.

FINNER, H.; STRASSBURGER, K.; HEID, I.M.; HERDER, C.; RATHMANN, W.; GIANI, G.; DICKHAUS, T.; LICHTNER, P.; MEITINGER, T.; WICHMANN, H.E.; ILLIG, T.; GIEGERG, C. How to link call rate and p-values for Hardy–Weinberg equilibrium as measures of genome-wide SNP data quality. **Statistics in Medicine**, v.29, p.2347-2358, 2010.

FNP Consultoria & Comércio. **Anualpec 2012**: Anuário da Pecuária Brasileira. São Paulo, SP. 2012. 378p.

FORTES, M.R.S.; REVERTER, A.; ZHANG, Y.; COLLIS, E.; NAGARAJ, S.H.; JONSSON, N.N.; PRAYAGA, K.C.; BARRIS, W.; HAWKEN, R.J. Association weight matrix for the genetic dissection of puberty in beef cattle. **Proceedings of the National Academy of Science of the United States of America**, v.107, n.31, p.13642-13647, 2010.

FORTES, M.R.S.; LI, Y.; COLLIS, E.; ZHANG, Y.; HAWKEN, R.J. The IGF1 pathway genes and their association with age of puberty in cattle. **Animal Genetics**, 2012. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2052.2012.02367.x/full>>. Acesso em: 19 mai. 2012.

GABRIEL, S.B.; SCHAFFNER, S.F.; NGUYEN, H.; MOORE, J.M.; ROY, J.; BLUMENSTIEL, B.; HIGGINS, J.; DEFELICE, M.; LOCHNER, A.; FAGGART, M.; LIU-CORDERO, S.N.; ROTIMI, C.; ADEYEMO, A.; COOPER, R.; WARD, R.; LANDER, E.S.; DALY, M.J.; ALTSHULER, D. The structure of haplotype blocks in the human genome. **Science**, v.296, p.2225-2229, 2002.

GAO, L.; ZAN, L.S.; WANG, H.B. et al. Polymorphismo of somatostatin gene and its association with growth traits in Chinese cattle. **Genetics and Molecular Research**, v.10, n.2, p.703-711, 2011.

GE, W.; DAVIS, M.E.; HINES, H.C.; IRVIN, K.M.; SIMMEN, R.C. Association of a genetic marker with blood serum insulin-like growth factor-I concentration and growth traits in Angus cattle. **Journal of Animal Science**, v.79, p.1757-1762, 2001.

GODDARD, M.E.; WRAY, N.R.; VERBYLA, K.; VISSCHER, P.M. Estimating effects and making predictions from genome-wide marker data. **Statistical Science**, v.24, n.4, p. 517-529, 2009.

GRISOLIA, A.B.; D'ANGELO, G.T.; PORTO NETO, L.R.; SIQUEIRA, F.; GARCIA, J.F. Myostatin (GDF8) single nucleotide polymorphisms in Nellore cattle. **Genetics and Molecular Research**, v.8, n.2, p.822-830, 2009.

GUO, Y.; CHEN, H.; LAN, X. et al. Novel SNPs of the bovine LEPR gene and their association with growth traits. **Biochemical Genetics**, v.46, p.828-834, 2008.

GUTIÉRREZ-GIL, B.; WIENER, P.; NUTE, G.R.; BURTON, D.; GILL, J.L.; WOOD, J.D.; WILLIAMS, J.L. Detection of quantitative trait loci for meat quality traits in cattle. **Animal Genetics**, v.39, n.1, p.51–61, 2008.

HANOTTE, O.; RONIN, Y.; AGABA, M.; NILSSON, P.; GELHAUS, A.; HORSTMANN, R.; SUGIMOTO, Y.; KEMP, S.; GIBSON, J.; KOROL, A.; SOLLER, M.; TEALE, A. Mapping of quantitative trait loci controlling trypanotolerance in a cross of tolerant West Africa N'Dama and susceptible East African Boran cattle. **Proceedings of the National Academy of Science of the United States of America**, v.100, n.13, p.7443-7448, 2003.

