



**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**DESENVOLVIMENTO DE UM PAINEL DE MARCADORES SNP DE BAIXA  
DENSIDADE E ANÁLISE DE DESEQUILÍBRIO DE LIGAÇÃO DE  
POLIMORFISMOS EM GENES CANDIDATOS EM BOVINOS DA RAÇA  
GIROLANDO**

**RONYERE OLEGÁRIO DE ARAÚJO**

**TESE DE DOUTORADO EM CIÊNCIAS ANIMAIS**

**BRASÍLIA/DF  
DEZEMBRO DE 2014**



**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**DESENVOLVIMENTO DE UM PAINEL DE MARCADORES SNP DE BAIXA  
DENSIDADE E ANÁLISE DE DESEQUILÍBRIO DE LIGAÇÃO DE  
POLIMORFISMOS EM GENES CANDIDATOS EM BOVINOS DA RAÇA  
GIROLANDO**

**RONYERE OLEGÁRIO DE ARAÚJO**

**ORIENTADOR: ALEXANDRE RODRIGUES CAETANO, Ph.D.**

**CO-ORIENTADOR: SAMUEL REZENDE PAIVA, Ph.D.**

**TESE DE DOUTORADO EM CIÊNCIAS ANIMAIS**

**PUBLICAÇÃO: 122D/2014**

**BRASÍLIA/DF  
DEZEMBRO DE 2014**

## REFERÊNCIA BIBLIOGRÁFICA E CATALOGAÇÃO

ARAÚJO, R. O. **Desenvolvimento de um painel de marcadores snp de baixa densidade e análise de desequilíbrio de ligação de polimorfismos em genes candidatos em bovinos da raça Girolando.** Brasília: Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, 2014, 171p. Tese de Doutorado.

Documento formal, autorizando reprodução desta tese de doutorado para empréstimo ou comercialização, exclusivamente para fins acadêmicos, foi passado pelo autor à Universidade de Brasília e acha-se arquivado na Secretaria do Programa. O autor e seu orientador reservam para si os outros direitos autorais de publicação. Nenhuma parte desta tese de doutorado pode ser reproduzida sem a autorização por escrito do autor ou do seu orientador. Citações são estimuladas, desde que citada a fonte.

## FICHA CATALOGRÁFICA

ARAÚJO, Ronyere Olegário de. **Desenvolvimento de um painel de marcadores snp de baixa densidade e análise de desequilíbrio de ligação de polimorfismos em genes candidatos em bovinos da raça Girolando.** Brasília: Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, 2014, 171 p. Tese (Doutorado em Ciências Animais) – Faculdade de Agronomia e Medicina Veterinária. 2014.

1. Blocos de haplótipos. 2. Ferramentas genômicas 3. Marcadores moleculares. 4. TagSnps. I. Caetano, A. R. II. Ph.D.

**CDD ou CDU**  
**Agris/FAO**

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**DESENVOLVIMENTO DE UM PAINEL DE MARCADORES SNP DE BAIXA  
DENSIDADE E ANÁLISE DE DESEQUILÍBRIO DE LIGAÇÃO DE  
POLIMORFISMOS EM GENES CANDIDATOS EM BOVINOS DA RAÇA  
GIROLANDO**

**RONYERE OLEGÁRIO DE ARAÚJO**

**TESE DE DOUTORADO SUBMETIDA AO  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
CIÊNCIAS ANIMAIS, COMO PARTE DOS  
REQUISITOS NECESSÁRIOS À OBTENÇÃO DO  
GRAU DE DOUTOR EM CIÊNCIAS ANIMAIS.**

**APROVADA POR:**

---

Alexandre Rodrigues Caetano, Ph.D. (Embrapa Recursos Genéticos e Biotecnologia).  
(ORIENTADOR)

---

Concepta Margaret Mcmanus Pimentel, Ph.D (Universidade de Brasília).  
(EXAMINADOR INTERNO)

---

Francisco Ernesto Moreno Bernal, Ph.D (Universidade de Brasília).  
(EXAMINADOR INTERNO)

---

Tereza Cristina de Oliveira Borba, Dra. (Embrapa Arroz e Feijão).  
(EXAMINADOR EXTERNO)

---

Marta Fonseca Martins, Dra (Embrapa Gado de Leite).  
(EXAMINADOR EXTERNO)

**BRASÍLIA/DF – 15 de DEZEMBRO de 2014**

Aos maiores orientadores da minha vida, meu Pai José Maria e minha Mãe Mirna Olegário, DEDICO integralmente este trabalho

## AGRADECIMENTOS

À DEUS, fonte de inspiração e sabedoria,

Aos meus pais José e Mirna, irmãos Rogério, Dayane, Thauane, que sempre proporcionaram todo suporte e incentivos necessários para que eu pudesse prosseguir, e mesmo com a distância, sempre estiveram presentes apoiando-me irrestritamente na elaboração desta Tese,

À minha esposa Luciana por toda paciência, confiança, incentivo e amor durante todos esses anos de estudos,

Ao meu orientador, Ph.D. Alexandre Caetano, pela orientação, pela grande amizade, pelo incentivo e, sobretudo pela referência tanto como pessoa quanto profissional,

Ao Ph.D. Samuel Paiva, Ph.D. Danielle Paiva e a Professora Ph.D. Connie McManus, pelas orientações no desenvolvimento dessa Tese,

À Ph.D. Cintia Marcondes, a Professora Dra. Analia Garnero e ao Professor Dr. Ricardo Gunski pela amizade, apoio e incentivo a pesquisa durante toda minha vida acadêmica, minhas estimadas considerações,

Aos integrantes da banca, Dra. Tereza Borba, Dra. Marta Martins, Professora Ph.D. Connie McManus e Professor Ph.D. Francisco Bernal, declaro-me honrado por terem acrescentado novas idéias neste trabalho,

À Universidade de Brasília e à Embrapa Recursos Genéticos e Biotecnologia pelos efeitos de ambiente permanente que me proporcionaram na realização dessa Tese,

À Embrapa Gado de Leite, na pessoa da Dra. Marta Martins, pelas amostras de DNA dos animais que foram disponibilizados para a realização deste trabalho,

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro,

Aos meus amigos do Laboratório de Genética Animal da Embrapa Recursos Genéticos e Biotecnologia, em especial, Elizabete Cristina, Luciana Villela, Gleison Biazio, Patrícia Ianella e Carolina Celso, pela amizade, ensinamentos, conselhos e momentos de distração,

Ao Ph.D. Fernando Brito pela amizade, incentivo e pela valiosa colaboração nas análises estatísticas,

Aos amigos(as) do Laboratório de Melhoramento Animal (LMA – UFSM), Dra. Dionéia, Dr. Jader, Dra. Mariana, Dr. Tomás e Professor Ph.D. Paulo Rorato, pela amizade construída no Mestrado,

Aos grandes amigos(a) Elton, Fabrício, Juliana, Poliana, Ronaldo, Thompson e Wíris pela amizade consolidada na graduação,

Enfim, gostaria de agradecer aqueles de certa forma auxiliaram a realização deste trabalho, muito obrigado!

## INDICE

<b>RESUMO .....</b>	<b>xi</b>
<b>ABSTRACT .....</b>	<b>xiii</b>
<b>CAPÍTULO 1 .....</b>	<b>1</b>
<b>REVISÃO DE LITERATURA .....</b>	<b>1</b>
<b>1.1. INTRODUÇÃO .....</b>	<b>1</b>
<b>1.1.1. Justificativa .....</b>	<b>4</b>
<b>1.2. OBJETIVOS .....</b>	<b>4</b>
1.2.1. Objetivo geral .....	4
1.2.2. Objetivos específicos.....	5
<b>1.3. REVISÃO DE LITERATURA .....</b>	<b>6</b>
1.3.1. Características Lácteas .....	6
1.3.2. Sistemas de Cruzamento .....	7
1.3.3. Marcadores de polimorfismo de base única (SNP) .....	8
1.3.4. Tecnologia de genotipagem GoldenGate .....	11
1.3.5. Blocos de haplótipos e desequilíbrio de ligação.....	13
1.3.6. Medidas de desequilíbrio de ligação .....	14
<b>1.4. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>18</b>
<b>CAPÍTULO 2 .....</b>	<b>24</b>
<b>DESENVOLVIMENTO DE UM PAINEL DE BAIXA DENSIDADE DE MARCADORES SNP DERIVADOS DE GENES CANDIDATOS PARA CARACTERÍSTICAS DE PRODUÇÃO, GENES CAUSADORES DE DOENÇAS GENÉTICAS, E PARA CONTROLE DE PARENTESCO EM ANIMAIS DA RAÇA GIROLANDO .....</b>	<b>24</b>
<b>2.1. RESUMO .....</b>	<b>25</b>
<b>2.2. ABSTRACT .....</b>	<b>27</b>
<b>2.3. INTRODUÇÃO .....</b>	<b>29</b>



<b>2.4. MATERIAL E MÉTODOS.....</b>	<b>32</b>
2.4.1. Seleção de SNPs em Genes Candidatos .....	32
2.4.2. Seleção de amostras dos animais para Genotipagem .....	33
2.4.3. Quantificação das amostras de DNA genômico com PicoGreen .....	33
2.4.4. Genotipagem .....	34
2.4.5. Critérios para Controle de Qualidade dos Dados Gerados .....	37
2.4.6. Análise de Componentes Principais .....	37
2.4.7. Teste de Paternidade, de Herança Mendeliana e Reprodutibilidade dos Genótipos Obtidos .....	38
2.4.7. Avaliação dos Resultados de Controle de Qualidade dos Dados .....	38
2.4.8. Estudo do Equilíbrio de Hardy-Weinberg e Nível de Polimorfismo do SNP .....	39
<b>2.5. RESULTADOS.....</b>	<b>39</b>
<b>2.6. DISCUSSÃO.....</b>	<b>57</b>
<b>2.7. CONCLUSÃO .....</b>	<b>61</b>
<b>ANEXOS .....</b>	<b>62</b>
<b>Anexo 01: Painel inicial dos SNPs, Probabilidade de hibridização dos oligonucleotídeos estabelecidos pela empresa Illumina (Score Final e Designability Rank) e Parâmetros descritivos do resultado da genotipagem GoldenGate (GC50, CallFreq, MAF, HW, GenTrain). .....</b>	<b>62</b>
<b>Anexo 02: Exemplo do efeito da edição dos clusters sobre a dispersão dos genótipos.....</b>	<b>85</b>
<b>2.8. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>86</b>
<b>CAPÍTULO 3.....</b>	<b>90</b>
<b>ESTIMATIVAS DE DESEQUILÍBRIO DE LIGAÇÃO E AVALIAÇÕES DA ESTRUTURA DE BLOCOS HAPLOTÍPICOS EM GENES CANDIDATOS NOS CROMOSSOMOS BTA06 E BTA14.....</b>	<b>90</b>
<b>3.1. RESUMO .....</b>	<b>91</b>
<b>3.2. ABSTRACT .....</b>	<b>93</b>
<b>3.3. INTRODUÇÃO .....</b>	<b>95</b>
<b>3.4. MATERIAL E MÉTODOS.....</b>	<b>98</b>
3.4.1 Amostras e Seleção de SNPs em genes candidatos.....	98

3.4.2 Medida do Desequilíbrio de Ligação .....	99
<b>3.5. RESULTADOS.....</b>	<b>102</b>
3.5.1 – Análise descritiva dos polimorfismos genotipados .....	102
3.5.2 – Análises no cromossomo 06 (BTA06) .....	111
3.5.2.1. Desequilíbrio de Ligação e Blocos de Haplótipos .....	116
3.5.2.2. Estimativas de <i>TagSNPs</i> a partir dos Polimorfismos encontrados .....	126
3.5.3 – Análises do cromossomo 14 (BTA14) .....	128
3.5.3.1 - Desequilíbrio de Ligação e Blocos de Haplótipos .....	132
3.5.3.2. Estimativas de <i>TagSNPs</i> a partir dos Polimorfismos encontrados .....	137
<b>3.6. DISCUSSÃO.....</b>	<b>139</b>
<b>3.7. CONCLUSÃO .....</b>	<b>143</b>
<b>3.8. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>144</b>
<b>CAPÍTULO 4.....</b>	<b>148</b>
<b>CONSIDERAÇÕES FINAIS E PERSPECTIVAS FUTURAS.....</b>	<b>148</b>
<b>4.1. CONSIDERAÇÕES FINAIS .....</b>	<b>149</b>
4.1.1. Mudanças recentes nos procedimentos de avaliação genética e suas repercussões ....	149
4.1.2. Avaliação crítica da tecnologia Veracode (BeadXPress) – Perspectivas Futuras .....	151
<b>4.2. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>155</b>

**DESENVOLVIMENTO DE UM PAINEL DE MARCADORES SNP DE BAIXA DENSIDADE E ANÁLISE DE DESEQUILÍBRIO DE LIGAÇÃO DE POLIMORFISMOS EM GENES CANDIDATOS EM BOVINOS DA RAÇA GIROLANDO**

Autor: Ronyere Olegário de Araújo

Orientador: Alexandre Rodrigues Caetano, Ph.D.

**RESUMO**

Desenvolvimento de um painel de marcadores snp de baixa densidade e análise de desequilíbrio de ligação de polimorfismos em genes candidatos em bovinos da raça Girolando. Ronyere Olegário de Araújo e Alexandre Rodrigues Caetano – Pesquisador PhD, Embrapa Recursos Genéticos e Biotecnologia, Brasília/DF. A crescente evolução de tecnologias tornaram possível analisar o genoma de várias espécies, compreender suas funções dentro dos sistemas biológicos e, sobretudo, começar a entender os mecanismos que controlam as interações entre os genótipos e os efeitos ambientais que estão envolvidos com a expressão de características de interesse econômico. O objetivo deste trabalho foi desenvolver e validar um painel personalizado a partir de um ensaio de 384 marcadores SNP localizados em genes candidatos que afetam características de produção, doenças genéticas, e para controle de parentesco, em animais da raça Girolando. Um total de 576 animais foram testados com o painel de 384 SNPs. Os SNPs selecionados foram usados para construir um painel baseado na tecnologia *GoldenGate*®. Os dados brutos da genotipagem foram analisados com o Módulo de Genotipagem do programa *GenomeStudio*, V2010.1, utilizando parâmetros padrão para clusterização dos dados de cada SNP. Como critério de filtragem dos SNPs e das amostras foram adotados os parâmetros: Separação de Clusters, Frequência de Genotipagem e a Taxa de Genotipagem. Após aplicação destes critérios, o arquivo final ficou composto de 249 SNPs e 419 animais. Após aplicação dos critérios de filtragem, foi possível observar uma diminuição da proporção de SNPs polimórficos (53,4%). Deste painel, 72 e 47 SNPs, estavam localizados nos cromossomos BTA06 e BTA14, respectivamente, e foram utilizados para os estudos de desequilíbrio de ligação - DL (parâmetros  $D'$  e  $r^2$ ) e construção de blocos de haplótipos. As estimativas dos valores de DL entre os SNPs variaram de

moderada à baixa magnitude, entre si e entre os BTAs. Para o parâmetro  $D'$ , a menor estimativa foi obtida para o BTA15 (0,0954) e a maior para o BTA24 (0,5311). Por outro lado, o parâmetro  $r^2$  apresentou menor estimativa no BTA17 (0,0011) e a maior no BTA22 (0,0834). Observou-se na população estudada a existência de 32 haplótipos em seis regiões do BTA06, com frequências observadas variando de 0,011 a 0,970. O número de SNPs capturados (*TagSNPs*) no BTA6 variou de 2 a 4 entre as regiões estruturadas e a distância entre eles variou de 75 a 70.177 pb. Para o BTA14, a estimativa de DL revelou a existência de 13 haplótipos em 4 regiões deste cromossomo, com frequências observadas variando de 0,023 à 0,589. Em geral, o número de SNPs capturados entre as regiões estruturadas no BTA14 foi igual a 2 e a distância entre eles variou de 225 a 17.284 pb, o que reduz em 50% os esforços no processo de genotipagem para cada uma das regiões gênicas estruturadas. Os resultados deste trabalho permitiram a caracterização de estruturas de blocos de haplótipos, em regiões genômicas relacionados com características de interesse zootécnico em animais da raça Girolando e poderão ser utilizados em estudos de associação entre essas regiões com características produtivas.

**PALAVRAS-CHAVE:** blocos de haplótipos, ferramentas genômicas, marcadores moleculares rebanhos mestiços, *tagsnps*

## **DEVELOPMENT OF A LOW DENSITY SNP MARKER PANEL AND LINKAGE DISEQUILIBRIUM ANALYSIS OF POLYMORPHISMS IN CANDIDATE GENES IN GIROLANDO CATTLE**

Author: Ronyere Olegário de Araújo

Advisor: Alexandre Rodrigues Caetano, Ph.D.

### **ABSTRACT**

Development of a low density SNP marker panel and analysis of linkage disequilibrium of polymorphisms in candidate genes in Girolando cattle. Ronyere Olegário de Araújo and Alexandre Rodrigues Caetano – Research PhD, Embrapa Recursos Genéticos e Biotecnologia, Brasília/DF. Recent technological advances made it possible to analyze the genomes of several species, understand their function within biological systems, and above all allow for initial understanding of mechanisms that control interactions between genotypes and environmental effects that are involved in the expression of traits of economic interest. The objective of this study was to develop and validate a custom panel with 384 SNP markers located in candidate genes affecting production traits, genetic diseases, and that could be used for paternity control in Girolando cattle. A total of 576 animals were tested with 384 SNP panel using *GoldenGate*® technology on a BeadExpress platform. Raw genotyping data were analyzed with the Genotyping Module in the GenomeStudio platform v2010.1 using standard parameters for clustering the data from each SNP. Data quality control and filtering was performed using the following parameters: Clusters separation, Genotyping Frequency and Genotyping Rate. After application of those criteria the final dataset was composed of 249 SNPs and 419 animals. After applying QC criteria, we observed a decrease in the proportion of polymorphic SNPs (53.4%). The panel contained 72 and 47 SNPs from chromosomes BTA06 and BTA14, respectively. Linkage disequilibrium studies - LD (parameter  $D'$  and  $r^2$ ) and construction of haplotype blocks were performed. Estimates of LD values between SNPs ranged from moderate to low magnitude among themselves and between BTAs. For the parameter  $D'$ , the lowest estimate was obtained for the BTA15 (0.0954) and the highest estimated in BTA24 (0.5311). Conversely,  $r^2$  showed the lowest estimates in BTA17 (0.0011) and higher in BTA22 (0.0834). A total of 32 haplotypes in six regions of BTA06 with frequencies ranging from 0.011 to 0.970 were observed. The number of SNPs captured

(TagSNPs) in BTA6 ranged from 2 to 4 between the structured regions and the distance between them ranged from 75 to 70177 pb. For BTA14, estimated LD revealed 13 haplotypes in four regions of the chromosome with observed frequencies ranging from 0.023 to 0.589. In general, the number of SNPs captured between the structured regions in BTA14 was equal to 2 (the own *TagSNP* and the adjacent SNP) and the distance between them ranged from 225 to 17284 pb, which could result in reductions of 50% in genotyping efforts for each gene structured regions. These results allowed the characterization of haplotype block structures in genomic regions related to traits of interest in Girolando cattle and will be useful for performing association studies between the characterized regions with production traits.

**KEY-WORDS:** crossbred cattle, genomic tools, haplotype blocks, molecular markers, *tagsnps*

**CAPÍTULO 1**  
**REVISÃO DE LITERATURA**

## 1.1. INTRODUÇÃO

O setor lácteo tem grande importância para o segmento agropecuário brasileiro. Nos últimos anos, a indústria leiteira do Brasil tem apresentado crescimento contínuo e crescente. As projeções divulgadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE), estima que em 2014 a produção de leite no Brasil seja de 36,75 bilhões de litros de leite, valor este que corresponde a um aumento de 5% em comparação aos 35 bilhões de litros de leite produzidos em 2013 (IBGE, 2014). Com base nestas informações, o Brasil destaca-se entre os cinco maiores produtores mundiais de leite. Apesar dos consideráveis avanços observados no arranjo produtivo do leite nos últimos anos, aspectos associados à melhorias da qualidade do produto, bem como a produtividade dos rebanhos, ainda representam desafios a serem superados.

Tradicionalmente, cruzamentos entre animais das raças *Bos taurus* (européias) com *Bos indicus* (zebuínos) (Pereira, 2008) tem sido amplamente utilizados por diversos programas de avaliação e melhoramento genético, com destaque, por exemplo, para a formação do mestiço Girolando (Nanzer, 2010). A formação dessa raça foi fundamentada em atender a demanda de animais de alta produção de leite e adaptação ao clima tropical (McManus et al., 2008), unindo a rusticidade de animais da raça Gir (G) com a alta produção de leite de animais da raça Holandesa (H) (Facó et al., 2002). O direcionamento dos cruzamentos entre essas raças busca a fixação do padrão racial do bimestiço, no grau de 5/8H com 3/8G, objetivando um animal produtivo e padronizado. A média atual de produção por lactação (305 dias) das fêmeas de raça Girolando é de  $3.848 \pm 1.965$  kg (Silva et al., 2013), com o pico de sua produção estendendo-se até os 10 anos de idade (Facó et al., 2005).



Até a última década, grande parte dos estudos genômicos com espécies de interesse zootécnico visou a identificação genes ou blocos gênicos em regiões que controlam ou influenciam características de importância econômica, como os QTLs (*Quantitative Trait Loci*). Trabalhos iniciais com animais (Botstein et al., 1980; Maeda et al. 1989; Lonergan et al., 1995) utilizaram marcadores moleculares do tipo RFLP (*Restriction Length Polymorphism*) e posteriormente SSR (*Simple Sequence Repeats*) do tipo microsatélite para construção de mapas genéticos. A utilização de marcadores moleculares em estudos de desequilíbrio de ligação, associação e mapeamento genético, assim como em ensaios diagnósticos para detecção de doenças genéticas e/ou polimorfismos associados a características de produção, esteve limitada por muito tempo por restrições tecnológicas. Entretanto, segundo Caetano (2009), este cenário foi aperfeiçoado a partir dos avanços tecnológicos obtidos na última década que geraram novas metodologias de alto desempenho, com maior acessibilidade e baixo custo, para prospecção, caracterização e genotipagem de marcadores SNP (*Single Nucleotide Polimorphisms*). A utilização e adequação destas novas metodologias vêm revolucionando os limites do aproveitamento de dados moleculares nos estudos de prospecção de genes de interesse econômico (Caetano, 2009). Da mesma forma, abriram-se novos horizontes para utilização de dados moleculares nos processos de avaliação genética e melhoramento animal.

Uma forma direcionada de se trabalhar com este tipo de marcador na busca de resultados, com maior aplicação a curto prazo, é o estudo de polimorfismos diretamente em genes candidatos (Rothschild & Soller, 1997) que sabidamente desempenham alguma função biológica associada à característica de interesse. Vários genes candidatos em bovinos já foram identificados e seus efeitos testados via associação com características quantitativas, como os genes relacionados a proteínas do leite, resistência à patógenos e parasitas, produção e composição do leite (Khatkar et al., 2004). A associação de genes candidatos com características fenotípicas pode ser realizada com base em apenas um sítio polimórfico, dentro dos próprios genes, ou em haplótipos dos genes candidatos (Lagziel et al., 1996).

Os estudos com genes candidatos requerem conhecimentos da fisiologia da característica de interesse, que podem não estar disponíveis ou até mesmo estar incompletos na maioria das vezes. Em geral, existem diversos genes candidatos para características quantitativas, e pode ser difícil a seleção dos candidatos para realização dos estudos em questão. Além disto, genes que não são parte de importantes processos fisiológicos conhecidos podem ter efeito significativo sobre a expressão das características em questão. A

abordagem do gene candidato tende a ser eficiente somente quando a variação genética causativa de variação fenotípica foi realmente identificada (Khatkar et al., 2004). Logo, os resultados obtidos com essa metodologia devem ser interpretados com cautela, uma vez que associações significativas podem ser obtidas por desequilíbrio de ligação entre alelos do(s) marcador(es) utilizado(s) com o(s) alelos nos genes causativos adjacentes (Andersson, 2001).

Metodologias mais recentes para genotipagem de marcadores SNP apresentam altos níveis de automação e acurácia e proporcionaram reduções drásticas nos custos de geração de dados. Diferentes tecnologias já estão disponíveis para a realização de ensaios de genotipagem de SNPs de baixa densidade (de dezenas até centenas de marcadores), média densidade (de milhares a dezenas de milhares) e alta densidade (de centenas de milhares a milhões). Os custos de geração de dados variam de centavos (no caso de ensaios de baixa densidade) a uma fração de centavo (no caso de ensaios de alta densidade) por genótipo (Caetano, 2009). Chips validados pela comunidade científica internacional estão disponíveis para genotipagem de bovinos, ovinos, suínos, equinos, frangos, etc., e podem ser desenvolvidos para qualquer espécie de interesse (Illumina Inc., San Diego, Califórnia).

As tecnologias de genotipagem disponíveis atualmente permitem reduções expressivas, em relação às tecnologias anteriores, nas necessidades de infraestrutura para a geração de dados moleculares com marcadores SNP. Isso se deve, principalmente, ao alto nível de padronização das novas tecnologias, em que a geração dos dados genômicos podem ser adquiridos em laboratórios multiusuários ou através de empresas terceirizadas, especializados na geração desses dados.

O aumento sustentável da produtividade e qualidade dos produtos gerados são de vital importância para consolidar o país dentre um dos principais fornecedores de alimentos de origem láctea no mundo. Além disso, a crescente busca por alternativas que promovam a sustentabilidade do sistema de produção, definidas principalmente pelas questões ambientais, econômicas e sociais, tornam-se alvos das atuais pesquisas. Assim, estudos com a finalidade de melhor compreender os processos biológicos envolvidos na produção de leite são fundamentais para o desenvolvimento de soluções tecnológicas inovadoras (Mello, 2010). Identificar genes responsáveis pela produção e caracterizar sua variabilidade, é uma abordagem efetiva para o incremento genético da raça, e que deverá prover subsídios para aumentar a produtividade e a competitividade da pecuária leiteira nacional (Mello et al., 2012).

### **1.1.1. Justificativa**

Ganhos em produtividade e aumentos na escala produtiva são apontados como os mecanismos primordiais para garantia de eficiência e sustentabilidade dos sistemas de produção. Em busca da melhoria da produtividade do rebanho, os cruzamentos entre raças européias e zebuínas assumem grande importância, sendo amplamente utilizados para melhorar a produção de leite e a eficiência reprodutiva e adaptativa em ambientes tropicais e subtropicais.

A crescente evolução de tecnologias tornaram possível analisar o genoma de várias espécies, localizar e identificar especificidades na sequência do genoma, compreender suas funções dentro dos sistemas biológicos e, sobretudo, começar a entender os mecanismos que controlam as interações entre os genótipos e os efeitos ambientais (Furlan et al., 2007). Segundo estes mesmos autores, estudos sobre as bases moleculares da variabilidade fenotípica abriram uma nova abordagem científica, caracterizada pela multiplicidade das questões envolvidas, que resultou no surgimento de novas áreas de pesquisa, cujos conhecimentos estão sendo aplicados em diversos campos da biologia, inclusive na zootecnia. Assim, considerando o impacto que tais conhecimentos têm sobre a compreensão dos fenômenos biológicos, é oportuno conduzir uma avaliação das potencialidades de aplicação das abordagens genômicas, tais como estudos com genes candidatos em pesquisas relacionadas a produção leiteira.

## **1.2. OBJETIVOS**

### **1.2.1. Objetivo geral**

O objetivo deste trabalho foi desenvolver e validar um painel customizado a partir de um ensaio de 384 marcadores SNP localizados em genes candidatos que afetam características de produção, genes causadores de doenças genéticas, e para controle de parentesco em bovinos de aptidão leiteira.

### 1.2.2. Objetivos específicos

- a) Validar um painel com 384 SNPs derivados de genes candidatos e causadores de doenças genéticas na plataforma Illumina *BeadXpress*;
- b) Estimar frequências alélicas na raça Girolando;
- c) Caracterizar os blocos de haplótipos e o desequilíbrio de ligação dos SNPs localizados em um mesmo gene na raça Girolando;
- d) Estimar *TagSNPs* a partir dos polimorfismos encontrados nos blocos de haplótipos;
- e) Estimar o poder de identificação de Grupo Genético dos marcadores contidos no ensaio, na raça Girolando.
- f) Estimar o poder de exclusão de paternidade dos marcadores contidos no ensaio, na raça Girolando.

### **1.3. REVISÃO DE LITERATURA**

#### **1.3.1. Características Lácteas**

Nos últimos anos houve um aumento do interesse na qualidade nutricional diferenciada do leite e de seus derivados, principalmente devido à competitividade das indústrias de laticínios no mercado internacional. Essa qualidade está relacionada principalmente aos seus constituintes, tais como a concentração de proteína e de gordura (Ribeiro et al., 2000; Zanela et al., 2006; Rosa et al., 2012).

Os genes codificadores das caseínas, principais proteínas do leite, albuminas, lactoglobulinas e também proteínas séricas, são considerados genes candidatos para produção e qualidade do leite, seus componentes e, conseqüentemente, também para características de habilidade materna da vaca, como peso à desmama (Rosa & Fragoso, 2011). Além das proteínas e dos lipídeos, outra característica importante para estudos de associação genética é a mastite, que representa uma das enfermidades mais frequentes em bovinos de aptidão leiteira e uma das principais causadoras de quedas tanto na qualidade quanto na produção de leite total, causando significativas perdas monetárias para os produtores e laticínios no mundo inteiro (Sodeland et al., 2011).

Com o aumento da acessibilidade de tecnologias moleculares, os marcadores moleculares vem sendo utilizados como ferramenta auxiliar nos processos de seleção. Logo, a identificação de marcadores associados a regiões no genoma que contribuam para a variação das características quantitativas (QTL) como produção de leite e de gordura, permitem a

seleção de indivíduos com valores genéticos superiores, reduzindo assim, o custo e o tempo requerido para a sua reprodução e por consequência, aumento da velocidade de resposta dos programas de melhoramento. Portanto, os resultados de estudos de associações genéticas poderiam ser utilizados para pré-selecionar touros jovens, candidatos ao teste de progênie, aumentando assim os diferenciais de seleção, encurtando o intervalo de geração, o que aumentaria os ganhos genéticos obtidos.

No Brasil, grande parte do leite produzido provém de vacas leiteiras mestiças que apresentam alta fração de genes das raças zebuínas. Características quantitativas complexas, de mensuração tardia, baixa herdabilidade e limitadas ao sexo têm sido mais difíceis de serem melhoradas com base exclusivamente no emprego do melhoramento genético convencional. Nesses casos, os ganhos genéticos poderão ser acelerados mediante o uso dos marcadores moleculares associados a locos de características quantitativas (QTL).

### **1.3.2. Sistemas de Cruzamento**

A heterose, ou vigor híbrido, é um fenômeno no qual os indivíduos resultantes de cruzamentos apresentam desempenho superior ao desempenho médio dos pais de raças puras (Pereira, 2008). Segundo este autor, uma das finalidades da sua utilização é reunir em um só tipo biológico as características desejáveis de duas ou mais raças, como a rusticidade das raças zebuínas e o potencial de produção das raças taurinas.

Em busca de melhorias na produtividade, os cruzamentos entre raças taurinas e zebuínas assumem grande importância, sendo amplamente utilizados para melhorar a produção de leite e a eficiência reprodutiva e adaptativa em ambientes tropicais e subtropicais (Junior et al., 2014). O principal objetivo deste tipo de cruzamento é utilizar da complementariedade entre raças divergentes para a obtenção de animais mais adaptados e produtivos sob tais condições, com a maximização dos efeitos de heteroze.

O cruzamento alternado simples consiste na utilização de duas raças puras alterando a raça do touro a cada geração. Neste caso, animais com duas composições genéticas dentro do rebanho são obtidos, permitindo realizar a reposição de fêmeas com animais do próprio plantel. Esse sistema é indicado para produtores com propriedades geridas com baixo nível tecnológico (Lopes et al., 2005). A raça sintética Girolando é formada a

partir desse tipo de cruzamento, até a terceira geração, na qual se realiza o cruzamento com outro animal de mesma proporção sanguínea,  $\frac{5}{8}$  Holandês +  $\frac{3}{8}$  Gir originando o pool sintético (Silva et al., 2010). Silva et al. (2010), estudando mestiços dessa raça, verificaram um aumento na média de produção de leite (PL) de 1.288 kg/lactação no período entre 2000 e 2009, saindo de 3.648 kg/lactação para 4.936 kg/lactação, padronizados para uma duração de lactação (DL) de 298 dias e idade ao primeiro parto (IPP) de 35,2 meses, representando 35,3% de ganho. Facó et al. (2009) ressaltaram que esses índices produtivos da raça Girolando são fundamentados principalmente em melhorias ambientais, aumentando o desempenho produtivo e reprodutivo desses animais. Da mesma forma, Silva & Veloso (2011) destacaram que os animais Girolando são bem adaptados às condições brasileiras, com boa resposta, produtiva, quando criados em pastagem, principalmente se ocorrer um bom manejo nutricional.

### **1.3.3. Marcadores de polimorfismo de base única (SNP)**

A sigla SNP origina-se do inglês, *Single Nucleotide Polymorphism*, polimorfismo de base única, e é utilizada para representar a forma mais comum de variação genética entre os indivíduos de determinada população. Marcadores do tipo SNPs são, na maioria das vezes, bialélicos altamente estáveis, cuja estimação de frequência pode ser realizada com relativa facilidade (Picoult-Newberg et al., 1999).

Há que se considerar, inicialmente, a fim de embasamento para este capítulo, que a base dos marcadores SNP configura-se por meio de mutações da molécula de DNA. Tais mutações são elementares e remetem à cadeia de bases nitrogenadas, ou seja, Adenina (A), Citosina (C), Timina (T) e Guanina (G) (Caetano, 2009). Os primeiros estudos deste cunho datam da década de oitenta, apesar das limitações técnicas encontradas naqueles anos, quando a produção de um dado molecular demandava, em média, cinco dias trabalhados (Caetano, 2009). Com o avançar dos tempos e a evolução tecnológica e científica, estes estudos foram ampliados e atualmente já foram descritos no genoma bovino, cuja composição remete ao número aproximado de alguns milhões de SNPs em 3 bilhões de pares de bases.

Baudet et al. (2006) relataram em seus estudos que os SNPs podem ser polimorfismos bi, tri ou tetra alélicos, possuindo, portanto, duas, três ou quatro formas distintas, sendo mais raros aqueles que apresentam três ou quatro formas. Sua frequência de

variações remete à substituições de bases nitrogenadas cuja característica estrutural é semelhante. As substituições mais comuns ocorrem entre uma purina (A ou C) por outra purina, ou de uma pirimidina (G ou T) por outra pirimidina são denominadas de transições. Em caso mais raros, ocorrem substituições de purinas por pirimidinas, ou vice-versa, as quais essas mutações são denominadas de transversões.

O sequenciamento do genoma bovino permitiu a identificação de Polimorfismos de Base Única (SNP) com relativa facilidade e quantidade. A grande vantagem desse tipo de marcador molecular é a possibilidade de automatização da genotipagem de milhares de locos em paralelo (Lee et al., 2006). Dentre as várias consequências diretas do desenvolvimento desta tecnologia destaca-se a geração de mapas genéticos de alta densidade saturados com marcadores moleculares, e também do sequenciamento de alguns genomas. Hoje o Consórcio Internacional do Genoma Bovino apresenta um Banco de dados de SNP (cerca de 770.000 marcadores) genotipados para várias raças de bovinos, inclusive para as raças brasileiras, como o Nelore e o Gir (VanRaden, 2014).

Convém destacar que, a evolução científica possibilitou estudos mais detalhados e mais adequados à especificidade sobre este tema. Contudo, apesar de se considerar que os maiores avanços demandaram dos últimos anos, a constatação de SNPs no genoma não é recente. Os primeiros estudos de sequenciamento de DNA já evidenciaram esta forma de polimorfismo, porém não oferecerem maiores detalhes e/ou minúcias relacionadas à temática (Orita et al., 1989).

Segundo Carneiro (2010), a partir dos testes com marcadores é possível determinar os alelos presentes em genes que controlam características de interesse para a seleção. Para o autor, o desenvolvimento dessas técnicas representou avanços significativos não somente para a ciência, mas também para o produtor, oferecendo a este benefícios de suma relevância. Observa-se também ampla utilização de marcadores moleculares para a identificação de indivíduos portadores de alelos causadores de doenças hereditárias, permitindo, conseqüentemente controle e, até mesmo, erradicação de algumas doenças danosas à população leiteira.

Distribuídos ao longo do genoma, estes marcadores podem ser utilizados para rastrear a variabilidade nos genes responsáveis por controlar características de interesse da pecuária, as quais normalmente possuem arquitetura poligênica e natureza quantitativa. Estudos realizados com animais da raça Holandesa mostraram que são necessários painéis



com número mínimo de 50 mil SNPs para explicar 46,18% da variabilidade genética para certas características produtivas (Wang et al., 2012). No caso de raças bovinas com aptidão para corte, devido ao histórico de seleção, o número de marcadores necessários é ainda mais elevado (2º Workshop Internacional: Genômica Aplicada à Pecuária, 2012).

Os métodos contemporâneos de genotipagem de marcadores SNP têm sido amplamente utilizados em decorrência dos baixos custos associados a geração de dados em larga escala (Guimarães et al., 2010). Em seus estudos sobre Genoma Bovino, Brito (2011) relata a importância de mapear os SNPs e, mediante este mapeamento, explorá-los em conjunto com plataformas de genotipagem, possibilitando a inclusão da informação genômica em esquemas de melhoramento genético. Estes testes, segundo a autora, possibilitam a identificação de características de medição, bem como herdabilidade baixa ou cuja expressão seja sexo-específica ou tardia. Do mesmo modo, Brito (2011) destacou também a obrigatoriedade do processo de validação para uso em uma população diferente daquela onde foi descoberto o marcador. A partir dos avanços em pesquisas e análise de SNPs, alguns estudiosos optaram por escolher de maneira criteriosa estes marcadores espalhados e espaçados de forma igualitária no decorrer do genoma para, de posse destes dados, capturar o efeito dos genes responsáveis pelas características de interesse. Através dessa forma de exploração, buscou-se associação frequente de características populacionais e não restritas à famílias específicas. Denominou-se esta forma mais ampla de abordagem para Seleção Genômica.

Estudos com humanos, e com espécies de interesse zootécnico, mostram que pode haver milhões de polimorfismos do tipo SNP no genoma de um indivíduo (Elsik et al., 2009; Li et al., 2009). Além dos marcadores SNP serem abundantes, suas bases moleculares permitem que haja uma distribuição homogênea de SNPs pelo genoma.

As tecnologias de genotipagem de SNP foram aplicadas primeiramente para desenvolvimento de ensaios diagnósticos para doenças genéticas (Holloway et al., 1999). Nesses casos, a demanda da aplicação é testar um número baixo de polimorfismos em um grande número de amostras, com rapidez e alta acurácia. As soluções tecnológicas mais adequadas para essas aplicações são os ensaios *Taqman* ou de minisequenciamento. Questões como alta repetibilidade, robustez e facilidade de padronização dos ensaios trouxeram grandes vantagens para a implementação de diferentes testes com diferentes plataformas e em diferentes laboratórios (Caetano, 2009).

Já existem exemplos reais da aplicação dessas tecnologias para prospecção de SNPs. Van Tassell et al. (2008), em estudo envolvendo 66 animais que representavam três rebanhos bovinos, prospectaram um total de 62.042 SNPs, a um custo de US\$0.48 por SNP, utilizando o *Illumina/Solexa Genome Analyzer*. Embora muito mais eficiente que o sequenciamento tradicional, a estratégia utilizada ainda depende da disponibilidade de uma sequência referência do genoma.

Metodologias para testes de exclusão de paternidade e identificação individual também foram desenvolvidas com marcadores SNP (Heaton et al., 2002). Adicionalmente, painéis de baixa densidade foram gerados para rastreabilidade de amostras (Heaton et al., 2005) e também para diagnosticar polimorfismos em genes associados a características de produção (Chessa et al., 2007, Merial Igenity®). Essas aplicações requerem a genotipagem de algumas dezenas de marcadores em grandes números de amostras, demandando quesitos não oferecidos pelos ensaios *Taqman* e de minisequenciamento. As tecnologias geralmente utilizadas nessas aplicações foram as plataformas baseadas em espectrometria de massa (MALD-TOF, Heaton et al., 2005) e em microarranjos (Kaminski et al., 2006; Chessa et al., 2007). Uma das grandes vantagens dos marcadores SNP é a fácil conversão dos ensaios entre diferentes plataformas.

#### **1.3.4. Tecnologia de genotipagem GoldenGate**

As tecnologias para genotipagem de SNPs têm avançado rapidamente nos últimos anos. Esta é uma situação que pode ser explicada pelo aumento no interesse da aplicação destes marcadores genéticos tanto para o mapeamento de doenças complexas quanto para estudos de associação com características de produção. Como resultado desta conjugação de interesses, tem sido descrita uma grande variedade de protocolos para a sua caracterização. Rosa et al. (2007), ressaltaram que dois fatores foram fundamentais para essa evolução: 1) a disponibilidade de sistemas robotizados que permitiram a confecção de arranjos de alta densidade; e 2) a adoção do método de detecção óptica (fluorescência), que conferiu a sensibilidade ideal para se fazer as medidas de intensidade, como a tecnologia de genotipagem *GoldenGate* (Illumina – *BeadXPRESS*). Entretanto, o ponto em comum entre quaisquer plataformas é que os ensaios para SNPs específicos podem ser adaptados para todas

as tecnologias, permitindo que a mesma informação genotípica possa ser acessada, independentemente da tecnologia utilizada.

O ensaio *GoldenGate* permite um grau elevado de multiplexação de loci (96 à 1536 SNPs) durante as etapas de extensão e de amplificação, minimizando o tempo, os volumes de reagentes, e as necessidades de material, uma vez que todo processo ocorre em placa, permitindo que se faça simultaneamente genotipagem em 96 amostras. Os arranjos de DNA, também conhecidos como ensaios ou painéis de DNA, são segmentos de DNA (oligonucleotídeos pré-sintetizados) que estão distribuídos, ordenadamente, sobre uma superfície sólida, denominada *microbead* ou bastonete de sílica, com dimensões de 240µm de tamanho e 28µm de diâmetro. Segundo Chaudhuri (2005), considera-se como pré-requisito fundamental para qualquer tipo de arranjo de DNA a existência de um “endereço próprio” para cada SNP do painel. Cada um desses endereços nos bastonetes são chamados de *spot* (ou ponto), e contém uma pequena quantidade de DNA fixada, denominada sonda (ou *probe*). Cada poço da placa apresenta um *pool* de bastonetes, sendo que cada um contém um tipo específico de marcador inserido no painel e um código específico que identifica o bastonete (*VeraCode*).

Cada uma destas sondas tende a se ligar apenas a sua sequência complementar de nucleotídeos, mediante processo chamado hibridização (Jaluria et al., 2007). Essa sequência complementar, normalmente um DNA complementar (cDNA) produzido à partir de um RNA mensageiro (mRNA), representa apenas um único gene do genoma e é chamada “alvo” ou *target*. A hibridização de cada sonda de DNA com o seu correspondente alvo (cDNA) é um processo baseado na complementaridade das cadeias de nucleotídeos, ou seja, na propriedade que duas cadeias homólogas têm de parear suas bases complementares (A com T e C com G), mediante a formação de pontes de hidrogênio.

Após o processo de hibridização, os bastonetes são lavados para remoção dos “alvos” excedentes (que não se ligaram às sondas) e, em seguida, exposta à ação de raios laser (leitor *BeadArray* da Illumina) que excitam os corantes que foram incorporados aos “alvos”, fazendo com que estes emitam luz (fluorescência). O leitor de *BeadXpress* utiliza um sistema de detecção a laser de dupla cor para identificar o código holográfico incorporado em cada bastonete *VeraCode* e para detectar a intensidade de sinal associado em cada *bead*.

### 1.3.5. Blocos de haplótipos e desequilíbrio de ligação

Um haplótipo é uma combinação de alelos em loci adjacentes, que fazem parte do mesmo cromossomo e são transmitidos juntos. Um haplótipo pode ser formado por um ou vários alelos, ou até pelo cromossomo inteiro. O desequilíbrio de ligação (DL) representa uma medida de correlação entre alelos em locos diferentes, representando a falta de independência, de forma estatística, entre os alelos. O DL existe entre dois *loci* ligados quando alelos nestes ocorrem no mesmo haplótipo de maneira mais frequente do que o esperado na população. Estas associações alélicas são devidas principalmente à proximidade física, mas também são influenciados pela história da população e as forças evolutivas, como por exemplo a recombinação entre os polimorfismos (Remington, et al., 2001; Khatkar et al., 2008). Com isso, os principais fatores que aumentam a extensão do DL são os endocruzamentos, tamanho efetivo da população reduzido, isolamento genético entre populações, raças ou linhagens, subdivisão entre populações, baixa taxa de recombinação, e seleção natural e/ou artificial (Farnir et al., 2000). Por conseguinte, as fontes de variação que diminuem a extensão do DL são os cruzamentos, altas taxas de recombinação e mutação.

O termo Desequilíbrio de Ligação foi utilizado pela primeira vez no trabalho publicado por Lewontin & Kojima (1960). Segundo Hedrick (2010), a designação feita pelos últimos autores foi devido ao fato de que a ligação entre locos era fator determinante para uma população atingir o equilíbrio. Entretanto, uma vez que suas definições matemáticas ainda não estavam claras (Slatkin, 2008), as discussões entre os cientistas não se preocuparam com o quanto essa designação era apropriada.

Outro ponto que merece atenção diz respeito ao DL e ligação física serem termos comumente confundidos, pois, muitas vezes, são apresentados como sinônimos. Essa confusão decorre normalmente porque locos intimamente ligados podem estar em elevado DL, como ressaltados por Flint-Garcia et al. (2003). Contudo, o fato de dois locos quaisquer estarem em DL não significa que estão fisicamente ligados, da mesma forma que dois locos estarem fisicamente ligados em determinado cromossomo ou haplótipo não significa que estão em DL (Slatkin, 2008; Hedrick, 2010). Devido a essa confusão, muitos autores sugerem que se utilize termos como Desequilíbrio da Fase Gamética ou Desequilíbrio Gamético para substituir o DL (Hartl & Clark, 2007; Hamilton, 2009; Hedrick, 2010).

Dekkers (2004) sugere que, devido à extensão do DL observado no genoma inteiro em populações de animais, é possível que marcadores informativos a cada 1 ou 2 cM (centimorgans) possam ser suficiente para detectar a maioria dos QTLs. Na detecção de bloco de haplótipo, Gabriel et al. (2002) demonstraram que a densidade de um marcador a cada 7,8 Kb foi suficiente para estruturar os padrões de 51 blocos de haplótipos em todo o genoma humano.

Estudos do DL no genoma inteiro continuam a surgir com o aumento no número de polimorfismos disponíveis e a diminuição dos custos associados com os genotipagem. McKay et al. (2007) avaliaram os níveis de DL nas raças holandesa e Angus usando aproximadamente 3.000 SNPs no genoma. A medida do DL utilizado foi o  $r^2$  que tem sido ser o mais robusto entre as estatísticas utilizadas nestes estudos (McRae et al. 2002), uma vez que este parâmetro se ajusta melhor em estimativas de DL em populações pequenas e por ser pouco influenciado pelas frequências do alelo menor (Devlin & Risch, 1995; Du et al., 2007). Os resultados mostraram que um total de 30.000 marcadores seria ideal para captar a variação genética distribuída pelo genoma bovino vislumbrando análises futuras de associação. Entretanto, análises a partir de SNPs provenientes de genes candidatos vem se mostrando como metodologia viável para identificação de DL em regiões específicas do genoma (Simko et al., 2006).

### **1.3.6. Medidas de desequilíbrio de ligação**

Considerando que marcadores SNP são normalmente bi-alélicos, se considerarmos dois locos polimórficos (A/T e C/G), em uma geração  $N$  qualquer de uma determinada população, quatro diferentes gametas, ou haplótipos, podem ser identificados (AC, AG, TC, TG). Se esses locos estiverem em equilíbrio de ligação (DL = 0 ou EL = 1), ou seja, numa situação de associação aleatória entre os seus alelos, as frequências observadas dos haplótipos serão exatamente equivalentes às suas frequências esperadas, as quais correspondem ao produto das respectivas frequências alélicas. Por outro lado, se esses locos não estiverem em equilíbrio de ligação (DL  $\neq$  0), de modo que seus alelos estejam combinados de maneira não-aleatória, ou preferencial, as frequências dos haplótipos apresentarão um desvio (D) em relação às frequências esperadas (Lewontin & Kojima, 1960).

Neste sentido, os principais parâmetros utilizados para estimar o nível de DL em uma população são os desvios esperados ( $D'$ ), proposto por Lewontin (1964), e o coeficiente de determinação ( $r^2$ ), sugerido por Hill & Robertson (1968).

A estatística  $D'$  pode ser calculada da seguinte forma:

$$|D'| = \frac{|D|}{D^{\max}}$$

em que:  $D'$  é a medida do DL,  $D = fAB - fAfB$ ,  $D^{\max}$  é o máximo valor possível de  $D$ . Assim, se  $D_{ab} < 0$ , tem-se:

$$|D'| = \frac{(D_{ab})^2}{\min(fAfb, fafB)}$$

Contudo, quando  $D_{ab} > 0$ , tem-se:

$$|D'| = \frac{(D_{ab})^2}{\min(fAfb, fafB)}$$

em que:  $f$  é a frequência de cada alelo (A ou B);  $\min$  = é a menor frequência possível.

O desvio esperado ( $D'$ ) é uma medida relativa e pode variar entre 0 e 1 ( $0 \leq D' \leq 1$ ). Com  $D' = 1$ , pode-se dizer que pelo menos um dos quatro haplótipos está ausente, o que caracteriza uma situação de DL (Slatkin, 2008; Hedrick, 2010).

Outra medida de DL é o coeficiente de determinação ( $r^2$ ) que foi proposta por Hill & Robertson (1968). Este coeficiente representa a correlação ao quadrado entre a presença e ausência de alelos em diferentes loci. Esta medida pode ser descrita como:

$$r^2 = \frac{(D_{ab})^2}{fAfb, fafB}$$

em que:  $D = fAB - fAfB$ ;  $fA$ ,  $fa$ ,  $fB$  e  $fb$  são as frequências alélicas e  $fAB$ ,  $fAb$ ,  $faB$  e  $fab$  são as frequências dos haplótipos.

Da mesma forma que o  $D'$ ,  $r^2$  pode variar entre 0 e 1 ( $0 \leq r^2 \leq 1$ ). Por conseguinte, quando  $r^2$  tende a 1, tem-se indicativo que pelo menos um haplótipo está ausente na amostra, o que caracteriza uma situação de DL (Slatkin, 2008; Hedrick, 2010).

As estatísticas  $r^2$  e  $D'$  refletem diferentes aspectos de desequilíbrio de ligação e apresentam comportamentos diversos em condições distintas. Na estatística  $r^2$ , somente se

encontra um valor igual a 1, ou seja, desequilíbrio de ligação completo, quando os dois loci em questão apresentam frequências alélicas idênticas. Por outro lado, na estatística  $D'$ , sempre que a combinação de um par de alelos não estiver presente na população em estudo,  $D'$  será igual a 1 (Flint-Garcia et al., 2003). Hedrick (1987), em um estudo baseado com simulações, mostrou que a correlação entre as duas medidas ( $r^2$  e  $D'$ ) é muitas vezes baixa (valores  $<0,6$ ). Esta diferença nos valores encontrados no desequilíbrio de ligação quando se adota  $r^2$  ou  $D'$  é resultado da história evolutiva que cada uma destas estatísticas incorpora para mensurar o desequilíbrio de ligação, sendo que cada método tem suas vantagens e desvantagens (Rosa, 2011).

A estatística  $r^2$  resume as histórias de recombinação e de mutação que estes loci sofreram, enquanto que  $D'$  mede apenas a história de recombinação e, sendo assim, fornece uma medida mais exata para estimar diferenças de eventos de recombinação. Entretanto,  $D'$  é fortemente afetada por amostras de tamanho pequeno, resultando em um comportamento enviesado quando se compara loci com baixas frequências alélicas (Du et al., 2007). Por outro lado, quando se tem o objetivo de examinar a resolução de estudos associativos, geralmente é conveniente a estatística  $r^2$ , uma vez que ela é um indicativo de como os marcadores podem se correlacionar com o QTL de interesse (Flint-Garcia et al., 2003; Du et al., 2007).

Gabriel et al. (2002) definiram que um alto DL (pouca ou nenhuma evidência de recombinação no genoma) é dado quando os valores de  $D'$  para pares de alelos estejam com o valor de 95% do intervalo de confiança, ou seja, com limite superior acima de 0,98 e o limite inferior acima de 0,7. Por outro lado, nomearam como “forte evidência de recombinação histórica” pares de loci nos quais o valor de  $D'$  esteja com o limite superior abaixo de 0,9. Paralelamente, a regra dos quatro gametas (Wang et al., 2002) é uma variação que utiliza as frequências populacionais dos quatro possíveis haplótipos de cada par de marcador computado. Se todos os quatro são observados com frequência mínima de 0,01, a recombinação é considerada como ocorrida. Blocos são formados por marcadores consecutivos em que apenas três gametas são observados. Outra maneira de definir blocos de haplótipos é pela coluna sólida de desequilíbrio de ligação (*Solid Spine of linkage disequilibrium* - Barrett et al., 2005). Este método utiliza um algoritmo que procura por uma coluna de forte DL de um marcador a outro ao longo do segmento, o que representa que o primeiro e o último marcador estão em forte DL com todos os marcadores intermediários, mas estes não necessariamente se encontram em DL entre si (Barrett et al., 2005). Todos estes

métodos são utilizados estatisticamente para determinar um valor para a associação entre os alelos e, a partir desse valor, definir um conjunto que forme o bloco haplotípico.