HAYES, B.J.; LIEN, S.; NILSEN, H. OLSEN, H.G.; BERG, P.; MACEACHERN, S.; POTTER, S.; MEUWISSEN, T.H.E. The origin of selection signatures on bovine chromosome 6. **Animal Genetics**, v.39, p.105-111, 2008.

HEDRICK, F.W. Gametic disequilibrium measures: proceed with caution. **Genetics**, v.117, p.331-341, 1987.

HICKEY, J.M.; KINGHORN, B.P.; TIER, B.; WILSON, J.F.; DUNSTAN, N.; VAN DER WERF, J.H.J. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. **Genetics Selection Evolution**, v.43, n.12, 2011. Disponível em: <<http://www.gsejournal.org/content/43/1/12/>>. Acesso em: 18 nov. 2011.

HILL, W.G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **Theoretical and Applied Genetics**, v.38, n.6, p.226-231, 1968.

HOASHI, S.; HINENOYA, T.; TANAKA, A.; OHSAKI, H.; SASAZAKI, S.; TANIGUCHI, M.; OYAMA, K.; MUKAI, F.; MANNEN, H. Association between fatty acid compositions and genotypes of FABP4 and LXR-alpha in Japanese Black cattle. **BMC Genetics**, v.9, n.84, 2008. Disponível: <<http://www.biomedcentral.com/1471-2156/9/84/>>. Acesso em: 25 mai. 2012.

HOSACK, D.A.; DENNIS JR, G.; SHERMAN, B.T.; LANE, H.C.; LEMPICKI, R.A. Identifying biological themes within lists of genes with EASE. **Genome Biology**, v.4, n.10, 2003. Disponível em: <<http://genomebiology.com/2003/4/10/r70>>. Acesso em: 13 abr. 2012.

HU, Z.L.; PARK, C.A.; FRITZ, E.C.; REECY, J.M. QTLdb: A Comprehensive Database Tool Building Bridges between Genotypes and Phenotypes. In: **9th World Congress on Genetics Applied to Livestock Production**. Leipzig, Germany August 1-6, 2010.

HUANG, D.W.; SHERMAN, B.T.; TAN, Q.; COLLINS, J.R.; ALVORD, W.G.; ROAYAEI, J.; STEPHENS, R.; BASELER, M.W.; LANE, H.C.; LEMPICKI, R.A. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. **Genome Biology**, v.8, n.9, 2007. Disponível em: <<http://genomebiology.com/2007/8/9/R183>>. Acesso em: 13 abr. 2012.

INNAN, H.; KIM, Y. Pattern of polymorphism after strong artificial selection in a domestication event. **Proceedings of the National Academy of Science of the United States of America**, v.101, n.29, p.10667-10672, 2004.

INNAN, H.; KIM, Y. Detecting local adaptation using the joint sampling of polymorphism data in the parental and derived populations. **Genetics**, v.179, p.1713-1720, 2008.

JIANG, J.; JIANG, L.; ZHOU, B.; FU, W.; LIU, J.F.; ZHAN, Q. SnaT: a SNP annotation tool for bovine by integrating various sources of genomic information. **BMC Genetics**, v.12, n.85, 2011. Disponível em: <<http://www.biomedcentral.com/1471-2156/12/85>>. Acesso em: 26 jan. 2012.

KHATKAR, M.S.; COLLINS, A.; CAVANAGH, J.A.L.; HAWKEN, R.J.; HOBBS, M.; ZENGER, K.R.; BARRIS, W.; McCLINTOCK, A.E.; THOMSON, P.C.; NICHOLAS, F.W.; RAADSMA, H.W. A first-generation metric linkage disequilibrium map of bovine chromosome 6. **Genetics**, v.174, p.79-85, 2006.

KIM, Y.; NIELSEN, R. Linkage disequilibrium as a signature of selection sweeps. **Genetics**, v.167, p.1513-1524, 2004.

LAUREANO, M.M.M.; BOLIGON, A.A.; COSTA, R.B.; FORNI, S.; SEVERO, J.L.P.; ALBUQUERQUE, L.G. 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.