Outra vantagem da utilização de blocos de haplótipos para a caracterização genética é que regiões de alto DL exibem diversidade limitada (Altshuler et al., 2005). Assim, um pequeno número distinto de haplótipos é predito em cada bloco para a maior parte dos cromossomos nas populações em comparação ao número predito pela permutação de combinações entre eles, caso os alelos dos diferentes locos estudados se associem ao acaso, ou seja,  $2^n$ , em que  $n$  é o número de marcadores (Cardon & Abecasis, 2003). Portanto, o uso de blocos haplotípicos em estudo de associação por DL tem sido aplicado extensamente e revelado grande importância na associação de genes candidatos em doenças complexas. O avanço de tecnologias mais simples, rápidas e acuradas permite que projetos sejam feitos tanto em regiões gênicas de pequena extensão em estudos de associação indireta (Carlton et al., 2005; Fang et al., 2005), quanto em estudos de associação em genoma completo (Stranger et al., 2005; Sladek et al., 2007).



#### 1.4. REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSHULER, D.; BROOKS, L.D.; CHAKRAVARTI, A. et al. A haplotype map of the human genome. **Nature**, v. 437, n. 7063, p. 1299-320, 2005.
- ANDERSSON, L. Genetic dissection of phenotypic diversity in farm animals. **Nature Reviews Genetics**, v. 2, n. 2, p. 130–138, 2001.
- BARRETT, J.C.; FRY, B.; MALLER, J. et al. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v. 21, n. 2, p. 263-265, 2005.
- BAUDET, C.; GALVEZ, M.; ZANONI, D. **Comparação de métodos para determinação de SNPs com medidas de confiabilidade**. SP, Universidade Estadual de Campinas: Instituto de Computação, 2006. Disponível em: <<http://www.ic.unicamp.br/~reltech/2006/06-15.pdf>> Acesso em 06/10/2013.
- BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M. et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, v. 32, p. 314–331, 1980.
- BRITO, Fernanda Vanieri. **Diversidade Genética e Acurácia da Informação Genômica em Bovinos de Corte**. Rio Grande do Sul. Faculdade de Agronomia da Universidade Federal do Rio Grande do Sul, 2011. 144p. Tese (Doutorado em Zootecnia) - Faculdade de Agronomia da Universidade Federal do Rio Grande do Sul, 2011.
- CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v. 38, n.spe, p. 64-71, 2009.
- CARDON, L.R.; ABECASIS, G. R. Using haplotype blocks to map human complex trait loci. **Trends Genetics**, v. 19, n. 3, p. 135-40, 2003.
- CARLTON, V.E.; HU, X.; CHOKKALINGAM, A.P. et al. PTPN22 genetic variation: evidence for multiple variants associated with rheumatoid arthritis. **The American Journal of Human Genetics**, v. 77, n. 4, p.567-581, 2005.
- CARNEIRO, R.L.R. **Marcadores Genéticos, Seleção Genômica e outras tecnologias**. (2010) Portal Dia do Campo Disponível em: <<http://www.diadecampo.com.br/zpublisher/materias/Materia.asp?id=22108&secao=Manejo>> Acesso em 06/10/2013.

- CHAUDHURI, J.D. Genes arrayed out for you: the amazing world of microarrays. **Medical Science Monitor**, v. 11, n. 2, p. 52-62, 2005.
- CHESSA, S.; CHIATTI, F.; CERIOTTI, G. et al. Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms. **Journal of Dairy Science**, v. 90, n. 1, p. 451-464, 2007.
- DEKKERS, J.C.M. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. **Journal of Animal Science**, v. 82, n. 6, p. 312-328, 2004.
- DEVLIN, B.; N. RISCH, A comparison of linkage disequilibrium measures for fine-scale mapping. **Genomics**, v. 29, n. 2, p. 311-322, 1995.
- DU, F.X.; CLUTTER, A.C.; LOHUIS, M.M. Characterizing linkage disequilibrium in pig populations. **International Journal of Biological Sciences**, v. 3, n. 3, p. 166-178, 2007.
- ELSIK, C.G.; TELLAM, R.; WORLEY, K.C. et al. The Bovine Genome Sequencing and Analysis Consortium. The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution. **Science**, v. 324, n. 5926, p. 522-528, 2009.
- FACÓ, O.; LÔBO, R.N.B.; MARTINS FILHO, R. et al. Análise do Desempenho Produtivo de Diversos Grupos Genéticos Holandês x Gir no Brasil. **Revista Brasileira de Zootecnia**, v. 31, n. 5, p. 1944-1952, 2002.
- FACÓ, O.; LÔBO, R.N.B.; MARTINS FILHO, R. et al. Idade ao Primeiro Parto e Intervalo de Partos de Cinco Grupos Genéticos Holandês x Gir no Brasil. **Revista Brasileira de Zootecnia**, v.34, n. 6, p.1920-1926, 2005.
- FACÓ, O.; MARTINS FILHOS, R.; LÔBO, R.N.B. et al. Efeito da Redução da Variação da Duração de Lactação na Avaliação Genética de Bovinos Leiteiros Mestiços. **Revista Ciência Agronômica**, v. 40, n. 2, p. 287-292, 2009.
- FANG, Y.; VAN MEURS, J.B.; D'ALESIO, A. et al. Promoter and 3'-untranslated-region haplotypes in the vitamin d receptor gene predispose to osteoporotic fracture: the rotterdam study. **The American Journal of Human Genetics**, v. 77, n. 5, p. 807-823, 2005.
- FARNIR, F.; COPPIETERS, W.; ARRANZ, J.J. et al. Extensive genome-wide linkage disequilibrium in cattle. **Genome Research**, v. 10, n. 6, p. 220-227, 2000.
- FLINT-GARCIA, S.A.; THORNSBERRY, J.M.; BUCKLER IV, E.S. Structure of linkage disequilibrium in plants. **Annuals Reviews in Plant Biology**, v. 54, n. 6, p. 357-374, 2003.
- FURLAN, L.R.; FERRAZ, A.L.J.; BORTOLOSSI, J.C. A genômica funcional no âmbito da produção animal: estado da arte e perspectivas. **Revista Brasileira de Zootecnia**, v. 36, p. 331-341, 2007 (suplemento especial).
- GABRIEL, S.B.; SCHAFFNER, S.F.; NGUYEN, H. et al. The structure of haplotype blocks in the human genome. **Science**, v. 296, n. 5576, p. 2225-2229, 2002.
- GUIMARÃES, M.F.M.; MELLO, F.; FONSECA, I. et al. **Marcadores Genéticos no Melhoramento Animal**. (2010) Disponível em: <http://www.girolando.com.br/index.php?paginasSite/tecnico,36>. Acesso em 06/10/2013.
- HAMILTON, M.B. **Population genetics**. Oxford: Wiley-Blackwell, 2009. 407p.

- HARTL, D.L.; CLARK, A.G. **Principles of population genetics**. 4.ed. Sunderland: Sinauer Associates, 2007. 565 p.
- HEATON, M.P.; HARHAY, G.P.; BENNETT, G.L. et al. Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. **Mammalian Genome**, v. 13, n. 5, p. 272-81, 2002.
- HEATON, M.P.; KEEN, J.E.; CLAWSON, M.L. et al. Use of bovine single nucleotide polymorphism markers to verify sample tracking in beef processing. **Journal of the American Veterinary Medical Association**, v. 226, n. 8, p. 1311-1314. 2005.
- HEDRICK, P.W. Gametic disequilibrium measures: Proceed with caution. **Genetics**, v.117, n. 2, p.331-341, 1987.
- HEDRICK, P.W. **Genetics of populations**. 4.ed. Sudbury, MA: Jones and Bartlett Publishers, 2010. 675p.
- HILL, W.G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **Theoretical and Applied Genetics**, v. 38, n. 6, p. 226-231, 1968.
- HOLLOWAY, J.W.; BEGHÉ, B.; TURNER, S. et al. Comparison of three methods for single nucleotide polymorphism typing for DNA bank studies: sequence-specific oligonucleotide probe hybridisation, TaqMan liquid phase hybridisation, and microplate array diagonal gel electrophoresis (MADGE). **Human Mutations**, v. 14, n. 4, p. 340-7, 1999.
- IBGE, Instituto Brasileiro de Geografia e Estatística. Produção Pecuária Nacional, Rio de Janeiro, v.36, 2008. Disponível em: <[http://ibge.gov.br/home/estatistica/indicadores/agropecuaria/producaoagropecuaria/abat-e-leite-couro-ovos\\_201401\\_publ\\_completa.pdf](http://ibge.gov.br/home/estatistica/indicadores/agropecuaria/producaoagropecuaria/abat-e-leite-couro-ovos_201401_publ_completa.pdf)> Acesso em: 29/11/2014.
- JALURIA, P.; KONSTANTOPOULOS, K.; BETENBAUGH, M. et al. A perspective on microarrays: current applications, pitfalls, and potential uses. **Microbial Cell Factories**, v. 13, n. 6, p. 1-14, 2007.
- JÚNIOR, A. B.; PRATA, M. A.; MOREIRA, H. L. et al. Avaliação de desempenho produtivo e reprodutivo de animais mestiços do cruzamento Holandês x Gir. **Boletim de Indústria Animal**, v.71, n.4 p.357-364, 2014.
- KAMINSKI, S.; BRYM, P.; RUSC, A. et al. Associations between milk performance traits in Holstein cows and 16 candidate SNPs identified by arrayed primer extension (APEX) microarray. **Animal Biotechnology**, v. 17, n. 1, p. 1-11, 2006.
- KHATKAR, M.S.; THONSON, P.C.; TAMMEN, I. et al. Quantitative trait loci mapping in dairy cattle: review and meta-analysis. **Genetics, Selection and Evolution**, v. 36, n. 2, p. 163-190, 2004.
- KHATKAR, M.S.; NICHOLAS, F.W.; COLLINS, A.R. et al. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. **BMC Genomics**, v. 9, n. 187, p. 1-18, 2008.
- LAGZIEL, A.; LIPKIN, E.; SOLLER, M. Associations between SSCP haplotypes at the bovine growth hormone gene and milk protein percentage. **Genetics**, v. 142, n. 3, p. 945-951, 1996.
- LEE, M.A.; KEANE, O.M.; GLASS, B.C. et al. Establishment of a pipeline to analyse non-synonymous SNPs in *Bos taurus*. **BMC Genomics**, v. 7, n.298, p. 1-12, 2006.

- LEWONTIN, R.C. The Interaction of Selection and Linkage. Ii. Optimum Models. **Genetics**, v. 50, n. 10, p. 757-782, 1964.
- LEWONTIN, R.C.; KOJIMA, K. The evolutionary dynamics of complex polymorphisms. **Evolution**, v. 14, n. 4, p. 458-472, 1960.
- LI, Y.; WILLER, C.; SANNA, S.; ABECASIS, G. Genotype imputation. **Annual Reviews Genomics and Human Genetics**, v. 10, n. 9, p. 387-406, 2009.
- LONERGAN, S.M.; ERNST, C.W.; BISHOP, M.D. et al. Relationship of restriction fragment length polymorphisms (RFLP) at the bovine calpastatin locus to calpastatin activity and meat tenderness. **Journal of Animal Science**, v. 73, n. 12, p. 3608-3612, 1995.
- LOPES, P. S.; TORRES, R. A.; REIS FILHO, J. C. Consanguinidade e Cruzamento. In: LOPES, P. S. Teoria do Melhoramento Animal. 1.ed. Belo Horizonte: FEPMVZ-Editora, 2005. cap.4, p.82-93.
- MAEDA, M.; MURAYAMA, N.; ISHII, H. et al. A simple and rapid method for HLA-DQA1 genotyping by digestion of PCR amplified DNA with allele specific restriction endonucleases. **Tissue Antigens**, v.34, n.5, p.290-298, 1989.
- MCKAY S.D.; SCHNABEL R.D.; MURDOCH B.M. et al. Whole genome linkage disequilibrium maps in cattle. **BMC Genetics**, v. 8, n. 74, p. 1-12, 2007.
- MCMANUS, C.; TEIXEIRA, R.A.; DIAS, L.T. et al. Características Produtivas e Reprodutivas de Vacas Holandesas e Mestiças Holandês x Gir no Planalto Central. **Revista Brasileira de Zootecnia**, v. 37, n. 5, p. 819-823, 2008.
- MCRAE, A.F.; MCEWAN, J.C.; DODDS, K.G. et al. Linkage disequilibrium in domestic sheep. **Genetics**, v. 160, n. 3, p. 1113-1122, 2002.
- MELLO, F. de. **Diversidade genética e associação do gene de osteopontina com a produção de leite em bovinos da raça Girolando**. Dissertação (Mestrado - Produção Animal). Programa de Pós-Graduação em Zootecnia. Faculdade de Agronomia. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010.
- MELLO, F.; COBUCI, J.A.; MARTINS, M.F. et al. Association of the polymorphism g.8514C>T in the osteopontin gene (SPP1) with milk yield in the dairy cattle breed Girolando. **Animal Genetics**, v. 43, n. 5, p. 647-648, 2012.
- NANZER, T.A.D.T. 2010. Produção de leite no Brasil e participação da genética Girolando com ênfase em reprodução. Disponível em: <<http://www.girolando.com.br/site/noticia.php?id=1859>>. Acesso em: 20/11/2011.
- ORITA, M.; IWAHANA, H.; KANAZAWA, H. et al. Detection of polymorphisms of human DNA by gel electrophoresis as singlestrand conformation polymorphisms. **Proceedings of the National Academy of Sciences USA**, v. 86, n. 8, p. 2766-2770, 1989.
- PEREIRA, J. C. C. **Melhoramento Genético Aplicado à Produção Animal**. 5. ed. Belo Horizonte: FEP-MVZ, 2008. v. 1. 618 p.
- PICOULT-NEWNERG, L.; IDEKER, T.E.; POHL, M.G. et al. Mining SNPs from EST databases. **Genome Research**, v. 9, n. 2, p. 167-174, 1999.
- REMYINGTON, D. L.; THORNSBERRY, J. M.; MATSUOKA, Y. et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome. **Proceedings of the National Academy of Sciences of the United States of America**. v. 98, n. 20, p. 11479-11484, 2001.

- RIBEIRO, M. E. R.; STUMPF JÚNIOR, W.; BUSS, H. Qualidade de leite. In: BITENCOURT, D.; PEGORARO, L.M.C.; GOMES, J.F. (Ed.). Sistemas de pecuária de leite: uma visão na região de clima temperado. Pelotas: **Embrapa Clima Temperado**, 2000. p. 175-195.
- ROSA, A. J. M.; FRAGOSO, R.R. . Análise genômica aplicada a produção animal. In: Fábio Gelape Faleiro; Solange Rocha Monteiro de Andrade; Fábio Bueno dos Reis Junior. (Org.). (Org.). BIOTECNOLOGIA: estado da arte e aplicações na agropecuária.. 1ed.: , 2011, p. 569-653.
- ROSA, D.C.; TRENTIN, J.M.; PESSOA, G.A. et al. Qualidade do leite em amostras individuais e de tanque de vacas leiteiras. **Arquivos do Instituto Biológico**, v. 79, n. 4, p. 485-493, 2012.
- ROSA, G.J.M.; ROCHA, L.B.; FURLAN, L.R. Estudos de expressão gênica utilizando-se microarrays: delineamento, análise, e aplicações na pesquisa zootécnica. **Revista Brasileira de Zootecnia**, v. 36, p. 185-209, 2007 (supl. especial).
- ROSA, J.R.B.F. **Análise do desequilíbrio de ligação e da estrutura populacional do germoplasma brasileiro de cana-de-açúcar**. 2011. 98f. Dissertação (Mestrado em Ciências). Escola Superior de Agricultura “Luiz de Queiroz” – Piracicaba, Universidade de São Paulo. 2011.
- ROTHSCHILD, M.F.; SOLLER, M. Candidate gene analysis to detect traits of economic importance in domestic livestock. **Probe**, v. 8, p. 13-21, 1997.
- SILVA, J. C. P. M.; VELOSO, C. M. **Raças de Gado Leiteiro**. 1.ed. Viçosa: Aprenda Fácil, 2011. 149 p.
- SILVA, M.V.G.B.; FREITAS, A.F.; PAIVA, L.C. et al. **Programa de Melhoramento Genético da Raça Girolando – Teste de Progênie, Sumário de Touros**. Juiz de Fora: EMBRAPA Gado de Leite, 2010. 48p.
- SILVA, M.V.G.B.; MARTINS, M.F.; PAIVA, L.C. et al. **Programa de Melhoramento Genético da Raça Girolando – Teste de Progênie, Sumário de Touros**. Juiz de Fora: EMBRAPA Gado de Leite – Documentos 163, 2013. 52p.
- SIMKO, I. HAYNES, K. G.; JONES, R. W. Assessment of linkage disequilibrium in potato genome with single nucleotide polymorphism markers. **Genetics**, v. 173, n. 4, p. 2237-2245, 2006.
- SLADEK, R.; ROCHELEAU, G.; RUNG, J. et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. **Nature**, v. 445, n. 7130, p. 881-885, 2007.
- SLATKIN, M. Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. **Nature Reviews Genetics**, v. 9, n. 6, p. 477-485, 2008.
- SODELAND, M.; GROVE, H.; KENT, M. et al. Molecular characterization of a long range haplotype affecting protein yield and mastitis susceptibility in Norwegian Red cattle. **BMC Genetics**, v.12, n.8, p.1-10, 2011.
- STRANGER, B.E.; FORREST, M.S.; CLARK, A.G. et al. Genome-wide associations of gene expression variation in humans. **PLoS Genetics**, v. 1, n. 6, p. e78, 2005.
- VANRADEN, P. M. United States Department of Agriculture - Agricultural Research Service. Disponível em:<  
<http://www.ars.usda.gov/pandp/people/people.htm?personid=5781>>. Acesso em: 29/11/2014.

- VAN TASSELL, C.P.; SMITH, T.P.L.; MATUKUMALLI, L.K. et al. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. **Nature Methods**, v. 5, n. 3, p. 247–252. 2008.
- WANG, N.; AKEY, J.M.; ZHANG, K. et al. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. **The American Journal of Human Genetics**, v. 71, n. 5, p. 1227-1234, 2002.
- WANG, X.; WURMSER, C.; PAUSCH, H. et al. Identification and Dissection of Four Major QTL Affecting Milk Fat Content in the German Holstein-Friesian Population. **PLoS ONE**, v. 7, n. 7, p. e40711, 2012.
- ZANELA M.B.; FISCHER, V.; RIBEIRO, M.E.R. et al. Qualidade do leite em sistemas de produção na região sul do Rio Grande do Sul. **Pesquisa Agropecuária Brasileira**, v. 41, n. 1, p. 153-159, 2006.

## **CAPÍTULO 2**

**DESENVOLVIMENTO DE UM PAINEL DE BAIXA DENSIDADE DE  
MARCADORES SNP DERIVADOS DE GENES CANDIDATOS PARA  
CARACTERÍSTICAS DE PRODUÇÃO, GENES CAUSADORES DE DOENÇAS  
GENÉTICAS, E PARA CONTROLE DE PARENTESCO EM ANIMAIS DA RAÇA  
GIROLANDO**

## 2.1. RESUMO

A utilização de marcadores moleculares tem contribuído de forma eficaz nos programas de melhoramento, isto porque possibilita a aplicação direta das informações de DNA na seleção, auxiliando a aumentar a eficiência e a ter maior rapidez na obtenção de ganhos genéticos com a seleção. Contudo, para a realização de estudos de associações mais acuradas, um dos grandes desafios consiste na qualidade dos dados genotípicos gerados, ou seja, provenientes de amostras de DNA de boa qualidade e sem erros de genotipagem. Neste sentido, o objetivo do presente trabalho foi desenvolver e validar um painel de baixa densidade a partir de um ensaio de 384 marcadores SNP derivados de genes candidatos para características de produção, genes causadores de doenças genéticas, e marcadores para controle de parentesco em animais da raça Girolando. Um painel de 384 marcadores SNP foi desenvolvido a partir de informações de genes candidatos relacionados principalmente à produção e qualidade do leite, a doenças hereditárias e teste de paternidade. A população investigada estava composta por 576 animais, sendo 204 machos e 372 fêmeas. Os critérios estabelecidos para a filtragem dos SNPs/animais foram: Separação de Clusters; Frequência de Genotipagem e Taxa de Genotipagem. Após aplicação destes critérios, foram eliminados 128 SNPs e 71 animais. Em seguida, um total de 5 SNPs foram eliminados do painel filtrado por apresentarem resultados inconsistentes com o padrão de herança mendeliana. Adicionalmente, outros 2 SNPs foram eliminados por apresentarem erros de reprodutibilidade. A proporção geral de SNPs polimórficos observados neste estudo ( $N = 576$  animais) considerando todos os 384 SNPs foi estimada em 71,6%. Após aplicação dos critérios de filtragem, foi possível observar uma redução da proporção de SNPs polimórficos (53,4%). O painel validado ficou composto de 249 SNPs (3 relacionados com doenças genéticas, 37 para testes de paternidade e 209 para



características de produção de leite) distribuídos em 69 genes candidatos. Este painel pode ser utilizado com diferentes propósitos na raça Girolando. Com este painel, pode-se realizar estudos de estruturação genética, estimar o grau de diversidade dentro e entre populações, identificar de QTL e regiões do genoma associadas a características de produção leiteira aptas à seleção. Adicionalmente, recomenda-se aumentar o número de amostras genotipadas com este painel para incrementar a acurácias das estimativas.

**PALAVRAS-CHAVE:** beadxpress, controle de qualidade, genotipagem, rebanhos mestiços, tecnologia goldengate.

## 2.2. ABSTRACT

The use of molecular markers has contributed effectively in breeding programs, because it enables direct application of DNA information in the selection, helping to increase efficiency and to have greater speed in obtaining genetic gains with selection. However, to carry out studies of more accurate associations, a major challenge is the quality of the generated genotypic data, that is, from good quality DNA samples and without genotyping errors. In this sense, the objective of this study was to develop and validate a low density panel from an assay of 384 SNP markers derived from candidate genes for production traits, genes that cause genetic diseases, and markers for paternity control in animals race Girolando. A 384SNP marker panel was developed from information of candidate genes related mainly to the production and quality of milk, hereditary diseases and paternity test. The population studied was comprised of 576 animals, being 204 males and 372 females. The criteria for filtering of SNPs / animals were: Clusters of separation; Frequency of Genotyping and Genotyping Rate. After applying these criteria, 128 SNPs were eliminated and 71 animals. Then, a total of five SNPs were removed from the filtrate panel for presenting inconsistent results with standard Mendelian inheritance. Additionally, two other SNPs have been eliminated due to reproducibility errors. The overall proportion of polymorphic SNPs observed in this study (N = 576 animals), considering all 384 SNPs was estimated at 71.6%. After applying the filtering criteria, it was possible to observe a reduction in the proportion of polymorphic SNPs (53.4%). The validated panel was composed of 249 SNPs (3-related genetic diseases, 37 to paternity test and 209 for milk production traits) distributed in 69

candidate genes. This panel can be used for different purposes in Girolando breed. With this panel, one can perform genetic structure studies estimate the degree of diversity within and between populations and identify QTL and regions of the genome associated with milk production traits suitable for selection. In addition, it is recommended to increase the number of samples genotyped with this panel to increase the accuracies of the estimates.

**KEY-WORDS:** beadxpress, crossbred cattle, genotyping, goldengate technology, quality control.

### 2.3. INTRODUÇÃO

O setor pecuário apresentou crescimento acelerado nas últimas décadas em diversas regiões rurais do Brasil, decorrente da elevada competitividade entre os produtores e dos aumentos nas exigências do mercado consumidor. De acordo com Souza & Buainain (2013), a elevada competitividade entre os produtores, associada a aumentos nas exigências do mercado consumidor gera demandas para adoção de tecnologias mais avançadas para manejo e principalmente melhoramento genético dos rebanhos, que permitam ganhos em produtividade, lucratividade e de qualidade dos produtos de origem animal. Para os produtores, estas alterações levaram a impactos que remetem a maior pressão para ampliação da escala de produção e maior exigência em relação à qualidade do leite produzido.

Recentes inovações e aprimoramentos nas tecnologias de sequenciamento de DNA e de genotipagem de marcadores moleculares favoreceram a geração de dados precisos, com custos reduzidos, tornando possível a utilização de marcadores moleculares como ferramenta nos processos de avaliação, melhoramento e manejo genético animal (Reis Filho et al., 2010). A utilização de informações moleculares nos procedimentos de avaliação genética permite maior compreensão de processos biológicos, favorecendo, conseqüentemente, o desenvolvimento de ferramentas e soluções tecnológicas, que visem a identificação e caracterização de fatores genéticos envolvidos na produção do leite. Conforme relatado por Mello et al. (2012), os estudos de polimorfismos associados a produção de leite têm sido divulgados e empregados com maior amplitude em rebanhos bovinos. Atualmente, os marcadores de escolha para atingir este objetivo são os polimorfismos de base única – SNP, por serem abundantes, amplamente distribuídos ao longo do genoma e com tecnologia

passível de automação. Busca-se seleção baseada não apenas em características de processos envolvidos na produção leiteira, mas também na identificação de genes responsáveis pela expressão diferencial do fenótipo e que interfiram nesta produção.

Inúmeros trabalhos relatam a associação entre SNPs e proteínas do leite (Cohen-Zinder et al., 2005; Wozuk et al., 2008; Schennink et al., 2008; Bionaz & Loor, 2011; Wang et al., 2012). Estudos identificaram também SNPs localizados em regiões codificantes de genes (Grisart et al., 2004; Chakravarty et al., 2005; Signorelli et al., 2009; Chatterjee et al., 2009; Jiang et al., 2010), sem que houvesse impacto aparente sobre o fenótipo, entretanto tratam-se de SNPs que podem influenciar o rendimento da expressão gênica, podendo ser utilizados como marcadores de regiões genômicas desconhecidas (Jacob & Puhon, 1992).

Lien et al. (2010) destacam que, a partir do monitoramento de genes bovinos através de painéis de SNPs, foi possível a identificação daqueles que encontravam-se associados à produção de leite. Em se tratando de ferramentas específicas a este fim, os experimentos através de *microchips*, segundo Ramsay (1998), ofereceram a abertura para estudos sobre a complexidade do genoma bovino. Considerando a associação de SNPs com a biossíntese protéica do leite, Kaminski et al. (2005) utilizaram *microchips* (*MilkProtChip*) de SNPs para genotipar simultaneamente 71 SNPs, localizados em 42 genes candidatos. Estes autores destacaram que os determinantes genéticos que interferem no teor do leite são objeto de estudo há mais de 50 anos, principalmente a partir da descoberta do polimorfismo na  $\beta$ -lactoglobulina. Com isso, a caracterização de um grande número de variantes genéticas relacionadas a proteínas do leite tornou-se factível. Em relação as proteínas do leite, estas podem ser classificadas em dois grupos: as caseínas *S1*  $\alpha$ -*CSN1S1*, *S2*  $\alpha$ -*CSN1S2*,  $\beta$ -*CSN2* e  $\kappa$ -*CSN3*; e as proteínas do soro  $\beta$ -lactoglobulina - *CGL* e a  $\alpha$ -*lactoalbumina* - *LALBA*.

Conforme Chessa et al. (2007), a detecção de polimorfismos em fragmentos de DNA por meio de ensaios adequados permite a avaliação dos genes associados às proteínas do leite, o que favorece o desenvolvimento de programas de seleção em rebanhos de bovinos leiteiros. Estes autores destacam a importância do mapeamento de SNPs, para que sejam identificados genes associados a doenças multifatoriais e também a características quantitativas (QTL). Estes mesmos autores destacam que mais de 95% das proteínas contidas no leite são codificadas no cromossomo 6 (genes estruturados), cromossomo 2 (genes *LALBA* e *CGL*) e cromossomo 4 (*CN*,  $\alpha$ *S1-CN*,  $\alpha$ *S2-CN*,  $\beta$ -*CN* e  $\kappa$ -*CN*). Os polimorfismos no grupo CN favorecem o entendimento da produção de caseína em diferentes locus, indicando a

participação dos haplótipos como elementos determinantes dessa característica associada à produção.

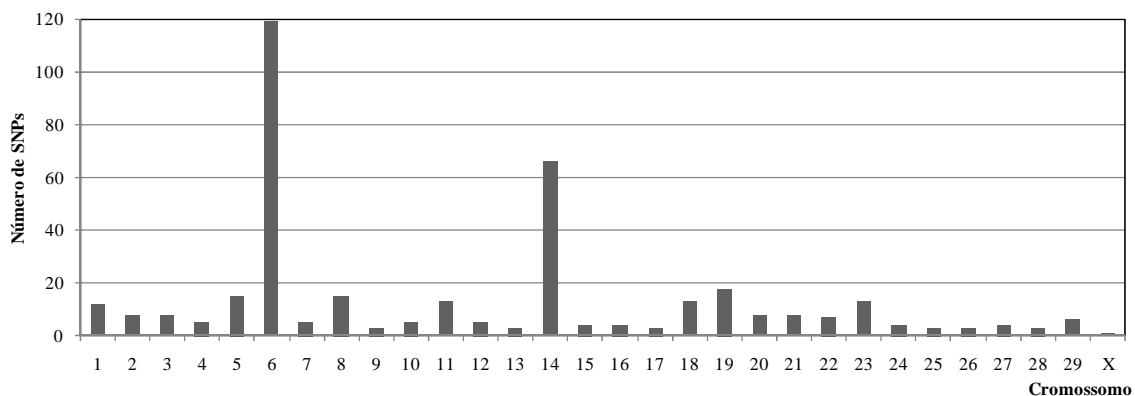
Em decorrência da elevada quantidade de SNPs envolvida em processo de genotipagem através da metodologia *GoldenGate*, mesmo mediante todos os avanços tecnológicos pertinentes ao procedimento, há que se primar por critérios que atribuam maior segurança na filtragem dos dados (Grattapaglia et al., 2011). Busca-se portanto, o desenvolvimento de painéis que permitam expandir a resolução e taxas de transferências, a fim de que resultados sejam validados e possam servir de fundamentação e/ou base em demais experimentos. Segundo Laurie et al. (2010), a qualidade dos dados de genotipagem, em relação a reprodutibilidade e a acurácia, representa um fator de grande importância para a geração de resultados confiáveis. Neste contexto, para os procedimentos de genotipagem de SNPs, várias estratégias que buscam padronizar o controle e garantir a qualidade (QC) na filtragem de dados foram desenvolvidos. Como se pode observar é preciso que todo o procedimento seja regrado, de modo que se configure confiabilidade aos resultados alcançados. Todavia, apesar de ser uma metodologia eficiente do ponto de vista da geração de um grande número de dados em um curto intervalo de tempo, de baixo custo e alto benefício por amostra genotipada, pouco se sabe sobre a qualidade e a confiabilidade dos genótipos gerados. Alguns estudos têm apresentado resultados satisfatórios quanto a intensidade dos genótipos gerados (Baird et al., 2008; Huang et al., 2009; Myles et al., 2010; Elshire et al., 2011), no entanto, os autores não são claros quanto aos parâmetros de qualidade utilizados na seleção dos marcadores.

Sendo assim, objetivou-se neste capítulo desenvolver e validar um painel de baixa densidade a partir de um ensaio de 384 marcadores SNP derivados de genes candidatos para características de produção, genes causadores de doenças genéticas, e marcadores para controle de parentesco em animais da raça Girolando.

## 2.4. MATERIAL E MÉTODOS

### 2.4.1. Seleção de SNPs em Genes Candidatos

Um conjunto de 384 SNPs foi obtido a partir da mineração de artigos publicados na literatura científica, bancos de dados públicos (*GeneBank*) e com base em dados gerados pela Embrapa para compor um ensaio GoldenGate a ser validado. Os SNPs selecionados foram derivados de genes candidatos relacionados principalmente à produção e qualidade do leite, com 66 e 119 SNPs concentrados nos cromossomos 6 e 14, respectivamente, e a doenças hereditárias (Anexo 01) com SNPs distribuídos em todos os cromossomos (Figura 2.1). Adicionalmente, também foram incluídos neste painel um total de 96 SNPs com alto nível de polimorfismo em diferentes raças de bovinos taurinos, específicos para condução de testes de paternidade (Heaton et al., 2002; Elsik et al., 2009; Matukumalli et al., 2009).



**Figura 2.1** Distribuição dos 384 SNPs nos 30 cromossomos bovino que compõe o ensaio desenhado para animais da raça Girolando.

#### 2.4.2. Seleção de amostras dos animais para Genotipagem

#### 2.4.3. Quantificação das amostras de DNA genômico com PicoGreen

Um total de 510 animais da raça Girolando (173 machos e 337 fêmeas) foram analisados com o painel de 384 SNPs em questão. As amostras referentes a estes animais foram disponibilizados pela Embrapa Gado de Leite de fazendas mineiras participantes do programa de melhoramento da raça Girolando. Os materiais utilizados na extração de DNA foram sangue (vacas) e sêmem (touros). Adicionalmente, 15 amostras correspondentes a 5 trios (touro, vaca e produto) foram utilizadas para os testes de herança mendeliana, e 43 amostras foram analisadas em duplicata para testar o poder de reprodutibilidade do painel. Essas repetições correspondem a duplicatas de 8 animais da raça Curraleira Pé Duro (banco de DNA do Laboratório de Genética Animal – Embrapa Cenargen) e de 35 animais da raça Girolando, totalizando 576 animais analisados.

O DNA total extraído foi quantificado por fluorimetria (*BioTek microplate spectrophotometer, Take3*) utilizando-se o reagente PicoGreen® (*PicoGreenDNA Assay Kit, Invitrogen*), seguindo o protocolo do fabricante. O *PicoGreen* é um fluoróforo ultra-sensível utilizado para a quantificação de DNA genômico em fita dupla (dsDNA). Os comprimentos de onda utilizados são de 485 nm e 520 nm para a excitação e para a emissão, respectivamente. A determinação da concentração de soluções de DNA é realizada por intermédio do estabelecimento de retas de calibração com concentração do  $\lambda$ DNA conhecida



(0 ng/ml à 1.000 ng/ml. Todas as amostras dos animais foram quantificadas e normalizadas no intervalo de concentração de 50 a 70 ng/ $\mu$ l.

#### 2.4.4. Genotipagem

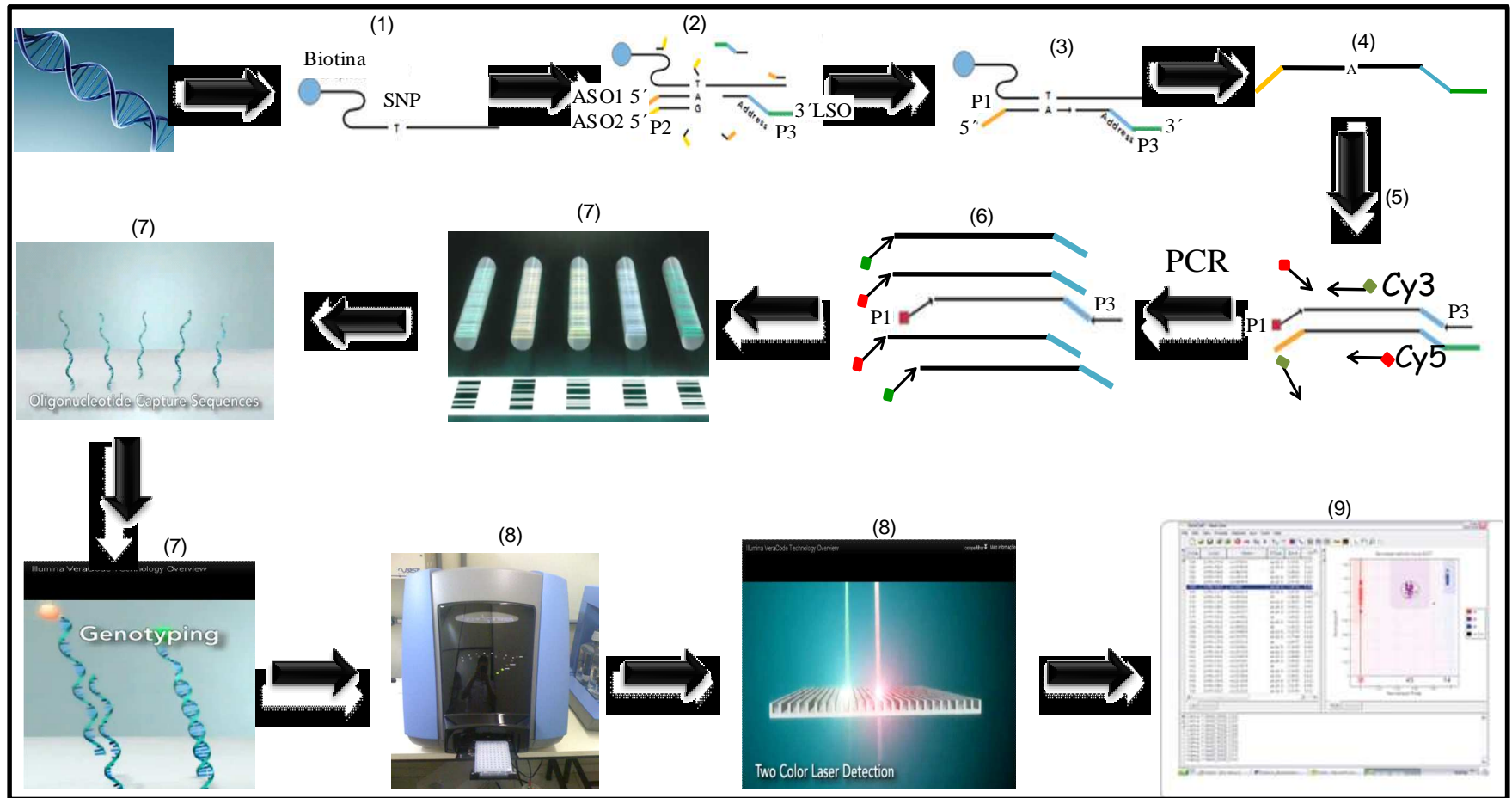
Os SNPs selecionados foram usados para construir um painel baseado na tecnologia *VeraCode* (Illumina Inc., San Diego, Califórnia), utilizando os procedimentos padrão indicados pelo fabricante. A genotipagem dos SNPs foi realizada em equipamento *BeadXpress*<sup>®</sup> Illumina (Illumina, San Diego, CA) de acordo com o protocolo recomendado pelo fabricante (Fan et al., 2003).

Na etapa inicial do ensaio *GoldenGate* a ativação do DNA Genômico (gDNA) foi executada através da incorporação de biotina, para que este possa ser ligado posteriormente à partículas paramagnéticas conjugadas com estreptavidina (Figura 2.2 – passo 1). Para este processo de ativação é necessário uma concentração inicial mínima de 250 ng/5 $\mu$ l de gDNA para cada amostra.

Três oligonucleotídeos (Óligos) foram desenhados para cada SNP do ensaio (Figura 2.2 – passo 2). Dois óligos são específicos para cada um dos alelos do sítio SNP (extremidade 3'), denominados de óligos alelo-específicos (ASO). O terceiro óligo (Óligo Locus Específico - LSO) é complementar à 50 bases a partir da posição do SNP (extremidade 5'). Na próxima etapa do ensaio, os óligos, o tampão de hibridização, e as partículas paramagnéticas foram combinados com o gDNA ativado (Figura 2.2 – passo 2). Quando ocorre o alinhamento perfeito entre o óligo específico com o gDNA ativado, o processo de extensão/ligação até o óligo locus-específico (Figura 2.2 – passo 3) é executado pelas enzimas (MEL). Posteriormente, o passo de purificação da reação hibridizada foi realizado (Figura 2.2 – passo 4).

Na próxima etapa, foram adicionados óligos alelo-específicos marcados com dois fluorocromos (Figura 2.2 – passo 5), o Cyanine-3 (Cy3 – verde) e o Cyanine-5 (Cy5 – vermelho), para realização da reação de PCR (Figura 2.2 – passo 5). Após a amplificação, os produtos de PCR marcados (produtos de ensaio *GoldenGate* - Figura 2.2 – passo 6) foram hibridizados com bastonetes de sílica *Veracode*<sup>®</sup> (Figura 2 – passo 7), contendo óligos complementares e um código de barras, específicos para cada SNP avaliado no ensaio.

Após a hibridização (Figura 2.2 – passo 7), os bastonetes de sílica *Veracode*® foram lidos com o equipamento *BeadXpress*® (Figura 2.2 – passo 8), onde um laser é utilizado para excitar os fluorocromos (Figura 2.2 – passo 8), permitindo analisar/detectar o sinal da fluorescência emitido por cada bastonete *Veracode*, de cada amostra testada. O equipamento capta simultaneamente os dois diferentes comprimentos de onda emitidos pelos fluorocromos utilizados (Cy3 e Cy5). Indivíduos homozigotos, para um determinado marcador SNP, emitem apenas um sinal de Cy3 ou Cy5, ao passo que indivíduos heterozigotos emitem sinal nos dois comprimentos de onda, simultaneamente (Figura 2).



**Figura 2.2** Fluxograma ilustrativo da técnica de ligação de oligonucleotídeos empregada na detecção de alelos em sítios SNP – Método *VeraCode* (fonte: [www.illumina.com](http://www.illumina.com)).

Os dados brutos da genotipagem foram analisados com o Módulo de Genotipagem do programa *GenomeStudio*, V2010.1, (Illumina, San Diego, CA), utilizando parâmetros padrão para clusterização dos dados de cada SNP (Figura 2.2 – passo 8). Os resultados de cada SNP foram reavaliados manualmente e os clusters editados (Anexo 03). Os ajustes manuais foram importantes para corrigir possíveis erros de genotipagem de amostras presentes em regiões de interface entre dois genótipos. Estimativas dos parâmetros de qualidade das genotipagens obtidas foram gerados individualmente para os animais e os SNPs testados. Os parâmetros de qualidade gerados foram utilizados posteriormente para filtrar os dados brutos e eliminar resultados com alta probabilidade de erro, conforme descrito no item 2.4.5.

#### 2.4.5. Critérios para Controle de Qualidade dos Dados Gerados

O primeiro critério de filtragem aplicado foi baseado no parâmetro Separação de Clusters (*Cluster Separation*), o qual avalia a separação dos clusters dos três genótipos possíveis em uma dimensão teta ( $\theta$ ), com valores que variam entre 0 e 1. Marcadores com valores de Separação de Clusters inferiores a 0,15 foram excluídos das análises subsequentes. Em seguida, marcadores com Frequência de Genotipagem (*Call Frequency*) inferior a 85% foram excluídos. Finalmente, amostras que apresentaram uma Taxa de Genotipagem (*Call Rate*) inferior a 95% foram excluídas das demais análises.

#### 2.4.6. Análise de Componentes Principais

Na população estudada estavam presentes animais de seis grupos genéticos (GGs) diferentes: 1/2Holandês+1/2Gir; 1/4Holandês+3/4Gir; 7/8Holandês+1/8Gir; 13/16Holandês+3/16Gir; 3/4Holandês+1/4Gir; 5/8Holandês+3/8Gir, além de animais que não possuíam informação. Para avaliar a estruturação destes GGs com base nos genótipos dos SNPs incluídos no painel foi aplicado a análise de componentes principais, utilizando a função *prcomp* do *software* R (R Core Team, 2012). A análise de componentes principais foi realizada a partir de uma matriz de distância cujo os elementos foram definidos como coeficiente de parentesco (*kinship*). O coeficiente de parentesco  $\theta_{AB}$  é a probabilidade de que um alelo tomado ao acaso no animal *A* seja idêntico por descendência a um alelo também tomado ao acaso no animal *B*. Dessa forma, a matriz de parentesco representa o grau de

relacionamento entre cada possível par de animais em uma amostra (Hedrick, 2010). Neste trabalho, a matriz de parentesco foi estimada através da função *kinship* disponível na biblioteca *kinship2* e a análise de componentes principais foi realizada usando a função *cmdscale* ambos implementados no *software* estatístico R (R Core Team, 2012).

#### **2.4.7. Teste de Paternidade, de Herança Mendeliana e Reprodutibilidade dos Genótipos Obtidos**

Os dados dos marcadores selecionados foram submetidos à análise no programa CERVUS 3.0 (MARSHALL et al. 1998), para gerar as estimativas de Heterozigosidade observada ( $H_o$ ) e esperada ( $H_e$ ) e as probabilidades combinadas de não exclusão do primeiro pai (NE-1P), do segundo pai (NE-2P) e a combinação de ambos pais (NE-PP). O padrão de transmissão alélica de cada marcador SNP foi avaliado com a utilização dos dados de cinco trios (touro, vaca e produto) genotipados, considerando-se os dados filtrados. Os testes de herança Mendeliana foram realizados com o programa *GenomeStudio*, versão 2010.1, para identificação de divergências no padrão de transmissão esperado, decorrentes de erros de genotipagem. Estes testes foram aplicados para avaliar o poder dos critérios adotados na geração de dados confiáveis. SNPs que apresentaram erros em algum desses testes foram excluídos manualmente do arquivo final (painel validado).

#### **2.4.7. Avaliação dos Resultados de Controle de Qualidade dos Dados**

A confiabilidade dos resultados obtidos para cada um dos SNPs testado, após a aplicação dos filtros de Controle de Qualidade, foi avaliada com os parâmetros *GenTrain* e *GenCall50*, conforme procedimentos indicados pelo fornecedor do equipamento. O parâmetro *GenTrain* estimado para cada SNP varia entre 0 e 1 e é baseado tanto nas formas dos clusters de cada genótipo (homozigotos e heterozigoto), quanto na distância entre os clusters. Valores mínimos de *GenTrain*  $\geq 0,40$  foram considerados satisfatórios.

O parâmetro *GenCall50* foi estimado para cada amostra em função dos resultados obtidos para cada um dos SNPs contidos no ensaio. Este critério avalia a genotipagem de cada amostra em relação ao centro do cluster dos genótipos de um determinado SNP. Genótipos com menor escore de *GenCall50* estão mais distantes do centro

do cluster do genótipo, e por conseguinte, apresentam uma confiabilidade mais baixa. Esta medida corresponde à probabilidade de 50% das amostras, para um dado SNP, estarem dentro de um dos clusters possíveis. Genótipos que apresentaram valores de *GenTrain* e de *GeneCall50*  $\leq 0,40$  foram eliminados dos resultados finais. A comparação entre os resultados obtidos antes e após a aplicação dos critérios de filtragem, foi realizado através do teste de *Wilcoxon*, utilizando o *software* SAS (SAS, 2010).

#### 2.4.8. Estudo do Equilíbrio de Hardy-Weinberg e Nível de Polimorfismo do SNP

Os desvios das frequências observadas em relação às esperadas para o equilíbrio de Hardy-Weinberg foram analisadas pelo teste de probabilidade, de acordo com Raymond & Rousset (1995). O nível de polimorfismo de cada marcador SNP inserido no ensaio foi avaliado pelo parâmetro convencional MAF (Frequência do Alelo Menor, do inglês *Minor Allele Frequency*). Um SNP foi considerado polimórfico se o valor da MAF estimado foi  $\geq 0,05$ . As estimativas de MAF foram feitas considerando apenas os dados obtidos com as amostras da raça Girolando, após a aplicação dos filtros de qualidade descritos acima.

## 2.5. RESULTADOS

A aplicação do primeiro critério de filtragem dos dados (eliminação de marcadores com Separação de Clusters  $< 0,15$ ) resultou na exclusão de 63 SNPs (Tabela 2.1).

**Tabela 2.1** Resultados da filtragem dos dados com base no parâmetro de Separação dos Clusters (*Cluster Separation*).

Performance do Ensaio	Antes da Filtragem	Depois da Filtragem
Número de Amostras	576	576
SNPs Testados	384	321
SNP com <i>Separação de Cluster</i> $< 0,15$	63	0
SNPs com <i>Frequência de Genotipagem</i> $< 90\%$	76	65

Média da <i>Taxa de Genotipagem</i>	86%	92%
Amostras com <i>Taxa de Genotipagem</i> $\geq 95\%$	316	418
% de Amostras com $\geq 95\%$	62	82*
Média Escore <i>GenTrain</i>	59%	71%
SNPs com <i>GenTrain</i> $\geq 0,40$	317	321
% de SNPs com <i>GenTrain</i> $\geq 0,40$	83	100*
Média do <i>GenCall50%</i>	69%	65%
SNPs com <i>GenCall50%</i> $\geq 0,40$	314	317
% dos SNPs com <i>GenCall 50%</i> $\geq 0,40$	82	98*

\*Aumento significativo ( $p < 0,001$ ) a partir dos testes *Wilcoxon*

Aumentos dos parâmetros *Taxa de Genotipagem* (20%), *GenTrain* (17%) e *GenCall50%* (16%), significativos a  $p < 0,001$  foram observados como resultado da aplicação do primeiro critério de filtragem dos dados. Adicionalmente, 100% dos SNPs apresentaram resultados acima do limite mínimo estabelecido para o parâmetro *GenTrain* ( $\geq 0,40$ ), após a filtragem dos dados. O principal fator para esse aumento pode ser associado a edição dos clusters para correção de possíveis erros de genotipagem. Com o aumento/redução do diâmetro do cluster e/ou deslocamento do mesmo, ocorreu o aumento proporcional da estimativa do parâmetro *gentrain* até um valor máximo. A partir deste valor, quaisquer edições acarretaram na diminuição de sua estimativa. Esta característica é específica para cada SNP e permitiu que as edições configurassem uma distribuição dos clusters otimizada para cada marcador.

A aplicação do segundo critério de filtragem (exclusão de SNPs com *Frequência de Genotipagem*  $< 0,90$ ) resultou na eliminação de 65 SNPs do ensaio (Tabela 2.2). Este filtro representou a exclusão de SNPs que falharam na genotipagem em mais de 10% das amostras.

**Tabela 2.2** Resultados da filtragem dos dados com base no parâmetro de *Frequência de Genotipagem*.

Performance do Ensaio	Antes da Filtragem	Depois da Filtragem
Número de Amostras	576	576
SNPs Testados	321	256
SNP com <i>Separação de Cluster</i> $< 0,15$	0	0
SNPs com <i>Frequência de Genotipagem</i> $< 0,90$	65	0

Média da <i>Taxa de Genotipagem</i>	92%	95%
Amostras com <i>Taxa de Genotipagem</i> $\geq 95\%$	418	447
% de Amostras com $\geq 95\%$	82	88
<hr/>		
Média Escore <i>GenTrain</i>	71%	71%
SNPs com <i>GenTrain</i> $\geq 0,40$	321	256
% de SNPs com <i>GenTrain</i> $\geq 0,40$	100	100
<hr/>		
Média do <i>GenCall50%</i>	65%	69%
SNPs com <i>GenCall50%</i> $\geq 0,40$	317	256
% dos SNPs com <i>GenCall 50%</i> $\geq 0,40$	98	100

\*Aumento significativo ( $p < 0,001$ ) a partir dos testes *Wilcoxon*

Aumentos percentuais de 6% e de 2% (Tabela 2.2) foram observados para os parâmetros *Taxa de Genotipagem* e *GenCall*, respectivamente. O parâmetro *GenTrain* se manteve estável, com média de 71% e com 100% das amostras acima do limite mínimo estabelecido ( $\geq 0,40$ ).

A aplicação do terceiro critério de filtragem, que representa um controle de qualidade da genotipagem para as amostras dos animais, levou à exclusão de 91 animais do conjunto amostral inicial (Tabela 2.3).



**Tabela 2.3** Resultado da exclusão de amostras com Taxa de Genotipagem < 0,95.

Performance do Ensaio	Antes da Filtragem	Depois da Filtragem
Número de Amostras	576	485
SNPs Testados	256	256
SNP com <i>Separação de Cluster</i> < 0,15	0	0
SNPs com <i>Frequência de Genotipagem</i> < 0,90	0	0
Média da <i>Taxa de Genotipagem</i>	95%	95%
Amostras com <i>Taxa de Genotipagem</i> ≥ 95%	485	485
% de Amostras com ≥ 95%	88*	100*
Média Escore <i>GenTrain</i>	71%	71%
SNPs com <i>GenTrain</i> ≥ 0,40	256	256
% de SNPs com <i>GenTrain</i> ≥ 0,40	100	100
Média do <i>GenCall50%</i>	69%	69%
SNPs com <i>GenCall50%</i> ≥ 0,40	256	256
% dos SNPs com <i>GenCall50%</i> ≥ 0,40	100	100

\*Aumento significativo (p<0,001) a partir dos testes *Wilcoxon*

No conjunto de dados final, pode ser observado que 100% das amostras e dos SNPs testados apresentaram valores superiores aos limites mínimos de *GenTrain* e *GenCall50* indicados.

Um total de 5 SNPs foram eliminados do painel filtrado por apresentarem resultados inconsistentes com o padrão de herança mendeliana (Tabela 2.4), a partir da informação dos 5 trios.

**Tabela 2.4** Informações do teste de paternidade após critérios de filtragem.

Teste de Paternidade						
Código do Filho	Código do Parental1	Código do Parental2	SNPs após Filtragens	Nº de SNPs com erro	SNPs Validados	Frequência de Herança*
CK01C3014_R001_C012	CK01C3014_R002_C012	CK01C3014_R003_C012	256	0	251	100
CK01C3007_R001_C001	CK01C3007_R002_C001	CK01C3007_R003_C001	256	0	251	100
CK01C3023_R001_C012	CK01C3023_R002_C012	CK01C3023_R003_C012	256	0	251	100
CK01C3015_R001_C012	CK01C3015_R002_C012	CK01C3015_R003_C012	256	0	251	100
CK01C3014_R001_C012	CK01C3007_R002_C001	CK01C3007_R003_C001	256	5	251	100

\*100% de herança para o painel validado

Em seguida, a reprodutibilidade de cada marcador restante foi avaliada com a comparação dos resultados obtidos para as amostras testadas em 86 duplicatas (43 amostras) e dois SNPs foram eliminados por apresentarem erros de reprodutibilidade (Tabela 2.5).