LAURIE, C.C.; DOHENY, K.F.; MIREL, D.B.; PUGH, E.W.; BIERUT, L.J.; BHANGALE, T.; BOEHM, F.; CAPORASO, N.E.; CORNELIS, M.C.; EDENBERG, H.J.; GABRIEL, S.B.; HARRIS, E.L.; HU, F.B.; JACOBS, K.B.; KRAFT, P.; LANDI, M.T.; LUMLEY, T.; MANOLIO, T.A.; MCHUGH, C.; PAINTER, I.; PASCHALL, J.; RICE, J.P.; RICE, K.M.; ZHENG, X.; WEIR, B.S. Quality control and quality assurance in genotypic data for genome-Wide association studies. **Genetic Epidemiology**, v.34, p.591–602, 2010.

LENSTRA, J.A.; BRADLEY, D.G. Systematics and phylogeny of cattle. In: R. FRIES, R.; RUVINSKY, A. (ed) **Genetics of cattle**, p.15-32, New York, USA: Ed CABI Publishing, 1999.

LEWONTIN, R.C. The interaction of selection and linkage. I. General considerations; Heterotic models. **Genetics**, v. 49, p.49-67, 1964.

LI, C.; BASARAB, J.; SNELLING, W.M.; BENKEL, B.; KNEELAND, J.; MURDOCH, B.; HANSEN, C.; MOORE, S.S. Identification and fine mapping of quantitative trait loci for backfat on bovine chromosomes 2, 5, 6, 19, 21, and 23 in a commercial line of *Bos taurus*. **Journal of Animal Science**, v.82, p.967–972, 2004.

LIRON, J.P.; PRANDO, A.J.; FERNANDEZ, M.E.; RIPOLI, M.V.; ROGBERG-MUNOZ, A.; GOSZCZYNSKI, D.E.; POSIK, D.M.; PERAL-GARCIA, P.; BALDO, A.; GIOVAMBATTISTA, G. Association between GNRHR, LHR and IGF1 polymorphisms and timing of puberty in male Angus cattle. **BMC Genetics**, v.13, n.26, 2012. Disponível em: <<http://www.biomedcentral.com/1471-2156/13/26>>. Acesso em: 19 mai. 2012.

LIU, Y.; QIN, X.; SONG, X.Z.H.; JIANG, H.; SHEN, Y.; DURBIN, K.J.; LIEN, S.; KENT, M.P.; SODELAND, M.; REN, Y.; ZHANG, L.; SODERGREN, E.; HAVLAK, P.; WORLEY, K.; WEINSTOCK, G.M.; GIBBS, R.A. *Bos taurus* genome assembly. **BMC Genomics**, v.10, n.180, 2009. Disponível em: <<http://www.biomedcentral.com/1471-2164/10/180>>. Acesso em: 06 jun. 2011.

LOUREIRO, B.; OLIVEIRA, L.J.; FAVORETO, M.; HANSEN, P. Colony-stimulating factor 2 inhibits induction of apoptosis in the bovine preimplantation embryo. **American Journal of Reproductive Immunology**, v.65, p.578-588, 2011.

LUIKART, G.; ENGLAND, P.R.; TALLMON, D.; JORDAN, S.; TABERLET, P. The power and promise of population genomics: from genotyping to genome typing. **Nature Reviews Genetics**, v.4, p. 981-994, 2003.

MACNEIL, M.D.; GROSZ, M.D. Genome-wide scans for QTL affecting carcass traits in Hereford x composite double backcross populations. **Journal of Animal Science**, v. 80, p.2316-2324, 2002.

MARTIN, J.B.; MILLARD, W. Brain regulation of growth hormone secretion. **Journal of Animal Science**, v.63, p.11-26, 1986.

MARTINEZ, M.L.; MACHADO, M.A.; NASCIMENTO, C.S.; SILVA, M.V.G.B.; TEODORO, R.L.; FURLONG, J.; PRATA, M.C.A.; CAMPOS, A.L.; GUIMARÃES, M.F.M.; AZEVEDO, A.L.S.; PIRES, M.F.A.; VERNEQUE, R.S. Association of BoLA-DRB3.2 alleles with tick (*Boophilus microplus*) resistance in cattle. **Genetics and Molecular Research**, v.5, n.3, p.513-524, 2006.