**Tabela 2.5** Análise descritiva do Teste de Reprodutibilidade para o ensaio *GoldenGate* com o painel de 251 SNPs.

Teste de Reprodutibilidade					
Amostra 1	Repetição da Amostra 1	SNPs após Filtragens	Nº de SNPs com erro	SNPs Validados	Frequência de Reprodutibilidade*
CK01C3014_R004_C012	CK01C3014_R005_C012	251	0	249	100
CK01C3007_R004_C012	CK01C3007_R005_C012	251	0	249	100
CK01C3023_R004_C012	CK01C3023_R005_C012	251	0	249	100
CK01C3015_R004_C012	CK01C3015_R005_C012	251	0	249	100
CK01BY029_R004_C012	CK01BY029_R005_C012	251	0	249	100
CK01C3014_R004_C012	CK01C3007_R004_C012	251	2	249	100

\*100% de herança para o Painel Validado

A partir desse painel validado (249 SNPs) foram aplicados os testes para a geração das estatísticas descritivas do teste de paternidade (Tabela 2.6). Para o conteúdo de informação polimórfica (PIC) foram obtidos valores médios de 0,260, e para a heterozigiosidade observada ( $H_o$ ) e esperada ( $H_e$ ), os valores de 0,370 e 0,329, respectivamente. As estimativas das probabilidades combinadas de não exclusão (Tabela 2.6) foram abaixo de 0,00% para todos os parâmetros adotados, o que representa uma estimativa acima de 99% para a confirmação de possível paternidade, assim como a identificação individual, utilizando este painel.

**Tabela 2.6** Estimativas de diversidade independentes para 249SNPs selecionados para os testes de paternidade para todo o conjunto de animais testados (N); K=número de alelos observados;  $H_{Obs}$ =heterozigiosidade observada;  $H_{Exp}$ =heterozigiosidade esperada; PIC=conteúdo de informação polimórfica; NE-1P=probabilidade média de não-exclusão para um pai candidato; NE-2P=probabilidade média de não-exclusão para um pai candidato dado o genótipo de um pai conhecido do sexo oposto; NE-PP= probabilidade média de não-exclusão com genótipo de ambos pais conhecidos; NE-I=probabilidade média de não-exclusão para a identidade de dois indivíduos não relacionados; NE-SI=probabilidade média de não-exclusão para a identidade de dois irmãos.

SNP	k	N	$H_{Obs}$	$H_{Exp}$	PIC	NE-1P	NE-2P	NE-PP	NE-I	NE-SI
ABCG2	2	401	0,202	0,182	0,165	0,984	0,917	0,855	0,686	0,831
ARMC1	2	396	0,566	0,462	0,355	0,894	0,823	0,731	0,397	0,618

ARS_BFGL_BAC_1250	2	384	0,326	0,492	0,371	0,879	0,815	0,722	0,38	0,599
ARS_BFGL_BAC_2063	2	414	0,452	0,43	0,338	0,908	0,831	0,742	0,417	0,639
ARS_BFGL_NGS_1310	2	419	0,532	0,487	0,368	0,882	0,816	0,723	0,382	0,603
ARS_BFGL_NGS_2153	2	419	0,53	0,5	0,374	0,876	0,813	0,719	0,376	0,594
ARS_BFGL_NGS_2217	2	417	0,369	0,38	0,307	0,928	0,846	0,76	0,457	0,675
ARS_BFGL_NGS_233	2	360	0,161	0,206	0,185	0,979	0,908	0,84	0,652	0,81
ARS_BFGL_NGS_2983	2	419	0,544	0,473	0,361	0,889	0,82	0,728	0,39	0,611
ARS_BFGL_NGS_3405	2	417	0,525	0,452	0,35	0,898	0,825	0,734	0,403	0,625
ARS_BFGL_NGS_6765	2	416	0,505	0,436	0,34	0,905	0,83	0,74	0,414	0,636
ARS_BFGL_NGS_6819	2	417	0,441	0,412	0,327	0,915	0,836	0,748	0,431	0,652
ARS_BFGL_NGS_7055	2	381	0,383	0,317	0,266	0,95	0,867	0,785	0,518	0,721
ARS_BFGL_NGS_7838	2	419	0,604	0,493	0,371	0,879	0,814	0,721	0,379	0,598
ARS_BFGL_NGS_8390	2	381	0,491	0,454	0,351	0,897	0,825	0,734	0,401	0,624
ARS_BFGL_NGS_8418	2	398	0,374	0,349	0,288	0,939	0,856	0,772	0,485	0,697
ARS_BFGL_NGS_8461	2	419	0,508	0,487	0,368	0,881	0,816	0,723	0,382	0,602
ARS_BFGL_NGS_9102	2	417	0,434	0,413	0,328	0,915	0,836	0,747	0,43	0,651
ARS_USMARC_11	2	419	0,413	0,37	0,301	0,932	0,849	0,763	0,466	0,682
AY77615	2	419	0,554	0,487	0,368	0,882	0,816	0,723	0,382	0,602
AY84247	2	419	0,489	0,483	0,366	0,884	0,817	0,724	0,384	0,605
AY84496	2	416	0,587	0,499	0,374	0,876	0,813	0,719	0,376	0,595
BCL2	2	411	0,109	0,108	0,102	0,994	0,949	0,906	0,802	0,897
BCL2	2	417	0,007	0,007	0,007	1	0,996	0,993	0,986	0,993
BIG1	2	418	0,531	0,499	0,374	0,876	0,813	0,719	0,376	0,595
BOLA	2	415	0,4	0,32	0,269	0,949	0,866	0,783	0,514	0,718
BOLA	2	412	0,597	0,498	0,373	0,877	0,813	0,72	0,377	0,596
BOLA	2	387	0,92	0,497	0,373	0,877	0,813	0,72	0,377	0,596
BOLA_RALY	2	408	0,615	0,501	0,375	0,875	0,813	0,719	0,375	0,594
BTA6_0272	2	417	0,271	0,255	0,222	0,968	0,889	0,814	0,588	0,77
BTA6_10792	2	389	0,486	0,43	0,337	0,908	0,831	0,742	0,418	0,64
BTA6_10792	2	418	0,567	0,473	0,361	0,889	0,82	0,728	0,39	0,612
BTA6_7705	2	419	0,611	0,489	0,369	0,881	0,815	0,722	0,381	0,601
BTA6_7705	2	403	0,556	0,465	0,357	0,892	0,822	0,73	0,394	0,616
BTA6_7706	2	400	0,215	0,245	0,215	0,97	0,893	0,819	0,601	0,778
BTA6_7706	2	419	0,606	0,499	0,374	0,876	0,813	0,719	0,376	0,595
BTB_0034683	2	389	0,458	0,461	0,355	0,894	0,823	0,731	0,397	0,619
BTB_0157696	2	418	0,572	0,499	0,374	0,876	0,813	0,719	0,376	0,595
BTB_0173944	2	368	0,416	0,499	0,374	0,876	0,813	0,719	0,376	0,595
BUT	2	406	0,483	0,485	0,367	0,883	0,816	0,724	0,383	0,604
C4ORF7_60	2	375	0,112	0,143	0,133	0,99	0,934	0,881	0,745	0,865
C4ORF7_91	2	417	0,59	0,491	0,37	0,88	0,815	0,722	0,38	0,6
CARD1	2	380	0,584	0,476	0,362	0,887	0,819	0,726	0,388	0,609
CARD1	2	402	0,52	0,485	0,367	0,883	0,816	0,724	0,383	0,604
CARD1	2	413	0,523	0,472	0,36	0,889	0,82	0,728	0,391	0,612
CCL2	2	418	0,533	0,457	0,352	0,896	0,824	0,733	0,399	0,622
CCR2	2	416	0,45	0,416	0,329	0,914	0,835	0,746	0,428	0,649
CD18	2	419	0,017	0,017	0,016	1	0,992	0,984	0,967	0,984

COXP1	2	415	0,255	0,281	0,241	0,961	0,879	0,801	0,557	0,749
CRH_TRIM5	2	418	0,548	0,47	0,359	0,89	0,82	0,728	0,392	0,613
CSN2	2	394	0,378	0,326	0,272	0,947	0,864	0,781	0,508	0,715
CSN2	2	323	0,183	0,181	0,165	0,984	0,918	0,856	0,687	0,831
CSN2	2	418	0,005	0,005	0,005	1	0,998	0,995	0,99	0,995
CSN2	2	351	0,006	0,006	0,006	1	0,997	0,994	0,989	0,994
CSN3	1	418	0	0	0	1	1	1	1	1
CSN3	2	380	0,668	0,497	0,373	0,877	0,814	0,72	0,377	0,596
CSN3	2	414	0,621	0,455	0,351	0,897	0,824	0,733	0,401	0,623
CSN3	2	419	0,518	0,466	0,357	0,892	0,821	0,73	0,394	0,616
CSN3	2	419	0,048	0,047	0,046	0,999	0,977	0,956	0,91	0,954
CSN1S1	2	418	0,132	0,123	0,115	0,992	0,942	0,895	0,777	0,883
CSN1S1_BMC_579	2	419	0,138	0,157	0,145	0,988	0,928	0,871	0,723	0,852
CSN1S2_BMC_1021	2	405	0,474	0,394	0,316	0,923	0,842	0,754	0,445	0,665
CSN1S2_BMC_1043	2	413	0,567	0,479	0,364	0,885	0,818	0,725	0,386	0,607
CSN1S2_BMC_1125	2	404	0,49	0,397	0,318	0,921	0,841	0,753	0,443	0,663
CSN1S2_BMC_1803	2	407	0,047	0,046	0,045	0,999	0,978	0,957	0,912	0,955
CSN1S2_BMC_191	2	375	0,472	0,4	0,32	0,92	0,84	0,752	0,44	0,66
CSN1S2_BMC_2196	1	419	0	0	0	1	1	1	1	1
CSN1S2_BMC_254	2	408	0,086	0,082	0,079	0,997	0,961	0,926	0,846	0,92
CSN1S2_BMC_510	2	419	0,031	0,031	0,03	1	0,985	0,971	0,94	0,97
CSN2_11	2	374	0,088	0,084	0,081	0,996	0,96	0,924	0,842	0,918
CSN2_BMC_236	2	391	0,504	0,447	0,347	0,9	0,827	0,736	0,406	0,628
CSN2_BMC_26	2	417	0,285	0,265	0,23	0,965	0,885	0,809	0,576	0,761
CSN2_BMC_34	2	419	0,265	0,254	0,222	0,968	0,889	0,814	0,589	0,77
CSN2_BMC_439	2	418	0,031	0,031	0,03	1	0,985	0,971	0,94	0,97
CSN2_BMC_577	2	419	0,563	0,497	0,373	0,877	0,813	0,72	0,377	0,596
CSN2_BMC_664	2	419	0,535	0,46	0,354	0,894	0,823	0,732	0,398	0,62
CSN2_BMC_947	2	412	0,558	0,5	0,375	0,875	0,813	0,719	0,375	0,594
CSN3_BMC_1141	2	367	0,365	0,309	0,261	0,952	0,869	0,788	0,526	0,727
CSN3_BMC_1285	2	357	0,443	0,409	0,325	0,917	0,838	0,749	0,434	0,654
CSN3_BMC_1391	2	395	0,6	0,493	0,371	0,879	0,815	0,721	0,379	0,599
CSN3_BMC_1462	2	378	0,675	0,498	0,374	0,876	0,813	0,719	0,376	0,595
CSN3_BMC_1477	2	418	0,249	0,229	0,202	0,974	0,899	0,828	0,621	0,791
CSN3_BMC_1482	2	413	0,186	0,215	0,192	0,977	0,904	0,836	0,64	0,803
CSN3_BMC_192	2	411	0,574	0,473	0,361	0,889	0,82	0,728	0,39	0,611
CSNK2A	2	411	0,002	0,002	0,002	1	0,999	0,998	0,995	0,998
CVM	2	419	0,017	0,017	0,016	1	0,992	0,984	0,967	0,984
CXCL1	2	399	0,298	0,258	0,224	0,967	0,888	0,813	0,585	0,768
DCT	2	417	0,014	0,014	0,014	1	0,993	0,986	0,972	0,986
DCT	2	408	0,034	0,034	0,033	0,999	0,983	0,968	0,934	0,967
DGAT	2	357	0,521	0,5	0,375	0,875	0,813	0,719	0,375	0,594
DGAT	2	365	0,071	0,074	0,071	0,997	0,964	0,933	0,861	0,928
DGAT	2	417	0,014	0,014	0,014	1	0,993	0,986	0,972	0,986
DGAT	2	419	0,289	0,277	0,238	0,962	0,881	0,803	0,561	0,752
DQ38115	2	419	0,539	0,499	0,374	0,876	0,813	0,719	0,376	0,595

DQ40415	2	393	0,483	0,496	0,373	0,877	0,814	0,72	0,377	0,597
DQ42294	2	419	0,458	0,441	0,343	0,903	0,828	0,738	0,41	0,633
DQ48541	2	417	0,48	0,454	0,351	0,897	0,825	0,734	0,401	0,624
DQ50095	2	419	0,563	0,459	0,353	0,895	0,823	0,732	0,398	0,62
DQ64718	2	418	0,481	0,451	0,349	0,898	0,825	0,734	0,403	0,625
DQ64718	2	418	0,414	0,38	0,308	0,928	0,846	0,759	0,457	0,674
DQ65063	2	405	0,499	0,456	0,352	0,896	0,824	0,733	0,4	0,622
DQ78676	2	417	0,535	0,501	0,375	0,875	0,813	0,719	0,375	0,594
DQ78676	2	419	0,492	0,481	0,365	0,885	0,818	0,725	0,385	0,606
DQ78676	2	419	0,37	0,342	0,283	0,942	0,858	0,774	0,492	0,702
DQ83764	2	418	0,474	0,434	0,34	0,906	0,83	0,74	0,415	0,637
DQ84669	2	382	0,529	0,496	0,373	0,877	0,814	0,72	0,378	0,597
DQ86681	2	419	0,527	0,478	0,363	0,886	0,818	0,726	0,387	0,608
DQ86681	2	416	0,486	0,428	0,336	0,908	0,832	0,742	0,419	0,641
DQ88831	2	419	0,427	0,391	0,314	0,924	0,843	0,755	0,448	0,667
DQ88831	2	405	0,477	0,442	0,344	0,902	0,828	0,737	0,409	0,631
DQ98482	2	417	0,59	0,495	0,372	0,878	0,814	0,721	0,378	0,597
DQ99083	2	419	0,363	0,356	0,293	0,937	0,854	0,769	0,478	0,692
DQ99597	2	411	0,54	0,501	0,375	0,875	0,813	0,719	0,375	0,594
DUMPS	2	415	0,005	0,005	0,005	1	0,998	0,995	0,99	0,995
EF02608	2	415	0,506	0,439	0,343	0,904	0,829	0,738	0,411	0,633
EF02608	2	402	0,493	0,482	0,365	0,884	0,817	0,725	0,385	0,606
EF02807	2	402	0,505	0,466	0,357	0,891	0,821	0,73	0,394	0,616
EF03408	2	419	0,537	0,496	0,372	0,877	0,814	0,72	0,378	0,597
EF03408	2	419	0,537	0,5	0,375	0,875	0,813	0,719	0,375	0,594
EF03408	2	419	0,496	0,463	0,355	0,893	0,822	0,731	0,396	0,618
EF04209	2	375	0,491	0,417	0,33	0,913	0,835	0,746	0,427	0,648
EF09351	2	416	0,567	0,474	0,361	0,888	0,819	0,727	0,389	0,611
EF14110	2	412	0,573	0,495	0,372	0,878	0,814	0,72	0,378	0,597
EF16480	2	419	0,518	0,485	0,367	0,883	0,816	0,724	0,383	0,604
FASN	2	341	0,516	0,415	0,329	0,914	0,836	0,747	0,429	0,65
FGF2	2	418	0,266	0,258	0,224	0,967	0,888	0,812	0,584	0,767
FNTA	2	419	0,339	0,309	0,261	0,952	0,869	0,788	0,525	0,727
GHR	1	418	0	0	0	1	1	1	1	1
GHRH	2	326	0,215	0,211	0,188	0,978	0,906	0,838	0,645	0,806
GHRH	2	418	0,512	0,448	0,347	0,9	0,826	0,736	0,406	0,628
Hapmap23359_BTA_14372	2	411	0,482	0,499	0,374	0,876	0,813	0,719	0,376	0,595
Hapmap23813_BTA_12845	2	411	0,353	0,321	0,269	0,949	0,865	0,783	0,513	0,718
Hapmap25366_BTA_6328	2	383	0,29	0,256	0,223	0,967	0,889	0,814	0,587	0,769
Hapmap27208_BTA_15750	2	418	0,426	0,441	0,344	0,903	0,828	0,738	0,41	0,632
Hapmap27732_BTA_12300	2	419	0,265	0,244	0,214	0,97	0,893	0,82	0,602	0,779
Hapmap29197_BTA_9714	2	418	0,555	0,498	0,374	0,876	0,813	0,72	0,376	0,596
Hapmap30796_BTA_13432	2	418	0,373	0,307	0,26	0,953	0,87	0,789	0,528	0,729
Hapmap31189_BTA_16111	2	419	0,332	0,299	0,254	0,955	0,873	0,793	0,537	0,735
Hapmap34164_BES10_Contig538_100	2	418	0,51	0,476	0,362	0,887	0,819	0,727	0,388	0,609
Hapmap35453_SCAFFOLD669_1513	2	398	0,548	0,474	0,361	0,888	0,819	0,727	0,389	0,611

Hapmap44212_BTA_12209	2	362	0,591	0,49	0,37	0,88	0,815	0,722	0,38	0,6
Hapmap46428_BTA_5450	2	397	0,499	0,5	0,375	0,875	0,813	0,719	0,375	0,594
Hapmap50821_BTA_9175	2	403	0,278	0,274	0,236	0,962	0,882	0,804	0,565	0,754
Hapmap51475_BTA_6217	2	414	0,449	0,416	0,329	0,914	0,835	0,746	0,428	0,649
Hapmap51741_BTA_5375	2	417	0,549	0,499	0,374	0,876	0,813	0,719	0,376	0,595
Hapmap53185_rs2901029	2	419	0,561	0,49	0,369	0,88	0,815	0,722	0,381	0,601
hypothetical_LOC53901	2	403	0,598	0,5	0,375	0,875	0,813	0,719	0,375	0,594
hypothetical_LOC78303	2	411	0,37	0,374	0,304	0,93	0,848	0,762	0,463	0,679
IGF1	2	418	0,258	0,26	0,226	0,966	0,887	0,812	0,582	0,766
IGF1	2	370	0,641	0,486	0,368	0,882	0,816	0,723	0,383	0,603
IGF1	2	390	0,562	0,457	0,352	0,896	0,824	0,733	0,4	0,622
IL8	2	419	0,513	0,489	0,369	0,881	0,815	0,722	0,381	0,601
IL8	2	399	0,659	0,499	0,374	0,876	0,813	0,719	0,376	0,595
IL8	1	417	0	0	0	1	1	1	1	1
ITGB2	2	402	0,57	0,458	0,353	0,895	0,824	0,732	0,399	0,621
KCBN2	2	419	0,468	0,421	0,332	0,912	0,834	0,745	0,424	0,646
KCBN2	2	406	0,468	0,425	0,334	0,91	0,833	0,743	0,421	0,643
KCBN2	2	410	0,622	0,5	0,374	0,876	0,813	0,719	0,376	0,594
KCBN2	2	418	0,498	0,468	0,358	0,891	0,821	0,729	0,393	0,614
KCNB2	2	402	0,682	0,495	0,372	0,878	0,814	0,72	0,378	0,597
KIT	2	393	0,631	0,488	0,368	0,881	0,816	0,723	0,382	0,602
KIT	2	416	0,012	0,017	0,017	1	0,992	0,984	0,967	0,983
KIT	2	419	0,014	0,014	0,014	1	0,993	0,986	0,972	0,986
KIEL	2	416	0,005	0,005	0,005	1	0,998	0,995	0,99	0,995
LALBA	2	419	0,332	0,317	0,266	0,95	0,867	0,785	0,518	0,721
LEP	2	405	0,56	0,465	0,357	0,892	0,822	0,73	0,395	0,616
LGB	2	405	0,007	0,007	0,007	1	0,996	0,993	0,985	0,993
LGB	1	415	0	0	0	1	1	1	1	1
LHB	2	393	0,071	0,069	0,066	0,998	0,967	0,937	0,87	0,933
LHB	2	419	0,005	0,005	0,005	1	0,998	0,995	0,991	0,995
LOC51267	2	393	0,529	0,469	0,359	0,89	0,821	0,729	0,392	0,614
LOC51267	2	417	0,082	0,087	0,083	0,996	0,958	0,922	0,837	0,916
LOC51282	2	411	0,586	0,469	0,358	0,89	0,821	0,729	0,393	0,614
LOC51558	2	414	0,44	0,406	0,323	0,918	0,838	0,75	0,436	0,656
LOC51970	2	418	0,213	0,194	0,175	0,981	0,912	0,848	0,668	0,82
LOC52955	2	392	0,383	0,357	0,293	0,936	0,854	0,768	0,478	0,691
LOC53681	2	418	0,522	0,492	0,371	0,879	0,815	0,721	0,379	0,599
LOC53681	2	419	0,191	0,196	0,176	0,981	0,912	0,847	0,666	0,819
LOC61713	2	416	0,226	0,24	0,211	0,971	0,894	0,822	0,606	0,782
LOC783036_LOC79086	2	414	0,109	0,103	0,098	0,995	0,951	0,91	0,81	0,901
LOC783036_LOC79086	2	419	0,568	0,483	0,366	0,884	0,817	0,724	0,384	0,605
LOC7847	2	404	0,498	0,405	0,322	0,918	0,839	0,751	0,437	0,657
LTF	2	418	0,385	0,34	0,282	0,942	0,859	0,775	0,494	0,704
LTF	2	416	0,596	0,501	0,375	0,875	0,813	0,719	0,375	0,594
LTF	2	415	0,337	0,312	0,263	0,952	0,869	0,787	0,523	0,725
MC1R	2	414	0,609	0,465	0,357	0,892	0,822	0,73	0,395	0,616

MC1R	2	414	0,007	0,007	0,007	1	0,996	0,993	0,986	0,993
MC1R	2	361	0,363	0,297	0,253	0,956	0,874	0,794	0,538	0,736
MGC12853	2	418	0,505	0,407	0,324	0,917	0,838	0,75	0,435	0,656
neaby_MCM	2	407	0,184	0,172	0,157	0,985	0,922	0,862	0,701	0,84
nearby_LOC78243	2	413	0,172	0,161	0,148	0,987	0,926	0,868	0,717	0,849
nearby_PYCR	2	418	0,404	0,366	0,299	0,933	0,851	0,765	0,47	0,685
no_hits	2	418	0,337	0,314	0,265	0,951	0,868	0,786	0,52	0,723
no_hits	2	413	0,094	0,094	0,09	0,996	0,955	0,916	0,825	0,909
no_hits	2	419	0,539	0,463	0,355	0,893	0,822	0,731	0,396	0,618
no_hits	2	418	0,019	0,019	0,019	1	0,991	0,981	0,963	0,981
no_hits	2	402	0,134	0,138	0,129	0,99	0,936	0,884	0,752	0,869
no_hits	2	410	0,088	0,084	0,08	0,996	0,96	0,924	0,843	0,919
no_hits	2	417	0,499	0,471	0,36	0,889	0,82	0,728	0,391	0,612
no_hits	2	417	0,182	0,17	0,155	0,986	0,922	0,863	0,704	0,841
no_hits	2	417	0,444	0,405	0,323	0,918	0,839	0,75	0,436	0,657
no_hits	2	419	0,243	0,242	0,213	0,971	0,894	0,821	0,604	0,78
no_hits	2	413	0,581	0,474	0,361	0,888	0,819	0,727	0,389	0,611
no_hits	2	417	0,429	0,409	0,325	0,916	0,837	0,749	0,433	0,654
NOS	2	398	0,216	0,197	0,177	0,981	0,911	0,846	0,665	0,818
ODC	2	419	0,115	0,108	0,102	0,994	0,949	0,906	0,802	0,896
OP	2	419	0,501	0,463	0,355	0,893	0,822	0,731	0,396	0,618
PI	2	413	0,257	0,45	0,349	0,899	0,826	0,735	0,404	0,626
PI	2	419	0,088	0,089	0,085	0,996	0,958	0,921	0,834	0,914
P51_23	2	402	0,5	0,47	0,359	0,89	0,82	0,728	0,392	0,613
PKD2_101	2	394	0,429	0,462	0,355	0,893	0,822	0,731	0,396	0,618
PKD2_124	2	410	0,01	0,024	0,024	1	0,988	0,977	0,953	0,976
PKD2_134	2	417	0,019	0,019	0,019	1	0,991	0,981	0,963	0,981
PKD2_275	2	419	0,01	0,01	0,009	1	0,995	0,991	0,981	0,991
PKD2_35	2	408	0,529	0,48	0,365	0,885	0,818	0,725	0,386	0,607
PKD2_37	2	369	0,154	0,191	0,173	0,982	0,914	0,849	0,672	0,823
PKD2_38	2	388	0,209	0,187	0,169	0,983	0,915	0,852	0,678	0,826
PKD2_390	2	413	0,518	0,417	0,33	0,913	0,835	0,746	0,427	0,649
PKD2_59	2	388	0,039	0,038	0,037	0,999	0,981	0,964	0,926	0,963
PKD2_65	2	419	0,084	0,085	0,081	0,996	0,96	0,924	0,842	0,918
PKD2_9714	2	419	0,172	0,165	0,151	0,986	0,924	0,866	0,711	0,845
PKIA	2	413	0,552	0,48	0,365	0,885	0,818	0,725	0,386	0,607
PLAT	2	415	0,578	0,434	0,34	0,906	0,83	0,74	0,415	0,637
POU1F	2	416	0,313	0,293	0,25	0,957	0,875	0,796	0,543	0,74
PPARGC1	2	388	0,369	0,338	0,281	0,943	0,86	0,776	0,496	0,705
PPARGC1	2	414	0,027	0,026	0,026	1	0,987	0,975	0,949	0,974
PPARGC1	2	419	0,437	0,38	0,307	0,928	0,846	0,76	0,457	0,675
PPARGC1	2	403	0,33	0,292	0,249	0,957	0,875	0,796	0,544	0,74
PRL	2	408	0,444	0,404	0,322	0,918	0,839	0,751	0,437	0,657
PRL	2	417	0,58	0,499	0,374	0,876	0,813	0,719	0,376	0,595
PRL	2	403	0,422	0,388	0,313	0,925	0,844	0,756	0,45	0,669
SDHA	2	378	0,23	0,22	0,196	0,976	0,902	0,833	0,633	0,798

STAT1	2	417	0,42	0,374	0,304	0,93	0,848	0,762	0,463	0,679
STAT5	1	419	0	0	0	1	1	1	1	1
STATH_175	2	398	0,425	0,368	0,3	0,933	0,85	0,764	0,468	0,683
TF	2	415	0,508	0,444	0,345	0,902	0,828	0,737	0,408	0,631
TLR4	1	419	0	0	0	1	1	1	1	1
TLR4	2	418	0,01	0,01	0,009	1	0,995	0,991	0,981	0,991
TLR4	2	416	0,087	0,083	0,079	0,997	0,96	0,925	0,845	0,92
TLR4	2	405	0,077	0,078	0,075	0,997	0,962	0,929	0,853	0,924
TLR4	2	374	0,003	0,003	0,003	1	0,999	0,997	0,995	0,997
TLR4	2	348	0,279	0,284	0,243	0,96	0,878	0,8	0,554	0,747
TNF	2	414	0,749	0,469	0,359	0,89	0,821	0,729	0,392	0,614
TYR	2	375	0,035	0,034	0,033	0,999	0,983	0,967	0,934	0,966
TYR	4	408	1	0,638	0,567	0,784	0,636	0,474	0,202	0,482
TYRP1	2	392	0,515	0,491	0,37	0,88	0,815	0,722	0,38	0,6
TYRP1	2	418	0,266	0,234	0,206	0,973	0,897	0,825	0,614	0,787
TYRP1	2	396	0,556	0,494	0,372	0,878	0,814	0,721	0,378	0,598
UA_IFASA_229	2	417	0,561	0,491	0,37	0,88	0,815	0,722	0,38	0,6
Total		1,97*	407*	0,370*	0,329*	0,260*	0,000**	0,000**	0,000**	0,000**

\*Valor Médio

\*\*Probabilidade combinada de não exclusão

Deste painel com 249 SNPs, foram extraídos 95 SNPs que apresentaram estimativas de Hardy-Weinberg e MAF acima de 0,001 selecionados para o conjunto de animais testados. Para o conteúdo de informação polimórfica (PIC) foram obtidos valores médios de 0,292, e para a heterozigosidade observada ( $H_o$ ) e esperada ( $H_e$ ), os valores de 0,395 e 0,368, respectivamente. Assim como o painel anterior (Tabela 2.6), as estimativas das probabilidades combinadas de não exclusão (Tabela 2.7) foram abaixo de 0,00% para todos os parâmetros adotados, o que representa uma estimativa acima de 99% para a confirmação de possível paternidade, assim como a identificação individual.

**Tabela 2.7** Estimativas de diversidade independentes para 95 SNPs mais distantes que apresentaram estimativas de Hardy-Weinberg e MAF acima de 0,001 selecionados para o conjunto de animais testados (N); K=número de alelos observados;  $H_{Obs}$ =heterozigosidade observada;  $H_{Exp}$ =heterozigosidade esperada; PIC=conteúdo de informação polimórfica; NE-1P=probabilidade média de não-exclusão para um pai candidato; NE-2P=probabilidade média de não-exclusão para um pai candidato dado o genótipo de um pai conhecido do sexo oposto; NE-I=probabilidade média de não-exclusão para a identidade de dois indivíduos não relacionados; NE-SI=probabilidade média de não-exclusão para a identidade de dois irmãos.



SNP	k	N	HObs	HExp	PIC	NE-1P	NE-2P	NE-PP	NE-I	NE-SI
ABCG2	2	401	0,202	0,182	0,165	0,984	0,917	0,855	0,686	0,831
ARS_BFGL_NGS_1310	2	419	0,532	0,487	0,368	0,882	0,816	0,723	0,382	0,603
ARS_BFGL_NGS_2217	2	417	0,369	0,38	0,307	0,928	0,846	0,76	0,457	0,675
ARS_BFGL_NGS_233	2	360	0,161	0,206	0,185	0,979	0,908	0,84	0,652	0,81
ARS_BFGL_NGS_6819	2	417	0,441	0,412	0,327	0,915	0,836	0,748	0,431	0,652
ARS_BFGL_NGS_8461	2	419	0,508	0,487	0,368	0,881	0,816	0,723	0,382	0,602
ARS_BFGL_NGS_9102	2	417	0,434	0,413	0,328	0,915	0,836	0,747	0,43	0,651
ARS_USMARC_11	2	419	0,413	0,37	0,301	0,932	0,849	0,763	0,466	0,682
AY84247	2	419	0,489	0,483	0,366	0,884	0,817	0,724	0,384	0,605
BOLA	2	415	0,4	0,32	0,269	0,949	0,866	0,783	0,514	0,718
BOLA_RALY	2	408	0,615	0,501	0,375	0,875	0,813	0,719	0,375	0,594
BTA6_0272	2	417	0,271	0,255	0,222	0,968	0,889	0,814	0,588	0,77
BTA6_7706	2	419	0,606	0,499	0,374	0,876	0,813	0,719	0,376	0,595
BTB_0034683	2	389	0,458	0,461	0,355	0,894	0,823	0,731	0,397	0,619
BUT	2	406	0,483	0,485	0,367	0,883	0,816	0,724	0,383	0,604
CARD1	2	380	0,584	0,476	0,362	0,887	0,819	0,726	0,388	0,609
CCL2	2	418	0,533	0,457	0,352	0,896	0,824	0,733	0,399	0,622
COXP1	2	415	0,255	0,281	0,241	0,961	0,879	0,801	0,557	0,749
CSN1S1_BMC_579	2	323	0,183	0,181	0,165	0,984	0,918	0,856	0,687	0,831
CSN1S2_BMC_1043	2	419	0,518	0,466	0,357	0,892	0,821	0,73	0,394	0,616
CSN1S2_BMC_1125	2	418	0,132	0,123	0,115	0,992	0,942	0,895	0,777	0,883
CSN2_BMC_26	2	419	0,138	0,157	0,145	0,988	0,928	0,871	0,723	0,852
CSN2_BMC_34	2	417	0,285	0,265	0,23	0,965	0,885	0,809	0,576	0,761
CSN2_BMC_633	2	419	0,265	0,254	0,222	0,968	0,889	0,814	0,589	0,77
CSN2_BMC_947	2	419	0,535	0,46	0,354	0,894	0,823	0,732	0,398	0,62
CSN2	2	412	0,558	0,5	0,375	0,875	0,813	0,719	0,375	0,594
CSN3_BMC_1285	2	357	0,443	0,409	0,325	0,917	0,838	0,749	0,434	0,654
CSN3_BMC_1477	2	418	0,249	0,229	0,202	0,974	0,899	0,828	0,621	0,791
CSN3_BMC_1482	2	413	0,186	0,215	0,192	0,977	0,904	0,836	0,64	0,803
DGATI	2	357	0,521	0,5	0,375	0,875	0,813	0,719	0,375	0,594
DQ38115	2	419	0,539	0,499	0,374	0,876	0,813	0,719	0,376	0,595
DQ40415	2	393	0,483	0,496	0,373	0,877	0,814	0,72	0,377	0,597
DQ42294	2	419	0,458	0,441	0,343	0,903	0,828	0,738	0,41	0,633
DQ48541	2	417	0,48	0,454	0,351	0,897	0,825	0,734	0,401	0,624
DQ64718	2	418	0,481	0,451	0,349	0,898	0,825	0,734	0,403	0,625
DQ64718	2	418	0,414	0,38	0,308	0,928	0,846	0,759	0,457	0,674
DQ65063	2	405	0,499	0,456	0,352	0,896	0,824	0,733	0,4	0,622
DQ78676	2	417	0,535	0,501	0,375	0,875	0,813	0,719	0,375	0,594
DQ78676	2	419	0,492	0,481	0,365	0,885	0,818	0,725	0,385	0,606
DQ83764	2	418	0,474	0,434	0,34	0,906	0,83	0,74	0,415	0,637
DQ84669	2	382	0,529	0,496	0,373	0,877	0,814	0,72	0,378	0,597
DQ86681	2	419	0,527	0,478	0,363	0,886	0,818	0,726	0,387	0,608
DQ88831	2	419	0,427	0,391	0,314	0,924	0,843	0,755	0,448	0,667
DQ99597	2	411	0,54	0,501	0,375	0,875	0,813	0,719	0,375	0,594
EF03408	2	419	0,537	0,496	0,372	0,877	0,814	0,72	0,378	0,597

EF04209	2	375	0,491	0,417	0,33	0,913	0,835	0,746	0,427	0,648
EF09351	2	416	0,567	0,474	0,361	0,888	0,819	0,727	0,389	0,611
EF16480	2	419	0,518	0,485	0,367	0,883	0,816	0,724	0,383	0,604
FGF2	2	418	0,266	0,258	0,224	0,967	0,888	0,812	0,584	0,767
FNTA	2	419	0,339	0,309	0,261	0,952	0,869	0,788	0,525	0,727
GHRH	2	418	0,512	0,448	0,347	0,9	0,826	0,736	0,406	0,628
Hapmap23359_BTA_14372	2	411	0,482	0,499	0,374	0,876	0,813	0,719	0,376	0,595
Hapmap23813_BTA_12845	2	411	0,353	0,321	0,269	0,949	0,865	0,783	0,513	0,718
Hapmap27208_BTA_15750	2	418	0,426	0,441	0,344	0,903	0,828	0,738	0,41	0,632
Hapmap27732_BTA_12300	2	419	0,265	0,244	0,214	0,97	0,893	0,82	0,602	0,779
Hapmap34164_BES10_Contig538_100	2	418	0,51	0,476	0,362	0,887	0,819	0,727	0,388	0,609
Hapmap35453_SCAFFOLD669_1513	2	398	0,548	0,474	0,361	0,888	0,819	0,727	0,389	0,611
Hapmap46428_BTA_5450	2	397	0,499	0,5	0,375	0,875	0,813	0,719	0,375	0,594
Hapmap50821_BTA_9175	2	403	0,278	0,274	0,236	0,962	0,882	0,804	0,565	0,754
Hapmap51475_BTA_6217	2	414	0,449	0,416	0,329	0,914	0,835	0,746	0,428	0,649
Hapmap51741_BTA_5375	2	417	0,549	0,499	0,374	0,876	0,813	0,719	0,376	0,595
IGF1	2	418	0,258	0,26	0,226	0,966	0,887	0,812	0,582	0,766
IGF1	2	370	0,641	0,486	0,368	0,882	0,816	0,723	0,383	0,603
IL8	2	419	0,513	0,489	0,369	0,881	0,815	0,722	0,381	0,601
KCBN2	2	419	0,468	0,421	0,332	0,912	0,834	0,745	0,424	0,646
LALBA	2	419	0,332	0,317	0,266	0,95	0,867	0,785	0,518	0,721
LHB	2	419	0,005	0,005	0,005	1,000	0,998	0,995	0,991	0,995
LOC51267	2	393	0,529	0,469	0,359	0,89	0,821	0,729	0,392	0,614
LOC51558	2	414	0,44	0,406	0,323	0,918	0,838	0,75	0,436	0,656
LOC52955	2	392	0,383	0,357	0,293	0,936	0,854	0,768	0,478	0,691
LOC53681	2	418	0,522	0,492	0,371	0,879	0,815	0,721	0,379	0,599
LOC536811	2	419	0,191	0,196	0,176	0,981	0,912	0,847	0,666	0,819
LOC783036_LOC79086	2	414	0,109	0,103	0,098	0,995	0,951	0,91	0,81	0,901
LOC783036_LOC790864	2	419	0,568	0,483	0,366	0,884	0,817	0,724	0,384	0,605
LTF	2	418	0,385	0,34	0,282	0,942	0,859	0,775	0,494	0,704
neaby_MCM	2	407	0,184	0,172	0,157	0,985	0,922	0,862	0,701	0,84
no_hits	2	418	0,337	0,314	0,265	0,951	0,868	0,786	0,52	0,723
no_hits	2	413	0,094	0,094	0,09	0,996	0,955	0,916	0,825	0,909
no_hits	2	419	0,539	0,463	0,355	0,893	0,822	0,731	0,396	0,618
no_hits	2	402	0,134	0,138	0,129	0,99	0,936	0,884	0,752	0,869
no_hits	2	410	0,088	0,084	0,08	0,996	0,96	0,924	0,843	0,919
no_hits	2	417	0,499	0,471	0,36	0,889	0,82	0,728	0,391	0,612
no_hits	2	417	0,182	0,17	0,155	0,986	0,922	0,863	0,704	0,841
ODC1	2	419	0,115	0,108	0,102	0,994	0,949	0,906	0,802	0,896
OPN	2	419	0,501	0,463	0,355	0,893	0,822	0,731	0,396	0,618
PI	2	413	0,257	0,45	0,349	0,899	0,826	0,735	0,404	0,626
PKD2_101	2	394	0,429	0,462	0,355	0,893	0,822	0,731	0,396	0,618
PKD2_38	2	388	0,209	0,187	0,169	0,983	0,915	0,852	0,678	0,826
PKD2_9714	2	419	0,172	0,165	0,151	0,986	0,924	0,866	0,711	0,845
POU1F	2	416	0,313	0,293	0,25	0,957	0,875	0,796	0,543	0,74

PPARGC1A	2	419	0,437	0,38	0,307	0,928	0,846	0,76	0,457	0,675
SDHA	2	378	0,23	0,22	0,196	0,976	0,902	0,833	0,633	0,798
STAT1	2	417	0,42	0,374	0,304	0,93	0,848	0,762	0,463	0,679
TF	2	415	0,508	0,444	0,345	0,902	0,828	0,737	0,408	0,631
TLR4	2	348	0,279	0,284	0,243	0,96	0,878	0,8	0,554	0,747
Total	2*	409	0,395*	0,368*	0,292*	0,000*	0,000*	0,000*	0,000*	0,000*

\*Valor Médio

\*\*Probabilidade combinada de não exclusão

Adicionalmente, foram extraídos 20 marcadores validados em estudos anteriores (Heaton et al., 2002) especificamente para testes de paternidade (Tabela 2.7). Para o conteúdo de informação polimórfica (PIC) foram obtidos valores médios de 0,347, e para a heterozigidade observada (Ho) e esperada (He), os valores de 0,486 e 0,449, respectivamente. As estimativas das probabilidades combinadas de não exclusão (Tabela 2.6) foram abaixo de 0,00% para a confirmação de paternidade, considerando o genótipo conhecido de ambos pais, e para os parâmetros relacionados com a identificação individual.

**Tabela 2.7** Estimativas de diversidade independentes para 20 SNPs validados para testes de paternidade selecionados para o conjunto de animais testados (N); K=número de alelos observado; HObs=heterozigidade observada; HExp=heterozigidade esperada; PIC=conteúdo de informação polimórfica; NE-1P=probabilidade média de não-exclusão para um pai candidato; NE-2P=probabilidade média de não-exclusão para um pai candidato dado o genótipo de um pai conhecido do sexo oposto; NE-PP= probabilidade média de não-exclusão com genótipo de ambos pais conhecidos; NE-I=probabilidade média de não-exclusão para a identidade de dois indivíduos não relacionados; NE-SI=probabilidade média de não-exclusão para a identidade de dois irmãos.

SNP	k	N	HObs	HExp	PIC	NE-1P	NE-2P	NE-PP	NE-I	NE-SI
DQ38115	2	419	0,539	0,499	0,374	0,876	0,813	0,719	0,376	0,595
DQ40415	2	393	0,483	0,496	0,373	0,877	0,814	0,72	0,377	0,597
DQ42294	2	419	0,458	0,441	0,343	0,903	0,828	0,738	0,41	0,633
DQ48541	2	417	0,48	0,454	0,351	0,897	0,825	0,734	0,401	0,624
DQ50095	2	419	0,563	0,459	0,353	0,895	0,823	0,732	0,398	0,62
DQ64718	2	418	0,481	0,451	0,349	0,898	0,825	0,734	0,403	0,625
DQ64718	2	418	0,414	0,38	0,308	0,928	0,846	0,759	0,457	0,674
DQ65063	2	405	0,499	0,456	0,352	0,896	0,824	0,733	0,4	0,622
DQ78676	2	419	0,492	0,481	0,365	0,885	0,818	0,725	0,385	0,606
DQ78676	2	419	0,37	0,342	0,283	0,942	0,858	0,774	0,492	0,702
DQ78676	2	417	0,535	0,501	0,375	0,875	0,813	0,719	0,375	0,594
DQ83764	2	418	0,474	0,434	0,34	0,906	0,83	0,74	0,415	0,637

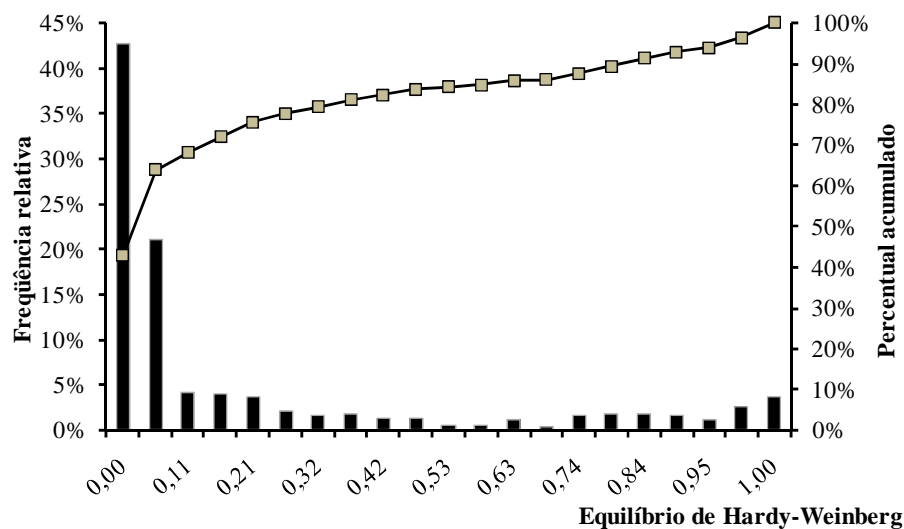
DQ84669	2	382	0,529	0,496	0,373	0,877	0,814	0,72	0,378	0,597
DQ86681	2	419	0,527	0,478	0,363	0,886	0,818	0,726	0,387	0,608
DQ86681	2	416	0,486	0,428	0,336	0,908	0,832	0,742	0,419	0,641
DQ88831	2	419	0,427	0,391	0,314	0,924	0,843	0,755	0,448	0,667
DQ88831	2	405	0,477	0,442	0,344	0,902	0,828	0,737	0,409	0,631
DQ98482	2	417	0,59	0,495	0,372	0,878	0,814	0,721	0,378	0,597
DQ99083	2	419	0,363	0,356	0,293	0,937	0,854	0,769	0,478	0,692
DQ99597	2	411	0,54	0,501	0,375	0,875	0,813	0,719	0,375	0,594
Total	2*	414*	0,486*	0,449*	0,347*	0,117	0,022	0,000	0,000	0,000

\*Valor Médio

\*\*Probabilidade combinada de não exclusão

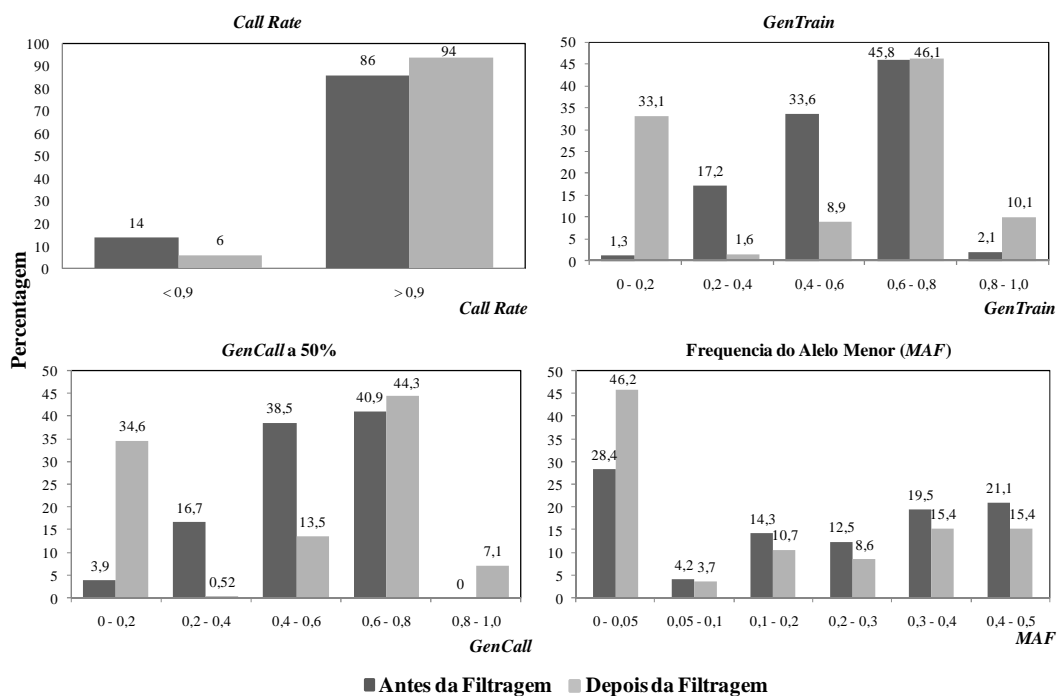
Considerando-se as 1920 transmissões mendelianas possíveis do ensaio original (5 trios x 384 SNPs), foram encontradas inicialmente 260 inconsistências concentradas em 52 SNPs (13,5%). Após as filtragens realizadas, o painel validado (249 SNPs) apresentou média de confirmação de paternidade e de reprodutibilidade de 100% (Tabelas 2.4 e 2.5).

Conforme esperado, os desvios em relação às proporções genóticas do equilíbrio de Hardy-Weinberg foram significativas ( $P < 0,01$ ), quando a população como um todo foi considerada (Figura 2.3). Esse desvio das proporções de equilíbrio poderia ser consequência da reunião de grupos genéticos diferentes com frequências alélicas diferentes em uma mesma população.



**Figura 2.3** Valor de  $P$  para o Equilíbrio de Hardy-Weinberg para o painel considerando todos os 384 SNPs.

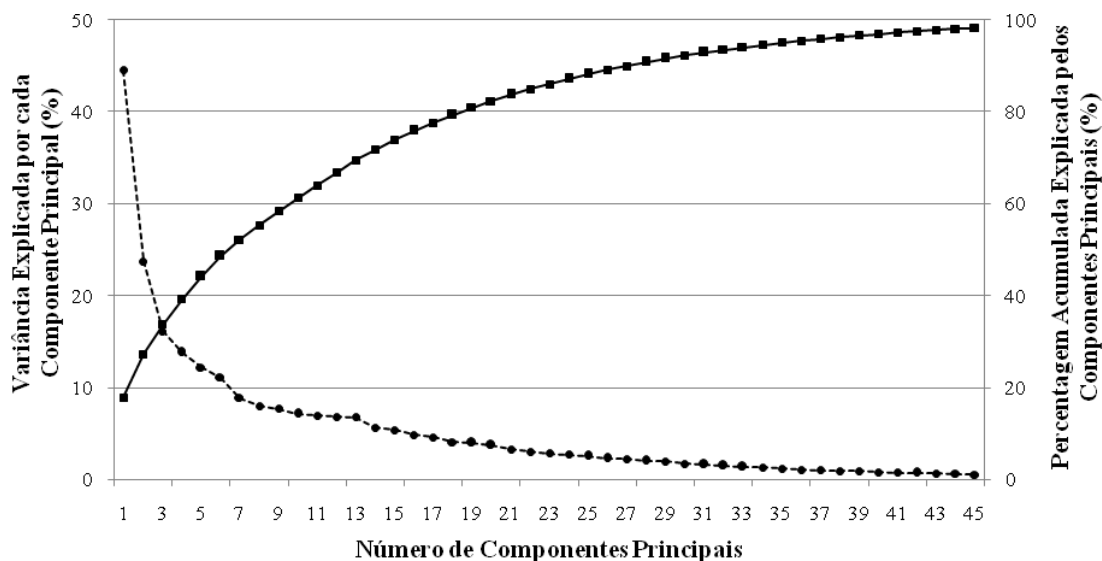
A proporção geral de SNPs polimórficos observados neste estudo (N = 576 animais) considerando todos os 384 SNPs foi estimada em 71,6%. Após aplicação dos critérios de filtragem, foi possível observar uma diminuição da proporção de SNPs polimórficos (53,4%, Figura 2.4). Animais aparentados (produtos dos cinco trios incluídos no estudo) e as amostras repetidas (66 amostras) foram excluídas dessas estimativas. Os 419 animais restantes (118 machos e 301 fêmeas) apresentaram Taxa de Genotipagem média de  $0,9699 \pm 0,017$ .



**Figura 2.4** Distribuição dos valores de Taxa de Genotipagem, *GenTrain* e *GenCall50* observados antes e depois da filtragem dos dados gerados com o painel de SNPs.

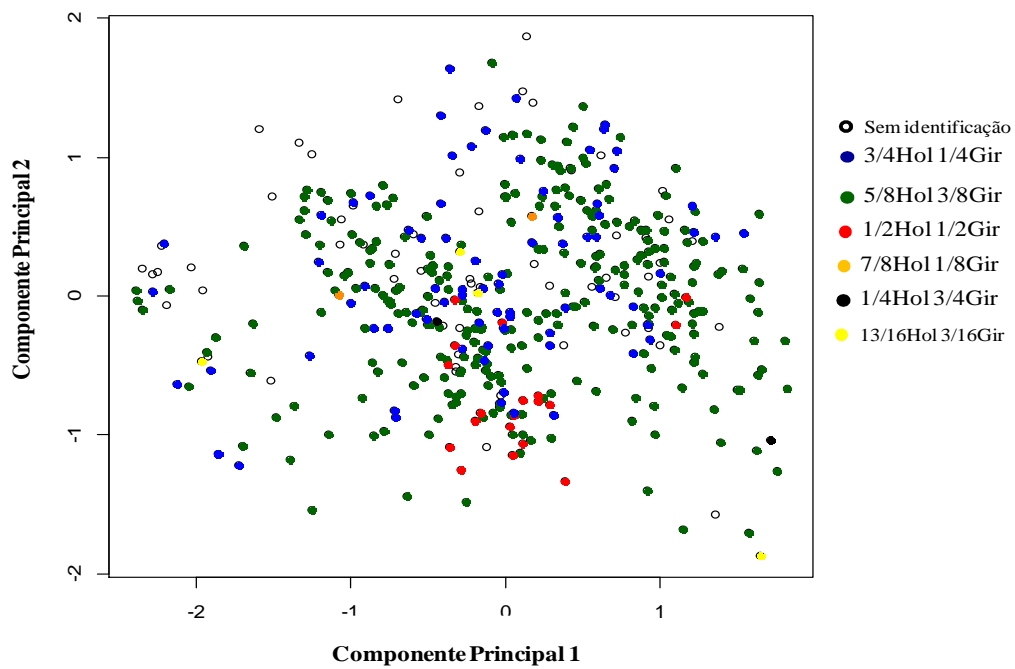
A análise de componentes principais não revelou estruturação de grupos genéticos na população estudada. O primeiro e segundo componentes principais (CPs) respondem individualmente por 45% e 24% da variação total entre os 249 SNPs examinados, e os primeiros dez CPs explicam mais de 61% da variação total (Figura 2.5). A partir do sétimo e décimo sexto componente principal, a variância explicada por cada componente foi abaixo de

10% e 5%, respectivamente, explicando, a partir desse ponto, uma quantidade cada vez menor da variação total (Figura 2.5).