MATUKUMALLI, L.K.; LAWLEY, C.T.; SCHNABEL, R.D.; TAYLOR, J.F.; ALLAN, M.F.; HEATON, M.P.; O'CONNELL, J.; MOORE, S.S.; SMITH, T.P.L.; SONSTEGARD, T.S.; VAN TASSEL, C.P. Development and characterization of a high density SNP genotyping assay for cattle. **PLoS ONE**, v.4, n.4, 2009. Disponível em: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0005350>>.

Acesso em: 06 jun. 2011.

MCCLURE, M.C.; MORSCI, N.S.; SCHNABEL, R.D.; KIM, J.M.; YAO, P.; ROLF, M.M.; MCKAY, S.D.; GREGG, S.J.; CHAPPLE, R.H.; NORTHCUTT, S.L.; TAYLOR, J.F. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. **Animal Genetics**, v.41, p.597–607, 2010.

MCKAY, S.D.; 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 Genetics**, v.8, n.74, 2007. Disponível em: <http://www.biomedcentral.com/1471-2156/8/74/>>. Acesso em: 23 abr. 2012.

MCPHERRON, A.C.; LEE, S.J. Double muscling in cattle due to mutations in the myostatin gene. **Proceedings of the National Academy of Science of the United States of America**, v.94, p.12457-12461, 1997.

MICHAL, J.J.; ZHANG, Z.W.; GASKINS, C.T.; JIANG, Z. The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. **Animal Genetics**, v.37, p.400-402, 2006.

MICKE, G.C.; SULLIVAN, T.M.; MCMILLEN, I.C.; GENTILI, S.; PERRY, V.E.A. Protein intake during gestation affects postnatal bovine skeletal muscle growth and relative expression of IGF1, IGF1R, IGF2 and IGF2R. **Molecular and Cellular Endocrinology**, v.332, p.234–241, 2011.

MIHM, M.; BAKER, P.J.; IRELAND, J.L.H.; SMITH, G.W.; COUSSENS, P.M.; EVANS, A.C.O.; IRELAND, J.J. Molecular evidence that growth of dominant follicles involves a reduction in follicle-stimulating hormone dependence and an increase in luteinizing hormone dependence in cattle. **Biology of Reproduction**, v.74, p.1051-1059, 2006.

MISSIO, R.S.; BRONDANI, I.L.; ALVES FILHO, D.C.; RESTLE, J.; ARBOITTE, M.Z.; SEGABINAZZI, L.R. Características da carcaça e da carne de tourinhos terminados em confinamento, recebendo diferentes níveis de concentrado na dieta. **Revista Brasileira de Zootecnia**, v.39, n.7, p.1610-1617, 2010.

MORENO-SÁNCHEZ, N.; RUEDA, J.; CARABAÑO, M.J.; REVERTER, A.; MCWILLIAN, S.; GONZÁLEZ, C.; DÍAZ, CL. Skeletal muscle specific genes networks in cattle. **Functional & Integrative Genomics**, v.10, p.609-618, 2010.

MORRIS, C.A.; BOTTEMA, C.D.K.; CULLEN, N.G.; HICKEY, S.M.; ESMAILIZADEH, A.K.; SIEBERT, B.D.; PITCHFORD, W.S. Quantitative trait loci for organ weights and adipose fat composition in Jersey and Limousin back-cross cattle finished on pasture or feedlot. **Animal Genetics**, v.41, p.589–596, 2010.

MORSCI, N.S.; SCHNABEL, R.D.; TAYLOR, J.F. Association analysis of adiponectin and somatostatin polymorphisms on BTA1 with growth and carcass traits in Angus cattle. **Animal Genetics**, v.37, p.554-562, 2006.

MUDADO, M.A.; REGITANO, L.C.A.; TIZIOTO, P.C.; SIQUEIRA, F.; ROSA, A.N.; HIGA, R. Finding a minimal set of SNPs for paternity identification in Nelore cattle. In: **X - meeting**, 2011.

MURDOCH, W.J.; VAN KIRK, E.A. Estrogenic upregulation of DNA polymerase β in oocytes of preovulatory ovine follicles. **Molecular Reproduction and Development**, v.58, p.417-423, 2001.