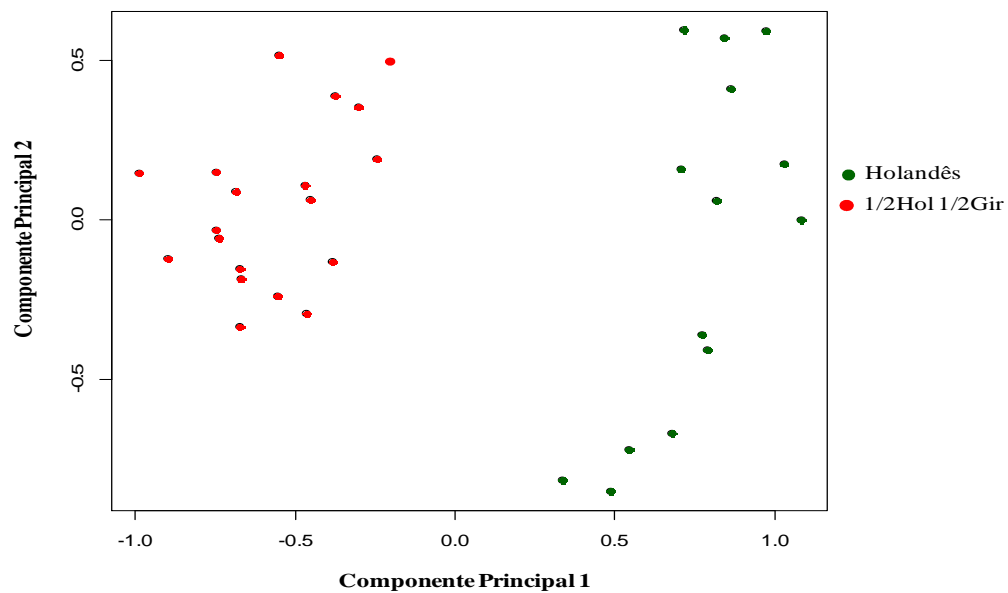


**Figura 2.5** Variância Individual e Acumulada dos Componentes Principais.

O gráfico de dispersão (Figura 2.6), foi obtido através dos escores dos componentes principais 1 e 2. Como pode ser observado, os CPs utilizados não identificaram subestruturação na população estudada, o que elimina possíveis erros nas estimativas associados ao efeito de GGs diferentes. Por outro lado, foi realizado análise de estruturação entre animais holandeses (excluídos no início das análises) de animais meio sangue ( $1/2$ Holandês+ $1/2$ Gir), para avaliar o poder de estruturação do painel validado. Como pode ser observado na Figura 2.7, o painel foi capaz de estruturar as duas populações, distinguindo os animais holandeses dos meio sangue.



**Figura 2.6** Estruturação dos grupos genéticos através da dispersão entre os escores do primeiro e segundo componentes principais.



## 2.6. DISCUSSÃO

Avanços tecnológicos obtidos na última década geraram novas metodologias para genotipagem de marcadores SNP, o que revolucionou os limites da utilização de dados moleculares nos estudos com genes candidatos de sequências reguladoras das proteínas do leite (Caetano, 2009). Conforme Kaminski et al (2005), em se tratando de rebanho bovino leiteiro em associação às suas características de interesse econômico, para uma maior exploração da variação de proteínas no teor do leite, um número maior de marcadores SNPs pode favorecer os resultados. Para estes autores, justifica-se o interesse em SNPs em decorrência de sua utilização na identificação de genes associados a doenças multifatoriais e QTL. Supõe-se que, em alelos próximos, os SNPs são herdados em conjunto com os QTLs através das gerações.

Os resultados apresentados neste capítulo comprovam a eficácia do ensaio *GoldenGate* para avaliação de genotipagem de SNPs em painéis de baixa densidade, evidenciando o amplo potencial na detecção e seleção de um grande número de marcadores SNPs. No mesmo sentido, González-Neira (2013) aponta para a viabilidade da metodologia *GoldenGate* em procedimentos de genotipagem em painéis personalizados de SNPs. Segundo a autora, esta ferramenta promove garantias na verificação de frequências dos alelos menores, bem como permite obtenção de maior probabilidade de sucesso na validação do experimento. Para esta autora, frente aos critérios de filtragem que excluem do experimento amostras e marcadores não confiáveis, é possível estabelecer painel de marcadores de SNPs de baixa densidade para verificação de critérios como reprodutibilidade e heranças genéticas. De acordo com Ziegler (2009), SNPs genotipados de forma inconsistente ou que não contribuirão



para a acurácia das avaliações genéticas devem ser excluídos no sentido de reduzir esforços computacionais, diminuir a quantidade de resultados falsos e melhorar a precisão das análises realizadas com os polimorfismos restantes.

Conforme Lin et al. (2009), é primordial a seleção de ferramentas que melhor se adaptem ao objetivo do trabalho, que ofereçam consistência no desempenho, flexibilidade em conteúdo, rapidez na geração de dados, e acessibilidade para estudos em larga escala. Para Kaminski et al. (2005), a genotipagem de SNPs em intervalos regulares, possibilita a detecção de regiões genômicas que apresentam variação nas frequências alélicas, auxiliando os processos de avaliação genética e melhoramento animal.

Também com intuito de desenvolver e validar um método rápido para a genotipagem das principais mutações das proteínas do leite bovino, Chessa et al. (2007) realizaram experimento com microchips, para a identificação simultânea de 22 SNPs localizados em 8 diferentes regiões de genes relacionados com a proteína do leite. A plataforma utilizada permitiu inserção e deleção polimorfismos para serem genotipados. Conforme os resultados deste estudo, através de *microchips*, foi constatada facilidade de genotipagem de diferentes regiões genômicas, que poderiam ser investigadas para estudos filogenéticos e de associação; avaliação das distâncias genéticas entre raças de bovinos; história da espécie bovina e identificação de parentesco.

Onteru et al. (2011) analisando características reprodutivas em porcas, utilizaram um limiar de  $MAF > 0,001$  e eficiência de genotipagem  $>80\%$  para verificar o EHW e excluir de SNP com erros de genotipagem. O *Gentrain Score*, utilizado nesta análise, baseou-se na análise de agrupamento dos genótipos, que foram realizadas por meio do *Gen Call score* da Illumina. Após este controle de qualidade foram removidos 6.814 SNPs (11%) do banco de dados genotípico. Estes autores foram menos rigorosos com os limiares estabelecidos nos critérios de qualidade dos marcadores/amostras de animais, em comparação aos adotados no presente capítulo.

Fan et al. (2011), avaliando características de composição corporal em suínos, utilizaram um limiar de  $MAF > 0,05$ , eficiência de genotipagem  $> 80\%$  e  $p\text{-valor} < 1 \times 10^{-6}$  no teste de  $\chi^2$  para verificar o EHW para exclusão de SNP. Neste estudo de Fan et al. (2011), excluíram-se inicialmente 5 amostras por apresentarem eficiência de genotipagem menor que 0,80, e no controle de qualidade dos SNPs foram removidos 2.286 com eficiência de genotipagem inferior a 90% e por falharem na genotipagem (*no call*). Desta forma, estes

autores alcançaram aproximadamente 50K SNPs, os quais foram utilizados para análises de associação.

Ankunov et al. (2009) aplicaram a tecnologia *VeraCode* para a genotipagem de SNPs em experimentos com trigo. A metodologia foi eficiente em determinar amostras polipóides do trigo, com uma média de 89% tetraplóides e 84% de genótipos hexaplóides. Neste experimento, após filtragem de SNPs de baixa qualidade, foi constatada precisão de genotipagem com média de 99% para hexaplóides e 100% para tetraplóides. De acordo com os estudos de Deulvot et al. (2010), em experimento realizado também com cultivo de vegetal (ervilha), o ensaio *GoldenGate* mostrou-se de mesma eficácia na validação do painel de SNPs. Este painel estava composto inicialmente de 384 SNPs, dos quais 356 SNPs (92,7%) permaneceram após a aplicação dos critérios de filtragem, mantendo nível acima de 94% para o parâmetro de qualidade *GenTrain*. Assim como na pesquisa foco deste estudo, a tecnologia *GoldenGate* foi utilizada nos experimentos de Ankunov et al. (2009) e Deulvot et al. (2010), os quais destacaram como ferramenta adequada para verificação de variações genética de locos relacionados com características complexas, apresentando resultados favoráveis a este tipo de cultura.

Os valores médios observados e esperados de heterozigosidade para os marcadores foram considerados dentro das expectativas, visto que os SNPs são marcadores bialélicos e que alguns iria apresentar variabilidade reduzida (SNPs relacionados a doenças genéticas). Estes números foram comparados e mostraram-se semelhantes aos obtidos em trabalhos que também tem utilizados marcadores para análises de paternidade em bovinos (Fisher et al., 2009; Hara et al., 2010). Os três painéis se mostraram eficientes na identificação de indivíduos não relacionados e na confirmação de paternidade quando se possui a informação de ambos os pais, uma vez que os valores de não exclusão foram abaixo de 0,0% (exclusão de 99%).

Verificou-se neste estudo que a metodologia *GoldenGate*, tanto otimiza a pesquisa quanto garante maior eficácia de dados obtidos, sendo, conseqüentemente, considerada ferramenta de sucesso para a validação do painel proposto. A proporção geral de SNPs polimórficos observados neste estudo (para 576 animais) considerando todos os 384 SNPs foi estimada em 71,6%. Após aplicação dos critérios de filtragem, foi possível observar uma redução da proporção de SNPs polimórficos (53,4%), porém estes marcadores foram considerados mais robustos. O painel validado de 249 SNPs pode ser utilizado com diferentes propósitos na raça Girolando. Os resultados obtidos através da análise de componentes

principais revelaram a eficiência que este painel, permitindo que seja utilizado em estudos de estruturação genética, estimação do grau de diversidade dentro e divergência genética entre populações, possível identificação de QTL e de regiões do genoma associadas a características de produção leiteira aptas à seleção. Mediante comparação do experimento aqui apresentado e demais estudos fundamentados nesta ferramenta, constata-se a eficácia do painel personalizado por meio da qualidade na geração dos genótipos dos SNPs, visando melhorias da qualidade e produtividade leiteira.

## 2.7. CONCLUSÃO

Os resultados demonstraram a adequação do ensaio *GoldenGate*, por meio da validação do painel personalizado com 249 SNPs, na geração de genótipos confiáveis relacionados a características de produção de leite, doenças hereditárias e teste de paternidade em animais da raça Girolando. O ensaio *GoldenGate* proporcionou, portanto, um grande número de marcadores SNPs capazes de detectar polimorfismo de DNA em rebanhos da raça Girolando, sendo estão indicado para estudos genômicos na presente raça.

## ANEXOS

**Anexo 01: Painel inicial dos SNPs, Probabilidade de hibridização dos oligonucleotídeos estabelecidos pela empresa Illumina (*Score Final e Designability Rank*) e Parâmetros descritivos do resultado da genotipagem GoldenGate (*GC50, CallFreq, MAF, HW, GenTrain*).**

Informações iniciais dos SNPs							Probabilidade de hibridização - Illumina		Parâmetros Descritivos do Resultado da Genotipagem					
ID	Referência	Nome do SNP/Gene/Acesso	BTA	Posição	Efeito	SNP	Score Final	Designability Rank	GC50	CallFreq	MAF	HW	GenTrain	Status
1	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19719788">http://www.ncbi.nlm.nih.gov/pubmed/19719788</a>	ABCG2/ ATP-binding cassette G2	6	12826341	Produção de Proteína	[A/C]	0,962	1	0,804	0,9529	0,1084	0,0148	0,8463	Permaneceu
2	<a href="http://www.biomedcentral.com/content/supplementary/1471-2156-8-32-S1.xls">http://www.biomedcentral.com/content/supplementary/1471-2156-8-32-S1.xls</a>	ABCG2_256	6	1917070863	Produção de Proteína	[A/G]	0,983	1	0,8257	0,8424	0,3774	0	0,8262	Excluído
3	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741677">http://www.ncbi.nlm.nih.gov/pubmed/15741677</a>	APOE/Apolipoprotein E	18	1917070866	Produção de Proteína	[G/C]	0,467	0	0	0	0	0	0	Excluído
4	<a href="http://www.ncbi.nlm.nih.gov/nucleotide/173991">http://www.ncbi.nlm.nih.gov/nucleotide/173991</a>	APOE/Apolipoprotein E	18	1917299522	Produção de Proteína	[T/G]	0,362	0	0,5342	0,8116	0,0569	0	0,5342	Excluído
5	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	ARMC1/BTA14	14	31708200	Produção de Proteína	[T/C]	0,911	1	0,639	0,9366	0,3617	0	0,639	Permaneceu
6	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_BAC_12507	13	15805830	-	[A/G]	0,927	1	0,5951	0,9094	0,4502	0	0,5951	Permaneceu
7	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_BAC_20631	16	16270590	-	[A/G]	0,941	1	0,7067	0,9909	0,3172	0,7795	0,7067	Permaneceu
8	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_102959	-	2021717225	-	[T/C]	0,892	1	0,879	0,8533	0,4841	0	0,879	Excluído
9	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_13104	20	2354405	Produção de Proteína	[T/C]	0,957	1	0,7405	1	0,4284	0,1336	0,7405	Permaneceu
10	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_21534	14	14021415	-	[A/G]	0,98	1	0,7408	1	0,471	0,3814	0,7408	Permaneceu
11	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_22171	21	17927012	-	[A/G]	0,803	1	0,7179	0,9928	0,2546	0,7915	0,7179	Permaneceu
12	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_2335	19	16437712	Produção de Proteína	[T/C]	0,533	0	0,6272	0,875	0,1501	0	0,6272	Permaneceu
13	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_24876/rs109952207	3	39599273	Produção de Proteína	[A/T]	0,872	1	0	0	0	0	0	Excluído

14	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_27620	17	2021717237	-	[T/C]	0,884	1	0,5579	0,8315	0,329	0	0,5579	Excluído
15	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_29837	18	4723921	-	[T/C]	0,903	1	0,6991	0,9982	0,3811	0,0004	0,6991	Permaneceu
16	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_34052 / NW_003104291	11	13597480	-	[C/G]	0,815	1	0,5494	0,9964	0,3518	0,0055	0,5494	Permaneceu
17	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_67659	25	15459573	-	[A/G]	0,86	1	0,7715	1	0,3197	0,0011	0,7715	Permaneceu
18	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_68191	25	73221850	-	[A/G]	0,953	1	0,8032	0,9964	0,3118	0,1976	0,8033	Permaneceu
19	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_70555	24	41289771	-	[T/C]	0,901	1	0,8215	0,8859	0,1779	0,0001	0,8215	Permaneceu
20	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_78387 / NW_003104190	9	39339420	-	[A/G]	0,834	1	0,6985	1	0,4692	0,0001	0,6985	Permaneceu
21	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_83908	28	9487826	-	[T/C]	0,894	1	0,7388	0,8605	0,3474	0,0157	0,7388	Permaneceu
22	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_84180 / NW_003103861	3	14674370	Produção de Proteína	[A/G]	0,897	1	0,6108	0,9312	0,2335	0,1299	0,6108	Permaneceu
23	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_84617 / NW_003104159	9	3524780	-	[T/C]	0,745	1	0,7765	0,9928	0,4453	0,8686	0,7765	Permaneceu
24	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_BFGL_NGS_91026 / rs110585730	2	11605190	-	[T/C]	0,953	1	0,7804	0,9964	0,3009	0,13	0,7804	Permaneceu
25	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	ARS_USMARC_116 / NW_003104062	7	98566451	Produção de Proteína	[A/G]	0,751	1	0,7064	0,9982	0,2396	0,0039	0,7064	Permaneceu
26	<a href="http://www.ncbi.nlm.nih.gov/ncore/NC_007301.3">http://www.ncbi.nlm.nih.gov/ncore/NC_007301.3</a>	ATPIA1/ ATPase, Na <sup>+</sup> /K <sup>+</sup> transporting, alpha 1 polypeptide	3	1917299525	Produção de Proteína	[A/C]	0,696	1	0,7305	0,7772	0,2727	0	0,7305	Excluído
27	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	AY776154 / NW_003103836	2	26997650	-	[A/G]	0,92	1	0,6853	1	0,4266	0,0009	0,6853	Permaneceu
28	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	AY842473 / NW_003103869	3	49703650	-	[A/G]	0,757	1	0,729	0,9982	0,4065	0,832	0,729	Permaneceu
29	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	AY844963 / NW_003103936	5	98102350	Produção de Proteína	[T/C]	0,956	1	0,5257	0,9909	0,4762	0,0004	0,5257	Permaneceu
30	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	AY857620	16	2021717266	-	[G/C]	0,625	1	0,8524	0,7862	0,2546	0	0,8524	Excluído
31	<a href="http://www.ncbi.nlm.nih.gov/ncore/AF515848">http://www.ncbi.nlm.nih.gov/ncore/AF515848</a>	BCL2/Bcl-2 protein	24	62103854	Produção de Proteína	[G/C]	0,724	1	0,5435	0,9891	0,0119	1	0,7764	Permaneceu
32	<a href="http://www.ncbi.nlm.nih.gov/ncore/AF515848">http://www.ncbi.nlm.nih.gov/ncore/AF515848</a>	BCL2/Bcl-2 protein	24	62104847	Produção de Proteína	[A/G]	0,918	1	0,4916	0,9764	0,0612	0,3733	0,4916	Permaneceu
33	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	BIG1/BTA14	14	33364431	Produção de Proteína	[T/A]	0,872	1	0,7446	0,9964	0,4773	0,4549	0,7446	Permaneceu

34	<a href="http://www.ncbi.nlm.nih.gov/nuccore/AB008590.1">http://www.ncbi.nlm.nih.gov/nuccore/AB008590.1</a>	BOLA/ MHC class I, heavy chain	23	1917299540	Produção de Proteína	[C/G]	0,355	0	0,7339	0,9909	0,1618	0	0,7725	Excluído
35	<a href="http://www.ncbi.nlm.nih.gov/pubmed/16704422">http://www.ncbi.nlm.nih.gov/pubmed/16704422</a>	BOLA/ vacciniarelatedkinase2	11	1917299535	Produção de Proteína	[A/C]	0,846	1	0	0	0	0	0	Excluído
36	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BOLA/beta-1,4-N-acetyl-galactosaminyl transferase 2	19	1917299537	Produção de Proteína	[T/C]	0,319	0	0	0	0	0	0	Excluído
37	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BOLA/beta-1,4-N-acetyl-galactosaminyl transferase 3 (2)	5	5027900	Produção de Proteína	[A/C]	0,529	0	0,4541	0,8822	0,4795	0,0002	0,4541	Permaneceu
38	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BOLA/MHC class I, heavy chain	23	27535221	Produção de Proteína	[A/G]	0,424	0	0,606	0,9221	0,4627	0	0,606	Permaneceu
39	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BOLA/pyruvate dehydrogenase complex, component X	15	27863560	Produção de Proteína	[T/C]	0,722	1	0,7216	0,9855	0,4651	0	0,7216	Permaneceu
40	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BOLA_RALYL/RALY RNA binding protein-like	14	80120315	Produção de Proteína	[C/G]	0,651	1	0,6433	0,9692	0,4813	0,0001	0,6433	Permaneceu
41	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTA_110801_no_rs/rs41575151	1	2021717268	-	[T/C]	0,82	1	0,6934	0,808	0,4922	0	0,6934	Excluído
42	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTA_63385_no_rs	27	2021717269	-	[A/C]	0,953	1	0,8287	0,7355	0,4803	0	0,8287	Excluído
43	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_02720/NW_003103979	6	86904512	Produção de Proteína	[T/C]	0,928	1	0,7187	0,9964	0,1618	0,1313	0,7188	Permaneceu
44	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_09701/NW_003103979	6	1917299550	Produção de Proteína	[A/T]	0,85	1	0	0	0	0	0,1479	Excluído
45	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_107922/UDP glucuronosyltransferase 2 family, polypeptide A3	6	86819253	Produção de Proteína	[A/G]	0,574	0	0,794	0,9149	0,302	0,0741	0,794	Permaneceu
46	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_107923/UDP glucuronosyltransferase 2 family, polypeptide A3	6	86819139	Produção de Proteína	[T/C]	0,883	1	0,6928	0,9964	0,3936	0,0017	0,6928	Permaneceu
47	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_77051/NW_003103978	6	87366114	Produção de Proteína	[A/G]	0,922	1	0,7495	0,9982	0,4138	0	0,7495	Permaneceu
48	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_77053	6	1917299558	Produção de Proteína	[A/C]	0,477	0	0,7121	0,75	0,413	0	0,7121	Excluído
49	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_77057/NW_003103979	6	87365840	Produção de Proteína	[A/G]	0,916	1	0,6072	0,9547	0,3615	0,0024	0,6072	Permaneceu
50	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_77060/NW_003103979	6	87325038	Produção de Proteína	[T/C]	0,849	1	0,3776	0,9837	0,1142	0,0117	0,5394	Permaneceu
51	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	BTA6_77062/NW_003103979	6	87324778	Produção de Proteína	[A/G]	0,986	1	0,5007	0,9946	0,4645	0,0003	0,5007	Permaneceu

52	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_00346830/ NW_003104101	8	49356880	-	[T/C]	0,939	1	0,7955	0,8678	0,3643	0,6955	0,7955	Permaneceu
53	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_00780234	20	2021717273	-	[A/G]	0,917	1	0,6621	0,587	0,5	0	0,6621	Excluído
54	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_01052867/ NW_003104083	8	2021717275	Produção de Proteína	[A/G]	0,739	1	0,7913	0,8406	0,2338	0	0,7913	Excluído
55	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_01576968	15	5027900	-	[A/G]	0,766	1	0,7189	0,9928	0,4818	0,0003	0,7189	Permaneceu
56	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_01739441/ NW_003103915	5	10560050	-	[T/C]	0,817	1	0,7034	0,9167	0,4615	0,0005	0,7034	Permaneceu
57	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_01898747/NW_003 104294	11	2021717281	-	[A/G]	0,969	1	0,8646	0,8587	0,4926	0	0,8646	Excluído
58	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	BTB_02057682	29	2021717284	-	[T/C]	0,689	1	0,5381	0,9764	0,1076	0	0,7688	Excluído
59	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/8678998">http://www.ncbi.nlm.nih.gov/pu bmed/8678998</a>	BUT/Butyrophilin	23	1917299566	Produção de Gordura	[T/C]	0,563	0	0,769	0,7862	0,174	0	0,769	Excluído
60	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/8678998">http://www.ncbi.nlm.nih.gov/pu bmed/8678998</a>	BUT/Butyrophilin	23	31343712	Produção de Gordura	[T/G]	0,84	1	0,5944	0,9728	0,4134	0,9107	0,5944	Permaneceu
61	<a href="http://www.gsejournal.org/conte&lt;br/&gt;nt/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/conte nt/pdf/1297-9686-41-24.pdf</a>	C4ORF7_601/NW_00310 3979	6	87355308	Produção de Proteína	[A/G]	0,814	1	0,6929	0,8678	0,0804	0,0001	0,7293	Permaneceu
62	<a href="http://www.gsejournal.org/conte&lt;br/&gt;nt/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/conte nt/pdf/1297-9686-41-24.pdf</a>	C4ORF7_914	6	87355621	Produção de Proteína	[T/C]	0,782	1	0,6151	0,9928	0,4352	0,0009	0,6151	Permaneceu
63	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/18005441">http://www.ncbi.nlm.nih.gov/pu bmed/18005441</a>	CARD15/caspase recruitment domain 15 protein	18	19210680	Produção de Proteína	[T/A]	0,97	1	0,7429	0,8986	0,3992	0	0,7429	Permaneceu
64	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/18005441">http://www.ncbi.nlm.nih.gov/pu bmed/18005441</a>	CARD15/caspase recruitment domain 15 protein	18	19210721	Produção de Proteína	[A/C]	0,97	1	0,7961	0,9783	0,3852	0,0503	0,7961	Permaneceu
65	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/21054834">http://www.ncbi.nlm.nih.gov/pu bmed/21054834</a>	CARD15/caspase recruitment domain 15 protein	18	19212150	Produção de Proteína	[T/A]	0,797	1	0,7701	0,9438	0,4223	0,238	0,7701	Permaneceu
66	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/17433017">http://www.ncbi.nlm.nih.gov/pu bmed/17433017</a>	CCL2/Chemokine (C-C motif) ligand 2	19	1917299579	Produção de Proteína	[A/G]	0,784	1	0	0	0	0	0	Excluído
67	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/17433017">http://www.ncbi.nlm.nih.gov/pu bmed/17433017</a>	CCL2/Chemokine (C-C motif) ligand 2	19	1917299581	Produção de Proteína	[T/C]	0,977	1	0,662	0,6359	0,1809	0	0,662	Excluído
68	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/17433017">http://www.ncbi.nlm.nih.gov/pu bmed/17433017</a>	CCL2/Chemokine (C-C motif) ligand 2	19	1917299584	Produção de Proteína	[A/G]	0,915	1	0	0	0	0	0	Excluído
69	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/17433017">http://www.ncbi.nlm.nih.gov/pu bmed/17433017</a>	CCL2/Chemokine (C-C motif) ligand 2	19	14674417	Produção de Proteína	[T/C]	0,933	1	0,7405	0,9964	0,3691	0,0001	0,7405	Permaneceu
70	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/17433017">http://www.ncbi.nlm.nih.gov/pu bmed/17433017</a>	CCL3/Chemokine (C-C motif) ligand 3	19	1917299587	Produção de Proteína	[T/C]	0,989	1	0	0	0	0	0,2419	Excluído
71	<a href="http://www.ncbi.nlm.nih.gov/pu&lt;br/&gt;bmed/17433017">http://www.ncbi.nlm.nih.gov/pu bmed/17433017</a>	CCR2/chemokine (C-C motif) receptor 2	22	53613712	Produção de Proteína	[A/G]	0,896	1	0,7977	0,9855	0,307	0,2248	0,7977	Permaneceu



72	<a href="http://www.ncbi.nlm.nih.gov/pubmed/9300371">http://www.ncbi.nlm.nih.gov/pubmed/9300371</a>	CD14/ Monocyte differentiation antigen CD14 precursor	7	1917299591	Doença Genética	[A/G]	0,656	1	0	0	0	0	0	Excluído
73	<a href="http://www.ncbi.nlm.nih.gov/pubmed/9361228">http://www.ncbi.nlm.nih.gov/pubmed/9361228</a>	CD18/BLAD	1	145114963	Doença Genética	[A/G]	0,925	1	0,7625	1	0,01	0,9976	0,8026	Permaneceu
74	<a href="http://www.ncbi.nlm.nih.gov/pubmed/9363619/">http://www.ncbi.nlm.nih.gov/pubmed/9363619/</a>	COXP1/ Cytochrome oxidase	5	1917299595	Produção de Proteína	[T/C]	0,945	1	0,5139	0,7699	0,1071	0	0,5139	Excluído
75	<a href="http://www.ncbi.nlm.nih.gov/pubmed/9363619/">http://www.ncbi.nlm.nih.gov/pubmed/9363619/</a>	COXP1/ Cytochrome oxidase	5	111053501	Produção de Proteína	[A/G]	0,86	1	0,7719	0,9801	0,1858	0,1393	0,7719	Permaneceu
76	<a href="http://www.ncbi.nlm.nih.gov/pubmed/18059556">http://www.ncbi.nlm.nih.gov/pubmed/18059556</a>	CRH and TRIM55/BTA14	14	32153693	Produção de Proteína	[A/G]	0,928	1	0,6907	0,9891	0,3782	0,0013	0,6907	Permaneceu
77	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN1S1/ Casein, alpha s1	6	1917299601	Produção de Proteína	[A/G]	0,986	1	0,563	0,9837	0,3517	0	0,563	Excluído
78	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN1S1/ Casein, alpha s1	6	87157212	Produção de Proteína	[A/G]	0,887	1	0,7212	0,5054	0,1703	0	0,7752	Permaneceu
79	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S1_BMC_16680/ Casein, alpha s1	6	1917299605	Produção de Proteína	[C/G]	0,868	1	0	0	0	0	0	Excluído
80	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S1_BMC_17969/ Casein, alpha s1	6	1917299608	Produção de Proteína	[T/C]	0,737	1	0	0	0	0	0	Excluído
81	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S1_BMC_1995/ Casein, alpha s1	6	1917299609	Produção de Proteína	[C/G]	0,859	1	0	0	0	0	0	Excluído
82	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S1_BMC_5798/ Casein, alpha s1	6	87145250	Produção de Proteína	[A/G]	0,879	1	0,735	0,9964	0,37	0,13	0,735	Permaneceu
83	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN1S2/ Casein, alpha s2	6	1917299613	Produção de Proteína	[T/C]	0,758	1	0,8092	0,7699	0,4659	0	0,8092	Excluído
84	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_10215/ Casein, alpha s2	6	87269042	Produção de Proteína	[T/C]	0,888	1	0,7412	1	0,0263	0,5959	0,7412	Permaneceu
85	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_10431/ Casein, alpha s2	6	87269258	Produção de Proteína	[A/G]	0,982	1	0,7618	0,9982	0,0735	0,1029	0,802	Permaneceu
86	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_11255/ Casein, alpha s2	6	87270082	Produção de Proteína	[T/C]	0,815	1	0,7319	0,9964	0,0855	0,005	0,7319	Permaneceu
87	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_14361/ Casein, alpha s2	6	1917299622	Produção de Proteína	[T/A]	0,821	1	0	0	0	0	0	Excluído
88	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_17192/ Casein, alpha s2	6	1917299623	Produção de Proteína	[T/C]	0,929	1	0,865	0,6141	0,469	0	0,8675	Excluído
89	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_18035/ Casein, alpha s2	6	87276866	Produção de Proteína	[A/G]	0,681	1	0,8024	0,9837	0,2928	0,0001	0,8024	Permaneceu
90	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_1915/ Casein, alpha s2	6	87260743	Produção de Proteína	[T/C]	0,909	1	0,69	0,9801	0,4122	0,0012	0,69	Permaneceu

91	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_21257/Casein, alpha s2	6	1917299630	Produção de Proteína	[T/C]	0,678	1	0,7956	0,7572	0,3325	0	0,8026	Excluído
92	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_21965/Casein, alpha s2	6	87280796	Produção de Proteína	[A/G]	0,41	0	0,5983	0,9601	0,2623	0	0,5983	Permaneceu
93	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_22088/Casein, alpha s2	6	1917299634	Produção de Proteína	[T/C]	0,865	1	0	0	0	0	0	Excluído
94	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_2544/Casein, alpha s2	6	87261372	Produção de Proteína	[T/C]	0,668	1	0,6354	0,9692	0,0206	0,6238	0,6688	Permaneceu
95	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_3312/Casein, alpha s2	6	1917299637	Produção de Proteína	[A/G]	0,898	1	0,7229	0,6703	0,2959	0	0,8091	Excluído
96	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_5105/Casein, alpha s2	6	87263933	Produção de Proteína	[T/C]	0,792	1	0,7369	0,8732	0,2873	0,0022	0,7369	Permaneceu
97	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_7008/Casein, alpha s2	6	1917299642	Produção de Proteína	[A/G]	0,6	1	0,8058	0,8261	0,261	0	0,8058	Excluído
98	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN1S2_BMC_8316/Casein, alpha s2	6	1917299643	Produção de Proteína	[T/C]	0,811	1	0	0	0	0	0	Excluído
99	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741664">http://www.ncbi.nlm.nih.gov/pubmed/15741664</a>	CSN2/Beta-casein	6	1917299645	Produção de Proteína	[A/C]	0,717	1	0	0	0	0	0	Excluído
100	<a href="http://www.ncbi.nlm.nih.gov/pubmed/3271384">http://www.ncbi.nlm.nih.gov/pubmed/3271384</a>	CSN2/Beta-casein	6	87181471	Produção de Proteína	[A/C]	0,945	1	0,5685	0,9964	0,0055	1	0,8121	Permaneceu
101	<a href="http://www.ncbi.nlm.nih.gov/pubmed/3271384">http://www.ncbi.nlm.nih.gov/pubmed/3271384</a>	CSN2/Beta-casein	6	87181542	Produção de Proteína	[G/C]	0,845	1	0,4663	1	0,0136	0,9854	0,6661	Permaneceu
102	<a href="http://www.ncbi.nlm.nih.gov/pubmed/3271384">http://www.ncbi.nlm.nih.gov/pubmed/3271384</a>	CSN2/Beta-casein	6	87181619	Produção de Proteína	[0/G]	0,673	1	0,7772	0,9746	0,0381	0,4559	0,8181	Permaneceu
103	<a href="http://www.ncbi.nlm.nih.gov/pubmed/3271384">http://www.ncbi.nlm.nih.gov/pubmed/3271384</a>	CSN2/Beta-casein	6	87183031	Produção de Proteína	[T/G]	0,914	1	0,6427	0,904	0,2044	0,0017	0,6427	Permaneceu
104	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_110/Beta-casein	6	87181489	Produção de Proteína	[T/G]	0,845	1	0,7728	0,8551	0,0413	0,4376	0,8135	Permaneceu
105	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_516/Beta-casein	6	1917299656	Produção de Proteína	[I/D]	0,619	1	0,6132	0,9203	0,1053	0,1577	0,6454	Excluído
106	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_2364/Beta-casein	6	87187581	Produção de Proteína	[T/C]	0,722	1	0,756	0,8986	0,3427	0,0268	0,756	Permaneceu
107	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_269/Beta-casein	6	87189676	Produção de Proteína	[A/G]	0,887	1	0,6081	0,9964	0,1582	0,0459	0,6081	Permaneceu
108	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_3118/Beta-casein	6	1917299662	Produção de Proteína	[T/C]	0,891	1	0,527	0,6993	0,4326	0	0,527	Excluído
109	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_344/Beta-casein	6	87189601	Produção de Proteína	[T/C]	0,935	1	0,7273	0,9982	0,1497	0,1688	0,7281	Permaneceu

110	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_4393/Beta-casein	6	87185552	Produção de Proteína	[T/C]	0,854	1	0,6805	0,9964	0,0136	0,7313	0,6805	Permaneceu
111	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_5775/Beta-casein	6	87184171	Produção de Proteína	[A/G]	0,943	1	0,65	1	0,4565	0,0163	0,65	Permaneceu
112	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_6334/Beta-casein	6	1917299670	Produção de Proteína	[T/C]	0,635	1	0,5923	0,8859	0,3476	0,1195	0,5923	Excluído
113	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_6640/Beta-casein	6	87183306	Produção de Proteína	[T/G]	0,799	1	0,7434	0,9946	0,3616	0,0035	0,7434	Permaneceu
114	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_9215/Beta-casein	6	1917299673	Produção de Proteína	[T/G]	0,931	1	0	0	0	0	0	Excluído
115	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN2_BMC_9472/Beta-casein	6	87180474	Produção de Proteína	[A/G]	0,891	1	0,7212	0,9801	0,4861	0,0618	0,7212	Permaneceu
116	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	1917299677	Produção de Proteína	[A/G]	0,823	1	0	0	0	0	0	Excluído
117	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	1917299679	Produção de Proteína	[T/A]	0,667	1	0	0	0	0	0,2421	Excluído
118	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	87390459	Produção de Proteína	[A/G]	0,739	1	0,5772	0,9221	0,0059	0,8425	0,8245	Permaneceu
119	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	87390573	Produção de Proteína	[G/A]	0,6	1	0,5623	0,9964	0,0018	1	0,8033	Permaneceu
120	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	87390576	Produção de Proteína	[A/G]	0,674	1	0,5473	0,9982	0,0027	1	0,7818	Permaneceu
121	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	1917071027	Produção de Proteína	[T/C]	0,602	1	0	0	0	0	0	Excluído
122	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	87390632	Produção de Proteína	[A/G]	0,834	1	0,6079	0,8587	0,1234	0,89	0,6079	Permaneceu
123	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSN3/Kappa-casein	6	87390673	Produção de Proteína	[G/A]	0,981	1	0,7813	0,8913	0,4797	0	0,7813	Permaneceu
124	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_11414/Kappa-casein	6	87387208	Produção de Proteína	[A/G]	0,768	1	0,6443	0,8732	0,2075	0,0001	0,6443	Permaneceu
125	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_12853/Kappa-casein	6	88481137	Produção de Proteína	[A/G]	0,788	1	0,7004	0,9058	0,314	0,2492	0,7004	Permaneceu
126	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_13916/Kappa-casein	6	87390993	Produção de Proteína	[A/G]	0,573	0	0,7783	0,9185	0,4339	0	0,7783	Permaneceu
127	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_13941/Kappa-casein	6	1923344359	Produção de Proteína	[T/C]	0,732	1	0	0	0	0	0	Excluído
128	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_14124/Kappa-casein	6	1923344360	Produção de Proteína	[T/C]	0,676	1	0	0	0	0	0	Excluído

129	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_14131/Kappa-casein	6	1923344361	Produção de Proteína	[T/G]	0,676	1	0,463	0,5543	0,2729	0	0,7363	Excluído
130	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_14406/Kappa-casein	6	1923344362	Produção de Proteína	[T/G]	0,891	1	0,5933	0,8225	0,1211	0	0,5933	Excluído
131	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_14624/Kappa-casein	6	87391701	Produção de Proteína	[A/G]	0,794	1	0,7692	0,8859	0,4479	0	0,7692	Permaneceu
132	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_14771/Kappa-casein	6	87391848	Produção de Proteína	[A/G]	0,905	1	0,7593	0,9909	0,1316	0,0441	0,7593	Permaneceu
133	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_14826/Kappa-casein	6	87391903	Produção de Proteína	[A/G]	0,943	1	0,8117	0,9728	0,1164	0,0058	0,8544	Permaneceu
134	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_15266/Kappa-casein	6	1923344370	Produção de Proteína	[G/C]	0,654	1	0,781	0,846	0,1981	0	0,781	Excluído
135	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_1923/Kappa-casein	6	87377716	Produção de Proteína	[A/G]	0,461	0	0,812	0,9656	0,3677	0,0018	0,812	Permaneceu
136	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_2090/Kappa-casein	6	1923344375	Produção de Proteína	[A/C]	0,519	0	0,5729	0,4203	0,3987	0	0,7164	Excluído
137	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_2178/Kappa-casein	6	1923344376	Produção de Proteína	[T/C]	0,889	1	0,7348	0,779	0,4628	0	0,7348	Excluído
138	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_2220/Kappa-casein	6	1923344379	Produção de Proteína	[A/G]	0,533	0	0	0	0	0	0	Excluído
139	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	CSN3_BMC_955/Kappa-casein	6	1923344380	Produção de Proteína	[T/A]	0,979	1	0,8205	0,8406	0,1692	0	0,8205	Excluído
140	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	CSNK2A2/Casein kinase 2, alpha prime polypeptide	18	26109950	Produção de Proteína	[A/G]	0,918	1	0,689	0,9692	0,0019	0,9636	0,7253	Permaneceu
141	<a href="http://genome.cshlp.org/content/16/1/97.full.pdf+html">http://genome.cshlp.org/content/16/1/97.full.pdf+html</a>	CVM/ Complex Vertebral Malformation	3	43412410	Doença Genética	[A/C]	0,931	1	0,7529	0,9946	0,0091	0,82	0,7926	Permaneceu
142	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	CXCL10/ Chemokine (C-X-C motif) ligand 10	6	92623585	Produção de Proteína	[A/G]	0,88	1	0,6213	0,942	0,1548	0,0005	0,6213	Permaneceu
143	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	DCT/dopachrome tautomerase	12	69534625	Produção de Proteína	[A/G]	0,651	1	0,5354	0,9692	0,0271	0,8434	0,7648	Permaneceu
144	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	DCT/dopachrome tautomerase	12	69544294	Produção de Proteína	[T/G]	0,983	1	0,681	0,9891	0,0156	0,8554	0,7169	Permaneceu
145	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT/ Acyl-CoA:1,2-diacylglycerol O-transferase	14	1802193	Produção de Proteína	[A/C]	0,375	0	0,8232	0,9692	0,486	0,5743	0,8232	Permaneceu
146	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT/ Acyl-CoA:1,2-diacylglycerol O-transferase	14	1923344394	Produção de Proteína	[A/G]	0,375	0	0	0	0	0	0	Excluído

147	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT/Acyl-CoA:1,2-diaclyglycerol O-transferase	14	1923344396	Produção de Proteína	[A/G]	0,382	0	0,8206	0,8442	0,4571	0	0,8206	Excluído
148	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT/Acyl-CoA:1,2-diaclyglycerol O-transferase	14	1923344398	Produção de Proteína	[A/C]	0,381	0	0,8223	0,8152	0,4489	0	0,8224	Excluído
149	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT1/Acyl-CoA:1,2-diaclyglycerol O-transferase	14	1795176	Produção de Proteína	[C/G]	0,629	1	0,7303	0,9964	0,0118	0,9998	0,7688	Permaneceu
150	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT1/Acyl-CoA:1,2-diaclyglycerol O-transferase	14	1804175	Produção de Proteína	[G/C]	0,797	1	0,6645	0,942	0,0433	0,537	0,6995	Permaneceu
151	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	DGAT1/Acyl-CoA:1,2-diaclyglycerol O-transferase	14	1805963	Produção de Proteína	[G/C]	0,645	1	0,7357	1	0,1667	0,2392	0,7365	Permaneceu
152	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ381153/NW_003103793	1	3249120	Tese de Parentesco	[A/C]	0,922	1	0,6674	1	0,4774	0,0831	0,6674	Permaneceu
153	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ404150/NW_003103810	1	2021717288	Tese de Parentesco	[A/C]	0,822	1	0,8254	0,7663	0,4326	0	0,8254	Excluído
154	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ404151/NW_003103826	1	151349518	Tese de Parentesco	[T/C]	0,607	1	0,7897	0,962	0,4482	0,516	0,7897	Permaneceu
155	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ404152/NW_003103830	2	2021717290	Tese de Paternidade	[T/C]	0,498	0	0	0	0	0	0	Excluído
156	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ422949/NW_003103841	2	65069060	Tese de Paternidade	[A/G]	0,724	1	0,7244	0,9964	0,3145	0,3228	0,7244	Permaneceu
157	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ422950/NW_003103855	3	2021717294	-	[T/G]	0,815	1	0,6227	0,7681	0,3656	0	0,6227	Excluído
158	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ435443/NW_003103870	3	2021717296	-	[T/G]	0,885	1	0,6561	0,7717	0,0857	0	0,6561	Excluído
159	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ485413/NW_003103904	4	94176200	Teste de Paternidade	[T/C]	0,678	1	0,6822	0,9928	0,3668	0,4384	0,6822	Permaneceu
160	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ500958/NW_003103919	5	27702800	Teste de Paternidade	[T/C]	0,947	1	0,7285	0,9982	0,3666	0	0,7285	Permaneceu
161	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ647186/NW_003103887	4	17200580	Teste de Paternidade	[A/G]	0,982	1	0,7923	0,9964	0,3491	0,1793	0,7923	Permaneceu
162	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ647189/NW_003103925	5	63273286	Teste de Paternidade	[T/C]	0,619	1	0,758	0,9946	0,2596	0,065	0,758	Permaneceu
163	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ650635/NW_003104036	7	55116250	Teste de Paternidade	[A/G]	0,897	1	0,8174	0,942	0,3558	0,0514	0,8174	Permaneceu
164	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ650636/NW_003104087	8	2021717305	Teste de Paternidade	[T/C]	0,651	1	0,6544	0,7518	0,1675	0	0,6544	Excluído

165	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ786759/NW_003104059	7	2021717306	Teste de Paternidade	[A/G]	0,951	1	0	0	0	0	0	Excluído
166	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ786763	12	11824650	Teste de Paternidade	[A/T]	0,988	1	0,7472	1	0,3949	0,1587	0,7473	Permaneceu
167	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ786764	12	25669980	Teste de Paternidade	[T/C]	0,601	1	0,5992	0,9946	0,2341	0,0342	0,5993	Permaneceu
168	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ786766	10	3278580	Teste de Paternidade	[A/G]	0,86	1	0,7452	0,9964	0,4818	0,3778	0,7452	Permaneceu
169	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ789028/NW_003103970	6	2021717313	Teste de Paternidade	[T/C]	0,92	1	0	0	0	0	0	Excluído
170	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ837643/NW_003104295	1	2021717315	Teste de Paternidade	[A/G]	0,949	1	0	0	0	0	0	Excluído
171	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ837646	11	1703600	Teste de Paternidade	[A/G]	0,999	1	0,6944	0,9982	0,3122	0,1282	0,6944	Permaneceu
172	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ846689/NW_003104193	9	2021717319	Teste de Paternidade	[T/C]	0,549	0	0,6389	0,8188	0,1515	0	0,6389	Excluído
173	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ846695	16	80747550	Teste de Paternidade	[T/C]	0,735	1	0,5687	0,8986	0,4446	0,6225	0,5687	Permaneceu
174	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ866817	15	38078720	Teste de Paternidade	[T/C]	0,949	1	0,74	0,9982	0,3938	0,0088	0,74	Permaneceu
175	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ866818	15	79187250	Teste de Paternidade	[A/C]	0,956	1	0,6557	0,9964	0,3145	0,0022	0,6557	Permaneceu
176	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ888310	17	887316	Teste de Paternidade	[T/G]	0,905	1	0,6113	0,9982	0,2577	0,0569	0,6118	Permaneceu
177	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ888313	20	17837650	Teste de Paternidade	[T/C]	0,782	1	0,7456	0,9529	0,3441	0,7492	0,7456	Permaneceu
178	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ984825/NW_003104285	10	2021717329	Teste de Paternidade	[A/G]	0,624	1	0,8448	0,8315	0,3224	0	0,8448	Excluído
179	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ984826	14	27751888	Teste de Paternidade	[T/A]	0,959	1	0,7691	0,9946	0,4508	0,0001	0,7691	Permaneceu
180	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ990834	14	59582201	Teste de Paternidade	[T/C]	0,762	1	0,7746	0,9982	0,2314	0,7793	0,7747	Permaneceu
181	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ995976	21	59582201	Teste de Paternidade	[T/C]	0,927	1	0,7968	0,9529	0,4876	0,1018	0,7968	Permaneceu
182	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	DQ995977	24	2021717336	Teste de Paternidade	[T/C]	0,675	1	0,6055	0,7319	0,4691	0	0,6055	Excluído
183	<a href="http://www.ncbi.nlm.nih.gov/pubmed/8486364">http://www.ncbi.nlm.nih.gov/pubmed/8486364</a>	DUMPS/ Uridine 5'-Monophosphate Synthase Deficiency	1	69756880	Doença Genética	[T/C]	0,963	1	0,4068	0,9873	0,0138	0,9276	0,5812	Permaneceu

184	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF026084	19	51400125	Teste de Paternidade	[A/G]	0,962	1	0,8614	0,9511	0,4124	0,3738	0,8614	Permaneceu
185	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF026087	13	1982210	Teste de Paternidade	[A/G]	0,716	1	0,7068	0,9873	0,3358	0,0032	0,7068	Permaneceu
186	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF028073	18	56026970	-	[A/G]	0,627	1	0,7834	0,9348	0,3779	0,167	0,7834	Permaneceu
187	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF034080	29	2021717342	Teste de Paternidade	[A/G]	0,704	1	0,7704	0,9982	0,3702	0,1691	0,7704	Permaneceu
188	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF034083	25	15459573	Teste de Paternidade	[A/C]	0,856	1	0,6853	0,9964	0,4527	0,198	0,6853	Permaneceu
189	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF034085	28	2021717347	Teste de Paternidade	[T/C]	0,896	1	0,695	0,75	0,3756	0	0,695	Excluído
190	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF034086	26	19605923	Teste de Paternidade	[A/G]	0,871	1	0,7413	0,9982	0,4719	0,4705	0,7413	Permaneceu
191	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF042091	28	44262045	Teste de Paternidade	[T/C]	0,871	1	0,81	0,8822	0,3018	0,0002	0,81	Permaneceu
192	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF089234	23	2021717353	Teste de Paternidade	[T/A]	0,745	1	0,8199	0,8732	0,4907	0	0,8222	Excluído
193	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF093509	22	2021717354	Teste de Paternidade	[T/A]	0,695	1	0,631	0,8496	0,3731	0	0,631	Excluído
194	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF093511	21	59582397	Teste de Paternidade	[A/C]	0,864	1	0,7043	0,9891	0,3819	0,0002	0,7043	Permaneceu
195	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF141102	27	37514025	Teste de Paternidade	[T/C]	0,9	1	0,7672	0,9801	0,4492	0,0025	0,7672	Permaneceu
196	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	EF164803	19	55174230	Teste de Paternidade	[T/C]	0,841	1	0,6237	1	0,4167	0,1948	0,6238	Permaneceu
197	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19719788">http://www.ncbi.nlm.nih.gov/pubmed/19719788</a>	FASN/Fatty Acid Synthase	19	1923344410	Produção de Gordura	[A/G]	0,595	0	0,7493	0,8225	0,1641	0	0,7673	Excluído
198	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19719788">http://www.ncbi.nlm.nih.gov/pubmed/19719788</a>	FASN/Fatty Acid Synthase	19	15345300	Produção de Gordura	[A/G]	0,531	0	0,7759	0,9783	0,2611	0	0,7759	Permaneceu
199	<a href="http://www.journalofdairysciences.org/article/S0022-0302(08)71199-8/abstract">http://www.journalofdairysciences.org/article/S0022-0302(08)71199-8/abstract</a>	FGF2/ Fibroblast Growth Factor 2	17	887316	Produção de Proteína	[A/G]	0,936	1	0,7453	0,9964	0,1627	0,294	0,7453	Permaneceu
200	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	FNTA/Farnesyltransferase e alpha subunit	27	36746945	Produção de Proteína	[T/C]	0,902	1	0,7776	0,9982	0,2042	0,0047	0,7777	Permaneceu
201	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741671">http://www.ncbi.nlm.nih.gov/pubmed/15741671</a>	GH/Growth Hormone	19	1923344417	Produção de Proteína	[A/G]	0,592	0	0	0	0	0	0	Excluído
202	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741670">http://www.ncbi.nlm.nih.gov/pubmed/15741670</a>	GH/Growth Hormone	19	1923344419	Produção de Proteína	[C/G]	0,414	0	0	0	0	0	0	Excluído

203	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741672">http://www.ncbi.nlm.nih.gov/pubmed/15741672</a>	GH/Growth Hormone	19	16437712	Produção de Proteína	[T/C]	0,905	1	0,5635	0,9928	0,0018	1	0,8049	Permaneceu
204	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	GHR/Growth Hormone Receptor	20	2354405	Produção de Proteína	[A/G]	0,998	1	0,7326	0,9819	0,1624	0,3351	0,7326	Permaneceu
205	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	GHRH/ Growth hormone-releasing hormone receptor	4	65801620	Homônio de Crescimento	[C/G]	0,826	1	0,5182	0,9964	0,3327	0,0693	0,5182	Permaneceu
206	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap23359_BTA_143727	6	3614332	Produção de Proteína	[T/G]	0,748	1	0,8151	0,9783	0,4778	0,2641	0,8151	Permaneceu
207	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap23813_BTA_128455	13	52496690	-	[A/C]	0,958	1	0,6517	0,9583	0,2051	0,1357	0,6517	Permaneceu
208	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap25366_BTA_63287/ NW_003104277	10	30632530	-	[T/C]	0,961	1	0,806	0,8859	0,1503	0,0253	0,8484	Permaneceu
209	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap27208_BTA_157501/ NW_003103871	3	71860050	-	[A/G]	0,942	1	0,6887	0,9946	0,3434	0,2852	0,6887	Permaneceu
210	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap27732_BTA_123008/ NW_003104020	7	27345560	-	[T/C]	0,621	1	0,6697	0,9964	0,1364	0,2166	0,6697	Permaneceu
211	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap29197_BTA_97143/ NW_003103804	1	37430750	Produção de Proteína	[T/A]	0,949	1	0,7096	0,9982	0,4546	0,0093	0,7096	Permaneceu
212	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap30796_BTA_134326/ NW_003103846	2	94984600	Produção de Proteína	[T/C]	0,909	1	0,547	0,9964	0,1827	0,0119	0,547	Permaneceu
213	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap31189_BTA_161111	12	19140225	Produção de Proteína	[A/G]	0,974	1	0,5892	1	0,1839	0,1813	0,5892	Permaneceu
214	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap31226_BTA_68802	23	2021717379	Produção de Proteína	[T/C]	0,94	1	0,5393	0,6413	0,1638	0	0,5393	Excluído
215	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap34164_BES10_Contig538_1008/ NW_003104285	10	90150330	-	[T/C]	0,976	1	0,7368	0,9964	0,4036	0,1844	0,7368	Permaneceu
216	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap35453_SCAFFOLD669_15136/ NW_003104266	10	5934530	-	[T/C]	0,94	1	0,7421	0,9293	0,3869	0,0002	0,7421	Permaneceu
217	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap44212_BTA_122090/ NW_003103900	4	69782720	-	[T/C]	0,767	1	0,7094	0,9004	0,4044	0	0,7094	Permaneceu
218	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap46428_BTA_54503	22	53543195	-	[T/A]	0,983	1	0,862	0,9094	0,4841	0,9834	0,862	Permaneceu
219	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap50821_BTA_91753	6	55697300	Produção de Proteína	[A/G]	0,94	1	0,6858	0,9384	0,168	0,1585	0,6858	Permaneceu
220	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap51475_BTA_62174	26	38233329	-	[T/C]	0,779	1	0,8358	0,9909	0,2998	0,0238	0,8358	Permaneceu
221	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap51741_BTA_53756	22	53613712	-	[T/C]	0,65	1	0,5568	0,9873	0,4743	0,0286	0,5568	Permaneceu