OLIPHANT, A.; BARKER, D.L.; STUELPNAGEL, J.R.; CHEE, M.S. BeadArrayTM Technology: Enabling an accurate, cost-effective approach to high-throughput genotyping. **BioTechniques**, v.32, p.56-61, 2002.

OLIVEIRA, L.J.; MCCLELLAN, S.; HANSEN, P.J. Differentiation of the endometrial macrophage during pregnancy in the cow. **PLoS ONE**, v.5, n.10, 2010. Disponível em: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0013213>>. Acesso em: 05 jun. 2012.

PAREEK, C.S.; SMOCZYNSKI, R.; PIERZCHALA, M.; CZAMIK, U.; TRETYN, A. From genotype to phenotype in bovine functional genomics. **Briefings in Functional Genomics**, v.10, n.3, p.165-171, 2011.

PIPER, E.K.; JONSSON, N.N.; GONDRO, C.; LEW-TABOR, A.E.; MOOLHUIJZEN, P.; VANCE, M.E.; JACKSON, L.A. Immunological profiles of Bos taurus and Bos indicus cattle infested with the cattle tick, Rhipicephalus (Boophilus) microplus. **Clinical and Vaccine Immunology**, v.16, n.7, p.1074-1086, 2009.

POLINENI, P.; ARAGONDA, P.; XAVIER, S.R.; FURUTA, R.; ADELSON, D.L. The bovine QTL viewer: a web accessible database of bovine Quantitative Trait Loci. **BMC Bioinformatics**, v.7, 2006. Disponível em: <http://www.biomedcentral.com/1471-2105/7/283>>. Acesso em: 02 fev. 2012.

QANBARI, S.; PIMENTEL, E.C.G.; TETENS, J.; THALLER, G.; LICHTNER, P.; SHARIFI, A.R.; SIMIANER, H. A genome-wide scan for signatures of recent selection in

Holstein cattle. **Animal Genetics**, 2010. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2052.2009.02016.x/full>. Acesso em: 03 jun. 2011.

QANBARI, S.; GIANOLA, D.; HAYES, B.; SCHENKEL, F.; MILLER, S.; MOORE, S.; THALLER, G.; SIMIANER, H. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. **BMC Genomics**, v.12, 2011. Disponível em: <http://www.biomedcentral.com/1471-2164/12/318>. Acesso em: 03 jun. 2011.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. ISBN 3-900051-07-0, 2012. Disponível em: <http://www.R-project.org>.

REED, J.L.; FAMILI, I.; THIELE, I.; PALSSON, B.O. Towards multidimensional genome annotation. **Nature Reviews Genetics**, v.7, p.130-141, 2006.

ROGBERG-MUÑOZ, A.; MELUCCI, L.; PRANDO, A.; VILLEGAS-CASTAGNASSO, E.E.; RIPOLI, M.V.; PERAL-GARCÍA, P.; BALDO, A.; AÑON, M.C.; GIOVAMBATTISTA, G. Association of bovine chromosome 5 markers with birth and weaning weight in Hereford cattle raised under extensive conditions. **Livestock Science**, v.135, p124–130, 2011.

ROUZÉ, P.; PAVY, N.; ROMBAUTS, S. Genome annotation: which tools do we have for it? **Current Opinion in Plant Biology**, v.2, p.90–95, 1999.

RYU, J.; KIM, Y.; KIM, C.; KIM, J.; LEE, C. Association of bovine carcass phenotypes with genes in an adaptive thermogenesis pathway. **Molecular Biology Reports**, v.39, n.2, p.1441-1445, 2012.

SABETI, P.C.; REICH, D.E.; HIGGINS, J.M.; LEVINE, H.Z.P.; RICHTER, D.J.; SCHAFFNER, S.F.; GABRIEL, S.B.; PLATKO, J.V.; PATTERSON, N.J.; MCDONALD, G.J.; ACKERMAN, H.C.; CAMPBELL, S.J.; ALTSHULER, D. COOPER, R.;

KWIATKOWSKI, D.; WARD, R.; LANDER, E.S. Detecting recent positive selection in the human genome from haplotype structure. **Nature**, v.419, p.832-837, 2002.

SAHANA, G.; GULDBRANDTSEN, B.; BENDIXEN, C.; LUND, M.S. Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. **Animal Genetics**, v.41, p.579-588, 2010.

SALIH, H.; ADELSON, D.L. QTL global meta-analysis: are trait determining genes clustered? **BMC Genomics**, v.10, n.184, 2009. Disponível em: <http://www.biomedcentral.com/1471-2164/10/184/>. Acesso em: 06 jun. 2011.

SHACKELFORD, S.D.; WHEELER, T.L.; KOOHMARAIE, M. Relationship between shear force and trained sensory panel tenderness ratings of 10 major muscles from *Bos indicus* and *Bos taurus* cattle. **Journal of Animal Science**, v.73, p.3333-3340, 1995.

SHEN, W.J.; SRIDHAR, K.; BERNLOHR, D.A.; KRAEMER, F.B. Interaction of rat hormone-sensitive lipase with adipocyte lipid-binding protein. **Proceedings of the National Academy of Science of the United States of America**, v.96, p.5528-5532, 1999.

SHERMAN, E.L.; NKRUMAH, J.D.; LI, C.; BARTUSIAK, R.; MURDOCH, B.; MOORE, S.S. Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle. **Journal of Animal Science**, v.87, p.37-45, 2009.

SILVEIRA, L.G.G.; FURLAN, L.R.; CURI, R.A.; FERRAZ, A.L.J.; ALENCAR, M.M.; REGITANO, L.C.A.; MARTINS, C.L.; ARRIGONI, M.B.; SUGUISAWA, L.; SILVEIRA, A.C.; OLIVEIRA, H.N. Growth hormone 1 gene (GH1) polymorphisms as possible markers of the production potential of beef cattle using the Brazilian Canchim breed as a model. **Genetics and Molecular Biology**, v.31, n.4, p.874-879, 2008.

SIMIANER, H.; QANBARI, S.; GIANOLA, D. Detection of selection signatures within and between cattle populations. In: **9th world congress on genetic applied to livestock production**, 1-6.8, Leipzig, Germany, 2010.

SLEER, L.S.; TAYLOS, C.C. Platelet-derived growth factors and receptors in the rat corpus luteum: localization and identification of an effect on luteogenesis. **Biology of Reproduction**, v.76, p.391-400, 2007.

SMITH, S.L.; EVERTS, R.E.; TIAN, X.C.; DU, F.; SUNG, L.Y.; RODRIGUEZ-ZAS, S.L.; JEONG, B.S.; RENARD, J.P.; LEWIN, H.A.; YANG, X. Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. **Proceedings of the National Academy of Science of the United States of America**, v.102, n.49, p.17582-17587, 2005.

STEIN, L. Genome annotation: From sequence to biology. **Nature Review Genetics**, v.2, p.493-503, 2001.

STEIN, L.D.; MUNGALL, C.; SHU, S.; CAUDY, M.; MANGONE, M.; DAY, A.; NICKERSON, E.; STAJICH, J.E.; HARRIS, T.W.; ARVA, A.; LEWIS, S. The generic genome browser: A building block for a model organism system database. **Genome Research**, v.12, p.1599-1610, 2002.

STELLA, A.; AJMONE-MARSAN, P.; LAZZARI, B.; BOETTCHER, P. Identification of selection signatures in cattle breeds selected for dairy production. **Genetics**, v.185, p.1451-1461, 2010.

TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, v.123, p.585-595, 1989.

TAN, S.H.; REVERTER, A.; WANG, Y.; BYRNE, K.A.; MCWILLIAM, S.M.; LEHNERT, S.A. Gene expression profiling of bovine in vitro adipogenesis using cDNA microarray. **Functional & Integrative Genomics**, v.6, p.235-249, 2006.

TARTAGLIA, L.A. The leptin receptor. **The Journal of Biological Chemistry**, v.272, n.10, p.6093-6096, 1997.

TAYLOR, C.C. Platelet-derived growth factor activates porcine thecal cell phosphatidylinositol-3-kinase-akt/PKB and ras-extracellular signal-regulated kinase-1/2 kinase signaling pathways via the platelet-derived growth factor- β receptor. **Endocrinology**, v.141, n.4, p.1545-1553, 2000.