222	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	Hapmap53185_rs29010294/ NW_003104134	8	89564380	-	[A/G]	0,748	1	0,7159	1	0,4239	0,0005	0,7159	Permaneceu
223	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	hypothetical LOC539014/BTA14	14	18160174	-	[A/C]	0,91	1	0,7363	0,9583	0,5	0,7984	0,7363	Permaneceu
224	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	hypothetical LOC783036/BTA14	14	66256529	-	[A/C]	0,902	1	0,7715	0,9891	0,2482	0,0004	0,7715	Permaneceu
225	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	IFNG/ Interferon Gamma	5	1923344431	Produção de Proteína	[A/C]	0,986	1	0	0	0	0	0,3437	Excluído
226	<a href="http://support.illumina.com/downloads/bovinesnp50v2.html">http://support.illumina.com/downloads/bovinesnp50v2.html</a>	IFNG/ Interferon Gamma	5	1923344433	Produção de Proteína	[A/C]	0,93	1	0	0	0	0	0	Excluído
227	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741672">http://www.ncbi.nlm.nih.gov/pubmed/15741672</a>	IGF1/ Insulin like growth factor-I	5	66548270	Produção de Proteína	[T/C]	0,946	1	0,6978	0,9982	0,1588	0,9712	0,6978	Permaneceu
228	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	IGF1/ Insulin like growth factor-I	5	66605021	Produção de Proteína	[T/C]	0,898	1	0,7772	0,913	0,3512	0	0,7772	Permaneceu
229	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19785908">http://www.ncbi.nlm.nih.gov/pubmed/19785908</a>	IGF2/Exon 10	29	1923344438	Produção de Proteína	[T/G]	0,811	1	0,8062	0,9475	0,2132	0	0,8062	Excluído
230	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19785908">http://www.ncbi.nlm.nih.gov/pubmed/19785908</a>	IGF2/Exon 2	29	6352711	Produção de Proteína	[T/C]	0,574	0	0,5429	0,8895	0,4043	0	0,5429	Permaneceu
231	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	IL8/Interleukin 8 (neutrophil activating peptide 1)	6	90560064	Produção de Proteína	[A/G]	0,891	1	0,7379	0,9982	0,431	0,3242	0,7379	Permaneceu
232	<a href="http://www.ncbi.nlm.nih.gov/pubmed/11252171">http://www.ncbi.nlm.nih.gov/pubmed/11252171</a>	IL8/Interleukin 8 (neutrophil activating peptide 1)	6	90561109	Produção de Proteína	[A/G]	0,472	0	0,6199	0,9873	0,4642	0	0,6199	Permaneceu
233	<a href="http://www.ncbi.nlm.nih.gov/pubmed/21054834">http://www.ncbi.nlm.nih.gov/pubmed/21054834</a>	IL8/Interleukin 8 (neutrophil activating peptide 1)	6	90562529	Produção de Proteína	[G/A]	0,916	1	0,6043	0,9928	0,0018	1	0,8633	Permaneceu
234	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	IL8RA(CXCR1) /Interleukin 8 Receptor	2	1923344451	Produção de Proteína	[T/C]	0,412	0	0,7806	0,7029	0,1611	0	0,7806	Excluído
235	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	IL8RA(CXCR1)/Interleukin 8 Receptor (chemokine receptor 1)	2	1923344448	Produção de Proteína	[A/G]	0,328	0	0	0	0	0	0	Excluído
236	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	ITGA2/Integrin Alpha 2 Subunit	20	1923344453	Produção de Proteína	[A/C]	0,977	1	0	0	0	0	0	Excluído
237	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	ITGB2/ Integrin	1	145113594	Produção de Proteína	[A/G]	0,527	0	0,6123	0,9493	0,3626	0	0,6123	Permaneceu
238	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	KCBN2/BTA14	14	1923344456	Produção de Proteína	[C/G]	0,934	1	0	0	0	0	0	Excluído
239	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+</a>	KCBN2/BTA14	14	38453058	Produção de Proteína	[C/G]	0,712	1	0,5926	0,9764	0,4685	0	0,5926	Permaneceu



253	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LALBA/ Alpha lactalbumin	5	1923344486	Produção de Proteína	[A/G]	0,421	0	0	0	0	0	0,2079	Excluído
254	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LALBA/ Alpha lactalbumin	5	31347880	Produção de Proteína	[A/G]	0,918	1	0,6107	0,9982	0,1942	0,636	0,6108	Permaneceu
255	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741676">http://www.ncbi.nlm.nih.gov/pubmed/15741676</a>	LEP/ Leptin	4	93262050	Produção de Proteína	[T/C]	0,736	1	0,7559	0,962	0,3606	0,0003	0,7559	Permaneceu
256	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741666">http://www.ncbi.nlm.nih.gov/pubmed/15741666</a>	LGB/Lactoglobulin, beta	11	1923344492	Produção de Proteína	[T/C]	0,698	1	0	0	0	0	0	Excluído
257	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LGB/Lactoglobulin, beta	11	1923344494	Produção de Proteína	[A/G]	0,423	0	0	0	0	0	0	Excluído
258	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LGB/Lactoglobulin, beta	11	103304720	Produção de Proteína	[A/G]	0,483	0	0,5663	0,9692	0,0037	0,9451	0,809	Permaneceu
259	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LGB/Lactoglobulin, beta	11	1923344496	Produção de Proteína	[G/C]	0,428	0	0	0	0	0	0	Excluído
260	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LGB/Lactoglobulin, beta	11	1923344497	Produção de Proteína	[T/G]	0,694	1	0,6957	0,5833	0,1009	0	0,7784	Excluído
261	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LGB/Lactoglobulin, beta	11	103305450	Produção de Proteína	[T/C]	0,426	0	0,4956	0,9728	0,0019	1	0,708	Permaneceu
262	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LGB/Lactoglobulin, beta	11	1923344500	Produção de Proteína	[T/G]	0,531	0	0	0	0	0	0	Excluído
263	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LHB/Luteinizing hormone beta polypeptide	18	14758485	Produção de Proteína	[A/G]	0,362	0	0,5438	0,9982	0,0073	1	0,7768	Permaneceu
264	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LHB/Luteinizing hormone beta polypeptide	18	19210680	Produção de Proteína	[T/G]	0,292	0	0,6432	0,9257	0,0538	0,3208	0,6771	Permaneceu
265	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC512677/BTA14	14	42234114	-	[T/G]	0,843	1	0,6278	0,9293	0,3704	0,0576	0,6278	Permaneceu
266	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC512677/BTA14	14	42314058	-	[A/G]	0,986	1	0,7501	0,9964	0,0473	0,2849	0,7501	Permaneceu
267	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC512826/BTA14	14	2576623	-	[C/G]	0,848	1	0,6198	0,9764	0,3961	0	0,6198	Permaneceu
268	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC515582/BTA14	14	64298207	-	[A/G]	0,948	1	0,7539	0,9982	0,2877	0,3801	0,7539	Permaneceu
269	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/</a>	LOC519708/BTA14	14	29859550	-	[A/C]	0,925	1	0,8063	0,9891	0,1117	0,0249	0,8063	Permaneceu

	1/Ferreira+Marques_Elisa_Fall+2009.pdf														
270	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC526726/BTA14	14	1923344516	-	[A/G]	0,944	1	0,8687	0,8152	0,3089	0	0,8687	Excluído	
271	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC529552/BTA14	14	1923344518	-	[A/T]	0,971	1	0,7327	0,8967	0,296	0	0,7327	Excluído	
272	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC529552/BTA14	14	31014368	-	[T/C]	0,912	1	0,6531	0,9004	0,2435	0,7752	0,6531	Permaneceu	
273	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC536811/BTA14	14	39640732	-	[A/G]	0,997	1	0,8224	1	0,1087	0,6946	0,8224	Permaneceu	
274	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC536811/BTA14	14	39641502	-	[A/G]	0,997	1	0,6089	0,9946	0,4262	0,8164	0,6089	Permaneceu	
275	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC617133/BTA14	14	66494815	-	[A/G]	0,944	1	0,8216	0,9873	0,145	0,0862	0,8216	Permaneceu	
276	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC783036 and LOC790864/BTA14	14	66194561	-	[A/G]	0,966	1	0,6572	0,9855	0,0551	0,2232	0,6918	Permaneceu	
277	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC783036 and LOC790864/BTA14	14	66211845	-	[A/G]	0,953	1	0,6895	1	0,4049	0,0046	0,6895	Permaneceu	
278	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	LOC78474/BTA14	14	61168426	-	[A/G]	0,929	1	0,785	0,971	0,2882	0	0,785	Permaneceu	
279	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LTF/Lactoferrin	22	41142978	Produção de Proteína	[A/G]	0,684	1	0,8175	0,9873	0,1817	0,1545	0,8175	Permaneceu	
280	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LTF/Lactoferrin	22	53522045	Produção de Proteína	[C/G]	0,686	1	0,5955	0,9909	0,489	0,0238	0,5955	Permaneceu	
281	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	LTF/Lactoferrin	22	53522973	Produção de Proteína	[A/C]	0,914	1	0,7722	0,9964	0,2136	0,0092	0,7722	Permaneceu	
282	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	MC1R/melanocortin 1 receptor	18	19210721	Produção de Proteína	[A/C]	0,335	0	0,4097	0,9891	0,0055	1	0,5853	Permaneceu	

283	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	MC1R/melanocortin 1 receptor	18	19212150	Produção de Proteína	[A/G]	0,744	1	0,6982	0,8659	0,1496	0,0002	0,6982	Permaneceu
284	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	MC1R/melanocortin 1 receptor	18	26109950	Produção de Proteína	[T/C]	0,696	1	0,5837	0,9837	0,3619	0	0,5837	Permaneceu
285	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	MDBK/ Epithelial cell inflammatory protein-1	6	1923344546	Produção de Proteína	[A/G]	0,887	1	0	0	0	0	0	Excluído
286	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	MGC128538/BTA14	14	45004053	-	[A/G]	0,958	1	0,5027	0,9946	0,296	0	0,5027	Permaneceu
287	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	MGC157244/BTA14	14	1923344549	-	[A/C]	0,581	0	0,6393	0,8315	0,4804	0	0,6393	Excluído
288	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	neaby MCM4/minichromosome maintenance complex component 4	14	21280522	Produção de Proteína	[T/C]	0,85	1	0,7202	0,9601	0,0943	0,0715	0,7202	Permaneceu
289	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	nearby LOC536186 CRISPLD1/cysteine-rich secretory protein LCCL domain containing 1	14	1923344553	Produção de Proteína	[G/C]	0,947	1	0,7073	0,8225	0,3546	0	0,7073	Excluído
290	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	nearby LOC782438/BTA14	14	43086130	Produção de Proteína	[A/G]	0,849	1	0,6571	0,9819	0,0996	0,1321	0,6571	Permaneceu
291	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	nearby LOC783431/BTA14	14	1923344558	Produção de Proteína	[A/G]	0,959	1	0	0	0	0	0	Excluído
292	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	nearby PYCRL	14	2276443	Produção de Proteína	[A/C]	0,822	1	0,682	0,9982	0,2468	0,0222	0,682	Permaneceu
293	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/ BTA14	14	1923344561	Produção de Proteína	[C/G]	0,987	1	0,6719	0,9964	0,1936	0	0,6727	Excluído
294	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	1923344563	Produção de Proteína	[T/C]	0,993	1	0,7509	0,9764	0,0464	0	0,7904	Excluído
295	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	6696523	Produção de Proteína	[T/C]	0,936	1	0,692	0,9928	0,3677	0	0,692	Permaneceu

296	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	23054179	Produção de Proteína	[G/C]	0,882	1	0,8215	0,9928	0,01	0,7383	0,8647	Permaneceu
297	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	1923344569	Produção de Proteína	[T/C]	0,612	1	0	0	0	0,9681	0	Excluído
298	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	25857110	Produção de Proteína	[A/G]	0,869	1	0,6534	0,9438	0,0845	0,0009	0,6534	Permaneceu
299	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	26626394	Produção de Proteína	[A/G]	0,922	1	0,7909	0,9583	0,0406	0,0271	0,8325	Permaneceu
300	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	34930139	Produção de Proteína	[T/G]	0,855	1	0,7187	0,9964	0,3673	0,0792	0,7187	Permaneceu
301	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	36837792	Produção de Proteína	[A/C]	0,901	1	0,5714	0,9964	0,1	0,3576	0,5714	Permaneceu
302	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	40142071	Produção de Proteína	[A/G]	0,92	1	0,7308	0,9928	0,2974	0,3552	0,7309	Permaneceu
303	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	40146422	Produção de Proteína	[A/G]	0,877	1	0,7916	0,9982	0,1452	0,3134	0,7916	Permaneceu
304	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	1923344583	Produção de Proteína	[A/G]	0,891	1	0,5107	0,9801	0,3826	0	0,5107	Excluído
305	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	40142071	Produção de Proteína	[T/C]	0,832	1	0,7245	0,7138	0,4898	0,0007	0,7245	Permaneceu
306	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	40146422	Produção de Proteína	[T/G]	0,963	1	0,5606	0,6286	0,1816	0,6191	0,5606	Permaneceu
307	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	1923344590	Produção de Proteína	[A/G]	0,976	1	0	0	0	0	0	Excluído

308	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	41374052	Produção de Proteína	[T/G]	0,919	1	0,7739	0,9964	0,2745	0,8199	0,7739	Permaneceu
309	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	no_hits/BTA14	14	65511994	Produção de Proteína	[A/C]	0,991	1	0,7772	0,8152	0,2433	0,7667	0,7772	Permaneceu
310	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	NOS2/Inducible nitric oxide synthase	19	1923344596	Produção de Proteína	[G/C]	0,88	1	0	0	0	0	0	Excluído
311	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	NOS2/Inducible nitric oxide synthase	19	19958180	Produção de Proteína	[T/G]	0,566	0	0,6156	0,9511	0,1095	0,1073	0,6156	Permaneceu
312	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	ODC1/ Ornithine decarboxylase	11	87181040	Produção de Proteína	[A/G]	0,754	1	0,5768	0,9982	0,0626	0,1807	0,5768	Permaneceu
313	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	OLR1/ Oxidized low density lipoprotein receptor	5	1923344601	Produção de Gordura	[A/C]	0,943	1	0	0	0	0	0	Excluído
314	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	OPN/Osteopontina	6	12941633	Produção de Proteína	[T/C]	0,715	1	0,7896	0,9982	0,3566	0,142	0,7896	Permaneceu
315	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	OPN_607/ Osteopontina	6	1923344605	Produção de Proteína	[D/I]	0,502	0	0	0	0	0	0	Excluído
316	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	P51_239/ NW_003103978	6	87362727	Produção de Proteína	[T/A]	0,856	1	0	0	0	0,9073	0	Permaneceu
317	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19393038">http://www.ncbi.nlm.nih.gov/pubmed/19393038</a>	PI/Protease Inhibitor alpha-1-antitrypsin	21	3088873	Produção de Proteína	[A/G]	0,689	1	0,7496	0,7808	0,3422	0	0,7496	Permaneceu
318	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19393038">http://www.ncbi.nlm.nih.gov/pubmed/19393038</a>	PI/Protease Inhibitor alpha-1-antitrypsin	21	1923344611	Produção de Proteína	[T/C]	0,409	0	0,5627	0,7844	0,4007	0	0,5627	Excluído
319	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19393038">http://www.ncbi.nlm.nih.gov/pubmed/19393038</a>	PI/Protease Inhibitor alpha-1-antitrypsin	21	17927012	Produção de Proteína	[C/G]	0,466	0	0,6098	0,9547	0,3491	0,1649	0,6098	Permaneceu
320	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19393038">http://www.ncbi.nlm.nih.gov/pubmed/19393038</a>	PI/Protease Inhibitor alpha-1-antitrypsin	21	1923344614	Produção de Proteína	[T/C]	0,753	1	0	0	0	0	0	Excluído
321	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19393038">http://www.ncbi.nlm.nih.gov/pubmed/19393038</a>	PI/Protease Inhibitor alpha-1-antitrypsin	21	1923344616	Produção de Proteína	[A/C]	0,945	1	0,7703	0,9801	0,3392	0	0,7704	Excluído
322	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	PIGR/Polymeric immunoglobulin receptor	16	2021717105	Produção de Proteína	[A/G]	0,717	1	0,8273	0,8442	0,2049	0	0,8273	Excluído
323	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_1013/ polycystic kidney disease 2	6	12882278	Doença Genética	[A/G]	0,492	0	0,8791	0,8859	0,3558	0,0102	0,8791	Permaneceu
324	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_1175/ polycystic kidney disease 2	6	2021717110	Doença Genética	[I/D]	0,719	1	0	0	0	0	0	Excluído

325	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_1241/ polycystic kidney disease 2	6	12871456	Doença Genética	[T/C]	0,819	1	0,6703	0,9837	0,0276	0	0,7056	Permaneceu
326	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_1349/ polycystic kidney disease 2	6	12858905	Doença Genética	[A/T]	0,675	1	0,7978	0,9928	0,0228	0,8376	0,8398	Permaneceu
327	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_1451/ polycystic kidney disease 2	6	2021717115	Doença Genética	[A/G]	0,45	0	0,8736	0,8152	0,2333	0	0,874	Excluído
328	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_2759/ polycystic kidney disease 2	6	12872974	Doença Genética	[A/G]	0,985	1	0,5273	0,9982	0,0073	1	0,7533	Permaneceu
329	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_349/ polycystic kidney disease 2	6	2021717120	Doença Genética	[A/C]	0,958	1	0	0	0	0	0	Excluído
330	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_353/ polycystic kidney disease 2	6	12859901	Doença Genética	[T/C]	0,783	1	0,8263	0,9493	0,396	0,2736	0,8263	Permaneceu
331	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_377/ polycystic kidney disease 2	6	12870592	Doença Genética	[T/A]	0,928	1	0,6201	0,8768	0,1116	0,0004	0,6201	Permaneceu
332	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_383/ polycystic kidney disease 2	6	12863683	Doença Genética	[T/C]	0,768	1	0,6207	0,9203	0,0925	0,0332	0,6207	Permaneceu
333	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_3909/ polycystic kidney disease 2	6	12874124	Doença Genética	[A/T]	0,992	1	0,7152	0,971	0,2957	0,0031	0,7152	Permaneceu
334	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_447/ polycystic kidney disease 2	6	2021717129	Doença Genética	[A/G]	0,962	1	0,6191	0,7935	0,1553	0	0,6191	Excluído
335	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_597/ polycystic kidney disease 2	6	12882694	Doença Genética	[A/G]	0,599	0	0,4549	0,9275	0,0273	0,7046	0,6499	Permaneceu
336	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_610/ polycystic kidney disease 2	6	2021717133	Doença Genética	[A/T]	0,845	1	0	0	0	0	0	Excluído
337	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_611/ polycystic kidney disease 2	6	2021717134	Doença Genética	[T/C]	0,901	1	0,7371	0,6884	0,4829	0	0,8246	Excluído
338	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_650/ polycystic kidney disease 2	6	12859604	Doença Genética	[A/G]	0,937	1	0,5509	1	0,0498	0,9778	0,5509	Permaneceu
339	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_901/ polycystic kidney disease 2	6	2021717138	Doença Genética	[T/C]	0,564	0	0,6507	0,8062	0,4742	0	0,6507	Excluído
340	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_953/ polycystic kidney disease 2	6	2021717140	Doença Genética	[A/G]	0,922	1	0	0	0	0	0,093	Excluído

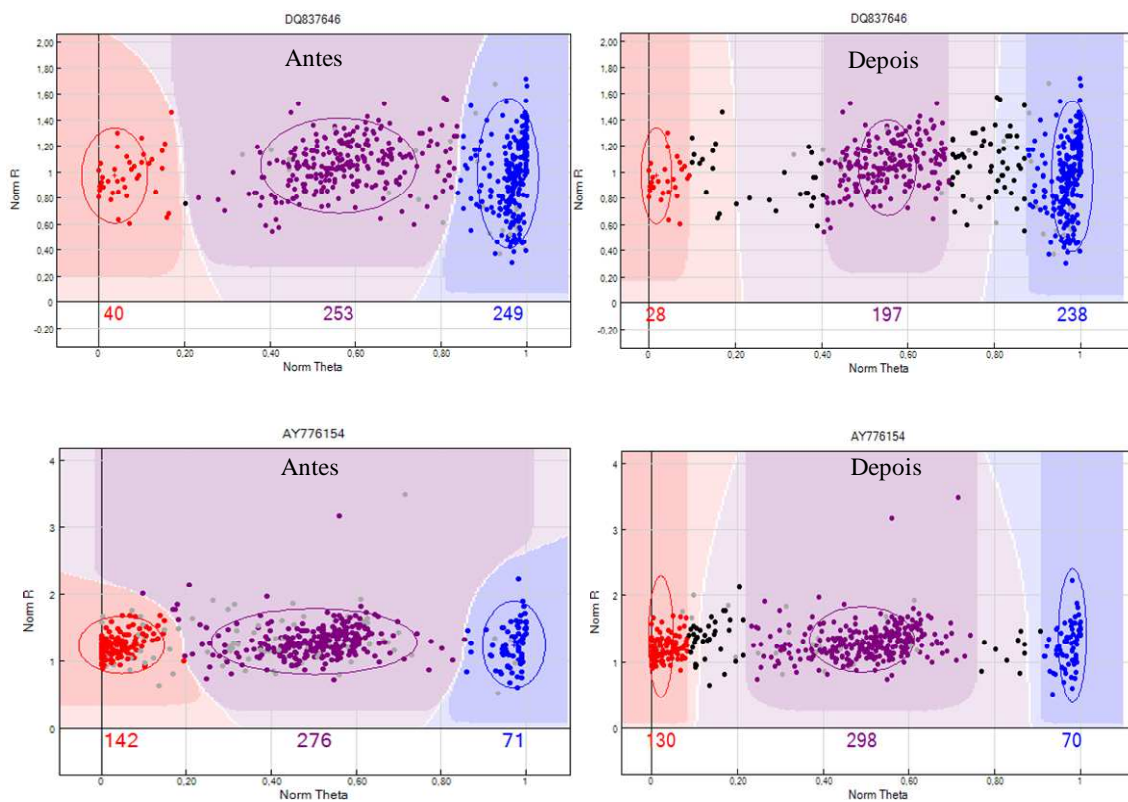


	471-2156-8-32.pdf													
341	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924865/pdf/1471-2156-8-32.pdf</a>	PKD2_97141/ polycystic kidney disease 2	6	12874405	Doença Genética	[T/C]	0,899	1	0,5753	0,9982	0,0917	0,2371	0,5754	Permaneceu
342	<a href="http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf">http://repository.library.ualberta.ca/dspace/bitstream/10048/494/1/Ferreira+Marques_Elisa_Fall+2009.pdf</a>	PKIA/BTA14	14	43920557	Produção de Proteína	[A/C]	0,889	1	0,754	0,9801	0,3983	0,0011	0,754	Permaneceu
343	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	PLAT/Tissue plasminogen activator	27	37414719	Produção de Proteína	[A/G]	0,742	1	0,736	0,9909	0,3108	0	0,736	Permaneceu
344	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741674">http://www.ncbi.nlm.nih.gov/pubmed/15741674</a>	POU1F1/ Pit1 transcription factor	1	35009085	Produção de Proteína	[A/G]	0,757	1	0,6877	0,9946	0,184	0,2055	0,6878	Permaneceu
345	<a href="http://physiolgenomics.physiology.org/content/21/1/1.full.pdf+html">http://physiolgenomics.physiology.org/content/21/1/1.full.pdf+html</a>	PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator	6	44855003	Produção de Proteína	[A/G]	0,876	1	0,7449	0,9112	0,2137	0,0364	0,7449	Permaneceu
346	<a href="http://physiolgenomics.physiology.org/content/21/1/1.full.pdf+html">http://physiolgenomics.physiology.org/content/21/1/1.full.pdf+html</a>	PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha	6	2021717150	Produção de Proteína	[A/G]	0,813	1	0	0	0	0	0	Excluído
347	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha	6	44875251	Produção de Proteína	[A/G]	0,937	1	0,6168	0,9402	0,1773	0,0225	0,6168	Permaneceu
348	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha	6	44875315	Produção de Proteína	[G/A]	0,616	1	0,754	0,9964	0,26	0,0068	0,754	Permaneceu
349	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha	6	44876517	Produção de Proteína	[A/C]	0,778	1	0,5081	0,9783	0,0194	0,9771	0,7259	Permaneceu
350	<a href="http://www.ncbi.nlm.nih.gov/pubmed/2395650">http://www.ncbi.nlm.nih.gov/pubmed/2395650</a>	PRL/Prolactina	23	31343712	Produção de Proteína	[A/G]	0,888	1	0,7176	0,9457	0,2567	0,1062	0,7176	Permaneceu
351	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	PRL/Prolactina	23	2021717161	Produção de Proteína	[T/C]	0,894	1	0	0	0	0	0	Excluído
352	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	PRL/Prolactina	23	2021717163	Produção de Proteína	[A/G]	0,51	0	0,716	0,8243	0,2692	0	0,716	Excluído
353	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	PRL/Prolactina	23	35105336	Produção de Proteína	[T/C]	0,528	0	0,7361	0,9928	0,4681	0,0014	0,7361	Permaneceu
354	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19719788">http://www.ncbi.nlm.nih.gov/pubmed/19719788</a>	PRL/Prolactina	23	35106256	Produção de Proteína	[T/C]	0,787	1	0,7331	0,9583	0,2977	0,0158	0,7331	Permaneceu

355	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	PRLR/ Prolactin Receptor	20	2021717167	Produção de Proteína	[A/C]	0,867	1	0,4751	0,6667	0,1658	0	0,4751	Excluído
356	<a href="http://www.journalofdairyscienc.e.org/article/S0022-0302(07)71908-2/abstract">http://www.journalofdairyscienc.e.org/article/S0022-0302(07)71908-2/abstract</a>	SCD/Stearoyl-CoA desaturase variant A (SCD)	26	2021717169	Produção de Gordura	[T/C]	0,863	1	0	0	0	0	0	Excluído
357	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	SDHA/Succinate dehydrogenase flavoprotein subunit A	20	2021717172	Produção de Proteína	[G/C]	0,517	0	0,3469	0,8243	0,233	0	0,514	Excluído
358	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	SDHA/Succinate dehydrogenase flavoprotein subunit A	20	17837650	Produção de Proteína	[T/C]	0,654	1	0,7333	0,9076	0,1168	0,3657	0,7333	Permaneceu
359	<a href="http://www.biomedcentral.com/content/pdf/1471-2164-9-151.pdf">http://www.biomedcentral.com/content/pdf/1471-2164-9-151.pdf</a>	SERPINA1/Serpin peptidase inhibitor, clade A, member 1	14	2021717104	-	[T/G]	0,794	1	0	0	0	0	0	Excluído
360	<a href="http://download.journals.elsevierhealth.com/pdfs/journals/0022-0302/PIIS0022030206724912.pdf">http://download.journals.elsevierhealth.com/pdfs/journals/0022-0302/PIIS0022030206724912.pdf</a>	STAT1/signal transducer and activator of transcription 1	2	79888605	Produção de Proteína	[A/G]	0,663	1	0,768	0,9891	0,2573	0,0096	0,768	Permaneceu
361	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	STAT5A/Signal transducer and activator of transcription 5A	19	43050422	Produção de Proteína	[T/A]	0,693	1	0,5343	0,9946	0	1	0,7632	Permaneceu
362	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15741668">http://www.ncbi.nlm.nih.gov/pubmed/15741668</a>	STAT5A/Signal transducer and activator of transcription 5A	19	2021717179	Produção de Proteína	[T/C]	0,585	0	0,7984	0,8388	0,3337	0	0,7984	Excluído
363	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	STATH_1754/NW_003103978	6	87231324	Produção de Proteína	[T/C]	0,775	1	0,7602	0,9149	0,2436	0,0237	0,7602	Permaneceu
364	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	STATH_2068/NW_003103978	6	2021717183	Produção de Proteína	[A/G]	0,5	0	0,8722	0,7935	0,2409	0	0,8722	Excluído
365	<a href="http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf">http://www.gsejournal.org/content/pdf/1297-9686-41-24.pdf</a>	STATH_BMC2_1578/NW_003103978	6	2021717185	Produção de Proteína	[A/G]	0,957	1	0	0	0	0	0	Excluído
366	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20449817">http://www.ncbi.nlm.nih.gov/pubmed/20449817</a>	TF/ Transferrin	1	2021717187	Produção de Gordura	[A/G]	0,894	1	0	0	0	0	0	Excluído
367	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20449817">http://www.ncbi.nlm.nih.gov/pubmed/20449817</a>	TF/ Transferrin	1	136709164	Produção de Gordura	[T/C]	0,914	1	0,7785	0,9891	0,3269	0,0006	0,7785	Permaneceu
368	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	2021717191	Produção de Proteína	[A/G]	0,484	0	0	0	0	0	0	Excluído
369	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	108834070	Produção de Proteína	[A/G]	0,823	1	0,7512	0,8732	0,0114	0,9809	0,7908	Permaneceu
370	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	108837080	Produção de Proteína	[C/G]	0,944	1	0,5438	0,9964	0,0082	1	0,7769	Permaneceu

371	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	108837360	Produção de Proteína	[T/C]	0,86	1	0,7207	0,9438	0,0393	0,7417	0,7586	Permaneceu
372	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	108837700	Produção de Proteína	[T/C]	0,866	1	0,868	0,8641	0,2044	0,5045	0,868	Permaneceu
373	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	108837800	Produção de Proteína	[A/G]	0,895	1	0,5801	1	0,0009	1	0,8287	Permaneceu
374	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	108838610	Produção de Proteína	[A/C]	0,964	1	0,5356	0,9928	0,0602	0,2148	0,5356	Permaneceu
375	<a href="http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf">http://www.biomedcentral.com/content/pdf/1471-2156-11-99.pdf</a>	TLR4/ Toll-like receptor 4 precursor	8	2021717204	Produção de Proteína	[A/C]	0,934	1	0,7431	0,8424	0,0699	0	0,7822	Excluído
376	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TNF/Prolactina	23	35110924	Produção de Proteína	[A/G]	0,989	1	0,4358	0,9855	0,3796	0	0,4358	Permaneceu
377	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TNF/Prolactina	23	2021717208	Produção de Proteína	[T/C]	0,587	0	0	0	0	0	0,2379	Excluído
378	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TYR/Tyrosinase	29	26907320	Produção de Proteína	[T/C]	0,405	0	0,3896	0,8514	0,0202	0,7342	0,5566	Permaneceu
379	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TYR/Tyrosinase	29	28647916	Produção de Proteína	[A/G]	0,978	1	0,6653	0,9656	0,4418	0	0,6653	Permaneceu
380	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TYRP1/tyrosinase-relatedprotein1	8	31710798	Produção de Proteína	[G/C]	0,91	1	0,7614	0,9982	0,1325	0,0048	0,7614	Permaneceu
381	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TYRP1/tyrosinase-relatedprotein1	8	31717680	Produção de Proteína	[T/C]	0,942	1	0,7667	0,9384	0,4402	0,0899	0,7667	Permaneceu
382	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	TYRP1/tyrosinase-relatedprotein1	8	31721357	Produção de Proteína	[T/C]	0,726	1	0,8129	0,9275	0,4375	0,0569	0,8129	Permaneceu
383	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	UA_IFASA_2295	11	97754170	Produção de Proteína	[A/G]	0,819	1	0,7449	0,9928	0,4234	0,0063	0,7449	Permaneceu
384	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>	UA_IFASA_6966	14	2021717399	-	[T/C]	0,835	1	0	0	0	0	0	Excluído

ID = Identificação do SNP; BTA = Cromossomo; GC50 = GenCall a 50%; CallFreq = Frequência de Genotipagem; MAF = Frequência do Alelo Menor; HW = Frequência do Equilíbrio de Hardy e Weinberg.

**Anexo 02: Exemplo do efeito da edição dos clusters sobre a dispersão dos genótipos.**

## 2.8. REFERÊNCIAS BIBLIOGRÁFICAS

- ANKUNOV, E.; NICOLET, C.; DVORAK, J. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. **Theoretical and Applied Genetics**, v. 119, n. 3, p. 507–517, 2009.
- BAIRD, N. A.; ETTER, P.D.; ATWOOD, T.S. et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers. **PLoS ONE**, v. 3, n. 10, p. e3376, 2008.
- BIONAZ, M.; LOOR, J.J. Gene Networks Driving Bovine Mammary Protein Synthesis During the Lactation Cycle. **Bioinform Biol Insights**, v. 5, p. 83–98, 2011.
- CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v. 38, n. spe, p. 64-71, 2009.
- CHAKRAVARTY K.; CASSUTO, H.; RESHEF, L. et al. Factors that control the tissue-specific transcription of the gene for phosphoenolpyruvate carboxykinase-C. **Critical Reviews in Biochemistry and Molecular Biology**, v. 40, n. 3, p. 129–154. 2005.
- CHATTERJEE, S.; SZUSTAKOWSKI, J.D.; NANGUNERI, N.R. et al. Identification of Novel Genes and Pathways Regulating SREBP Transcriptional Activity. **PLoS ONE**, v. 4, n. 4, p. e5197, 2009.
- CHESSA, S.; CHIATTI, F.; CERIOTTI, G. et al. Development of a Single Nucleotide Polymorphism Genotyping Microarray Platform for the Identification of Bovine Milk Protein Genetic Polymorphisms. **Journal of Dairy Science**, v. 90, n. 1, p. 451 – 464, 2007.
- COHEN-ZINDER, M.; SEROUSSI, E.; LARKIN, D. M., et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. **Genome Research**, v. 15, n. 7, p. 936-944, 2005.
- DEULVOT, C.; CHARREL, H.; MARTY, A. et al. Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. **BMC Genomics**, v. 11, n. 468, p. 1-10, 2010.

- ELSHIRE, R.J.; GLAUBITZ, J.C.; SUN, Q. et al. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. **PLoS One**, v. 6, n. 5, p. e19379, 2011.
- ELSIK, C.G.; TELLAM, R.; WORLEY, K.C. et al. The Bovine Genome Sequencing and Analysis Consortium. The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution, **Science**, v. 324, n. 5926, p. 522-528, 2009.
- FAN, B.; ONTERU, S.K.; DU, Z.Q. et al. Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs. **PLoS One**, v. 6, n. 2, p.e14726, 2011.
- FAN, J.B.; OLIPHANT, A.; SHEN, R. et al. Highly parallel SNP genotyping. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 68, p. 69-78, 2003.
- FISHER, P.J.; MALTHUS, B.; WALKER, M.C. et al. The number of single nucleotide polymorphisms and on-farm data required for whole-herd parentage testing in dairy cattle herds. **Journal of Dairy Science**, v. 92, p. 369-374, 2009.
- GONZÁLEZ-NEIRA, A. The GoldenGate genotyping assay: custom design, processing, and data analysis. **Methods in Molecular Biology**, v. 1015, n. 5, p. 147-153, 2013.
- GRATTAPAGLIA, D.; ORZENIL, B.S.J.; KIRST, M. et al. High-throughput SNP genotyping in the highly heterozygous genome of Eucalyptus: assay success, polymorphism and transferability across species. **BMC Plant Biology**, v.11, n. 65, p. 1-18, 2011.
- GRISART, B.; FARNIR, F.; KARIM, L. et al. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition **Proceedings National Academy of Sciences**, v. 101, n. 8, p. 2398–2403, 2004.
- HARA, K.; WATANABE, S.; MUKAI, H.; MANNEN, H. Development of SNP markers for individual identification and parentage test in a Japanese Black cattle population. **Journal of animal Science**, v. 81, p. 152-157, 2010.
- HEATON, M.P.; HARHAY, G.P.; BENNETT, G.L. et al. Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. **Mammalian Genome**, v. 13, n. 5, p. 272–281, 2002.
- HEDRICK, P.W. **Genetics of populations**. 4ed. Sudbury,MA: Jones and Bartlett Publishers, 2010. 675 p.
- HUANG, D.W.; SHERMAN, B.T.; LEMPICKI, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. **Nucleic Acids Research**, v. 37, n. 1, p. 1–13, 2009.
- JACOB, E.; PUHAN, Z. Technological properties of milk as influenced by genetic polymorphism of milk proteins (a review). **International Dairy Journal**, v. 2, n. 3, p.157–178, 1992.
- JIANG, L.; LIU, J.; SUN, D. et al. Genome Wide Association Studies for Milk Production Traits in Chinese Holstein Population. **PLoS ONE**, v.5, n. 10, p. e13661, 2010.
- KAMINSKI, S.; AHMAN, A.; RUSC, A. et al. MilkProtChip—A microarray of SNPs in candidate genes associated with milk protein biosynthesis—Development and validation. **Journal of Applied Genetics**, v. 46, n. 1, p.45–58. 2005

- LAURIE, C.C.; DOHENY, K.F.; MIREL, D.B. et al. Quality control and quality assurance in genotypic data for genome-wide association studies. **Genetic Epidemiology**, v. 34, n. 6, p. 591-602, 2010.
- LIEN, K.Y.; CHUANG, Y.H.; HUNG, L.Y. et al. Rapid isolation and detection of cancer cells by utilizing integrated microfluidic systems. **Lab Chip**, v. 10, n. 21, p. 2875–2886, 2010.
- LIN, C.H.; YEAKLEY, J.M.; MCDANIEL, T.K. et al. Medium- to High-Throughput SNP Genotyping Using VeraCode Microbeads. **Methods in Molecular Biology**, v. 496, p. 129-142, 2009.
- MATUKUMALLI, L.K.; LAWLEY, C.T.; SCHNABEL, R.D. et al. Development and characterization of a high density SNP genotyping assay for cattle. **PLoS One**, v. 4, n. 4, p. e5350, 2009.
- MELLO, F.; COBUCCI, J.A.; MARTINS, M.F. et al. Association of the polymorphism g.8514C>T in the osteopontin gene (SPP1) with milk yield in the dairy cattle breed Girolando. **Animal Genetics**, v. 43, n. 5, p. 647–648, 2012.
- MYLES, S.; CHIA, J.M.; HURWITZ, B. et al. Rapid genomic characterization of the genus *vitis*. **PLoS One**, v. 13, n. 5, p. e8219, 2010.
- ONTERU, S.K.; FAN, B.; NIKKILA, et al. Whole-genome association analyses for lifetime reproductive traits in the pig. **Journal of Animal Science**, v. 89, n. 4, p. 988-995, 2011.
- R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. Vienna, 2012. Disponível em: <<http://www.R-project.org>>.
- RAMSAY, G. DNA chips: State of the art. **Nature Biotechnology**, v. 16, n. 1, p. 40–44, 1998.
- RAYMOND, M.; ROUSSET, F. GENEPOP (version 1.2) population genetic software for exact tests and ecumenicism. **Journal of Heredity**, v.86, n. 3, p. 248-249, 1995.
- REIS FILHO, J.C.; LOPES, P.S.; VERNEQUE, R.S. et al. Population structure of Brazilian Gyr dairy cattle. **Revista Brasileira de Zootecnia**, v. 39, p. 2640-2645, 2010
- SAS Institute Inc. SAS/STAT User's Guide: ver. 9.22. SAS Institute Inc; Cary, NC, USA: 2010.
- SCHENNINK, A.; HECK J.M.L.; BOVENHUIS H. et al. Milk fatty acid unsaturation: genetic parameters and effects of stearoyl-CoA desaturase (SCD1) and acyl CoA: diacylglycerol acyltransferase 1 (DGAT1). **Journal of Dairy Science**, v. 91, n. 5, p. 2135–2143, 2008.
- SIGNORELLI, F.; ORRU, L.; NAPOLITANO, F. et al. Exploring polymorphisms and effects on milk traits of the DGAT1, SCD1 and GHR genes in four cattle breeds. **Livestock Science**, v.125, n. 2, p.74–79, 2009.
- SOUZA, P.R.; BUAINAIN, A.M. O ambiente tecnológico e a competitividade da agricultura familiar produtora de leite em Passo Fundo e Região (Brasil) **Espacios**, v. 34, n. 10, p.16-36, 2013.
- WANG, X.; WURMSER, C.; PAUSCH, H. et al. Identification and Dissection of Four Major QTL Affecting Milk Fat Content in the German Holstein-Friesian Population. **PLoS ONE**, v. 7, n. 7, p. e40711, 2012.

WOSZUK, J.N.; NOSKOWIAK, A.; STRABEL, T. et al. An effect of the DGAT1 gene polymorphism on breeding value of Polish Holstein-Friesian sires. **Animal Science Papers and Reports**, v. 26, n. 1, p.17-23, 2008.

ZIEGLER, A. Genome-Wide Association Studies: Quality Control and Population-Based Measures. **Genetic Epidemiologic**, v.33, n. Suppl 1, p. S45–S50, 2009.



### **CAPÍTULO 3**

**ESTIMATIVAS DE DESEQUILÍBRIO DE LIGAÇÃO E AVALIAÇÕES DA  
ESTRUTURA DE BLOCOS HAPLOTÍPICOS EM GENES CANDIDATOS NOS  
CROMOSSOMOS BTA06 E BTA14**

### 3.1. RESUMO

Objetivou-se com o presente capítulo, estimar o desequilíbrio de ligação e identificar a estrutura dos blocos de haplótipos utilizando SNPs presentes em genes candidatos nos cromossomos BTA06 e BTA14 em animais da raça Girolando. Foram utilizados para a análise, 419 animais (118 machos e 301 fêmeas) de grupos genéticos diferentes da raça Girolando e 249 SNPs. Deste painel, 72 e 47 SNPs, localizados nos BTA06 e BTA14, respectivamente, foram utilizados para os estudos de desequilíbrio de ligação (DL) e construção de blocos de haplótipos. Para estimar o nível de DL foram utilizados os parâmetros de desvio esperado ( $D'$ ) e o coeficiente de determinação ( $r^2$ ). Tanto estes parâmetros quanto as estatísticas descritivas para cada par de SNP foram calculados utilizando os programas *gPlink* e as figuras dos blocos de haplótipos e frequências haplotípicas foram obtidos através do programa *Haploview* v 4.2. Considerando os cromossomos BTA06 e BTA14, o parâmetro  $D'$  variou de 0,83 a 1,00 e 0,88 a 1,00, respectivamente. Para o parâmetro  $r^2$ , as estimativas oscilaram entre 0,11 a 0,85 (BTA06) e 0,04 a 0,59 (BTA14). Em relação ao BTA14, as estimativas médias para  $D'$  e  $r^2$  foram de  $0,29 \pm 0,27$  e  $0,02 \pm 0,05$ , respectivamente. Um total de 24 SNPs do BTA06 foram excluídos das análises de DL com a utilização dos seguintes critérios: SNPs fora do equilíbrio de Hardy-Weinberg, com frequência genotípica inferior a 25% e com frequência alélica inferior a 1%. Assim, dos 72 polimorfismos genotipados no BTA06, somente 48 SNPs foram utilizados para os cálculos de DL e para construção dos blocos haplotípicos. Observou-se na população estudada a existência de 32 haplótipos em seis regiões do BTA06, com frequências

observadas variando de 0,011 a 0,970. Houve também, grande variação das frequências esperadas em relação às frequências observadas dos haplótipos. O número de SNPs capturados (*TagSNPs*) no BTA6 variaram de 2 a 4 entre as regiões estruturadas e a distância entre eles variou de 75 a 70.177 pb. Dos 47 polimorfismos genotipados no BTA14, 33 SNPs foram utilizados para os cálculos de DL e para construção dos blocos haplotípicos. A estimativa de DL revelou a existência de 13 haplótipos em 4 regiões deste cromossomo, com frequências observadas variando de 0,023 à 0,589. Em geral, o número de SNPs capturados entre as regiões estruturadas no BTA14 foi igual a 2 (o próprio *TagSNP* e o SNP adjacente) e a distância entre eles variou de 225 a 17.284 pb, o que reduz em 50% os esforços no processo de genotipagem para cada uma das regiões gênicas estruturadas. Estes resultados permitiram a caracterização de estruturas de blocos de haplótipos, em regiões genômicas relacionados com características de interesse zootécnico em animais da raça Girolando. Este resultado torna viável sua utilização em estudos de associação entre essas regiões com características produtivas.

**PALAVRAS-CHAVE:** Diversidade genética, rebanhos mestiços, chips personalizados, seleção genômica ampla, ensaio goldengate.

### 3.2. ABSTRACT

The aim of this study was to estimate the linkage disequilibrium and identify the structure of haplotype blocks using SNPs present in candidate genes in BTA06 and BTA14 chromosomes in animals Girolando breed. Were used for the analysis, 419 animals (118 males and 301 females) of different genetic groups in Girolando breed and 249 SNPs. In this panel, 72 and 47 SNPs located in BTA06 and BTA14, respectively, were used for the linkage disequilibrium studies (LD) and construction of haplotype blocks. To estimate the LD level were utilized the parameters expected deviation ( $D'$ ) and the coefficient of determination ( $r^2$ ). Both these parameters as the descriptive statistics for each pair of SNP were calculated using the *gPLink* programs and the figures of haplotype blocks and haplotype frequencies were obtained through the program *Haploview* v 4.2. Considering the BTA06 and BTA14 chromosomes, the parameter  $D'$  ranged from 0.83 to 1.00 and 0.88 and 1.00 respectively. For the parameter  $r^2$ , estimates ranged from 0.11 to 0.85 (BTA06) and 0.04 to 0.59 (BTA14). In relation to BTA14, the average estimates for  $D'$  and  $r^2$  was  $0.29 \pm 0.27$  and  $0.02 \pm 0.05$ , respectively. A total of 24 SNPs were excluded from the BTA06 LD analyzes using the following criteria: SNPs out of Hardy-Weinberg equilibrium, with genotypic frequencies below 25% and allele frequency less than 1%. Thus, of the 72 polymorphisms were genotyped in BTA06, only 48 SNPs were used for the calculations LD and construction of haplotypic blocks. It was observed in the study population the existence of 32 haplotypes in six regions of BTA06, with observed frequencies ranging from 0.011 to 0.970. There was also wide variation in expected frequencies in relation to the observed frequencies of the

haplotypes. The number of SNPs captured (*TagSNPs*) in BTA6 ranging from 2 to 4 between the structured regions and the distance between them ranging 75-70177 pb. Of the 47 polymorphisms genotyped in BTA14, 33 SNPs were used for the LD calculations and construction of haplotypic blocks. The estimated LD revealed the existence of 13 haplotypes in four regions of this chromosome with observed frequencies ranging from 0.023 to 0.589. Overall, the number of SNPs captured between the structured regions in BTA14 was equal to 2 (the own *TagSNP* and the adjacent SNP) and the distance between them ranging 225-17284 pb, which reduces by 50% the efforts in the process of genotyping for each structured gene regions. These results allowed, even though on a small scale, the characterization of haplotype blocks in genomic regions related to traits of interest in Girolando animals. This result can be used in association studies between these regions and production traits.

**KEY-WORDS:** crossbred cattle, custom chips, genetic diversity, genome wide selection, goldengate assay.

### 3.3. INTRODUÇÃO

A indústria leiteira do Brasil representa um importante segmento do agronegócio brasileiro e tem apresentado crescimento contínuo e crescente, produzindo em 2009 mais de 29 bilhões de litros de leite (USDA, 2009). As projeções divulgadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE), estima que em 2014 a produção de leite no Brasil seja de 36,75 bilhões de litros de leite, valor este que corresponde a um aumento de 5% em comparação aos 35 bilhões de litros de leite produzidos em 2013 (IBGE, 2014). Com base nestas informações, o Brasil destaca-se entre os cinco maiores produtores mundial de leite. Apesar dos consideráveis avanços observados no arranjo produtivo do leite nos últimos anos, aspectos associados à melhorias da qualidade do produto, bem como a produtividade dos rebanhos, ainda representam desafios a serem superados.

Para maior produtividade da indústria leiteira no Brasil, cruzamentos entre animais de raças taurinas (*Bos taurus*) e zebuínas (*Bos indicus*) têm sido amplamente utilizados em programas de avaliação e melhoramento genético (Pereira, 2008). Como exemplo principal temos a raça mestiça Girolando (Nanzer, 2010), a qual apresenta alta produção de leite e é adaptada ao clima tropical (Facó et al., 2002; McManus et al., 2008).

Neste contexto, para que se possa melhorar o agronegócio leiteiro no Brasil, são necessárias melhorias não apenas na produtividade, mas também na qualidade do leite produzido. Assim, além de estudos genético-quantitativos, estudos genômicos, com foco na identificação de genes ou blocos gênicos que controlam ou influenciam características de importância econômica são importantes. Diferentes estudos detectaram QTLs para produção

de leite, de teor de gordura e de proteína em todos os 29 cromossomos autossômicos do genoma bovino (Olsen et al., 2007; Khatkar et al., 2008; Sodeland et al., 2011). Logo, estudos de associação genômica e de desequilíbrio de ligação, principalmente em genes candidatos, podem agregar resultados positivos e satisfatórios para alavancar a pecuária leiteira no Brasil.

Segundo Altshuler et al. (2005), há probabilidade de que indivíduos portadores de determinado alelo em um loco também apresentem alelos específicos em loci próximos. Esta associação é denominada desequilíbrio de ligação (DL), e pode ser representada pela estimativa de correlação entre alelos em loci diferentes. A partir do surgimento de tecnologias de genotipagem de marcadores em massa do tipo SNP (*Single Nucleotide Polymorphism*) em animais de interesse zootécnico, estudos de DL ganharam grande importância para realização de análises de mapeamento genético (Amaral et al., 2008). Segundo Marques et al. (2008), estes marcadores são fundamentais no processo de verificação do DL, na construção de mapas de haplótipos e em pesquisas sobre genes candidatos.

Diversos trabalhos (McKay et al., 2007; Bohmanova et al., 2010; Lu et al., 2012; Espigolan et al., 2013) têm sido realizados sobre o DL em bovinos, entretanto este termo propriamente dito foi empregado pela primeira vez por Lewontin & Kojima (1960). Esta denominação foi utilizada por ter sido constatada que a ligação entre locos era essencial ao equilíbrio da população. Conforme Britto (2010), o DL reflete a habilidade dos alelos de um locus prever a presença de alelos em outros *loci*. Para Almeida et al. (2013) a compreensão dos padrões de DL representa a base para o mapeamento de genes e planejamento de estudos de associação.

O DL em uma população pode ser afetado por fatores como tamanho efetivo ( $N_e$ ), taxa de mutação, migração, recombinação e seleção. O DL tende a se reduzir com o aumento da ocorrência de durante a formação dos gametas (Amaral et al., 2008). Conforme Bohmanova et al. (2010), o conhecimento do padrão de DL no genoma bovino é importante no arranjo genético e nos estudos de associação genômica. Trata-se de uma informação fundamental para inferir sobre a distribuição de *crossing-over* (que ocorre quando existe troca de segmentos de cromossomos homólogos) e identificar regiões estruturadas no genoma, em comparação ao grau de diversidade entre as raças. Vallejo et al. (2003) afirmaram que as diferenças nos marcadores utilizados em estudos desta natureza são causas potenciais de constatação de DL.

Diferentes medidas podem ser utilizadas para estimar DL em uma população. A estatística  $D'$  concede parâmetros mais exatos para a verificação de diferenças de eventos

de recombinação, apesar de sofrer interferência em situações em que a amostra é reduzida. Já a estatística  $r^2$  favorece a confirmação do histórico de recombinação e indica como os marcadores correlacionam-se com as regiões cromossômicas que apresentam efeito na expressão de características de interesse econômico, ou seja, que contém QTLs (Du et al., 2007).

Outra questão abordada neste estudo foi a identificação de estruturas de blocos de haplótipos em genes candidatos, que apresentam aplicações importantes no delineamento e na análise de experimentos genéticos (Wade et al., 2002). Tal identificação aponta a descrição de padrões comuns na variação genética dos genomas, cujas aplicações são essenciais para o delineamento e análise adequados de experimentos genéticos (Khatkar et al., 2008). Dentro de um bloco haplotípico acontece pouca ou nenhuma recombinação e os SNPs situados nesta região podem ser herdados juntos em gerações seguintes. Também nessas regiões, um subconjunto pequeno de SNPs (*TagSNPs*) se faz suficiente na distinção de cada um dos haplótipos. Nesse sentido, para genotipagem de um bloco haplotípico, não há necessidade de usar todos os SNPs, usam-se apenas os *TagSNPs*.

Na identificação de estruturas de blocos de haplótipos em genes candidatos, há alguns elementos que necessitam de atenção (Guryev et al., 2006). Dentre tais elementos, a recombinação responde pela delimitação entre início e fim dos blocos de haplótipos. Já as informações referentes ao histórico da população determinam a extensão dos blocos de haplótipos. Em se tratando de amostra de indivíduos de uma população, é possível inferir os haplótipos a partir dos dados de DL e frequência alélica.

Em bovinos, o cromossomo 06 têm sido relacionados a importantes QTLs com efeito sobre características de produção de leite (Khatkar et al., 2008; Olsen et al., 2007). Já para o cromossomo 14, em pesquisa realizada com touros da raça holandesa, Marques et al. (2008) ressaltam sua relevância sobre características quantitativas, que refletem em efeitos sobre o percentual de gordura no leite, bem como sobre a produção de leite. Portanto, objetivou-se com o presente estudo, estimar o desequilíbrio de ligação e identificar a estrutura dos blocos de haplótipos utilizando SNPs presentes em genes candidatos nos cromossomos BTA06 e BTA14 em animais da raça Girolando.



### 3.4. MATERIAL E MÉTODOS

#### 3.4.1 Amostras e Seleção de SNPs em genes candidatos

Inicialmente, um conjunto de 384 SNPs foi obtido a partir da mineração de artigos publicados na literatura científica, bancos de dados públicos (*GeneBank*) e com base em dados gerados pela Embrapa, para compor o painel a ser validado. Os SNPs selecionados foram derivados de genes candidatos relacionados principalmente à produção e qualidade do leite, a doenças hereditárias e ideais para realização de testes de paternidade. Neste trabalho, foram testadas 510 amostras de animais da raça Girolando (173 machos e 337 fêmeas).

A genotipagem dos SNPs foi realizada com um equipamento *BeadXpress®* Illumina (Illumina, San Diego, CA), de acordo com o protocolo recomendado pelo fabricante (Fan et al., 2003). Foram utilizados três critérios para identificar e excluir os marcadores e amostras que apresentaram resultados inconsistentes: Separação de Clusters (*Cluster Separation*); Frequência de Genotipagem (*Call Frequency*) e Taxa de Genotipagem (*Call Rate*), além dos testes de paternidade e reprodutibilidade.

As frequências alélicas de todos os SNPs foram estimadas após aplicação do filtro de controle de qualidade, o qual excluiu das análises SNPs que apresentaram taxa de genotipagem inferior a 75% (Purcell et al., 2007). Após aplicação deste critério, o arquivo final foi composto por 419 animais (118 machos e 301 fêmeas) de grupos genéticos diferentes (22 eram 1/2Hol +1/2Gir; 2 eram 1/4Hol+3/4Gir; 2 eram 7/8Hol+1/8Gir; 4 eram 13/16Hol+3/16Gir; 77 eram 3/4Hol+1/4Gir; 273 eram 5/8Hol+3/8Gir; e para 39 animais não

foram encontrados a identificação) e 249 SNPs. Do painel original filtrado, 72 e 47 SNPs, localizados nos cromossomos BTA06 e BTA14, respectivamente, foram utilizados para os estudos de desequilíbrio de ligação e construção de blocos de haplótipos.

### 3.4.2 Medida do Desequilíbrio de Ligação

O Desequilíbrio de Ligação (DL) foi determinado pelo coeficiente de determinação ( $r^2$ ), sugerido por Hill & Robertson (1968). Este parâmetro representa o coeficiente de correlação no estado alélico entre os alelos no mesmo gameta (Hartl & Clark 1997), o qual é definido como:

$$r^2 = \frac{D_{ij}^2}{p_1 p_2 q_1 q_2}$$

em que:  $D_{ij} = P_{ij} - p_i q_j$ , é a medida de DL proposta por Lewontin (1964), sendo  $P_{ij}$  a frequência observada do haplótipo que apresenta de alelos i e j;  $p_i$  é a frequência do alelo i e  $q_j$  a frequência do alelo j, e  $p_1 p_2 q_1 q_2$  é o produto da frequência dos quatro alelos em dois loci.

Outro parâmetro, também utilizado neste estudo para se estimar o desequilíbrio de ligação com a finalidade de definir os blocos de haplótipos para quaisquer dois loci bialélico, foi o coeficiente de desvio padronizado ( $D'$ ), o qual é definido como (Lewontin, 1964):

$$D' = \sum_{i=1}^2 \sum_{j=2}^2 p_i q_j \frac{|D_{ij}|}{D_{max}}$$

em que:

$$D_{max} = \begin{cases} \min[p_i q_j, (1 - p_i)(1 - q_j)] & \text{se } D_{ij} < 0 \\ \min[p_i(1 - q_j), (1 - p_i)q_j] & \text{se } D_{ij} > 0 \end{cases}$$

Uma vez que o valor de  $D'$  é dependente das frequências alélicas, seus valores mínimo e máximo são assim definidos:  $D_{min}$  = maior valor de  $-p_1 p_2$  e  $-q_1 q_2$ ;  $D_{max}$  = maior valor de  $p_1 q_2$  e  $q_1 p_2$ . Por esta razão, a magnitude do desequilíbrio de ligação é descrita por  $D'$ , que é definido como:  $D' = D/D_{max}$  (quando D for positivo) e  $D' = D/D_{min}$  (quando D for negativo) (Hartl & Clark 1997).

Ambas as medidas,  $D'$  e  $r^2$  são utilizadas para descrever a magnitude do desequilíbrio de ligação porque representam diferentes aspectos das associações gaméticas. Assim, quando  $D'$  é próximo de zero,  $r^2$  também se aproxima de zero. Contudo, à medida que  $D'$  aumenta  $r^2$  pode apresentar qualquer valor entre zero e  $D'$  (Hartl & Clark 1997).

Os genótipos dentro dos blocos de haplótipos foram obtidos utilizando um algoritmo EM, conforme sugerido por Gabriel et al. (2002). Este algoritmo cria estimativas precisas da frequência dos haplótipos na população em função da probabilidade máxima determinada a partir dos genótipos que não estão em fase de ligação para cada SNP utilizado (Purcell et al., 2007). Considerando o  $D'$ , este algoritmo considera que um par de SNPs apresenta DL forte quando o DL for superior a 95% e o intervalo de confiança de  $D'$  estiver entre 0,70 e 0,98. Para estes autores os blocos de haplótipos são regiões sobre a qual uma proporção muito pequena (5%) de comparações entre pares SNP informativos mostram evidência de histórico de recombinação.

Os parâmetros  $r^2$ ,  $D'$  e as estatísticas descritivas para cada par de SNP foram calculados utilizando-se os programas *gPlink*, *Plink* (Purcell et al., 2007) e as figuras dos blocos de haplótipos e frequências haplotípicas foram obtidos através do programa *Haploview* v 4.2 (Barret et al., 2005). Para cada SNP, foram utilizados os valores *default* do *Plink* e do *Haploview* para os seguintes parâmetros: *cutoff* do *p-valor* do Equilíbrio de Hardy-Weinberg (0,001), porcentagem mínima de genótipos (75%) e frequência alélica mínima (0,001).

Após a determinação dos SNPs informativos para cada cromossomo, foram estimados para cada gene os blocos haplotípicos formados, a partir dos SNPs válidos, pela metodologia padrão do *Haploview* (Gabriel et al., 2002). Nesta metodologia, foram utilizadas informações do parâmetro  $D'$  para gerar intervalos de confiança a 95% para cada combinação de pares de SNPs. Esta combinação determinou se uma dada região apresentou “DL forte”, “DL inconclusivo” ou “Alta recombinação”. Os pares de SNPs indicam forte evidência de recombinação histórica quando o limite superior do intervalo de confiança de  $D'$  é menor do que 0,9 (Gabriel et al., 2002).

O programa *Haploview* permitiu representar o DL por meio de diferentes esquemas de cores. No mais tradicional deles, voltado para representação dos valores de  $D'$  e *LOD score*, branco indica  $LOD < 2$  e  $D' < 1$ ; azul indica  $LOD < 2$  e  $D' = 1$ ; tons de rosa/vermelho brilhante indicam  $LOD \geq 2$  e  $D' < 1$ ; e vermelho indica  $LOD \geq 2$  e  $D' = 1$ . Em um segundo esquema, voltado para a representação dos intervalos de confiança de  $D'$ , uma forte evidência de DL é representada pela cor cinza escuro enquanto que forte evidência de

recombinação é representada por branco; cinza claro indica uma comparação não informativa. No terceiro esquema, branco indica  $r^2 = 0$ , tons de cinza indicam  $0 < r^2 < 1$  e preto indica  $r^2 = 1$ .

Um valor de  $D'$  multialélico também foi calculado com o *Haploview*. Este consistiu em uma medida de DL entre dois blocos de haplótipos e representou o nível de recombinação entre os dois blocos. Quanto mais próximo de zero, maior número de recombinações históricas entre os blocos. O  $D'$  multialélico não é computado para os haplótipos raros, pois estes contribuem muito pouco para o valor geral (Gabriel et al., 2002). Com este mesmo programa foi realizada análise par a par entre os SNPs utilizados de modo a identificar os marcadores correlacionados ou em elevado DL (*TagSNPs*), extraíndo somente *TagSNPs* com  $r^2 > 0,5$ .

O declínio do desequilíbrio de ligação, em função da distância física entre pares de SNPs adjacentes foi estimado, em EXCEL, a partir de um modelo exponencial. Para ajuste deste modelo, optou-se por sua linearização por meio de uma transformação logarítmica. Assim, ajustou-se o modelo  $\ln(y) = a + b \ln(x)$ , em que  $y$  é a estimativa do desequilíbrio de ligação entre dois marcadores e  $x$ , a distância entre eles. Neste modelo, a constante  $a$  representa o valor inicial do DL e a constante  $b$  representa o valor de declínio do desequilíbrio de ligação (LDD).

### 3.5. RESULTADOS

#### 3.5.1 – Análise descritiva dos polimorfismos genotipados

O tamanho total do genoma autossômico foi de 2.512,08 Mb, com o menor comprimento localizado no cromossomo 25 (42,90 Mb) e o maior localizado no BTA1 (158,34 Mb) (Tabela 3.1). A densidade de SNPs oscilou entre os BTAs analisados, variando de 2 a 72SNP (Tabela 3.1). A distância física média dos marcadores genotipados foi de  $50,88 \pm 34,11$  Mb, variando de 0,89 a 151,35 Mb. Já a distância média entre os pares de SNPs adjacentes foi  $27,87 \pm 29,00$  Mb, variando de 0,001 Kb a 148,10 Mb (Tabela 3.1). Os parâmetros utilizados para estimar os valores de Desequilíbrio de Ligação (DL) entre os SNPs variaram entre si e entre os BTAs. Para o parâmetro  $D'$ , a menor estimativa foi obtida para o BTA15 (0,0954) e a maior estimada no BTA24 (0,5311). Por outro lado, o parâmetro  $r^2$  apresentou menor estimativa no BTA17 (0,0011) e a maior no BTA22 (0,0834).

Da mesma forma, observou-se oscilação entre as estimativas de frequência do alelo menor (MAF) entre os BTAs, variando de  $0,0828 \pm 0,0886$  a  $0,3502 \pm 0,1543$ , respectivamente para os BTAs24 e 25 (Tabela 3.1).

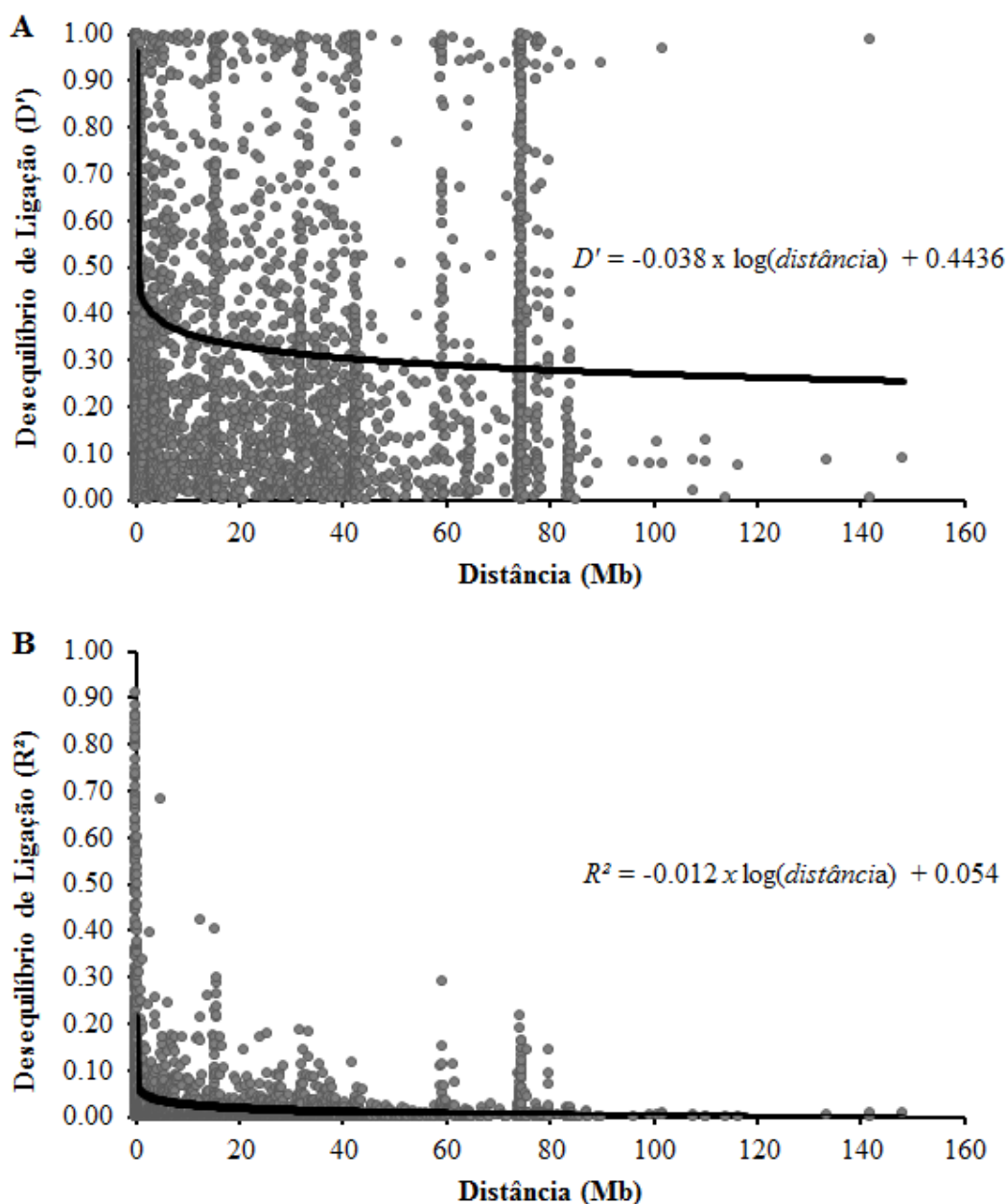
**Tabela 3.1** Estatísticas descritivas das análises dos SNPs para cada um dos 29 cromossomos autossômicos, considerando os 419 animais genotipados.

BTA	Tamanho <sup>a</sup>	SNP (n)	Distância média <sup>b</sup> (bM)	$D'$			$r^2$			MAF	
				Média	Desvio-padrão	Mediana	Média	Desvio-padrão	Mediana	Média	Desvio-padrão
1	158.34	9	75.35808	0.3583	0.3992	0.0911	0.0081	0.0173	0.0028	0.2355	0.1756
2	137.06	5	68.28342	0.2261	0.2501	0.1422	0.0123	0.0166	0.0072	0.2231	0.0382
3	121.43	4	28.73804	0.2649	0.2748	0.2287	0.0070	0.0088	0.0043	0.1639	0.1059
4	120.83	5	27.46043	0.1893	0.1767	0.1359	0.0244	0.0370	0.0051	0.2740	0.0326
5	121.19	8	44.44848	0.1795	0.2136	0.1133	0.0187	0.0517	0.0035	0.2139	0.0659
6	119.46	72	74.07817	0.4442	0.3452	0.3441	0.0452	0.1144	0.0055	0.1608	0.1189
7	112.64	3	43.45020	0.2064	0.3144	0.0276	0.0094	0.0159	0.0002	0.1929	0.0655
8	113.38	11	0.00301	0.3947	0.3395	0.1911	0.0106	0.0279	0.0030	0.1270	0.1164
9	105.71	2	35.81464	0.0985	---	0.0985	0.0085	---	0.0085	0.2745	0.0272
10	104.31	4	27.35395	0.1047	0.0926	0.0738	0.0031	0.0036	0.0012	0.2315	0.0642
11	107.31	6	0.00073	0.4327	0.4064	0.2798	0.0034	0.0037	0.0019	0.1535	0.1518
12	91.16	5	0.00967	0.4898	0.3588	0.3680	0.0084	0.0118	0.0044	0.1581	0.1593
13	84.24	3	13.82362	0.1277	0.0668	0.1322	0.0058	0.0033	0.0069	0.2027	0.0419
14	84.65	47	17.68679	0.2912	0.2690	0.2033	0.0231	0.0520	0.0051	0.2287	0.1433
15	85.30	4	61.26300	0.0954	0.0612	0.1082	0.0091	0.0096	0.0063	0.2894	0.0187
16	81.72	2	64.47696	0.1066	---	0.1066	0.0062	---	0.0062	0.2384	0.0276
17	75.16	2	34.35996	0.1243	---	0.1243	0.0011	---	0.0011	0.1949	0.0882
18	66.00	11	0.00004	0.4784	0.3871	0.3455	0.0533	0.1938	0.0025	0.2025	0.1665
19	64.06	8	36.72571	0.2406	0.1888	0.2096	0.0116	0.0112	0.0076	0.1558	0.1204
20	72.04	4	40.06076	0.1655	0.1212	0.1778	0.0091	0.0106	0.0050	0.1833	0.0791
21	71.60	5	0.00020	0.1432	0.1452	0.1151	0.0151	0.0232	0.0058	0.3107	0.0824
22	61.44	6	0.09074	0.3181	0.2854	0.1835	0.0834	0.2108	0.0177	0.3076	0.1411
23	52.53	7	0.00210	0.3511	0.3226	0.3145	0.0792	0.1195	0.0245	0.3057	0.1102
24	62.71	3	0.00099	0.5311	0.4954	0.6076	0.0167	0.0147	0.0226	0.0828	0.0886
25	42.90	2	12.33304	0.1035	---	0.1035	0.0039	---	0.0039	0.3502	0.1543
26	51.68	2	18.62741	0.1266	---	0.1266	0.0072	---	0.0072	0.2439	0.0263
27	45.41	3	0.66777	0.3572	0.2319	0.4660	0.0448	0.0360	0.0546	0.2455	0.0554
28	46.31	2	34.77422	0.1142	---	0.1142	0.0103	---	0.0103	0.2474	0.0037
29	51.51	4	23.13930	0.2990	0.2175	0.3086	0.0130	0.0233	0.0026	0.2300	0.1473

<sup>a</sup>Tamanho real do cromossomo, segundo Zimin et al. (2009), Genome Biology 10:R42 doi:10.1186/gb-2009-10-4-r42

<sup>b</sup>Distância média entre SNPs adjacentes

O declínio do desequilíbrio de ligação ( $D'$  e  $r^2$ ) em função da distância física (Mb) entre pares de bases de marcadores adjacentes, considerando todos os 249 marcadores moleculares, distribuídos ao longo de todos os 29 BTAs autossômicos, podem ser observados na Figuras 3.1A e 3.1B, respectivamente. Para a estimativa geral de  $D'$ , pode ser observado que o declínio iniciou a partir de 0,44, quando a distância física entre os SNPs adjacentes foi de 38 Kb (Figura 3.1A) e se estabilizou em 0,28, mesmo com o aumento da distância física. Em relação ao parâmetro  $r^2$ , o declínio iniciou em 0,054 quando a distância física entre os SNPs adjacentes atinge 12 Kb (Figura 3.1B) tendendo a valor zero com o aumento da distância física.



**Figura 3.1** Declínio do deesequilíbrio de ligação geral ( $D'$  em **A** e  $r^2$  em **B**) em função da distância física (Mb) considerando os 249 SNPs distribuídos nos 29 cromossomos autossômicos.

Foram estimados as estatísticas descritivas dos diferentes grupos genéticos (GG) representados nas amostras testadas nas estimativas obtidas para o declínio no deesequilíbrio de ligação. Dos 419 animais da raça Girolando utilizados neste trabalho, 22 eram  $1/2\text{Hol} + 1/2\text{Gir}$ ; 2 eram  $1/4\text{Hol} + 3/4\text{Gir}$ ; 2 eram  $7/8\text{Hol} + 1/8\text{Gir}$ ; 4 eram  $13/16\text{Hol} + 3/16\text{Gir}$ ; 77 eram  $3/4\text{Hol} + 1/4\text{Gir}$ ; 273 eram  $5/8\text{Hol} + 3/8\text{Gir}$ ; e para 39 animais não foram encontrados a identificação. Assim, foi separado um arquivo para análise descritiva

constituído somente por animais dos grupos genéticos 3/4Hol+1/4Gir e 5/8Hol+3/8Gir, os quais totalizaram 350 animais. As estatísticas descritivas obtidas para essas amostras podem ser observadas na tabela 3.2.

**Tabela 3.2** Estatísticas descritivas das análises dos SNPs para cada um dos 29 cromossomos autossômicos, considerando somente os grupos genéticos 3/4Hol+1/4Gir e 5/8Hol+3/8Gir.

BTA	Tamanho <sup>a</sup>	SNP (n)	Distância média <sup>b</sup> (bM)	<i>D'</i>	<i>r</i> <sup>2</sup>	MAF	
				Mediana	Mediana	Média	Desvio-padrão
1	158.34	9	70.62	0.18	0.00	0.23	0.17
2	137.06	5	43.93	0.17	0.01	0.22	0.04
3	121.43	4	29.64	0.25	0.00	0.16	0.11
4	120.83	5	36.28	0.11	0.00	0.27	0.04
5	121.19	8	41.59	0.11	0.01	0.21	0.07
6	119.46	72	28.03	0.36	0.01	0.16	0.12
7	112.64	3	47.48	0.03	0.00	0.20	0.06
8	113.38	11	38.68	0.32	0.00	0.13	0.12
9	105.71	2	35.81	0.07	0.00	0.27	0.03
10	104.31	4	47.55	0.10	0.00	0.23	0.07
11	107.31	6	52.51	0.41	0.00	0.15	0.15
12	91.16	5	33.17	0.28	0.00	0.16	0.16
13	84.24	3	33.68	0.08	0.00	0.19	0.04
14	84.65	48	24.44	0.22	0.01	0.23	0.14
15	85.30	4	41.78	0.12	0.01	0.28	0.01
16	81.72	2	64.48	0.09	0.00	0.23	0.03
17	75.16	2	34.36	0.16	0.00	0.20	0.09
18	66.00	11	19.04	0.37	0.00	0.20	0.16
19	64.06	8	20.85	0.34	0.02	0.15	0.12
20	72.04	4	37.14	0.10	0.00	0.19	0.08
21	71.60	5	30.93	0.08	0.01	0.31	0.08
22	61.44	6	18.87	0.12	0.01	0.30	0.15
23	52.53	7	4.23	0.33	0.06	0.31	0.08
24	62.71	3	13.88	0.42	0.00	0.08	0.08
25	42.90	2	12.33	0.21	0.02	0.34	0.16
26	51.68	2	18.63	0.12	0.01	0.25	0.03
27	45.41	3	0.51	0.46	0.04	0.24	0.06
28	46.31	2	34.77	0.10	0.01	0.24	0.01
29	51.51	4	22.14	0.32	0.00	0.23	0.14

<sup>a</sup>Tamanho real do cromossomo, segundo Zimin et al. (2009), Genome Biology 10:R42 doi:10.1186/gb-2009-10-4-r42

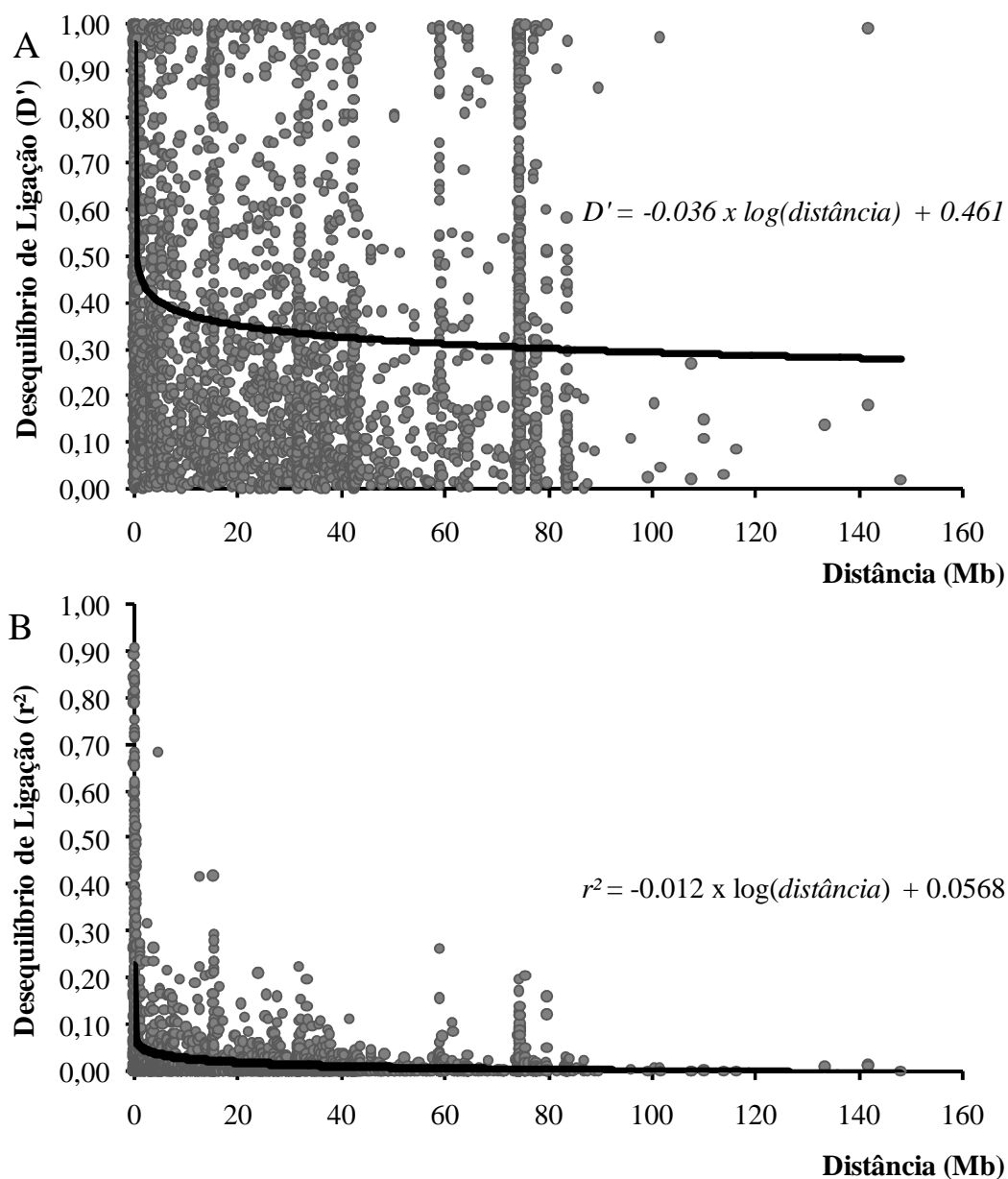
<sup>b</sup>Distância média entre SNPs adjacentes

Os parâmetros utilizados para estimar os valores de Desequilíbrio de Ligação (DL) entre os SNPs variaram entre si e entre os cromossomos (BTA), da mesma forma que ocorreu com as estimativas considerando a população total. A distância média entre os pares de SNPs adjacentes foi  $32,32 \pm 31,09$  Mb (Tabela 3.2). Para os parâmetros *D'* e *r*<sup>2</sup> a menor



estimativa foi obtida para os BTA13 ( $D'=0,09$ ;  $r^2=0,00$ ) e a maior estimada no BTA11 e BTA18 (0,48), para o parâmetro  $D'$ , e no BTA22 e BTA23 (0,09) para o parâmetro  $r^2$ .

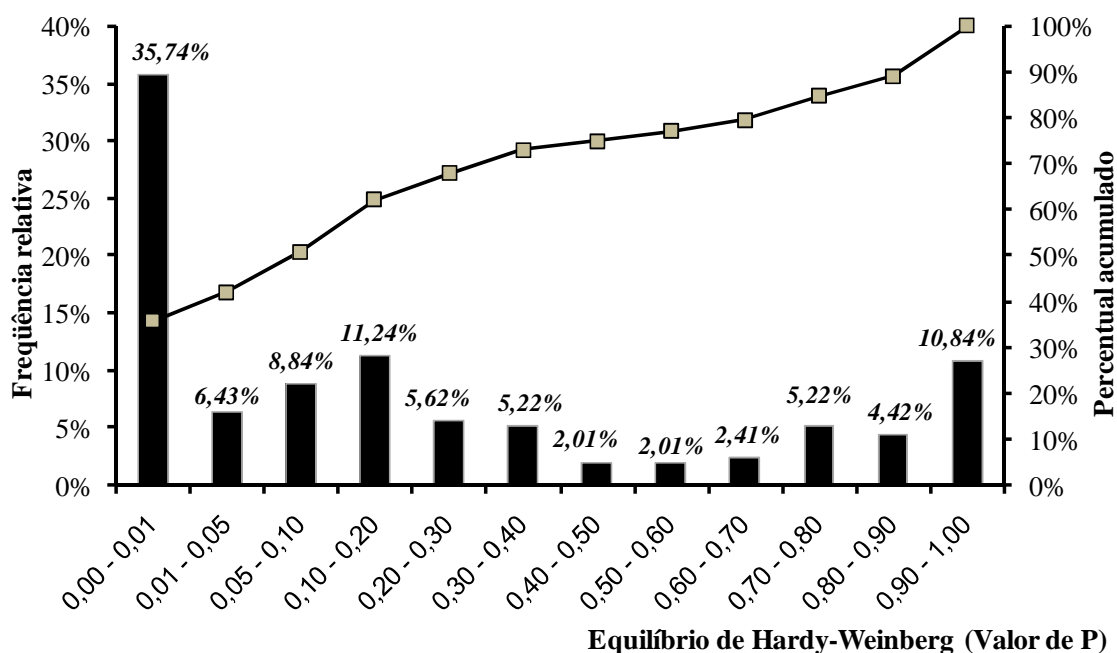
Da mesma forma, observou-se oscilação entre as estimativas de frequência do alelo menor (MAF) entre os BTA, variando de  $0,08\pm 0,08$  para o BTA24 e  $0,31\pm 0,08$ , para os BTA21 e BTA23 (Tabela 3.2).



**Figura 3.2** Declínio do desequilíbrio de ligação geral ( $D'$  em **A** e  $r^2$  em **B**) em função da distância física (Mb) considerando os 249 SNPs distribuídos nos 29 cromossomos autossômicos em animais dos grupos genéticos 3/4Holandês + 1/4Gir; 5/8Holandês + 3/8Gir.

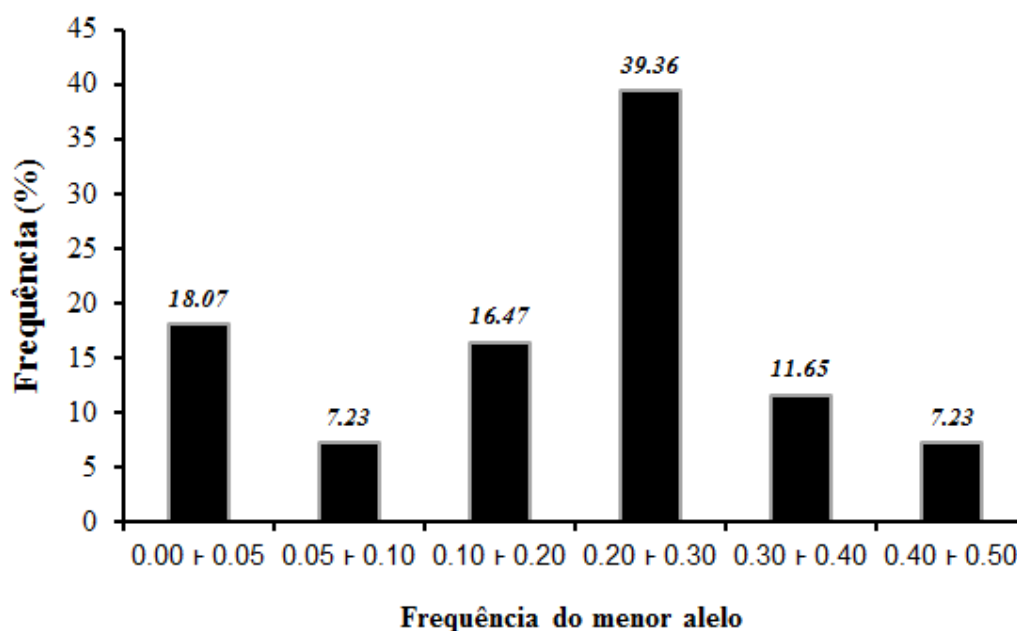
O declínio do desequilíbrio de ligação ( $D'$  e  $r^2$ ) em função da distância física (Mb) entre pares de bases de marcadores adjacentes, considerando todos os 249 marcadores moleculares, distribuídos ao longo de todos os 29 BTA autossômicos, considerando animais dos grupos genéticos 3/4Holandês + 1/4Gir; 5/8Holandês + 3/8Gir, podem ser visualizados nas Figuras 3.2A e 3.2B, respectivamente. Podem ser observados padrões semelhantes de declínio ao se considerar os parâmetros  $D'$  e  $r^2$  estimados para as populações estudadas. Para a estimativa geral de  $D'$ , considerando animais dos grupos genéticos 3/4Holandês + 1/4Gir; 5/8Holandês + 3/8Gir, pode ser observado que o declínio se iniciou a partir de 0,47, quando a distância física entre os SNPs adjacentes foi de 39 Kb (Figura 3.2A) e se estabilizou em 0,28, mesmo com o aumento da distância física. Em relação ao parâmetro  $r^2$ , o declínio se iniciou em 0,06 quando a distância física entre os SNPs adjacentes atingem 12 Kb (Figura 3.2B) tendendo a valor zero com o aumento da distância física. Neste sentido, embora existam oscilações nos valores das estimativas, o resultado geral se mostrou semelhante considerando os dois conjuntos de dados, se considerarmos os dois conjuntos de dados analisados.

Na distribuição das frequências relativas ao EHW, considerando-se a população total, pode se observar que aproximadamente 36% dos valores encontrados estavam fora do equilíbrio esperado (Figura 3.3). Este resultado era esperado, uma vez que os desvios das proporções de equilíbrio são decorrentes do agrupamento, em uma mesma população, de animais pertencentes a grupos genéticos que sabidamente não estão em equilíbrio HW (1/2Hol + 1/2Gir). De qualquer forma, o critério de Equilíbrio de Hard-Weinberg (EHW) não deve ser considerado de forma muito restritiva, já que se trata de população doméstica e sob seleção. Muitas vezes adota-se um critério permissivo para considerar que há desvio significativo ( $P < 0,001$ ) apenas como forma de eliminar os SNPs que possam apresentar viés de genotipagem.



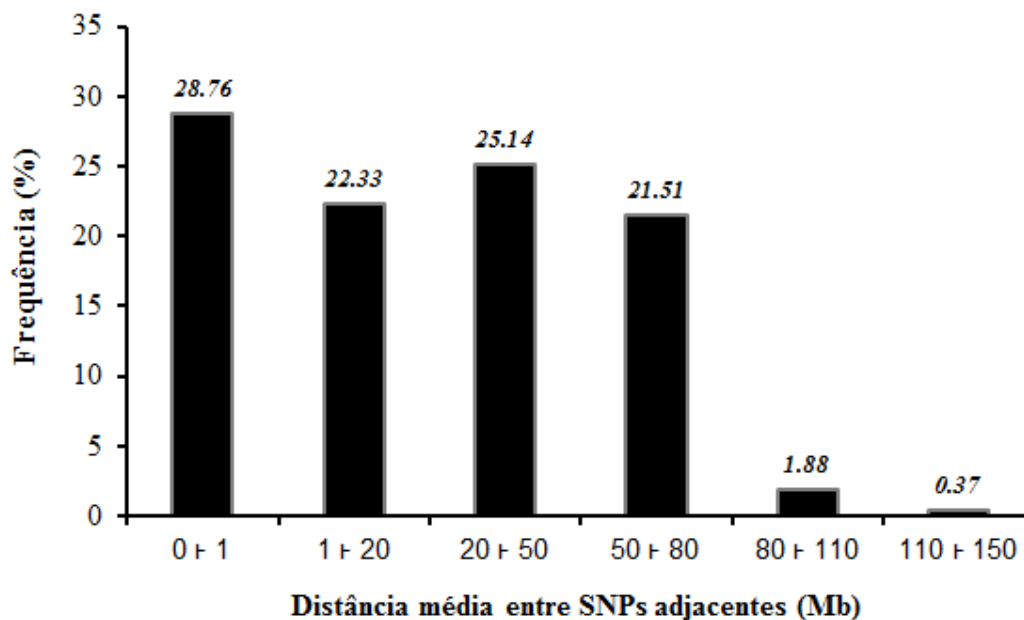
**Figura 3.3** Valor de  $P$  para a frequência relativa e a percentagem acumulada do Equilíbrio de Hardy-Weinberg (EHW) considerando os 249 SNPs distribuídos nos 29 cromossomos autossômicos.

A frequência do alelo menor (MAF) apresentou média de  $0,22 \pm 0,06$ , dos quais 81,93% dos SNPs apresentaram com MAF maior que 5% (Figura 3.4).



**Figura 3.4** Frequência do Menor alelo para os 249 SNPs distribuídos nos 29 cromossomos autossômicos.

A frequência observada entre os pares de SNPs adjacentes apresentou-se elevada entre as distâncias e 1Mb a 80 Mb, variando de 21 Mb a 28 Mb, neste intervalo (Figura 3.5). Distâncias superiores a 80 Mb foram observadas para poucos pares de bases, ou seja, pouco mais de 2% dos pares de SNPs apresentaram distâncias superiores a 80Mb.



**Figura 3.5** Distribuição do número de pares de SNPs adjacentes em função da distância física dos 29 cromossomos autossomos.

Na Tabela 3.3 está apresentado um resumo das equações de regressão de declínio dos desequilíbrios de ligação em função do logaritmo das distâncias físicas entre pares de SNPs, para cada BTA autossomo bovino. Nesta tabela, o *intercepto* estimado para cada BTA, representa o valor inicial do Declínio do DL (DDL). A partir desse ponto, início-se o DDL conforme a estimativa do  $\log(\text{distância})$ . Foram genotipados somente dois polimorfismos (SNPs) para os cromossomos BTA9, 16, 17, 25, 26 e 28, portanto para estes cromossomos não foram estimadas as equações de regressão do declínio do desequilíbrio de ligação.

**Tabela 3.3** Resumo das equações de regressão do desequilíbrio de ligação ( $D'$  e  $r^2$ ) em função do logaritmo das distâncias físicas entre pares de SNPs adjacentes para cada cromossomo (BTA) autossomo bovino.

BTA	Parâmetros	$D'$			$r^2$		
		Estimativa	Erro padrão	p-valor	Estimativa	Erro padrão	p-valor
1	<i>Intercepto</i>	0.5461	0.1349	0.0004	0.0145	0.0060	0.0229
	<i>log(distância)</i>	-0.0539	0.0325	0.1097	-0.0018	0.0014	0.2145
2	<i>Intercepto</i>	0.6668	0.4506	0.1771	0.0330	0.0308	0.3152
	<i>log(distância)</i>	-0.1224	0.1232	0.3495	-0.0057	0.0084	0.5148
3	<i>Intercepto</i>	1.2271	0.3561	0.0262	0.0116	0.0193	0.5804
	<i>log(distância)</i>	-0.3000	0.1086	0.0508	-0.0014	0.0059	0.8209
4	<i>Intercepto</i>	0.4936	0.0932	0.0007	0.0935	0.0165	0.0005
	<i>log(distância)</i>	-0.0987	0.0278	0.0074	-0.0224	0.0049	0.0019
5	<i>Intercepto</i>	0.5363	0.0622	0.0000	0.1161	0.0115	0.0000
	<i>log(distância)</i>	-0.1098	0.0174	0.0000	-0.0300	0.0032	0.0000
6	<i>Intercepto</i>	0.4729	0.0069	0.0000	0.0548	0.0023	0.0000
	<i>log(distância)</i>	-0.0348	0.0019	0.0000	-0.0116	0.0006	0.0000
7	<i>Intercepto</i>	2.3700	1.3138	0.3222	0.1183	0.0680	0.3320
	<i>log(distância)</i>	-0.5713	0.3451	0.3460	-0.0288	0.0179	0.3538
8	<i>Intercepto</i>	0.3979	0.0517	0.0000	0.0115	0.0042	0.0081
	<i>log(distância)</i>	-0.0041	0.0103	0.6907	-0.0012	0.0008	0.1566
10	<i>Intercepto</i>	0.1063	0.1277	0.4522	0.0066	0.0046	0.2248
	<i>log(distância)</i>	-0.0005	0.0353	0.9902	-0.0010	0.0013	0.4693
11	<i>Intercepto</i>	0.7121	0.4645	0.1638	0.0027	0.0043	0.5473
	<i>log(distância)</i>	-0.0772	0.1230	0.5477	0.0002	0.0011	0.8745
12	<i>Intercepto</i>	0.3739	0.1605	0.0482	0.0129	0.0051	0.0358
	<i>log(distância)</i>	0.0457	0.0448	0.3382	-0.0018	0.0014	0.2466
13	<i>Intercepto</i>	-0.1275	0.2195	0.6650	-0.0110	0.0001	0.0081
	<i>log(distância)</i>	0.0754	0.0640	0.4482	0.0049	0.0000	0.0052
14	<i>Intercepto</i>	0.3796	0.0140	0.0000	0.0472	0.0026	0.0000
	<i>log(distância)</i>	-0.0341	0.0045	0.0000	-0.0093	0.0008	0.0000
15	<i>Intercepto</i>	-0.0584	0.1604	0.7343	-0.0132	0.0255	0.6322
	<i>log(distância)</i>	0.0429	0.0441	0.3866	0.0062	0.0070	0.4254
18	<i>Intercepto</i>	0.5123	0.0537	0.0000	0.0904	0.0228	0.0002
	<i>log(distância)</i>	-0.0258	0.0131	0.0533	-0.0282	0.0056	0.0000
19	<i>Intercepto</i>	0.3944	0.0809	0.0003	0.0190	0.0051	0.0026
	<i>log(distância)</i>	-0.0662	0.0295	0.0429	-0.0032	0.0019	0.1096
20	<i>Intercepto</i>	-0.3096	0.2170	0.2267	-0.0088	0.0270	0.7608
	<i>log(distância)</i>	0.1378	0.0620	0.0904	0.0052	0.0077	0.5368
21	<i>Intercepto</i>	0.2160	0.0247	0.0000	0.0274	0.0028	0.0000
	<i>log(distância)</i>	-0.0341	0.0059	0.0004	-0.0058	0.0007	0.0000
22	<i>Intercepto</i>	0.3478	0.0441	0.0000	0.0951	0.0512	0.0859
	<i>log(distância)</i>	-0.0631	0.0122	0.0002	-0.0249	0.0141	0.1015
23	<i>Intercepto</i>	0.3507	0.0647	0.0000	0.0790	0.0239	0.0037
	<i>log(distância)</i>	-0.0465	0.0214	0.0430	-0.0172	0.0079	0.0427
24	<i>Intercepto</i>	0.5279	0.4016	0.4140	0.0162	0.0093	0.3296
	<i>log(distância)</i>	-0.0115	0.0855	0.9146	-0.0016	0.0020	0.5609
27	<i>Intercepto</i>	0.2488	0.2331	0.4793	0.0393	0.0422	0.5230
	<i>log(distância)</i>	-0.1092	0.1711	0.6384	-0.0056	0.0310	0.8871
29	<i>Intercepto</i>	0.3673	0.2839	0.2654	-0.0147	0.0270	0.6155
	<i>log(distância)</i>	-0.0246	0.0959	0.8102	0.0099	0.0091	0.3363

### 3.5.2 – Análises no cromossomo 06 (BTA06)

De acordo com as frequências alélicas relativas ao BTA06, observou-se que todos os animais apresentaram o alelo *G* fixado em polimorfismos nos genes *CSN2* (ID = 86), *CSN3* (ID = 106) e *IL8* (ID = 130). A frequência genotípica considerando todos os locos revelou-se bastante heterogênea, com médias de  $0,05 \pm 0,06$  para o genótipo homozigoto para o alelo menor,  $0,64 \pm 0,27$  para o genótipo homozigoto para o alelo maior e  $0,30 \pm 0,22$  para o genótipo do heterozigoto (Tabela 3.4).

**Tabela 3.4** Análise descritiva das frequências alélicas e genotípicas dos SNPs genotipados no cromossomo 06 de bovino da raça Girolando.

ID	SNP	NCHROBS	Frequência				
			Alélicas		Genotípicas		
			A1	A2	A1A1	A1A2	A2A2
48	<i>ABCG2/ ATP-binding cassette G2</i>	788	C 0,1042	A 0,8958	CC 0,0000	CA 0,2083	AA 0,7917
50	<i>BTA6_02720</i>	645	A 0,1517	G 0,8483	AA 0,0146	AG 0,2741	GG 0,7113
52	<i>BTA6_107922</i>	829	G 0,3043	A 0,6957	GG 0,0724	GA 0,4638	AA 0,4638
53	<i>BTA6_107923</i>	712	G 0,3998	A 0,6002	GG 0,1211	GA 0,5574	AA 0,3215
54	<i>BTA6_77051</i>	829	A 0,4115	G 0,5885	AA 0,1146	AG 0,5938	GG 0,2917
56	<i>BTA6_77057</i>	771	G 0,3610	A 0,6390	GG 0,0884	GA 0,5453	AA 0,3664
57	<i>BTA6_77060</i>	829	A 0,1272	G 0,8728	AA 0,0307	AG 0,1930	GG 0,7763
58	<i>BTA6_77062</i>	836	G 0,4646	A 0,5354	GG 0,1667	GA 0,5958	AA 0,2375
59	<i>C4ORF7_601</i>	805	G 0,0751	A 0,9249	GG 0,0188	GA 0,1127	AA 0,8685
60	<i>C4ORF7_914</i>	796	A 0,4425	G 0,5575	AA 0,1527	AG 0,5795	GG 0,2678
62	<i>CSNIS1/ Casein, alpha s1</i>	836	G 0,3383	A 0,6617	GG 0,0402	GA 0,5962	AA 0,3636
66	<i>CSNIS1_BMC_5798/ Casein, alpha s1</i>	712	A 0,3781	G 0,6219	AA 0,1229	AG 0,5104	GG 0,3667
68	<i>CSNIS2_BMC_10215/ Casein, alpha s2</i>	832	A 0,0240	G 0,9760	AA 0,0000	AG 0,0479	GG 0,9521
69	<i>CSNIS2_BMC_10431/ Casein, alpha s2</i>	824	G 0,0710	A 0,9290	GG 0,0000	GA 0,1420	AA 0,8580
70	<i>CSNIS2_BMC_11255/ Casein, alpha s2</i>	836	A 0,0833	G 0,9167	AA 0,0167	AG 0,1333	GG 0,8500

73	<i>CSNIS2_BMC_18035/ Casein, alpha s2</i>	836	A 0,2705	G 0,7295	AA 0,0345	AG 0,4720	GG 0,4935
74	<i>CSNIS2_BMC_1915/ Casein, alpha s2</i>	834	A 0,4019	G 0,5981	AA 0,1245	AG 0,5549	GG 0,3207
76	<i>CSNIS2_BMC_21965/ Casein, alpha s2</i>	836	A 0,2732	G 0,7268	AA 0,0346	AG 0,4773	GG 0,4881
78	<i>CSNIS2_BMC_2544/ Casein, alpha s2</i>	809	A 0,0225	G 0,9775	AA 0,0000	AG 0,0451	GG 0,9549
80	<i>CSNIS2_BMC_5105/ Casein, alpha s2</i>	825	G 0,2842	A 0,7158	GG 0,0487	GA 0,4710	AA 0,4803
84	<i>CSN2/Beta-casein</i>	806	A 0,2025	C 0,7975	AA 0,0158	AC 0,3733	CC 0,6109
88	<i>CSN2_110/Beta-casein</i>	813	C 0,0419	A 0,9581	CC 0,0000	CA 0,0837	AA 0,9163
90	<i>CSN2_BMC_2364/Beta- casein</i>	751	A 0,3393	G 0,6607	AA 0,0921	AG 0,4944	GG 0,4135
91	<i>CSN2_BMC_269/Beta- casein</i>	771	G 0,1569	A 0,8431	GG 0,0146	GA 0,2845	AA 0,7008
93	<i>CSN2_BMC_344/Beta- casein</i>	836	T 0,1490	C 0,8510	TT 0,0167	TC 0,2646	CC 0,7188
94	<i>CSN2_BMC_4393/Beta- casein</i>	813	A 0,0157	G 0,9843	AA 0,0000	AG 0,0313	GG 0,9687
95	<i>CSN2_BMC_5775/Beta- casein</i>	836	G 0,4563	A 0,5437	GG 0,1813	GA 0,5500	AA 0,2688
96	<i>CSN2_BMC_6334/Beta- casein</i>	728	A 0,3506	G 0,6494	AA 0,1365	AG 0,4282	GG 0,4353
97	<i>CSN2_BMC_6640/Beta- casein</i>	775	C 0,3667	A 0,6333	CC 0,1042	CA 0,5250	AA 0,3708
99	<i>CSN2_BMC_9472/Beta- casein</i>	832	A 0,4778	G 0,5222	AA 0,2072	AG 0,5412	GG 0,2516
86	<i>CSN2/Beta-casein</i>	832	0 0,0000	G 1,0000	00 0,0000	0G 0,0000	GG 1,0000
84	<i>CSN2/Beta-casein</i>	836	C 0,0395	A 0,9605	CC 0,0000	CA 0,0791	AA 0,9209
85	<i>CSN2/Beta-casein</i>	836	G 0,0156	C 0,9844	GG 0,0000	GC 0,0313	CC 0,9688
102	<i>CSN3/Kappa-casein</i>	740	G 0,1081	A 0,8919	GG 0,0108	GA 0,1946	AA 0,7946
108	<i>CSN3_BMC_11414/Kappa- casein</i>	836	A 0,1868	G 0,8132	AA 0,0071	AG 0,3593	GG 0,6336
109	<i>CSN3_BMC_12853/Kappa- casein</i>	830	G 0,2852	A 0,7148	GG 0,0691	GA 0,4321	AA 0,4988
110	<i>CSN3_BMC_13916/Kappa- casein</i>	832	A 0,4472	G 0,5528	AA 0,1416	AG 0,6112	GG 0,2472
115	<i>CSN3_BMC_14624/Kappa- casein</i>	765	A 0,4459	G 0,5541	AA 0,1152	AG 0,6613	GG 0,2235
116	<i>CSN3_BMC_14771/Kappa- casein</i>	737	A 0,1287	G 0,8713	AA 0,0063	AG 0,2448	GG 0,7490
117	<i>CSN3_BMC_14826/Kappa-</i>	706	G	A	GG	GA	AA

	<i>casein</i>		0,1205	0,8795	0,0275	0,1860	0,7865
119	<i>CSN3_BMC_1923/Kappa-casein</i>	798	A 0,3684	G 0,6316	AA 0,0913	AG 0,5541	GG 0,3546
103	<i>CSN3/Kappa-casein</i>	756	G 0,0021	A 0,9979	GG 0,0000	GA 0,0042	AA 0,9958
104	<i>CSN3/Kappa-casein</i>	830	A 0,0049	G 0,9951	AA 0,0000	AG 0,0098	GG 0,9902
106	<i>CSN3/Kappa-casein</i>	830	0 0,0000	G 1,0000	00 0,0000	0G 0,0000	GG 1,0000
107	<i>CSN3/Kappa-casein</i>	826	A 0,4681	G 0,5319	AA 0,1230	AG 0,6902	GG 0,1868
124	<i>CXCL10/ Chemokine (C-X-C motif) ligand 10</i>	796	A 0,1449	G 0,8551	AA 0,0022	AG 0,2854	GG 0,7124
126	<i>Hapmap23359_BTA_143727</i>	836	C 0,4735	A 0,5265	CC 0,2335	CA 0,4798	AA 0,2866
127	<i>Hapmap50821_BTA_91753</i>	796	A 0,1594	G 0,8406	AA 0,0239	AG 0,2711	GG 0,7050
128	<i>IL8/Interleukin 8 (neutrophil activating peptide 1)</i>	832	A 0,4292	G 0,5708	AA 0,1708	AG 0,5167	GG 0,3125
129	<i>IL8/Interleukin 8 (neutrophil activating peptide 1)</i>	779	G 0,4737	A 0,5263	GG 0,1425	GA 0,6623	AA 0,1952
130	<i>IL8/Interleukin 8 (neutrophil activating peptide 1)</i>	836	0 0,0000	G 1,0000	00 0,0000	0G 0,0000	GG 1,0000
131	<i>KIT/ proto-oncogene tyrosine-protein kinase Kit</i>	830	A 0,3996	G 0,6004	AA 0,1049	AG 0,5893	GG 0,3058
133	<i>KIT/ proto-oncogene tyrosine-protein kinase Kit</i>	836	A 0,0063	T 0,9937	AA 0,0000	AT 0,0125	TT 0,9875
134	<i>KIT/ proto-oncogene tyrosine-protein kinase Kit</i>	836	G 0,0021	A 0,9979	GG 0,0000	GA 0,0042	AA 0,9958
137	<i>OPN/Osteopontina</i>	763	T 0,3656	C 0,6344	TT 0,1146	TC 0,5021	CC 0,3833
139	<i>P51_239</i>	826	A 0,0438	T 0,9562	AA 0,0021	AT 0,0833	TT 0,9146
140	<i>PKD2_1013/ polycystic kidney disease 2</i>	830	A 0,3650	G 0,6350	AA 0,1579	AG 0,4142	GG 0,4279
142	<i>PKD2_1241/ polycystic kidney disease 2</i>	836	A 0,0138	G 0,9862	AA 0,0064	AG 0,0149	GG 0,9788
143	<i>PKD2_1349/ polycystic kidney disease 2</i>	812	A 0,0084	T 0,9916	AA 0,0000	AT 0,0167	TT 0,9833
145	<i>PKD2_2759/ polycystic kidney disease 2</i>	737	G 0,0052	A 0,9948	GG 0,0000	GA 0,0104	AA 0,9896
147	<i>PKD2_353/ polycystic kidney disease 2</i>	779	A 0,3863	G 0,6137	AA 0,1266	AG 0,5193	GG 0,3541
148	<i>PKD2_377/ polycystic kidney disease 2</i>	826	T 0,1120	A 0,8880	TT 0,0283	TA 0,1675	AA 0,8042
149	<i>PKD2_383/ polycystic kidney disease 2</i>	779	A 0,0931	G 0,9069	AA 0,0000	AG 0,1861	GG 0,8139
150	<i>PKD2_3909/ polycystic kidney disease 2</i>	836	T 0,2844	A 0,7156	TT 0,0381	TA 0,4926	AA 0,4693



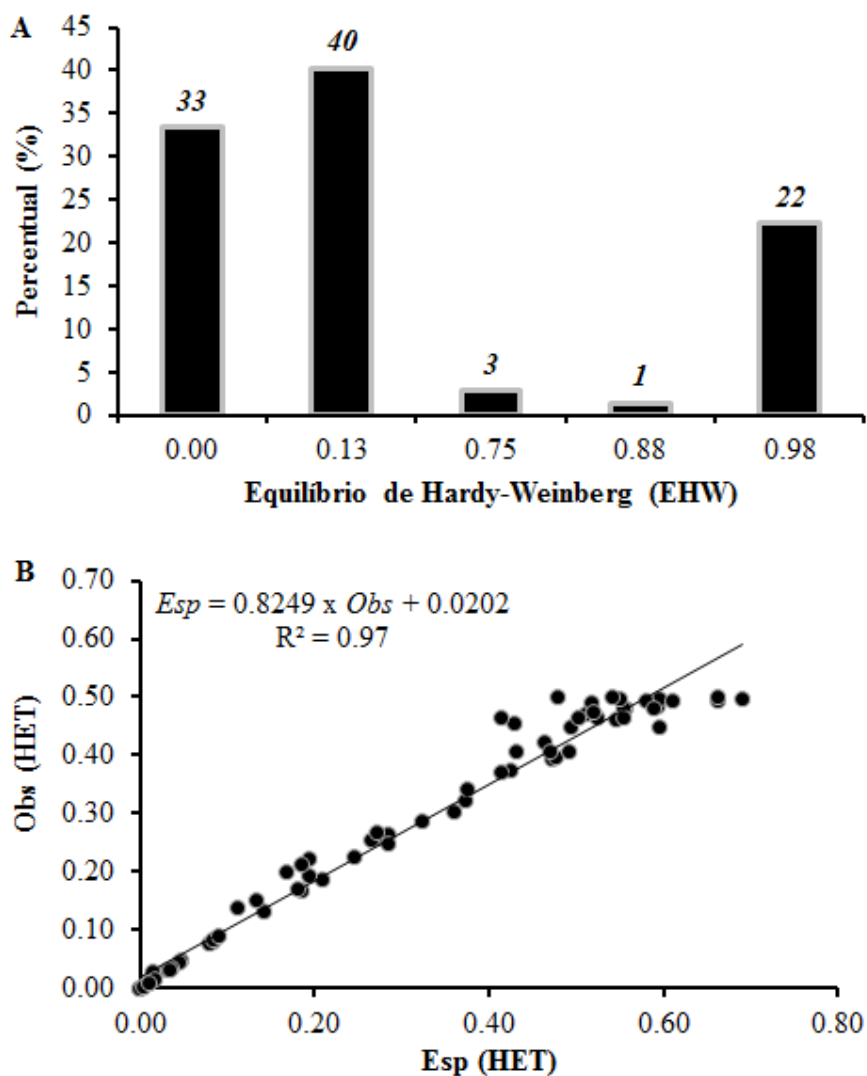
152	<i>PKD2_597/polycystic kidney disease 2</i>	836	A 0,0179	G 0,9821	AA 0,0000	AG 0,0357	GG 0,9643
155	<i>PKD2_650/polycystic kidney disease 2</i>	775	A 0,0469	G 0,9531	AA 0,0021	AG 0,0896	GG 0,9083
158	<i>PKD2_97141/polycystic kidney disease 2</i>	824	A 0,0948	G 0,9052	AA 0,0042	AG 0,1813	GG 0,8146
160	<i>PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha</i>	836	A 0,2191	G 0,7809	AA 0,0315	AG 0,3753	GG 0,5933
163	<i>PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha</i>	808	C 0,0170	A 0,9830	CC 0,0000	CA 0,0339	AA 0,9661
161	<i>PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha</i>	789	G 0,2479	A 0,7521	GG 0,0354	GA 0,4250	AA 0,5396
162	<i>PPARGC1A/peroxisome proliferative activated receptor, gamma, coactivator 1 alpha</i>	821	A 0,1724	G 0,8276	AA 0,0108	AG 0,3233	GG 0,6659
164	<i>STATH_1754</i>	803	G 0,2472	A 0,7528	GG 0,0397	GA 0,4150	AA 0,5453

ID = SNP = Polimorfismo de Base Única; NCHROBS = Número total de cromossomos observados na população para cada marcador; A1 = Alelo 1; A2 = Alelo 2

Em relação ao gene *ABCG2* (Tabela 3.4), entre os 419 animais com genótipos válidos, 80% foram homocigotos para o alelo mais frequente ( $AA=335$  animais) e os 20% restantes, heterocigotos ( $AC=84$  animais) para este polimorfismo, não sendo constatados animais com genótipo  $CC$  (Tabela 3.4). Para os polimorfismos identificados na região do gene *CSN1S1* (Tabela 3.4) houve a predominância de animais heterocigotos, variando entre 214 a 250 animais com o genótipo  $AG$ . Por outro lado, considerando-se os polimorfismos no gene *CSN1S2*, houve maior frequência de animais homocigotos para o alelo  $G$ . Entre os 13 polimorfismos genotipados na região do gene *CSN2* (Tabela 3.4), três apresentaram  $MAF < 0,05$ , e dentre os demais, foram encontrados três alelos diferentes com predominância para o alelo  $G$  ( $freqG > freqA > freqC$ ). Da mesma forma, grande variação nas frequências dos genótipos pode ser observada para polimorfismos presentes na região do gene *CSN3*, contudo, somente dois alelos (alelo  $A$  e  $G$ ) foram encontrados, com predominância para o alelo  $G$  (Tabela 3.4). A maior frequência do alelo  $G$  nos polimorfismos dos genes das caseínas está ligada a influência da raça taurina (Holandesa) na formação do mestiço Girolando. As frequências dos alelos  $T$  (0,37) e  $C$  (0,63) para o polimorfismo no gene *OPN* (Tabela 3.4) não

estão igualmente distribuídas na população, o que resultou em variações nas frequências dos genótipos, com predominância para animais heterozigotos (50%) (Tabela 3.4).