TAYLOR, J.F.; COUTINHO, L.L.; HERRING, K.L.; GALLAGHER JR, D.S.; BRENNEMAN, R.A.; BURNEY, N.; SANDERS, J.O.; TURNER, J.W.; SMITH, S.B.; MILLER, R.K.; SAVELL, J.W.; DAVIS, S.K. Candidate gene analysis of GH1 for effects on growth and carcass composition of cattle. **Animal Genetics** v.29, p.194-201, 1998.

TELLAM, R.L.; LEMAY, D.G.; VAN TASSEL, C.P.; LEWIN, H.A.; WORLEY, K.C.; ELSIK, C.G. Unlocking the bovine genome. **BMC Genomics**, v.10, n.193, 2009. Disponível em: <<http://www.biomedcentral.com/1471-2164/10/193>>. Acesso em: 06 jun. 2011.

TEO, Y.Y. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. **Current Opinion in Lipidology**, v.19, p.133–143, 2008.

THE BOVINE HAPMAP CONSORTIUM et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. **Science**, v.324, p.528-531, 2009.

THOMAS, P.D.; MI, H.; LEWIS, S. Ontology annotation: mapping genomic regions to biological function. **Current Opinion in Chemical Biology**, v.11, p.4–11, 2007.

VIERGUTZ, T.; LÖHRKE, B. Role of the platelet-activating factor and its receptor in the proliferative regulation of bovine ovarian granulosa cells. **Cell proliferation**, v.40, p.949-960, 2007.

VILLA-ANGULO, R.; MATUKUMALLI, L.K.; GILL, C.A.; CHOI, J.; VAN TASSEL, C.P.; GREFENSTETTE, J.J. High-resolution haplotype block structure in the cattle genome. **BMC Genetics**, v.10, n.19, 2009. Disponível em: <<http://www.biomedcentral.com/1471-2156/10/19>>. Acesso em: 04 jun. 2011.

VISSCHER, P.M.; HILL, W.G.; WRAY, N.R. Heritability in the genomics era – concepts and misconceptions. **Nature Reviews Genetics**, v.9, p255-266, 2008.

VOIGHT, B.F.; KUDARAVALLI, S.; WEN, X.; PRITCHARD, J.L. A map of recent positive selection in the human genome. **PLoS Biology**, v.4, n.3, 2006. Disponível em: <<http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0040072>>. Acesso em: 08 ago. 2011.

WEIR, B.S.; COCKERHAM, C.C. Estimating F-statistics for the analysis of population structure. **Evolution**, v.38, n.6, p.1358-1370, 1984.

WEISS, K.M.; CLARK, A.G. Linkage disequilibrium and the mapping of complex human traits. **Trends in Genetics**, v.18, n.1, p.19-24, 2002.

WIENER, P.; BURTON, D.; AJMORE-MARSAN, P.; DUNNER, S.; MOMMENS, G.; NIJMAN, I.; RODELLAR, C.; VALENTINI, A.; WILLIAMS, J.L. Signatures of selection? Patterns of microsatellite diversity on a chromosome containing a selected locus. **Heredity**, v.90, p.350-358, 2003.

WIENER, P.; WOOLLIAMS, J.A.; FRANK-LAWALE, A.; RYAN, M.; RICHARDSON, R.I.; NUTE, G.R.; WOOD, J.D.; HOMER, D.; WILLIAMS, J.L. The effects of a mutation in the myostatin gene on meat and carcass quality. **Meat Science**, v.83, p.127–134, 2009.

WOODS, A.M.; JUDD, A.M. Interleukin-4 increases cortisol release and decreases adrenal androgen release from bovine adrenal cells. **Domestic Animal Endocrinology**, v.34, p.372–382, 2008.

YANG, R.J.; LI, J.Y.; GAO, X.; ZHAO, Z.H.; ZHANG, L.P.; GAO, H.J.; XU, S.Z. Fas-associated death domain (FADD) mediated activation of apoptosis program in bovine follicular granulosa cells. **Journal of Animal and Veterinary Advances**, v.10, n.15, p.1994-2003, 2011.