Entre os 72 SNPs avaliados, 33% (Figura 3.6A) não estavam em EHW ( $P < 0,01$ ). Adicionalmente, observou-se que a diferença dos valores médios para a Heterozigosidade observada ( $0,31 \pm 0,22$ ) e esperada ( $0,28 \pm 0,18$ ) não foi significativa a 5% de probabilidade (Figura 3.6B).



**Figura 3.6** Valor de  $P$  para estimativa de Equilíbrio de Hardy Weinberg (EHW, **A**) e Heterozigosidade (HET) observada (Obs) x esperada (Esp) (**B**) estimados para SNPs do BTA06.

Observou-se que, mesmo quando considerou-se a população Girolando como um todo (constituída de grupos genéticos diferentes), as proporções de equilíbrio de HW foram

rejeitadas ( $P < 0,05$ ) para a maioria das combinações dos SNPs (Figura 3.6). Não foi observado um padrão que sugerisse um mecanismo comum de forças atuando sobre esses locus que pudesse ser considerado responsável pela ausência de equilíbrio. Adicionalmente cabe destacar que, naturalmente, raças puras apresentam locus fixados em alelos alternativos, e quanto mais distante for a origem da raça, maior será a variedade do alelo fixado. Assim, ao se estabelecer programas de cruzamento entre raças puras diferentes, estes locos estarão fora do equilíbrio de HW.

### 3.5.2.1. Desequilíbrio de Ligação e Blocos de Haplótipos

Um total de 24 SNPs do BTA06 foram excluídos das análises de desequilíbrio de ligação (DL) com a utilização dos seguintes critérios: SNPs fora do equilíbrio de Hardy-Weinberg, frequência genotípica inferior a 25% e com frequência alélica inferior a 1%. Assim, a partir dos 72 SNPs avaliados no cromossomo 06 (BTA06), somente 48 SNPs foram utilizados para os cálculos de DL e para construção dos blocos haplotípicos.

Observou-se, na população estudada, a existência de 32 haplótipos em seis regiões do BTA06, com frequências observadas variando de 0,011 a 0,970 (Tabela 3.5). Houve também, grande variação das frequências esperadas, a qual representa o valor da frequência do haplótipo em equilíbrio de HW, em relação às frequências observadas dos haplótipos.

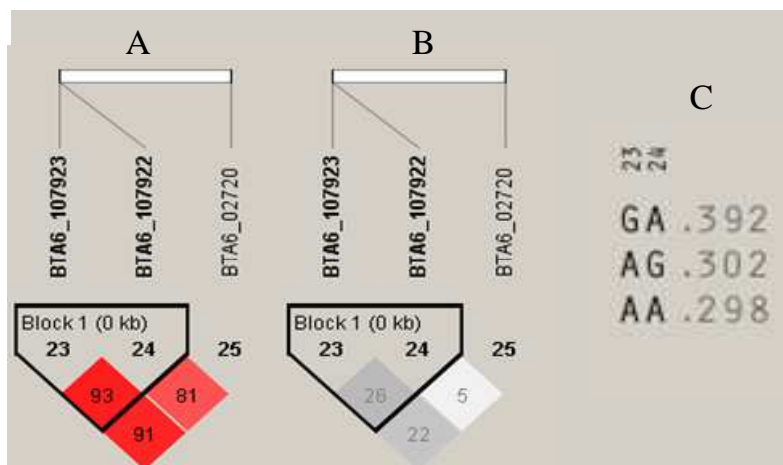
**Tabela 3.5** Identificação dos haplótipos (ID Hap), distribuição das frequências haplotípicas (Observada e Esperada) e fase alélica do gene na população – Cromossomo 06.

Região/Gene	Intervalo	ID Bloco	ID Hap	Frequência Obs	Frequência Esp	Fase alélica Bt	Bi
<i>BTA6</i>	86819139 a 86904512	1	Hap 01 – GA	0,392	0,403	GA	AG
			Hap 02 – AG	0,302	0,183		
			Hap 03 – AA	0,298	0,418		
			Hap -- – GG	0,000	0,122		
<i>CSNIS2</i>	87261372 a 87269042	1	Hap 04 – GG	0,970	0,335	AA	GG
			Hap 05 – AA	0,024	0,006		
			Hap -- – GA	0,000	0,382		
			Hap -- – AG	0,000	0,277		
			Hap 06 – GA	0,689	0,199		
<i>CSN2</i>	87276866 a 87280796	2	Hap 07 – AG	0,237	0,197	AA	GG
			Hap 08 – AA	0,041	0,074		
			Hap 09 – GG	0,033	0,530		
<i>CSN2</i>	87180474 a 87183306	1	Hap 10 – ACC	0,358	0,175	ACC	GAA
			Hap 11 – GCA	0,304	0,191		
			Hap 12 – GAA	0,211	0,331		

			Hap 13 – ACA	0,118	0,303		
	87184171	2	Hap 14 – GGA	0,444	0,385	GGA	AAG
	a		Hap 15 – AAG	0,140	0,085		
	87189676		Hap 16 – AGA	0,388	0,458		
			Hap 17 – GGG	0,012	0,072		
			Hap 18 – GA	0,482	0,246		
CSN3	87390993	1	Hap 19 – AG	0,472	0,248	AG	GA
	a		Hap 20 – GG	0,042	0,306		
	87391701		Hap -- – AA	0,000	0,199		
	87391848	2	Hap 21 – AG	0,112	0,016	AG	GA
	a		Hap 22 – GA	0,858	0,766		
	87391903		Hap 23 – AA	0,019	0,113		
			Hap 24 – GG	0,011	0,105		
IL8	90561109	1	Hap 25 – GA	0,496	0,300	AG	GA
	a		Hap 26 – AG	0,413	0,203		
	90562529		Hap 27 – GG	0,075	0,270		
			Hap 28 – AA	0,016	0,226		
PKD2	12871456	1	Hap 29 – GAA	0,437	0,018	AA	GG
	a		Hap 30 – AGG	0,375	0,014		
	12882694		Hap 31 – GAG	0,149	0,969		
			Hap 32 – AAG	0,033	0,000		

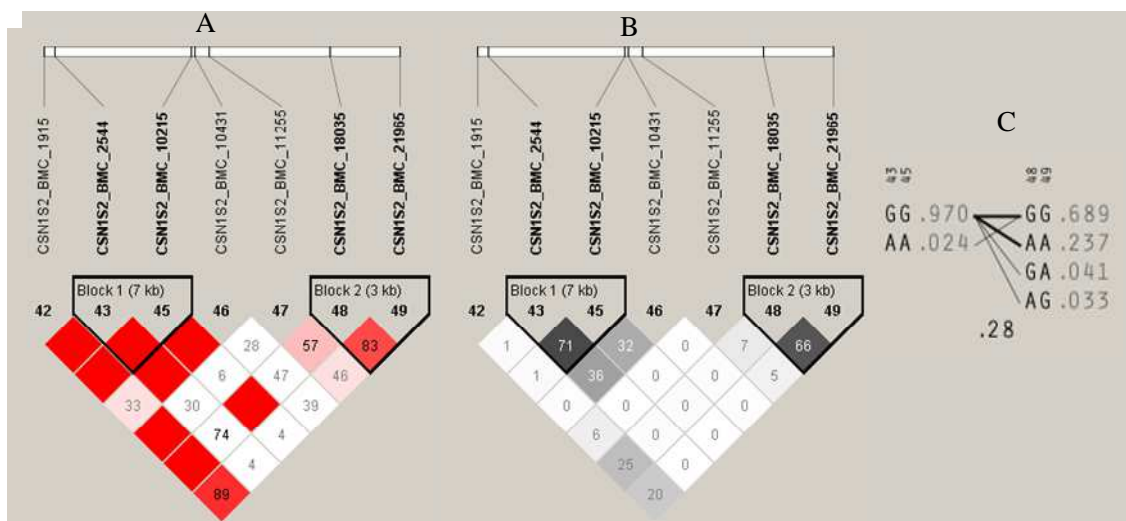
Bt= *Bos Taurus*; Bi= *Bos Indicus*

Na região *BTA6* foi observado um bloco haplotípico entre os SNPs *BTA6\_107923* (rs41610994) e *BTA6\_107922* (rs41610993) que compreendeu aproximadamente 114 pb (Figura 3.7). Este bloco apresentou três haplótipos possíveis, formados por dois alelos diferentes, com estimativas de frequências semelhantes, o que destaca a variabilidade desta região. Este haplótipo está localizado em uma região *intrônica* do gene que sintetiza a proteína da bilirrubina (Gene: *glucuronosyltransferase 2 family, polypeptide A3*; UGT2A3; ID: 511743), responsável por catalisar reações de metabolismo de substâncias tóxicas (veneno, drogas, etc.) tornando-as mais solúveis e assim mais fáceis de serem eliminadas do organismo. O valor do DL observado foi de 0,93, considerando o parâmetro  $D'$  (Figura 3.7A) e 0,26 para o parâmetro  $r^2$  (Figura 3.7B). A partir das frequências esperadas, observou-se que todos os haplótipos apresentaram desvios, indicando a ausência de equilíbrio de HW (Tabela 3.5; Figura 3.7C). Observou-se alto valor de  $D'$  entre os SNPs adjacentes *BTA6\_107923* e *BTA6\_02720* (rs29010267) (Figura 3.7A) distantes aproximadamente a 85Kpb. Entretanto, esta estimativa não foi validada devido a baixa estimativa observada para o parâmetro  $r^2$  (Figura 3.7B).



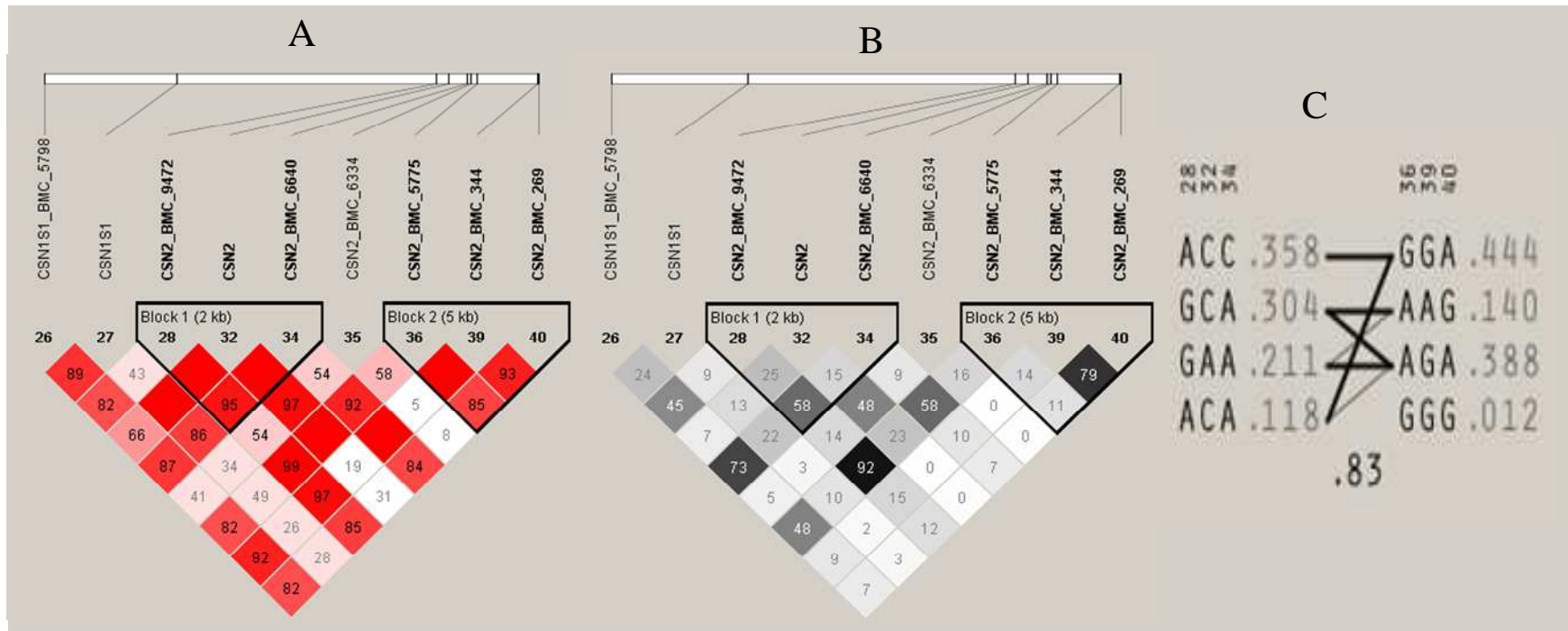
**Figura 3.7** Bloco estruturado na região gene *BTA6* (rs41610993; rs41610994): A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 06.

Observou-se no gene *CSNIS2* (Figura 3.8) a existência de seis haplótipos, divididos em dois blocos (Tabela 3.5), distantes entre si em cerca de 7.800pb. Para os haplótipos presentes no bloco 1, estruturado pelos SNPs *CSNIS2\_BMC\_2544* (ss86217860) e o *CSNIS2\_BMC\_10215* (ss86217865) houve predominância do haplótipo 4 (genótipo *GG*), com frequência de 97% dos haplótipos possíveis na população, o que indica a existência de uma seleção direcionada nos animais para esta região. O DL para este bloco foi igual a 1,0, considerando o parâmetro  $D'$  (Figura 3.8A), e 0,71 para o parâmetro  $r^2$ . Este bloco está localizada na região não codificadora do gene *CSNIS2* (5' UTR) e apresenta cerca de 7.670 nucleotídeos (Figuras 3.8A e 3.8B). Considerando o bloco 2 estruturado na região *intrônica* deste gene, entre os SNPs *CSNIS2\_BMC\_18035* (ss86217870) e *CSNIS2\_BMC\_21965* (ss86217872), observa-se a predominância do genótipo *GG* representado pelo haplótipo 06 (Tabela 3.5), presente em cerca de 69% dos animais, seguido do haplótipo 07 (genótipo *AA*). Entre os SNPs que estruturam o bloco 2 estão presentes cerca de 3.930 pb. Este bloco apresentou valor de  $D'$  igual a 0,83 (Figura 3.8A) e  $r^2$  igual a 0,66 (Figura 3.8B). As linhas entre os blocos (Figura 3.8C) representam o grau de DL entre os dois blocos, tratando cada haplótipo dentro de um bloco como um alelo dessa região. Assim, quanto maior a espessura da linha, maior o DL entre os blocos. Neste caso, observa-se baixa correlação entre os blocos estruturados nessa região, mas com predominância de apresentar o haplótipo *GG* no bloco 1 e o haplótipo *GG* no bloco 2.



**Figura 3.8** Blocos estruturados no gene *CSN1S2* (ss86217862): A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 06.

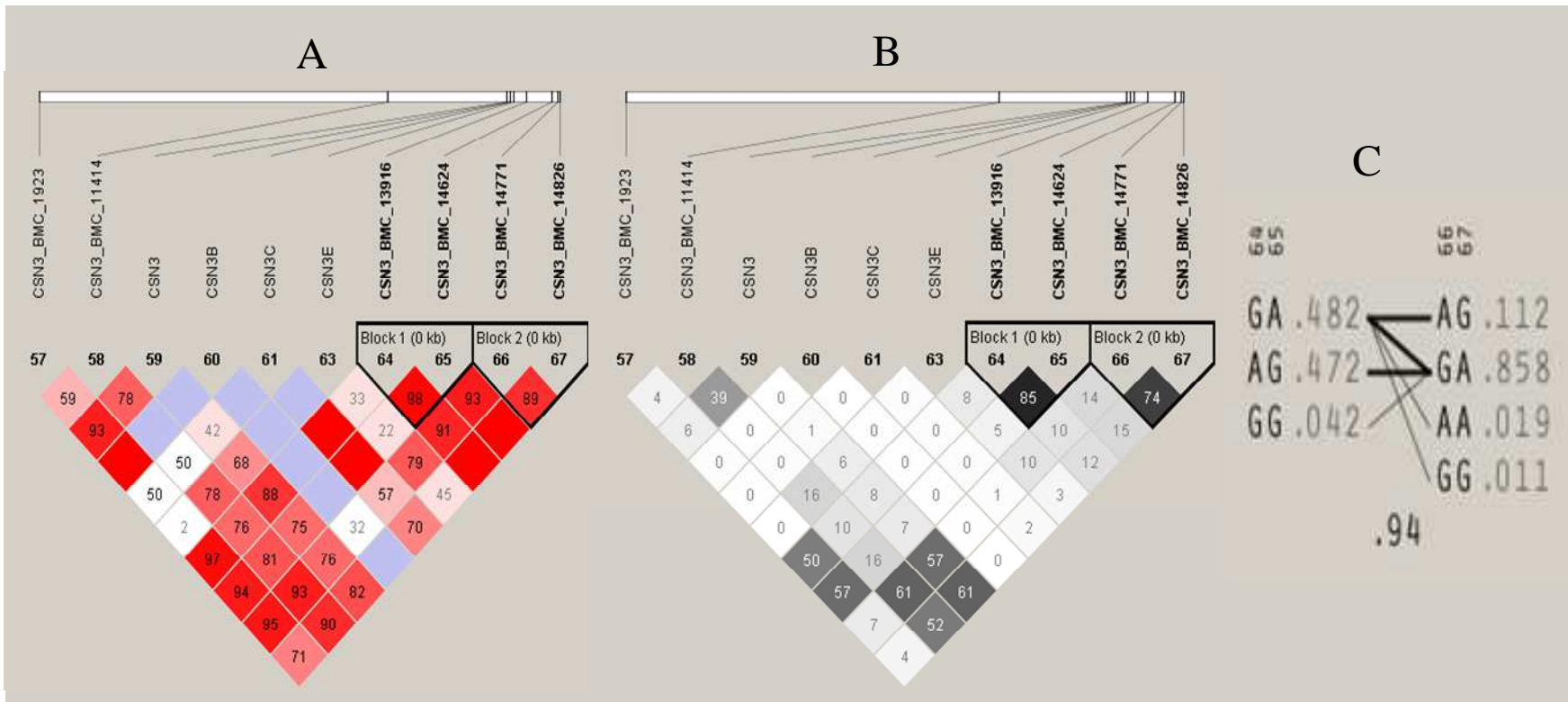
O gene *CSN2*, codificador da proteína Caseína a qual é produzida na glândula mamária e liberada no leite (Figura 3.9), apresentou 2 blocos e 8 haplótipos possíveis, ambos localizados na região *intrônica* deste gene (Tabela 3.5). Devido a maior densidade de SNPs nessa região, os haplótipos formados apresentaram 3 SNPs. Para o bloco 1 ( $D'=0,95$  e  $r^2=0,58$  para o DL), formado entre os SNPs *CSN2\_BMC\_9472* (ss86217844) e *CSN2\_BMC\_6640* (ss86217847), que possui dimensão de 2.832 nucleotídeos, houve menor variação das frequências entre os haplótipos encontrados, e conseqüentemente, maior variabilidade de haplótipos nesta região. Por outro lado, o bloco 2 ( $D'=0,85$  e  $r^2=0,11$ ) estruturado entre os SNPs *CSN2\_BMC\_5775* (ss86217849) e *CSN2\_BMC\_269* (ss86217855) que estão localizados no *exon* VII do gene *CSN2*, apresentou 5.505 nucleotídeos e maior frequência para os haplótipos 14 e 16, que juntos somam mais de 83% para os haplótipos observados nesta região (Figura 3.9C). Outro ponto que merece destaque é a alta correlação entre os SNPs *CSN2\_BMC\_9472* (SNP28) e o *CSN2\_BMC\_5775* (SNP36), apresentando DL de  $D'=0,99$  e  $r^2=0,92$  para o DL. Este resultado é corroborado pelo índice que mede o grau de DL entre os blocos estruturados, que para estes dois blocos foi estimado em 83%. Este resultado sugere forte evidência para considerar estes SNPs como *TagSNPs* dentro dos blocos que estão inseridos. Adicionalmente, observa-se uma correlação significativa entre o SNP26 (*CSN1S1\_BMC\_5798* - ss86217841) e o SNP34 (*CSN2\_BMC\_6640* - ss86217847), localizados em genes diferentes, com estimativas de  $D'=0,87$  e  $r^2=0,73$ .



**Figura 3.9** Blocos estruturados no gene *CSN2*: A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 06.

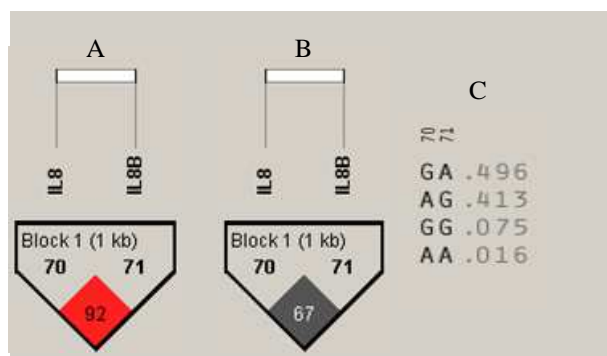
Os polimorfismos identificados na região *intrônica* do gene *CSN3* (Figura 3.10) estruturaram 2 blocos e 7 haplótipos com frequências bem distintas (Tabela 3.5). Para o bloco 1, formado entre os SNPs *CSN3\_BMC\_13916* (ss86217886) *CSN3\_BMC\_14624* (ss86217891), o valor do DL considerando o  $D'$  foi igual a 0,98 e 0,85 para o  $r^2$ . Este bloco possui a extensão de 708 nucleotídeos (Figuras 3.10A e 3.10B), e maior predominância dos haplótipos 18 e 19 (Tabela 3.5; Figura 3.10C) que juntos somam mais de 95% dos haplótipos observados nesta região. Em relação ao bloco 2, SNP *CSN3\_BMC\_14771* (ss86217892) e SNP *CSN3\_BMC\_14826* (ss86217893), este foi estruturado com o menor número de nucleotídeos (55pb), com mais de 85% dos animais genotipados apresentando o haplótipo 22, corroborando a frequência esperada. O DL para este bloco foi igual a 0,89, considerando o parâmetro  $D'$  (Figura 3.10A), e 0,74 para o parâmetro  $r^2$  (Figura 3.10B). Foi encontrado alto grau de DL (Figura 3.10C) entre esses dois blocos (0,94), o qual está diretamente relacionado com a proximidade física entre eles (147pb). Vale ressaltar a existência de correlação acima de 50% e 72% para  $r^2$  e  $D'$ , respectivamente, entre os blocos estruturados e outros SNPs da região *CSN3*, mostrando que os blocos estruturados representam toda a região do gene *CSN3*.





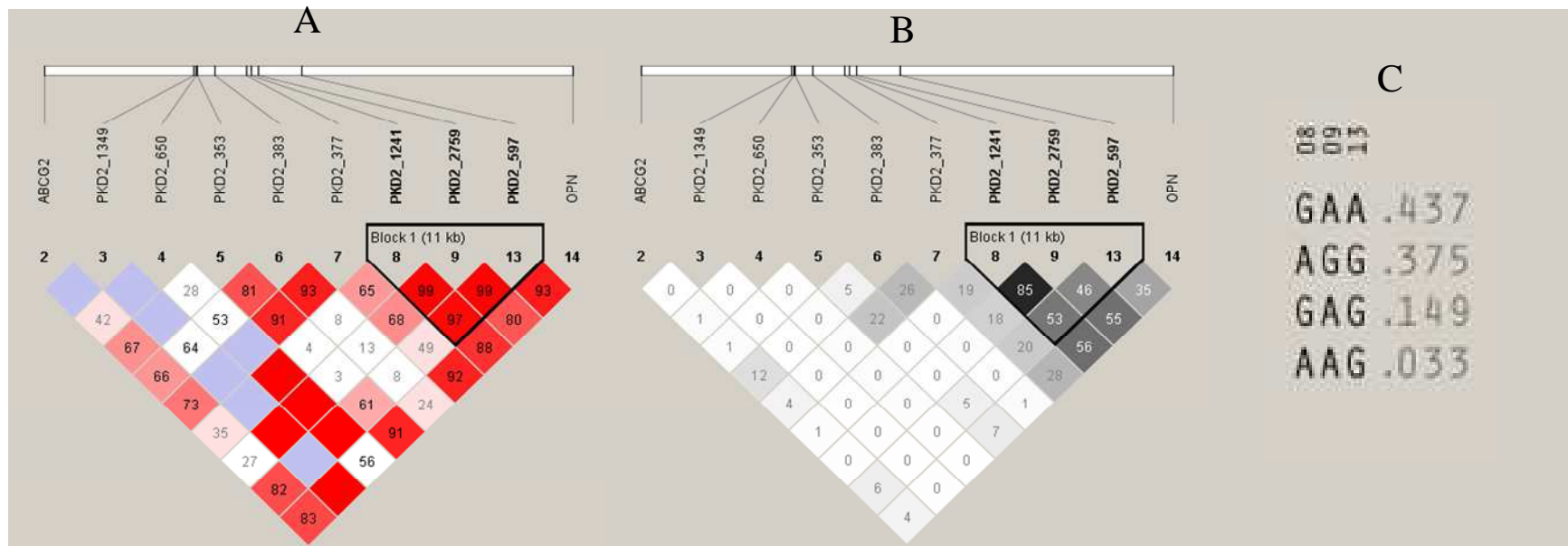
**Figura 3.10** Blocos estruturados no gene *CSN3*: A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 06.

Os polimorfismos genotipados na região *exon* do gene *IL8* (SNPs nas posições 90561109 a 90562529), que atua nos mecanismos de defesa do organismo contra infecções, revelaram a estruturação de 1 bloco (Figuras 3.11A e 3.11B) com 4 haplótipos possíveis (Figura 3.11C). Entre os quatro haplótipos, aqueles que apresentaram maior frequência foram 25 (*GA*) e 26 (*AG*) (Tabela 3.5; Figura 3.11C). O valor do DL para este bloco, o qual possui dimensão de 1.420 nucleotídeos, foi igual a 0,92 para o  $D'$  (Figuras 3.11A) e 0,67 para o  $r^2$  (Figuras 3.11B).



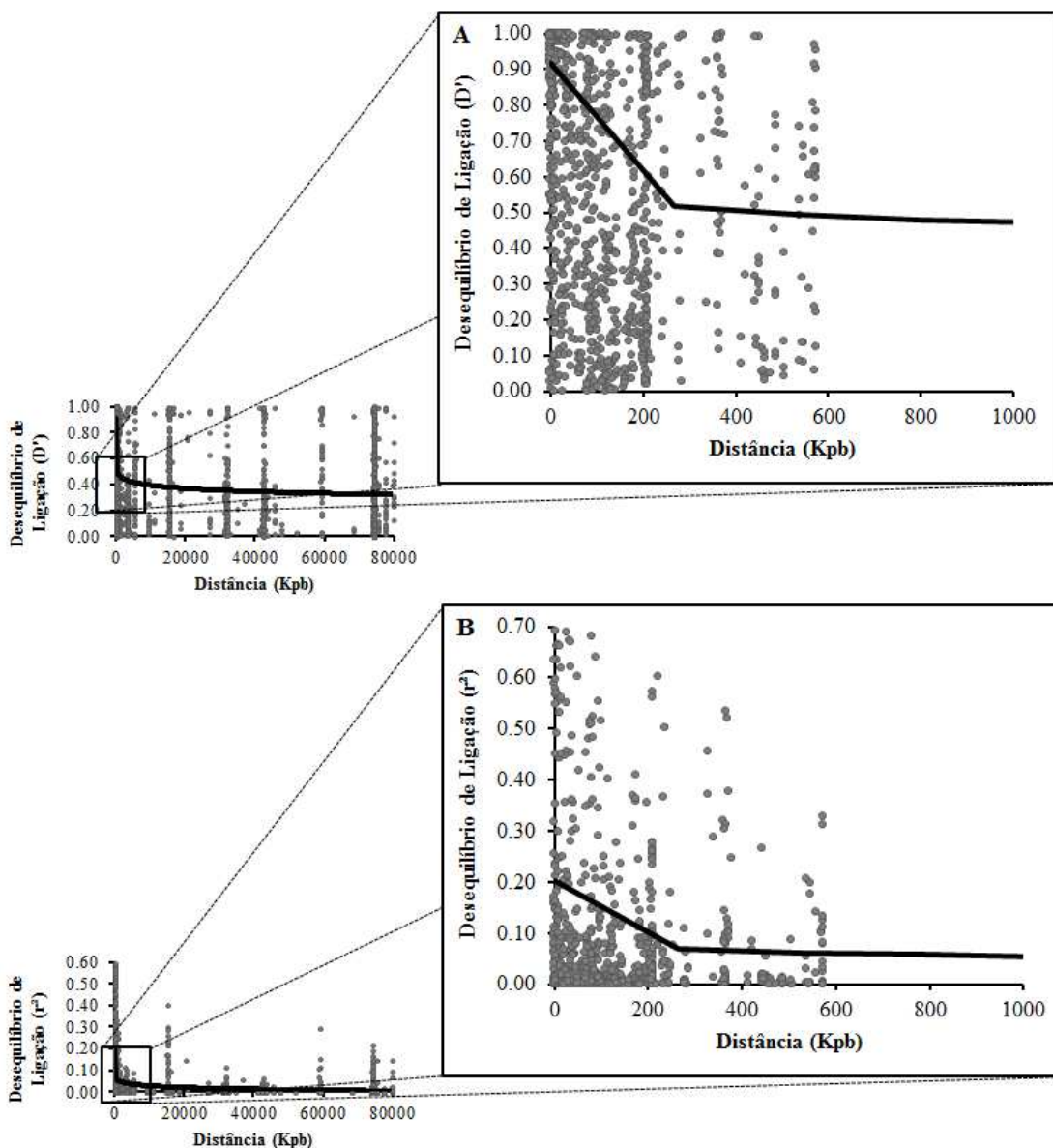
**Figura 3.11** Bloco estruturado no gene *IL8*: A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 06.

Na região do gene *PKD2* (Figura 3.12), cujas mutações são responsáveis pelo desenvolvimento de rins policísticos, foi encontrado o maior bloco estruturado no cromossomo 06, com dimensão de 11.238 nucleotídeos (entre SNPs nas posições 12871456 a 12882694). Este bloco apresentou valor de DL igual a 0,97 para o  $D'$  (Figura 3.12A) e 0,53 para  $r^2$  (Figura 3.12B). Neste bloco foram encontrados 4 haplótipos, com os haplótipos 29 (*GAA*) e 30 (*AGG*) (Tabela 3.5; Figura 3.12C) apresentando maior frequência. Adicionalmente, pode ser observado que existe um alto valor do DL estimado para  $D'$  e  $r^2$  entre os SNPs *PKD2\_1241* e *PKD2\_2759* com o SNP *OPN* (relacionado com a produção de leite), separados por 70.177 e 68.650 nucleotídeos, respectivamente. Esta alta correlação sugere que exista um *TagSNP* que englobe a região do gene *PKD2* e a região do gene *OPN*.



**Figura 3.12** Bloco estruturado no gene *PKD2*: A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 06.

Para o cromossomo 06, a distância média entre pares de marcadores adjacentes, que geraram estimativas de DL, foi de  $31,96 \pm 59,61$  Mb, variando de 3pb a 89Mb. O declínio do desequilíbrio de ligação ( $D'$  e  $r^2$ ), dos marcadores moleculares em função da distância física (Kpb) entre pares de bases de SNPs adjacentes, pode ser observado na Figura 3.13A e 3.13B, respectivamente. Considerando o parâmetro  $D'$ , observou-se que o declínio do desequilíbrio de ligação (DDL) iniciou em 0,92 e estabilizou a partir de 250 Kbp, mesmo com o aumento da distância física entre marcadores, com o valor de  $D' = 50$ . Para o parâmetro  $r^2$ , pode ser observada a mesma tendência, entretanto com estimativas inferiores, quando comparado com o parâmetro  $D'$ . O DDL considerando o parâmetro  $r^2$  iniciou com o valor de 0,20 e estabilizou, a partir de 250 Kbp, com valor de 0,08.



**Figura 3.13** Declínio do desequilíbrio ( $D'$  em A e  $r^2$  em B) de ligação geral em função da distância física (Kpb) para o cromossomo BTA6.

### 3.5.2.2. Estimativas de *TagSNPs* a partir dos Polimorfismos encontrados

Para os cálculos dos *TagSNPs*, foi utilizado um limiar de coeficiente de correlação ( $r^2$ ) igual a 0,5 (Tabela 3.6). Em geral, o número de SNPs capturados no cromossomo 6 variou de 2 a 4 entre as regiões estruturadas e a distância entre estes variou de 75 a 70.177 pb (Tabela 3.6).

**Tabela 3.6** SNPs capturados com a seleção de *TagSNPs* localizados no cromossomo 06 Bovino.

Gene/Região	TagSNP	SNP Capturado	$r^2$
BTA6	BTA6_107923/ rs41610993	BTA6_107922	0,62
		BTA6_107923	1,00
CSNIS2	CSNIS2_BMC_5105/ ss86217862	CSNIS2_BMC_21965	0,85
		CSNIS2_BMC_18035	0,86
		CSNIS2_BMC_5105	1,00
CSN2	CSN2_BMC_6640/ss86217847	CSN2_BMC_2364	0,92
		CSN2_BMC_9472	0,59
		CSN2_BMC_5775	0,59
		CSN2_BMC_6640	1,00
		CSN2_BMC_4393	0,64
CSN2	CSN2D/ss86217850	CSN2D	1,00
CSN2	CSN2_BMC_344/ss86217854	CSN2_BMC_269	0,80
		CSN2_BMC_344	1,00
CSN3	CSN3_BMC_14771/ss86217892	CSN3_BMC_11414	0,62
		CSN3_BMC_14826	0,75
		CSN3	0,57
		CSN3_BMC_14771	1,00
CSN3	CSN3_BMC_14624/ss86217891	CSN3_BMC_13916	0,85
		CSN3_BMC_1923	0,58
		CSN3_BMC_14624	1,00
IL8	IL8/ ID: 280828	IL8B	0,67
		IL8	1,00
PKD2	PKD2_1241/ID: 530393	PKD2_597	0,53
		PKD2_1241	1,00
		PKD2_2759	0,85
		OPN	0,57

Para os *TagSNPs* identificados na região do gene *CSN2*, *CSN3* e *BTA6* a captura de SNPs ocorreu somente em *íntrons* de cada região. Por outro lado, os *TagSNPs* identificados na região do gene *PKD2* e *IL8*, capturam SNPs em regiões de *exon* das respectivas regiões, com destaque para a região do gene *PKD2* que apresenta uma dimensão superior a 70 Kpb de extensão e captura também SNP em outro gene (*OPN*), corroborando os resultados das estimativas dos critério  $D'$  e  $r^2$  apresentados sobre a Figura 3.12. Em relação à região do gene *CSNIS2*, o *TagSNP* identificado captura SNP na região não codificadora 5' deste gene (Tabela 3.6). Os *TagSNPs* identificados nas regiões *BTA6*, *CSN2* e *IL8* reduzem em 50% o número de SNPs necessários para explicar a variabilidade genética observada em cada uma destas regiões gênicas. No caso dos *TagSNPs* para os genes *CSNIS2* e *CSN3* essa redução representa 67%, ao passo que no caso dos *TagSNPs* identificados para os genes

*PKD2*, *CSN3* e *CSN2* a redução representa 75% dos esforços necessários no processo de genotipagem em cada uma dessas regiões gênicas.

### 3.5.3 – Análises do cromossomo 14 (BTA14)

Na distribuição das frequências alélicas para os SNPs genotipados no cromossomo 14 (BTA14), pode ser observado que a frequência do alelo menor (MAF) apresentou média de  $0,26 \pm 0,16$ , com 15% dos SNPs apresentando  $MAF < 0,05$  (Tabela 3.7).

A análise de frequência genotípica revelou heterogeneidade entre os SNPs genotipados no BTA14. As médias e os desvios-padrão foram  $0,08 \pm 0,09$ ,  $0,55 \pm 0,25$  e  $0,37 \pm 0,20$  para o genótipo homocigoto de alelo menor, homocigoto para o alelo maior e para o heterocigoto, respectivamente.

**Tabela 3.7** Análise descritiva das frequências alélicas e genotípicas dos SNPs genotipados no cromossomo 14 de bovino da raça Girolando.

ID	SNP	NCHROBS	Frequência				
			Alélica		Genotípica		
			A1	A2	A1A1	A1A2	A2A2
217	<i>ARMCI/BTA14</i>	794	A 0,3629	G 0,6371	AA 0,0833	AG 0,5592	GG 0,3575
218	<i>ARS_BFGL_NGS_21534</i>	836	A 0,4740	G 0,5260	AA 0,2083	AG 0,5313	GG 0,2604
220	<i>BIG1/BTA14</i>	830	A 0,4770	T 0,5230	AA 0,2129	AT 0,5282	TT 0,2589
221	<i>BOLA_RALYL/RALY RNA binding protein-like</i>	821	C 0,4968	G 0,5032	CC 0,1962	CG 0,6013	GG 0,2026
222	<i>CRH_TRIM55/BTA14</i>	830	A 0,3685	G 0,6315	AA 0,1044	AG 0,5282	GG 0,3674
223	<i>DGAT/Acyl-CoA:1,2-diacylglycerol O-transferase</i>	712	A 0,4619	K 0,5381	AA 0,2015	AK 0,5209	KK 0,2776
227	<i>DGAT1/Acyl-CoA:1,2-diacylglycerol O-transferase</i>	731	A 0,0369	K 0,9631	AA 0,0024	AK 0,0690	KK 0,9286
228	<i>DGAT1/Acyl-CoA:1,2-diacylglycerol O-transferase</i>	830	A 0,0073	K 0,9927	AA 0,0000	AK 0,0146	KK 0,9854
229	<i>DGAT1/Acyl-CoA:1,2-diacylglycerol O-transferase</i>	836	A 0,1635	K 0,8365	AA 0,0188	AK 0,2896	KK 0,6917
230	<i>DQ984826</i>	830	A 0,4508	T 0,5492	AA 0,1569	AT 0,5879	TT 0,2552
258	<i>hypothetical_LOC539014/BTA14</i>	808	A 0,4892	C 0,5108	AA 0,1983	AC 0,5819	CC 0,2198
259	<i>hypotheticalLOC783036/BTA14</i>	819	A	C	AA	AC	CC

			0,2468	0,7532	0,0617	0,3702	0,5681
237	<i>KCBN2/potassium voltage-gated channel</i>	836	G 0,3052	A 0,6948	GG 0,0750	GA 0,4604	AA 0,4646
233	<i>KCBN2/potassium voltage-gated channel</i>	809	G 0,3140	A 0,6860	GG 0,0774	GA 0,4731	AA 0,4495
234	<i>KCBN2/potassium voltage-gated channel</i>	821	C 0,4618	G 0,5382	CC 0,1550	CG 0,6136	GG 0,2314
232	<i>KCBN2/potassium voltage-gated channel</i>	830	C 0,3873	A 0,6127	CC 0,1420	CA 0,4906	AA 0,3674
236	<i>KCNB2/potassium voltage-gated channel</i>	800	A 0,4346	G 0,5654	AA 0,1111	AG 0,6471	GG 0,2418
239	<i>KIEL/EST158/REV, QTL marker</i>	832	A 0,0105	C 0,9895	AA 0,0042	AC 0,0126	CC 0,9832
240	<i>LOC512677/BTA14</i>	788	C 0,3650	A 0,6350	CC 0,1084	CA 0,5133	AA 0,3783
241	<i>LOC512677/BTA14</i>	832	A 0,0450	G 0,9550	AA 0,0042	AG 0,0816	GG 0,9142
242	<i>LOC512826/BTA14</i>	821	C 0,3877	G 0,6123	CC 0,0890	CG 0,5975	GG 0,3136
243	<i>LOC515582/BTA14</i>	821	G 0,2937	A 0,7063	GG 0,0758	GA 0,4358	AA 0,4884
244	<i>LOC519708/BTA14</i>	832	C 0,1119	A 0,8881	CC 0,0021	CA 0,2197	AA 0,7782
246	<i>LOC529552/BTA14</i>	764	A 0,2426	G 0,7574	AA 0,0547	AG 0,3759	GG 0,5695
248	<i>LOC536811/BTA14</i>	832	G 0,4248	A 0,5752	GG 0,1691	GA 0,5115	AA 0,3194
249	<i>LOC536811/BTA14</i>	836	A 0,1146	G 0,8854	AA 0,0146	AG 0,2000	GG 0,7854
250	<i>LOC617133/BTA14</i>	832	G 0,1426	A 0,8574	GG 0,0294	GA 0,2264	AA 0,7442
251	<i>LOC783036_LOC790864</i>	827	A 0,0537	G 0,9463	AA 0,0000	AG 0,1074	GG 0,8926
252	<i>LOC783036_LOC790864</i>	836	A 0,4115	G 0,5885	AA 0,1333	AG 0,5563	GG 0,3104
253	<i>LOC78474/BTA14</i>	805	A 0,2866	G 0,7134	AA 0,0366	AG 0,5000	GG 0,4634
254	<i>MGC128538/BTA14</i>	830	G 0,2824	A 0,7176	GG 0,0335	GA 0,4979	AA 0,4686
260	<i>neaby_MCM4/minichromosome maintenance complex component 4</i>	808	G 0,0991	A 0,9009	GG 0,0022	GA 0,1940	AA 0,8039
262	<i>nearby_LOC782438/BTA14</i>	826	G 0,0888	A 0,9112	GG 0,0021	GA 0,1734	AA 0,8245
264	<i>nearby_PYCRL</i>	830	C 0,2411	A 0,7589	CC 0,0376	CA 0,4071	AA 0,5553
267	<i>no_hits/BTA14</i>	832	G 0,2004	C 0,7996	GG 0,0355	GC 0,3299	CC 0,6347
272	<i>no_hits/BTA14</i>	830	G	A	GG	GA	AA



			0,0476	0,9524	0,0021	0,0909	0,9070
270	<i>no_hits/BTA14</i>	832	A 0,3706	G 0,6294	AA 0,0981	AG 0,5449	GG 0,3570
273	<i>no_hits/BTA14</i>	832	C 0,0104	G 0,9896	CC 0,0000	CG 0,0209	GG 0,9791
278	<i>no_hits/BTA14</i>	798	G 0,0775	A 0,9225	GG 0,0066	GA 0,1419	AA 0,8515
275	<i>no_hits/BTA14</i>	818	A 0,0415	G 0,9585	AA 0,0000	AG 0,0830	GG 0,9170
274	<i>no_hits/BTA14</i>	832	C 0,3776	A 0,6224	CC 0,1360	CA 0,4833	AA 0,3808
277	<i>no_hits/BTA14</i>	832	C 0,0962	A 0,9038	CC 0,0021	CA 0,1883	AA 0,8096
268	<i>no_hits/BTA14</i>	832	G 0,2929	A 0,7071	GG 0,0649	GA 0,4561	AA 0,4791
269	<i>no_hits/BTA14</i>	836	A 0,1438	G 0,8562	AA 0,0188	AG 0,2500	GG 0,7313
271	<i>no_hits/BTA14</i>	822	G 0,3797	A 0,6203	GG 0,1034	GA 0,5527	AA 0,3439
280	<i>no_hits/BTA14</i>	830	C 0,2782	A 0,7518	CC 0,0669	CA 0,4226	AA 0,5105
256	<i>PKIA/BTA14</i>	822	C 0,4093	A 0,5907	CC 0,1266	CA 0,5654	AA 0,3080

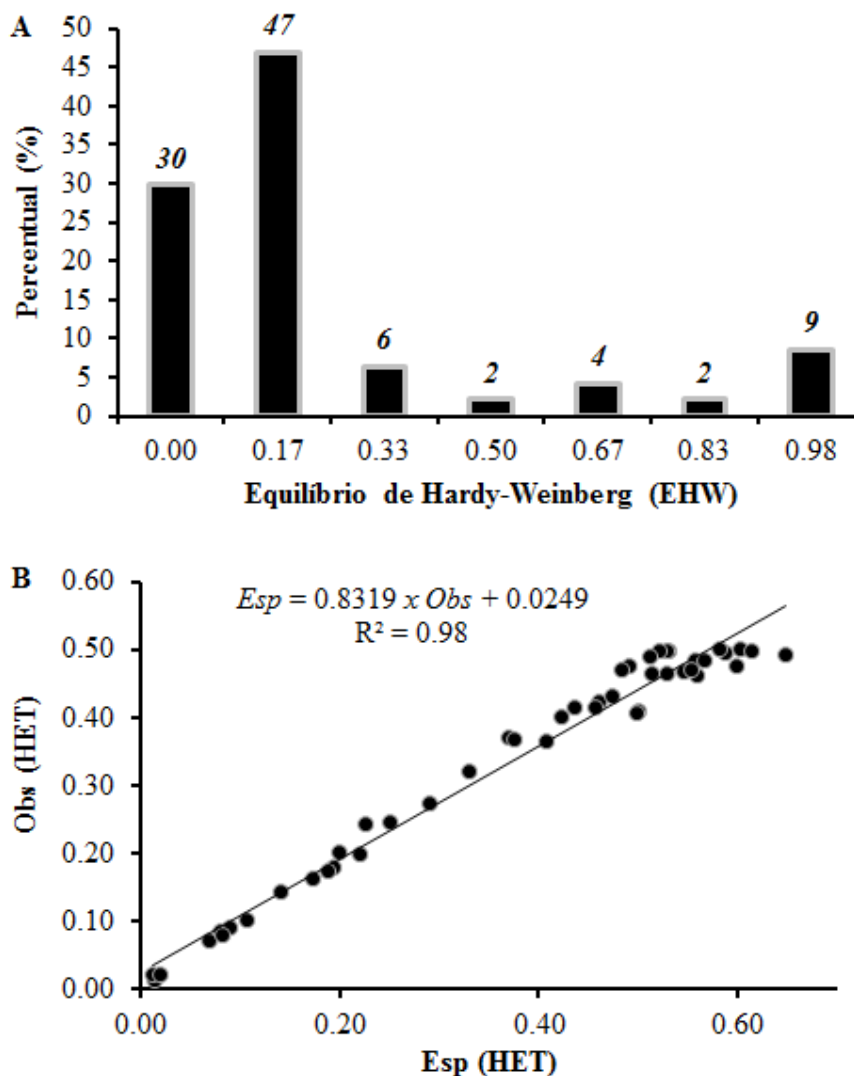
ID = Identificação do SNP; SNP = Polimorfismo de Base Única; NCHROBS = Número total de cromossomos observados na população para cada SNP; A1 = Alelo 1; A2 = Alelo 2

Considerando a região do gene *DGAT*, entre os sete polimorfismos inseridos no ensaio original, quatro permaneceram no painel validado. Este resultado era esperado, considerando as informações geradas pela empresa Illumina, visto que as sequências dos *óligos* de alguns dos polimorfismos desta região apresentavam baixa probabilidade de genotipagem (Anexo 01 – Capítulo 02 desta Tese). Os resultados na análise das frequências dos alelos do polimorfismo no gene *DGAT* (ID = 223) (Alelo A = Alelo A = 0,46; Alelo C = Alelo K = 0,54) apresentaram grande variabilidade, resultando em uma proporção maior de heterozigotos para este loco (52% dos animais da população). Considerando os outros três polimorfismos genotipados na região do gene *DGATI*, foram encontrados dois alelos (Tabela 3.7). Destes polimorfismos, um (ID = 227) apresentou predominância para o alelo G (0,96), que representa o Alelo K, sendo que um animal apresentou genótipo CC (genótipo AA), 29 animais heterozigotos (CG = AK) e 398 animais homozigotos para o alelo G (KK), com 7% de genótipos em falta. Para os outros dois polimorfismos (ID = 227; ID = 228), observou-se maior frequência para o alelo C, que representa o Alelo K, variando de 0,84 (região reguladora 3') à 0,99 (região do *exon* 15). De qualquer forma, essa maior frequência, representada pelos polimorfismos na região do gene *DGATI*, indica em um primeiro

momento reflexo de um processo de seleção que vem sendo intensificado na população estudada (Tabela 3.7).

Os polimorfismos utilizados na região do gene *KCBN2* apresentaram grande variação nas frequências alélicas, cujos valores de mínimo e máximo para MAF foram de 0,30 e 0,46 respectivamente. Da mesma forma, pode ser observado grande variação para os polimorfismos relacionados com a produção de proteínas no leite, representadas pelos SNPs *LOC512677* - *LOC78474*. Entre os 12 SNPs nesta região, foram encontrados três tipos de SNPs e somente um apresentou  $MAF < 0,05$ .

Entre os 47 SNPs analisados, 30% (Figura 3.14A) não estavam em EHW ( $P < 0,01$ ). Adicionalmente, a diferença dos valores médios para a Heterozigosidade observada ( $0,37 \pm 0,20$ ) e esperada ( $0,34 \pm 0,16$ ) não foram significativos a 5% (Figura 3.14B). Estes resultados foram semelhantes aos obtidos nas análises com o BTA06 (Figura 3.6), e mesmo considerando a população Girolando como um todo (constituída de grupos genéticos diferentes), as proporções de equilíbrio foram rejeitadas ( $P < 0,01$ ) para a maioria das combinações dos SNPs na população. Contudo, como destacado anteriormente, raças puras apresentam, naturalmente, locos fixados em alelos alternativos, e quanto mais distante for a origem da raça, maior será o contraste do alelo fixado. Assim, ao se estabelecer programas de cruzamento entre raças puras diferentes, é esperado que estes locos estejam fora do equilíbrio de HW.



**Figura 3.14** Valor de  $P$  para estimativa de Equilíbrio de Hardy Weinberg (EHW, **A**) e Heterozigosidade (HET) observada (Obs) x esperada (Esp) (**B**) estimados para SNPs do BTA14.

### 3.5.3.1 - Desequilíbrio de Ligação e Blocos de Haplótipos

Semelhante ao adotado para as análises do BTA06, foram identificados quais SNPs seriam efetivamente utilizados nas estimativas de desequilíbrio de ligação (DL). Foram utilizados os parâmetros previamente definidos (locos fora do equilíbrio de Hardy-Weinberg, locos com frequência genotípica inferior a 25% e com frequência alélica inferior a 1%) para excluir os SNPs. Assim, a partir dos 47 polimorfismos analisados no cromossomo 14

(BTA14), 33 SNPs foram utilizados para os cálculos de DL e para construção dos blocos haplotípicos.

Em relação aos polimorfismos genotipados no BTA14 da população estudada, a estimativa revelou a existência de 13 haplótipos em 4 regiões deste cromossomo, com frequências observadas variando de 0,023 à 0,589 (Tabela 3.8). Observou-se grande variação entre as frequências observadas e as esperadas dos haplótipos estruturados no cromossomo 14 (Tabela 3.8). Este resultado era esperado, uma vez que as frequências dos haplótipos estão diretamente ligados às frequências alélicas/genotípicas, e 30% destes polimorfismos não estavam em equilíbrio.

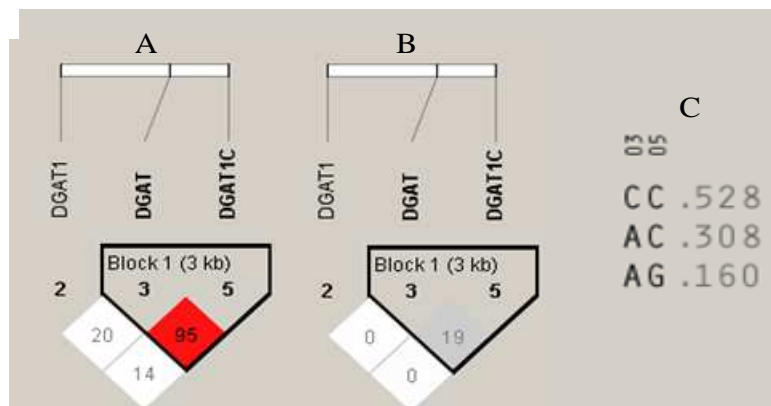
**Tabela 3.8** Identificação dos haplótipos (Hap ID), distribuição das frequências haplotípicas (Observada e Esperada) e fase alélica do gene na população – Cromossomo 14.

Gene/Região	Posição	ID Bloco	ID Hap	Frequência Obs	Frequência Esp	Fase alélica	
						Bt	Bi
<i>DGAT</i> (ID: AJ318490)	1802193 a	1	HAP 01 – CC	0,528	0,450	AG	CC
			HAP 02 – AC	0,308	0,386		
	HAP 03 – AG		0,160	0,076			
	HAP -- – CG		0,000	0,080			
<i>KCBN2</i> (ss107795102)	38453058 a	1	HAP 04 – AA	0,589	0,420	CG	AA
			HAP 05 – CG	0,302	0,122		
	HAP 06 – CA		0,085	0,266			
	HAP 07 – AG		0,023	0,192			
<i>LOC536811</i> (ss61480578)	39640732 a	1	HAP 08 – GA	0,575	0,509	AG	GA
			HAP 09 – GG	0,310	0,376		
	HAP 10 – AG		0,115	0,049			
	HAP -- – AA		0,000	0,066			
<i>LOC783036 – LOC790864</i> (ss61535152)	66194561 a	1	HAP 11 – GG	0,535	0,557	AG	GA
			HAP 12 – AG	0,411	0,389		
	HAP 13 – GA		0,054	0,032			
	HAP -- – AA		0,000	0,022			

Bt = *Bos Taurus*; Bi = *Bos Indicus*

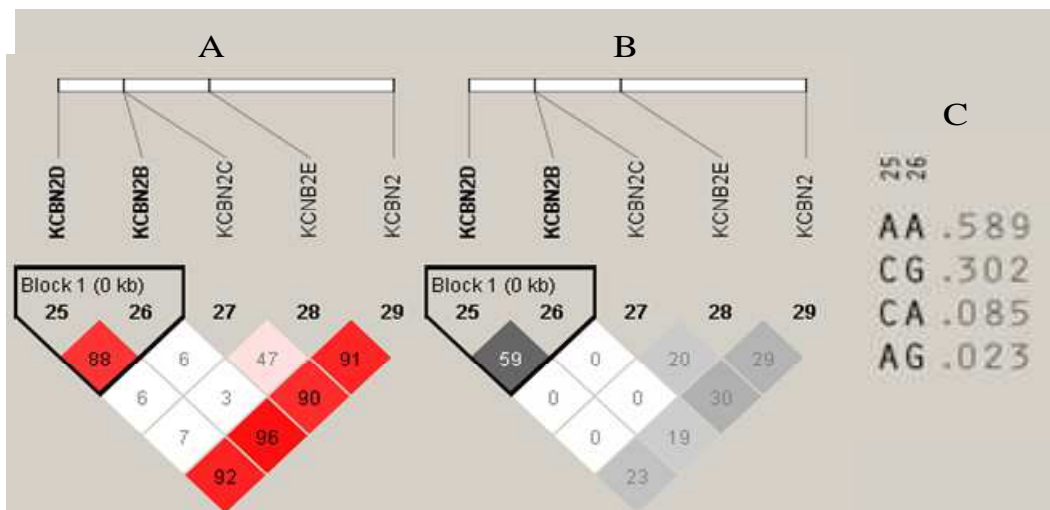
Em relação ao gene *DGAT* (ID: AJ318490), foi identificado um bloco com 3.770pb de extensão (Figuras 3.15) e 3 haplótipos possíveis (Tabela 3.8; Figura 3.15). Este bloco apresenta valor de DL igual a 0,95 para o parâmetro  $D'$  (Figura 3.15A) e 0,19 para o  $r^2$ . A frequência observada seguiu a mesma tendência da frequência esperada, em que o haplótipo homocigoto para o alelo C (que representa o Alelo K) apresentou maior frequência em relação aos demais. Este bloco está presente na região do *exon8* do gene *DGATI* e

apresenta a substituição A/A → C/G, a qual determina duas variantes: lisina (K) e alanina (A), as quais estão diretamente relacionado com a produção de gordura no leite.



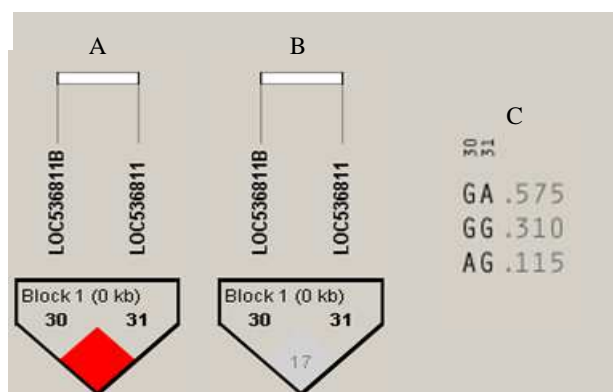
**Figura 3.15** Bloco estruturado na região do gene *DGAT*: A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 14.

Da mesma forma, pode ser observado que para o gene *KCBN2* (SNPs ss63780010 e ss107795102) houve a estruturação de um bloco (225pb de extensão – Figuras 3.15A e 3.15B) com 4 haplótipos possíveis (Tabela 3.8; Figura 3.16C). Este bloco apresentou DL igual a 0,88 para o  $D'$  (Figura 3.16A) e 0,59 para  $r^2$  (Figura 3.16B). Observa-se maior frequência para o haplótipo AA, uma vez que mais de 58% dos animais genotipados apresentaram este genótipo. Este bloco está integralmente localizado no *exon* do gene *KCBN2*. Outro ponto que merece destaque é o alto valor de  $D'$  entre o SNP29 (*KCBN2* - ss61516059) com os demais marcadores nesta região, sugerindo em um primeiro momento ser um candidato a *TagSNP* desta região.



**Figura 3.16** Bloco estruturado no gene *KCBN2* (ss107795102): A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 14.

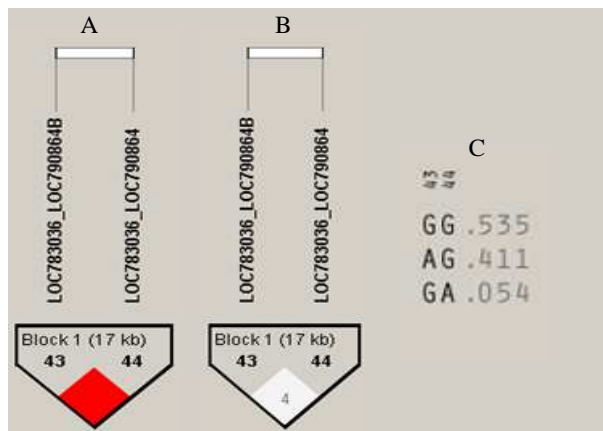
Para a região genômica que apresentou os polimorfismos *LOC536811* (SNPs ss61480576 e ss61480578), observou-se a estruturação de um bloco haplotípico com 770pb de extensão, com três haplótipos possíveis (Tabela 3.8; Figura 3.17). Este bloco, que está localizado na região *intron1* do gene *JPH1* relacionado com proteínas de membrana complexos juncionais das células, apresentou valor de DL igual a 1,0 para  $D'$  (Figura 3.17A) e 0,17 para  $r^2$  (Figura 3.17B).



**Figura 3.17** Bloco estruturado no gene *LOC536811* (SNPs ss61480576 e ss61480578): A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 14.

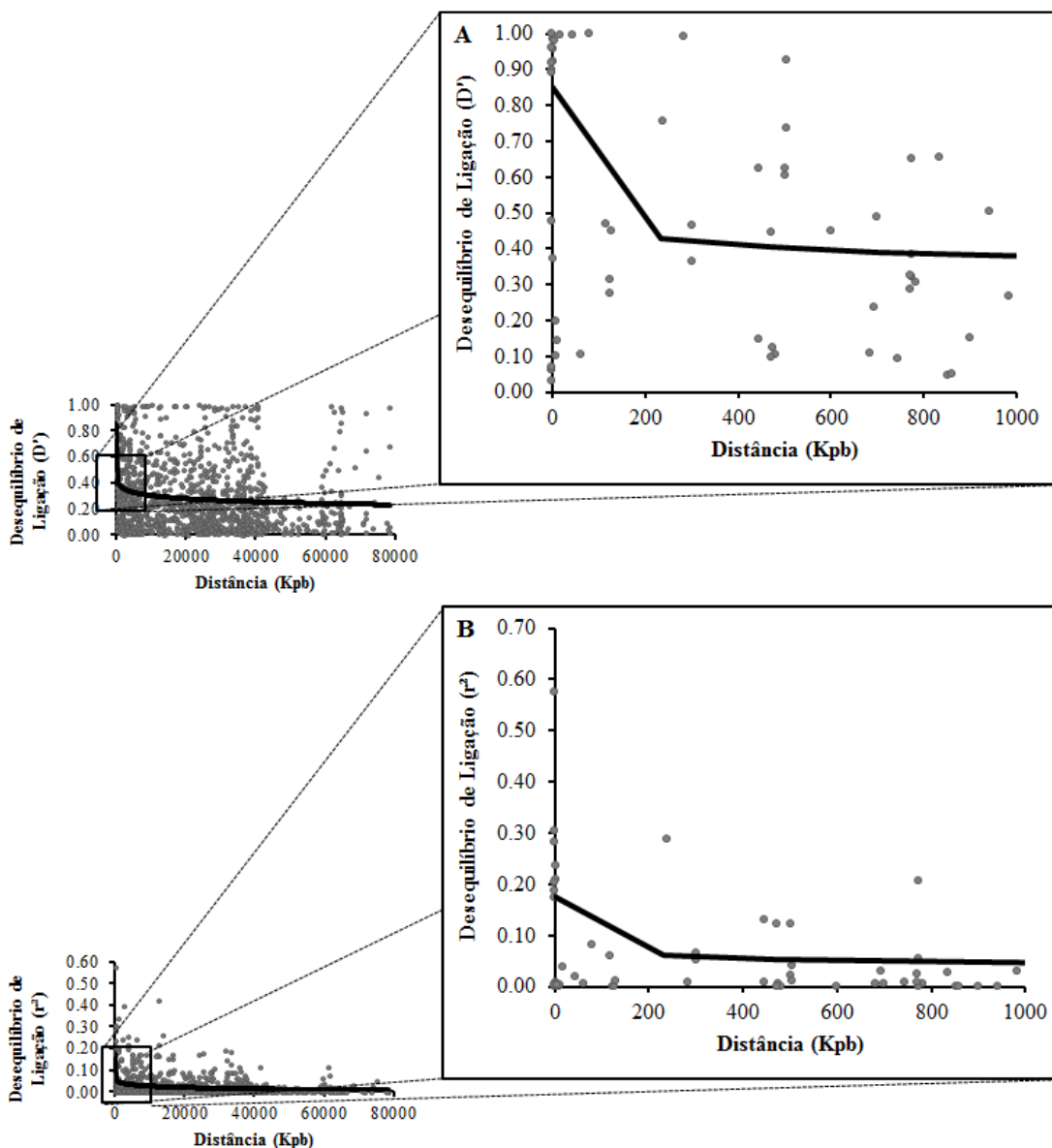
Para a região genômica que apresentou polimorfismos *LOC783036* (SNPs ss61535152 e ss38329727), foi identificado o maior bloco de haplótipo do cromossomo 14, com 17.284pb de extensão, com três haplótipos possíveis (Tabela 3.8; Figura 3.18). Este

bloco, que está localizado em região sem função conhecida, apresentou valor de DL igual a 1,0 para  $D'$  (Figura 3.18A) e 0,04 para  $r^2$  (Figura 3.18B).



**Figura 3.18** Bloco estruturado no gene *LOC783036* (ss61535152): A= valor de  $D'$ ; B =  $r^2$ ; C = frequências observadas dos haplótipos – Cromossomo 14.

Para o cromossomo 14, a distância média entre pares de marcadores adjacentes foi de  $24,44 \pm 18,30$  Mb, variando de 1pb a 78 Mb. O declínio do desequilíbrio de ligação ( $D'$  e  $r^2$ ), dos marcadores moleculares em função da distância física (Mb) entre pares de bases de SNPs (Figura 3.19). Considerando o parâmetro  $D'$ , pode ser observado que o declínio do desequilíbrio de ligação (DDL) iniciou-se com 0,87 e estabilizou-se a partir de 250Kbp, mesmo com o aumento da distância física entre marcadores, com o valor de 0,40. Para o parâmetro  $r^2$ , pode ser observada a mesma tendência, entretanto com estimativas bem abaixo quando comparado com o parâmetro  $D'$ . O DDL considerando o parâmetro  $r^2$ , inicia com o valor de 0,18 e estabiliza, a partir de 250Kbp, com valor de 0,05.



**Figura 3.19** Declínio do desequilíbrio ( $D'$  em **A** e  $r^2$  em **B**) de ligação em função da distância física (Kpb) para o cromossomo 14.

### 3.5.3.2. Estimativas de *TagSNPs* a partir dos Polimorfismos encontrados

Para se estimar os *TagSNPs*, foi utilizado um limiar de coeficiente de correlação ( $r^2$ ) igual a 0,5. Em geral, o número de SNPs capturados entre as regiões estruturadas no cromossomo 14 foi igual a 2 (o próprio *TagSNP* e o SNP adjacente) e a distância entre estes variou de 225 a 17.284 pb (Tabela 3.8). Com isso, estes *TagSNPs* podem



reduzir em 50% os esforços no processo de genotipagem para cada uma das regiões gênicas estruturadas.

**Tabela 3.9** SNPs Capturados com a seleção de *TagSNPs* localizados no cromossomo 14 Bovino.

Gene/Região	<i>TagSNP</i>	SNP Capturado	$r^2$
<i>DGAT</i>	<i>DGAT/ID: 282609</i>	<i>DGAT;</i>	1,00
		<i>DGAT1</i>	1,00
<i>KCBN2</i>	<i>KCBN2/ss107795102</i>	<i>KCBN2/ss63780014;</i>	0,60
		<i>KCBN2/ ss107795102</i>	1,00
<i>LOC536811</i>	<i>LOC536811/ss61480578</i>	<i>LOC536811/ss61480576;</i>	1,00
		<i>LOC536811/ss61480578.</i>	1,00
<i>LOC783036_</i> <i>LOC790864</i>	<i>LOC783036_LOC790864/</i> <i>ss61535152</i>	<i>LOC783036_LOC790864/ss</i>	1,00
		<i>38329727;</i> <i>LOC783036_LOC790864/ss</i> <i>61535152.</i>	1,00

Para o *TagSNP* identificado na região do gene *DGAT*, a captura do SNP identificado no *exon 2* ocorreu com SNP genotipado na região *intrônica*. Considerando o gene *KCBN2*, o *TagSNP* identificado está presente na região do *exon* deste gene, da mesma forma que o SNP capturado. Por outro lado, os *TagSNPs* identificados na região do SNP *LOC536811* e do SNP *LOC783036* estão localizados no *intron* de cada região, e os SNPs adjacentes capturados estão situados em regiões *intrônica*.

### 3.6. DISCUSSÃO

Um total de 249 SNPs identificados em 69 genes candidatos distribuídos em todos 29 cromossomos autossômicos do genoma bovinos foram analisados em 419 animais da raça Girolando, assim como 72 e 47 SNPs distribuídos ao longo dos cromossomos BTA06 e BTA14. O foco principal das análises realizadas foi o estudo do desequilíbrio de ligação e avaliações da estrutura de blocos haplotípicos em genes candidatos nos cromossomos BTA06 e BTA14, já que genes localizados nestes cromossomos são responsáveis pela produção de proteínas secretadas no leite e/ou que afetam os níveis de produção e/ou a qualidade do leite (Khatkar et al., 2008; Marques et al., 2008; Villa-Angulo et al., 2009).

Em processos diversos de simulação de estudos sobre as propriedades de DL, Wang et al. (2005) confirmaram a interferência do tamanho da amostra no resultado obtido. Tenesa et al. (2003) corroboram essa afirmativa destacando que o poder do mapeamento de DL varia conforme o tamanho da amostra. Em pesquisa relatada por estes autores em população de 50 bovinos leiteiros considerando os cromossomos 2 e 6, foi constatado DL significativo somente nas distâncias genéticas menores. Silva et al. (2010) também comprovaram a importância de uma amostra numerosa para a constatação real do DL. O estudos destes autores foram com animais da raça Girolando, contando com população de apenas 25 animais, utilizando chip comercial Illumina's BovineSNP50. Foi sugerido que amostras pequenas (abaixo de 100 indivíduos) tendem a superestimar a quantidade de DL porque podem não incluir haplótipos recombinantes raros (Tenesa et al., 2003), além de promover uma subestimativa da taxa de decréscimo do DL como aumento da distância física

(Dunning et al., 2000). Contextualizando essas informações com o presente trabalho, pode ser observado que os 419 animais utilizados, que geraram análise em média de 810 cromossomos, estão acima dos limiares estabelecidos para gerar informações não viesadas sobre a estrutura de blocos de haplótipos que conforme Reich et al. (2010) pode ser medido com 80 a 100 cromossomos. Conforme observado na grande maioria dos estudos (Tenesa et al., 2003; Wang et al., 2005; Villa-Angulo et al., 2009), a necessidade de uso de amostras numerosas (acima de 100 animais sem parentesco) é essencial para validação dos resultados e sua consideração em experimentos futuros.

Em relação ao polimorfismo *ABCG2* (ID = 48), Cohen-Zinder et al. (2005) sugeriram que este seria um QTL para produção de leite, por que foi constatado sua influência na secreção de substratos clinicamente importantes no leite. Vários SNPs já foram identificados no gene *ABCG2*, dentre estes o mais representativo refere-se a uma substituição do alelo *A* para o alelo *C* localizados no *exon* 14, causando uma mudança de aminoácido de uma tirosina para serina. A baixa frequência do alelo *C* encontrado na população de Girolando estudada, sugere que este polimorfismo seja incluído nos programas de seleção objetivando aumentar a produção de leite, sobretudo por não estar estruturado em nenhum bloco de haplótipo e não estar em DL com outros SNPs.

Para a região do gene da *Kappa Caseína* (*CSN3*) presente no BTA06, a frequência média dos genótipos dos SNPs genotipados foram semelhantes, variando de 0,30 (*AA*) a 0,37 (*GG*). Dentre os SNPs genotipados na região do *exon* IV deste gene (Tabela 3.4), pode ser observado que o loco representado por ID = 106, está fixado para o alelo *G* (100% *GG*). Este resultado é superior aos 0,63 e 0,67 obtidos por Stipp et al. (2013) e Botaro et al. (2009), respectivamente, para o genótipo *GG*, desta mesma região, em rebanhos da raça Girolando. Em animais da raça Holandesa, Stipp et al. (2013) obtiveram frequências similares para os três genótipos possíveis ( $\pm 0,30$ ), valores que também são superiores aos obtidos no presente estudo. Valente (1996) e Rodrigues (2006) ressaltaram que para raça Gir há uma alta frequência do genótipo *GG*. Assim, neste trabalho pode ser observada a influência da raça Gir na determinação das frequências alélicas para este loco específico em animais da raça Girolando. O alelo *A*, um dos alelos do gene em questão, está relacionado com maior produção de proteínas no leite, quando comparado ao alelo *G*.

Polimorfismos encontrados na região do *exon* VII do gene *CSN2* têm sido relatados como responsáveis por aumentar a produção de proteína no leite e, conseqüentemente, diminuir a concentração de gordura e rendimento do leite (Nilsen et al.,

2009; Oleński et al., 2012). Neste trabalho, o SNP *CSN2\_BMC\_344* (ID=93), localizado na região do *exon* VII, apresentou frequências alélicas iguais a 0,15 (Alelo *T*) e 0,85 (Alelo *C*), valores estes semelhante aos 0,11 (Alelo *T*) e 0,89 (Alelo *C*) relatados por Oleński et al. (2012) em rebanho da raça Holandesa. Nilsen et al. (2009) analisaram haplótipos caseína e suas associações com características de produção de leite e concluíram que o alelo *C* da beta-caseína pode ser utilizado em programas de seleção para aumentar a concentração de proteína no leite. Vale destacar este resultado, sobretudo se consideradas as prioridades atuais na seleção de gado leiteiro, que vem sendo direcionados para identificar animais com maior produção de proteína (consequentemente menos gordura) no leite.

Estudos sobre a região do gene *DGAT* demonstraram que existe uma tendência de raças taurinas apresentarem predominância para o Alelo *A* (associado ao aumento na produção de proteína e de leite), com frequências do *A* variando 0,35 a 0,7 (Spelman et al., 2002; Kaupe et al., 2004; Berry et al., 2010; Lešková et al., 2013). Berry et al. (2010) ressaltam que raças europeias têm altas frequências do alelo *A*, apesar de serem criados tradicionalmente para produção de carne e, como tal, este alelo é considerado parte do haplótipo ancestral do *DGATI*. Por outro lado, para as raças zebuínas, a tendência está na predominância do Alelo *K*, associado à diminuição da produção de proteína e aumento na produção de gordura no leite (Pappas et al., 2004; Egito et al., 2005; Lacorte et al., 2006; Komisarek et al., 2011). Em relação a animais mestiços, sua origem baseada geralmente em cruzamento entre raças com aptidões diferentes, resulta em uma maior segregação do locus, o que favorece a maior predominância do genótipo heterozigoto (Cardoso et al., 2011). Cardoso et al. (2011), apresentaram valores das frequências alélicas para o *DGAT* iguais a 0,46 e 0,54, para, respectivamente, o Alelo *A* e *K*, com o genótipo *AK* igual a 0,54. Estes resultados são iguais aos obtidos no presente trabalho (*DGAT* ID = 223). Entretanto, Santos et al. (2007) observaram tendência oposta em animais da raça Canchim, os quais encontraram valores iguais a 0,40 e 0,60 para os Alelos *K* e *A*, respectivamente. Todavia, um consentimento entre todos esses trabalhos consiste na importância do impacto do alelo *K* do gene *DGATI* sobre a produção total e na produção média diária de leite com um efeito aditivo, o que corrobora a iniciativa de realizar seleções assistida por marcadores em programas de melhoramento da raça Girolando.

Considerando os cromossomos 06 e 14, a extensão dos blocos de haplótipos variou de 55 à 17.284 pb, resultados semelhantes aos blocos com extensão de até 15.600 pb obtidos por Villa-Angulo et al. (2009) em bovinos de origem zebuína e taurina. Por outro

lado, Kim & Kirkpatrick (2009) obtiveram uma extensão média dos blocos de haplótipos variando de 25,4 Kb a 113,7 Kb para animais da raça Holandesa. Estes valores são superiores aos encontrados em estudos com animais de origem zebuína (Utsunomiya et al., 2013), em populações de animais bovinos mestiços (Mokry et al., 2013) e animais mestiços de outras espécies (Amaral et al., 2008).

Constatou-se que o coeficiente de DL variou de acordo com a proximidade entre os blocos de haplótipos, confirmando-se ser importante a identificação de blocos de haplótipos para estudos de associação, visto que a seleção de determinado número de SNPs define determinado haplótipo de interesse no processo de seleção (Zhang et al., 2002). Alguns estudos demonstraram extensões variáveis de DL entre 5 e pouco mais de 10 cM em populações de bovinos, e que existem vários fatores que determinam o DL, como a distância entre os marcadores, a pressão de seleção e o efeito fundador (Mckay et al., 2007; Sargolzaei et al., 2008; Espigolan et al., 2013).

Silva et al. (2010), trabalhando com animais da raça Gir, reportaram valores de DL, que variaram de 0,60 a 0,72 para o parâmetro  $D'$  e de 0,17 a 0,24 para o parâmetro  $r^2$ , abaixo dos apresentados neste trabalho, os quais variaram de 0,83 a 1,0 para o parâmetro  $D'$  e de 0,11 a 0,85 para o parâmetro  $r^2$ . Estudo semelhante, foi realizado por Amaral et al. (2008), em amostras de suínos de raças Chinesa e Européia, e a comparação entre as raças revelou grandes e significativas diferenças no grau e padrão de DL. Na pesquisa realizada por Bohmanova et al. (2010) com touros norte americanos da raça Holandesa, o DL foi estimado através das estatísticas  $D'$  e  $r^2$ , cujas médias foram 0,72 e 0,20, respectivamente, valores que foram inferiores aos obtidos no presente capítulo.

### **3.7. CONCLUSÃO**

Foi possível a caracterização de estruturas de blocos de haplótipos, em regiões genômicas relacionados com características de interesse zootécnico em animais da raça Girolando. Este resultado torna viável sua utilização em estudos de associação entre essas regiões com características produtivas. O desequilíbrio de ligação se estende por pequenas regiões gênicas, tendência observada para animais mestiços de várias outras espécies.

### 3.8. REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, F.S.; LOPES, C.M.; ORSI, M.L. et al. Genetic monitoring by RAPD markers for repopulation programs of *Salminus brasiliensis* (Pisces, Characiformes). **Acta Scientiarum - Animal Sciences**, v. 35, n. 2, p. 119-126, 2013
- ALTSHULER, D.; BROOKS, L.D.; CHAKRAVARTI, A. et al. A haplotype map of the human genome. **Nature**, v. 437, n. 7063, p. 1299-320, 2005.
- AMARAL, A.J.; MEGENS, H.J.; CROOIJMANS, R.P.M.A. et al. Linkage Disequilibrium Decay and Haplotype Block Structure in the Pig. **Genetics**, v. 179, n. 1, p. 569-579, 2008.
- 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.
- BERRY, D.P.; HOWARD, D.; O'BOYLE, P. et al. Associations between the K232A polymorphism in the diacylglycerol-O-transferase 1 (DGAT1) gene and performance in Irish Holstein-Friesian dairy cattle. **Irish Journal of Agricultural and Food Research**, v. 49, n. 1, p. 1-9, 2010.
- BOTARO, B. G.; LIMA, Y. V. R.; CORTINHAS, C.S. et al. Effect of the kappa-casein gene polymorphism, breed and seasonality on physicochemical characteristics, composition and stability of bovine milk. **Revista Brasileira de Zootecnia**, v.38, n.12, p.2447-2454, 2009.
- BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F.S. Characteristics of linkage disequilibrium in North American Holsteins. **BMC Genomics**, v.11, n. 421, p. 1-11, 2010.
- BRITTO, Fernanda V. **A Genômica aplicada aos rebanhos bovinos**. (2010). Disponível em: <http://www.beefpoint.com.br/radares-tecnicos/melhoramento-genetico/a-genomica-aplicada-aos-rebanhos-bovinos-65931/>. Acesso em 03 de Novembro de 2013.
- COHEN-ZINDER, M.; SEROUSSI, E.; LARKIN, D. M., et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6

- affecting milk yield and composition in Holstein cattle. **Genome Research**, v. 15, n. 7, p. 936-944, 2005.
- DU, F.X.; CLUTTER, A.C.; LOHUIS, M.M. Characterizing linkage disequilibrium in pig populations. **International Journal of Biological Sciences**, v. 3, n. 3, p. 166–178, 2007.
- DUNNING, A.M., DUROCHER, F., HEALEY, C.S. et al. The extent of linkage disequilibrium in four populations with distinct demographic histories. **American Journal of Human Genetics**, v. 67, n. 6, p. 1544–1554, 2000.
- ESPIGOLAN, R.; BALDI, F.; BOLIGON, A.A. et al. Study of whole genome linkage disequilibrium in Nellore cattle. **BMC Genomics**, v. 14, n. 305, p. 1-8, 2013.
- FACÓ, O.; LÔBO, R.N.B.; MARTINS FILHO, R. et al. Análise do Desempenho Produtivo de Diversos Grupos Genéticos Holandês x Gir no Brasil. **Revista Brasileira de Zootecnia**, v. 31, n. 5, p. 1944-1952, 2002.
- FAN, J.B.; OLIPHANT, A.; SHEN, R. et al. Highly parallel SNP genotyping. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 68, p. 69-78, 2003.
- GABRIEL, S.B.; SCHAFFNER, S.F.; NGUYEN, H. et al. The structure of haplotype blocks in the human genome. **Science**, v. 296, n. 5576, p. 2225-2229, 2002.
- GURYEV, V.; SMITS, B.M.G.; BELT, J.Van de. et al. Haplotype Block Structure Is Conserved across Mammals. **PLoS Genetics**, v. 2, n. 7, p. 1111-1118, 2006.
- HARTL, D.L.; CLARK, A.G. Principles of population genetics. Sinauer Associates Inc. Sunderland, 1997.
- HILL, W.G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **Theoretical and Applied Genetics**, v. 38, n. 6, p. 226-231, 1968.
- IBGE, Instituto Brasileiro de Geografia e Estatística. Produção Pecuária Nacional, Rio de Janeiro, v.36, 2008. Disponível em: <[http://ibge.gov.br/home/estatistica/indicadores/agropecuaria/producaoagropecuaria/abat-e-leite-couro-ovos\\_201401\\_publ\\_completa.pdf](http://ibge.gov.br/home/estatistica/indicadores/agropecuaria/producaoagropecuaria/abat-e-leite-couro-ovos_201401_publ_completa.pdf)> Acesso em: 29/11/2014.
- KAUPE, B.; WINTER, A.; FRIES, R. et al. DGAT1 polymorphism in Bos Indicus and Bos taurus cattle breeds. **Journal of Dairy Research**, v.71, n. 2, p. 182-187, 2004.
- KHATKAR, M.S.; NICHOLAS, F.W.; COLLINS, A.R. et al. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. **BMC Genomics**, v. 9, n. 187, p. 1-18, 2008.
- KIM, E.S.; KIRKPATRICK, B.W. Linkage disequilibrium in the North American Holstein population. **Animal Genetics**, v. 40, n. 3, p. 279–288, 2009.
- LEŠKOVÁ, L.; BAUER, M.; CHRENEK, P. Detection of DGAT1 gene polymorphism and its effect on selected biochemical indicators in dairy cows after calving. **Acta Veterinaria Brno**, v. 82, n. 3, p. 265–269, 2013.
- LEWONTIN, R.C.; KOJIMA, K. The evolutionary dynamics of complex polymorphisms. **Evolution**, v. 14, n. 4, p. 458-472, 1960.
- LEWONTIN, R.C. The Interaction of Selection and Linkage. ii. Optimum Models. **Genetics**, v. 50, n. 10, p. 757-782, 1964.
- LU, D.; SARGOLZAEI, M.; KELLY, M. et al. Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. **Frontiers in Genetics**, v. 3, n. 152, p. 1-10, 2012.



- MARQUES, E. F.; SCHNABLE, R. D.; STOTHARD, P. et al. High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle. **BMC Genetics**, v. 9, n. 45, p. 1-12, 2008.
- MCKAY S.D.; SCHNABEL R.D.; MURDOCH B.M. et al. Whole genome linkage disequilibrium maps in cattle. **BMC Genetics**, v. 8, n. 74, p. 1-12, 2007.
- MCMANUS, C.; TEIXEIRA, R.A.; DIAS, L.T. et al. Características Produtivas e Reprodutivas de Vacas Holandesas e Mestiças Holandês x Gir no Planalto Central. **Revista Brasileira de Zootecnia**, v. 37, n. 5, p. 819-823, 2008.
- MOKRY, F.B.; HIGA, R.H.; MARCONDES, C.R. et al. In the search of the polled locus in a *Bos taurus* x *Bos indicus* population. In: INTERNATIONAL SYMPOSIUM ON ANIMAL FUNCTIONAL GENOMICS, 5., 2013, Guarujá.
- NANZER, T.A.D.T. 2010. Produção de leite no Brasil e participação da genética Girolando com ênfase em reprodução. Disponível em: <<http://www.girolando.com.br/site/noticia.php?id=1859>>. Acesso em: 20 nov. 2011.
- NILSEN, H.; OLSEN, H. G.; HAYES, B. et al. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. **Genetics Selection Evolution**, v. 41, n. 24, p. 1-12, 2009.
- OLEŃSKI, K.; CIEŚLIŃSKA, A.; SUCHOCKI, T. Polymorphism in coding and regulatory sequences of beta-casein gene is associated with milk production traits in Holstein-Friesian cattle. **Animal Science Papers and Reports**, v. 30, n. 1, p. 5-12, 2012.
- OLSEN, H.G.; NILSEN, H.; HAYES, B. et al. Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. **BMC Genetics**, v. 8, n. 32, p. 1-9, 2007.
- PEREIRA, J. C. C. **Melhoramento Genético Aplicado à Produção Animal**. 5. ed. Belo Horizonte: FEP-MVZ, 2008. v. 1. 618 p.
- PURCELL, S.; NEALE, B.; TODD-BROWN, K. et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. **American Journal of Human Genetics**, v. 81, n. 3, p. 559-575, 2007.
- REICH, D.E.; CARGILL, M.; BOLK, S. et al. Linkage disequilibrium in the human genome. **Nature**, v. 411, n. 6834, p. 199–204, 2001.
- RODRIGUES, S.G. Estudo das frequências dos alelos A e B dos genes da kappa-caseína e beta-lactoglobulina e suas associações com produção de leite em bovinos F1 Girolando. 2006. 30f. Tese (Doutorado em Ciência Animal) – Universidade Federal de Minas Gerais, Belo Horizonte.
- SANTOS, L.P.; REGITANO, L.C.A.; PAPPAS, M. et al. Genotipagem e validação de SNPs no gene DGAT1 em animais da raça Canchim.. In: CONGRESSO BRASILEIRO DE GENÉTICA, 53., 2007, Águas de Lindóia. CONGRESSO BRASILEIRO DE GENÉTICA, 53, 2007.
- SARGOLZAEI, M.; SCHENKEL, F.S.; JANSEN, G.B. et al. Extent of linkage disequilibrium in Holstein cattle in North America. **Journal of Dairy Science**, v. 91, n. 5, p. 2106-2117, 2008.
- SILVA, C.R.; NEVES, H.H.R.; QUEIROZ, S.A. et al. **Extent Of Linkage Disequilibrium In Brazilian Gyr Dairy Cattle Based On Genotypes Of AI Sires For Dense SNP Markers**. (2010) Disponível em:

- <http://www.kongressband.de/wcgalp2010/assets/pdf/0633.pdf>. Acesso em 02 de Outubro de 2013.
- SODELAND, M.; GROVE, H.; KENT, M. et al. Molecular characterization of a long range haplotype affecting protein yield and mastitis susceptibility in Norwegian Red cattle. **BMC Genetics**, v. 12, n. 70, p. 1-10, 2011.
- SPELMAN, R.J.; FORD, C.A.; MCELHINNEY, P. et al. Characterization of the DGAT1 gene in the New Zealand dairy population. **Journal of Dairy Science**, v. 85, n. 12, p. 3514-3517, 2002.
- STIPP, A.T.; BIGNARDII, P.R.; POLI-FREDERICOI, R.C. et al. Polimorfismos genéticos da kappa-caseína e da beta-lactoglobulina e produção de leite em bovinos. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 65, n. 1, p. 275-280, 2013.
- TENESA, A.; KNOTT, S.A.; WARD, D. et al. Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes. **Journal of Animal Science**, v. 81, n. 3, p. 617-623, 2003.
- USDA - United States Department of Agriculture. Global Agriculture Information Network on line. 2009. Disponível em: <<http://gain.fas.usda.gov/Lists/Advanced%20Search/AllItems.aspx>>. Acesso em: 15 mai. 2012.
- UTSUNOMIYA, Y. T.; CARMO, A. S.; CARVALHEIRO, R. et al. Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. **BMC Genetics**, v. 14, n. 52, p. 1-12, 2013.
- VALENTE, E.P. Levantamento das frequências dos fenogrupos do sistema B de grupos sanguíneos e dos alelos A e B do gene da k-caseína em populações das raças Holandesa e Gir do Estado de Minas Gerais. 1996. 69f. Dissertação (Mestrado em Medicina Veterinária) - Escola de Veterinária da Universidade Federal de Minas Gerais, Belo Horizonte.
- VALLEJO, R.L.; LI, Y.L.; ROGERS, G.W. et al. Genetic diversity and background linkage disequilibrium in the North American Holstein cattle population. **Journal of Dairy Science**, v. 86, n. 12, p. 4137-4147, 2003.
- VILLA-ANGULO, R., MATUKUMALLI, L.K., GILL, C.A. et al. High-resolution haplotype block structure in the cattle genome. **BMC Genetics**, v.10, n. 19, p. 1-13, 2009.
- WADE, C.M.; KULBOKAS, E.J.; KIRBY, A.W. et al. The mosaic structure of variation in the laboratory mouse genome. **Nature**, v. 420, p. 574-578, 2002.
- WANG, Y.; ZHAO, L.P.; DUDOIT, S. **A Fine-Scale Linkage Disequilibrium Measure Based on Length of Haplotype Sharing**. University of California, Berkeley.U.C. Berkeley Division of Biostatistics Working Paper Series, 2005.
- ZHANG, K.; DENG, M.H.; CHEN, T. et al. A dynamic programming algorithm for haplotype block partitioning. **Proceedings of the National Academy of Sciences USA**, v. 99, n. 11, p. 7335-7339, 2002.
- ZIMIN, A.V.; DELCHER, A.L., FLOREA, L. et al. A whole-genome assembly of the domestic cow, *Bos Taurus*. **Genome Biology**, v.10, n. 4 (R42), p. 1-10, 2009.

## **CAPÍTULO 4**

### **CONSIDERAÇÕES FINAIS E PERSPECTIVAS FUTURAS**

## 4.1. CONSIDERAÇÕES FINAIS

### 4.1.1. Mudanças recentes nos procedimentos de avaliação genética e suas repercussões

Nos últimos anos, diferentes metodologias de genotipagem de DNA têm sido desenvolvidas e utilizadas amplamente, revolucionando os estudos genômicos e gerando grandes quantidades de informações para as mais diversas espécies. Grande parte dessa evolução só foi possível a partir do sequenciamento do genoma de várias espécies, que, aliadas à evolução metodológica, permitiram o desenvolvimento de uma variedade de marcadores moleculares. Estas técnicas têm sido utilizadas em estudos que , proporcionaram o desenvolvimento de ferramentas importantes para o melhoramento genético de animais de produção. Dentre as tecnologias disponíveis, vale destacar os ensaios de genotipagem de SNPs de baixa densidade (customizados ou comerciais, contendo de dezenas até centenas de marcadores), média densidade (de milhares a dezenas de milhares) e de alta densidade (de centenas de milhares a milhões).

Várias tecnologias foram apresentadas por diferentes empresas para a genotipagem de painéis de SNP de baixa densidade (Illumina – *BeadXPress*, ABI/BioTrove – OpenArray, ABI-SNPlex, ABI-TaqMan, etc.), cada uma apresentando suas especificidades em relação aos custos de desenvolvimento e aos procedimentos laboratoriais. Entretanto, segundo Caetano (2009), o ponto em comum entre qualquer uma dessas plataformas é que os ensaios para SNPs específicos podem ser adaptados para todas as tecnologias, permitindo que a mesma informação genotípica possa ser acessada, independentemente da tecnologia

utilizada. Os métodos contemporâneos de genotipagem de marcadores SNP apresentaram na última década altos níveis de automação e acurácia, ocasionando reduções drásticas nos custos de geração de dados em relação às tecnologias anteriores (Caetano, 2009). Segundo este mesmo autor, os custos de geração de dados variam de centavos (no caso de ensaios de baixa densidade) a uma fração de centavo (no caso de ensaios de alta densidade) por genótipo.

O trabalho descrito nessa tese teve como foco desenvolver uma ferramenta que permitisse agregar todos os principais marcadores moleculares descritos na literatura com as seguintes características: a) efeito sobre características de produção bovinos de aptidão leiteira; b) mutações responsáveis por causar doenças genéticas; e c) realização de testes de confirmação de paternidade; em um ensaio simples, de baixo custo e alta reprodutibilidade. O trabalho foi iniciado em 2008 em um projeto financiado pela Embrapa, em um contexto em que apenas um número reduzido de marcadores SNP estava sendo rotineiramente testado (Mello et al., 2011; Mello et al., 2012; Paiva et al., 2013) com a tecnologias que demandam muita mão de obra e não permitem grandes níveis de paralelização (PCR-RFLP com leitura em géis de agarose). A possibilidade de testar 384 SNPs em um único ensaio, realizado em paralelo com um total de 96 indivíduos, a um custo de US\$14,4 por amostra, representava na época um avanço tecnológico significativo, com perspectivas muito positivas de gerar um produto com grande aplicabilidade para o setor produtivo. Outros esforços similares haviam gerado produtos tecnológicos análogos (Kaminski et al., 2006; Chessa et al., 2007; Marques et al., 2008) e foram adotados como ferramenta padrão nos respectivos países.

Durante a execução do projeto, presenciamos uma verdadeira revolução nos programas de avaliação e melhoramento genético de bovinos de aptidão leiteira, na América do Norte, Europa e Oceania. Métodos desenvolvidos para estimar o valor genético de um indivíduo, considerando dezenas de milhares de marcadores genéticos distribuídos pelo genoma (a chamada Seleção Genômica Ampla – *Genome-Wide Selection*), foram desenvolvidos por (Meuwissen et al., 2001) e efetivamente postos em prática em 2008, quando ocorreu a publicação do primeiro sumário de touros para a raça Holandesa com Valores Genômicos, nos Estados Unidos.

Desde então, painéis com 50 mil SNPs são rotineiramente utilizados para genotipar touros e vacas com controle leiteiro, e os dados genotípicos utilizados para avaliação dos animais. Métodos para utilizar genótipos gerados com painéis de baixa densidade, com alguns milhares de marcadores, a custos mais baixos, na imputação dos genótipos do painel com 50 mil SNPs foram desenvolvidos e validados (Li et al., 2006;

Goddard e Hayes, 2009; Boichard et al., 2012) e incorporados nos procedimentos em uso. Nos Estados Unidos, as avaliações do gado Holandês apresentam estimativas do valor genômico desde 2008 (<https://www.cdc.us/eval.htm>). A base de dados de animais genotipados com o BovineSNP50 em 2011 já contava com mais de 85.000 animais genotipados, proporcionando ganhos médios de 35% na confiabilidade das avaliações de touros jovens para características produtivas (VanTassell, 2011). No primeiro trimestre deste ano, a mesma base de dados já possuía mais de 523.000 animais genotipados, e a correlação entre estimativas da PTA Genômica e tradicional para Mérito Total (produção de leite) de touros jovens, sem filhas, e com avaliação tradicional está >50% (VanRaden, 2014).

A Avaliação Genômica touro pode ser feita com uma fração dos recursos necessários para a realização de um teste de progênie. O teste tradicional de um touro em 2008 custava ~US\$50.000, durando por volta de seis anos a partir do nascimento do animal. Com os métodos atuais, é possível obter uma avaliação do animal com 70% de acurácia, mesmo antes do nascimento do animal (caso seja realizada uma biopsia do embrião), a um custo de <US\$70. Vale destacar que parte dos SNPs utilizados na presente tese foram incorporados aos chips comerciais atualmente disponíveis (50K).

As mudanças nos procedimentos de avaliação genética descritas trouxeram grandes benefícios para o setor produtivo e, por outro lado, descontextualizaram a aplicação dos resultados obtidos com a execução dessa tese. Embora os métodos de avaliação genômica ainda não tenha sido aplicados na avaliação genética da raça Girolando, estudos que visam estabelecer as bases de dados básicas necessárias para iniciar o processo estão em andamento. Assim, embora os resultados obtidos tenham grande aplicabilidade para identificar animais geneticamente superiores e não portadores de mutações responsáveis por causar doenças genéticas que foram incluídas no ensaio validado, o avanço tecnológico observado nos últimos anos deixou o painel de 384 SNPs testado sem utilização prática.

#### **4.1.2. Avaliação crítica da tecnologia Veracode (BeadXPress) – Perspectivas Futuras**

O ensaio *GoldenGate*, através da tecnologia *Veracode (BeadXPress)* possui um o sistema de leitura com capacidade para obtenção de 36.864 genótipos (384 SNPs X 96 amostras) em dois dias de trabalho. Este cenário representava significativo avanço na geração de dados moleculares quando comparados com os obtidos através de genotipagem com

microssatélites, os quais em situações ideais (*Multiplex* padronizado, 5 *primers*, 96 amostras, Sequenciador ABI 3730 com 16 placas) 7.680 “datapoint” poderiam ser obtidos no mesmo período de trabalho.

Em relação à estrutura e procedimentos necessários para a geração de dados genômicos, cabe destacar também algumas das dificuldades que encontramos no desenvolvimento deste trabalho, sobretudo sob o ponto de vista laboratorial. A tecnologia *Veracode (BeadXPress)* possui uma rotina complexa de laboratório, uma vez que o processo para obtenção destes resultados envolve a manipulação da amostra de DNA com 18 reagentes alternando em 5 placas diferentes. Essa complexidade se torna ainda maior quando se acrescenta o fator “transporte” desses reagentes até o usuário. Existem especificidades de temperatura entre esses reagentes, e variações podem influenciar as características químicas dos reagentes, enviesando os resultados. Outro limitante em relação ao transporte está nas placas de hibridização, em que os oligonucleotídeos pré-sintetizados estão fixados em bastonetes de sílica. Estes bastonetes, como descrito no capítulo I desta tese, são cilindros de vidro com dimensões micrométricas e qualquer vibração nas placas pode danificar os bastonetes e, conseqüentemente, dificuldade para que o leitor possa capturar a informação do genótipo. Diante dessa realidade, os serviços terceirizados de genotipagem e sequenciamento tornaram mais atrativo, seguro para o usuário/cliente.

Os equipamentos utilizados na rotina do laboratório têm que ser previamente aprovados pelos técnicos da Illumina e o usuário final capacitado para iniciar as atividades (treinamento oferecido pela Illumina). Além disso, calibrações anuais, desde as pipetas automáticas até o scaneador *BeadXPress*, são necessárias para que a empresa possa continuar fornecendo garantia sobre os dados gerados. Diante dessas explicações, pode-se perceber as dificuldades que foram encontradas utilizando a tecnologia *Veracode (BeadXPress)* na geração do dado genômico desta tese.

Paralelo ao exposto, e em reconhecimento de toda essa evolução tecnológica, a empresa Illumina anunciou no início do ano de 2013 que não seriam mais aceitos novos pedidos para o equipamento *BeadXPress*, descontinuando efetivamente a venda de ensaios comerciais e customizados de baixa densidade com essa tecnologia. O contexto tecnológico atual apresenta perspectivas positivas, com novos avanços incrementais sendo lançados frequentemente para as tecnologias já disponíveis para genotipagem e sequenciamento do genoma completo de várias espécies. Adicionalmente, novas tecnologias estão sendo desenvolvidas e aprimoradas, criando possibilidades para novas reduções de custo e

estabelecendo novos limites de aplicação, outrora restritos apenas pela inventividade humana (Oxford Nanopore Technologies®, Schneider e Dekker 2012). A alta cobertura do genoma com marcadores torna possível particionar 100% da variabilidade genética aditiva de uma característica por todo genoma, permitindo estimar se o valor de substituição de alelo em cada um dos *loci* envolvidos com a característica (Caetano, 2009). Esse procedimento permite então estimar o valor genético de um indivíduo com base nos genótipos de todos os marcadores associados com a característica, ou seja, o valor genômico, sem a coleta da medida fenotípica do indivíduo (Meuwissen et al., 2001).

Nesse sentido, GWS ou seleção genômica (GS), proposta por Meuwissen et al. (2001), é uma análise baseada na avaliação de grande número de marcadores amplamente distribuídos ao longo do genoma. Obtidos estes marcadores, são estimados os seus efeitos com base em dados fenotípicos de uma população de estimação. Estimados os efeitos, são obtidos os modelos de predição dos valores genéticos genômicos (VGG). As acurácias dos modelos são obtidas nas populações de validação e, em seguida, eles são aplicados em populações de seleção (Crossa et al., 2011). No melhoramento genético, uma vez preditos os valores genéticos genômicos, estes poderão ser utilizados para a avaliação dos valores genéticos de indivíduos em fases precoces de seleção, com a predição dos fenótipos futuros de indivíduos genotipados em etapas preliminares do programa de melhoramento. Tal estratégia acelera a seleção, tendo em vista que torna possível, em determinados ciclos de seleção, eleger genótipos de interesse sem necessidade de fenotipagem (Resende et al., 2008).

Já existe a disponibilidade de ensaios de baixa densidade (*Infinium BovineLD - 7K*) para aplicações como seleção genômica, teste de paternidade e rastreamento. Os usuários deste chip poderão comparar seus dados àqueles gerados em plataformas de maior densidade (imputação com 98% de confiabilidade), ampliando a aplicação das informações genômicas, já que o novo produto é complementar ao *Bovine SNP50* e outros ensaios semelhantes.

Como evidenciado, o uso de ferramentas genômicas para implementar processos de avaliação de mérito genético resultou em avanços tecnológicos significativos, podendo acarretar aumentos de eficiência e maiores lucros, como consequência de aumentos nas acurácias e reduções nos intervalos de geração. Além disso, ainda existem grandes perspectivas para lançamento de inovações tecnológicas para geração de dados moleculares a curto e médio prazo, com tendências de reduções de custos, inclusive para sequenciamento completo de genomas em várias espécies de interesse zootécnico, aumentando dessa forma os limites que as tecnologias podem fornecer. Essa conjuntura torna cada vez mais importante a



coleta de dados fenotípicos relacionados com as características que impactam a lucratividade dos sistemas produtivos e a qualidade dos produtos gerados, e que podem ser aprimoradas pela incorporação de avaliações genômicas nos programas de melhoramento. Portanto, a definição de objetivos de seleção, com base em índices econômicos definidos em conjunto com o setor produtivo, e a coleta de dados fenotípicos associados, são conseqüentemente de primordial importância para desenvolvimento competitivo e sustentável dos arranjos de produção animal.

#### 4.2. REFERÊNCIAS BIBLIOGRÁFICAS

- BOICHARD, D.; CHUNG, H.; DASSONNEVILLE, R. et al. Design of a Bovine Low-Density SNP Array Optimized for Imputation. **PLoS ONE**, v.7, n.3, p. e34130, 2012.
- CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v. 38, n.spe, p. 64-71, 2009.
- CHESSA, S.; CHIATTI, F.; CERIOTTI, G. et al. Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms. **Journal of Dairy Science**, v. 90, n. 1, p. 451-464, 2007.
- CROSSA, J.; PÉREZ, P.; CAMPOS, G. DE LOS. et al. Genomic Selection and Prediction in Plant Breeding. **Journal of Crop Improvement**, v.25, n.3, p. 239–261, 2011.
- GODDARD, M.; HAYES, B.J. Mapping genes for complex traits in domestic animals and their use in breeding programs. **Nature Reviews Genetics**, v.10, n.6, p.381-391, 2009.
- KAMINSKI, S.; BRYM, P.; RUSC, A. et al. Associations between milk performance traits in Holstein cows and 16 candidate SNPs identified by arrayed primer extension (APEX) microarray. **Animal Biotechnology**, v. 17, n. 1, p. 1-11, 2006.
- LI, Y.; DING, J.; ABECASIS, G.R. Mach 1.0: rapid haplotype reconstruction and missing genotype inference. **American Society of Human Genetics**, v.79, n.10, 2006.
- MARQUES, E. F.; SCHNABLE, R. D.; STOTHARD, P. et al. High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle. **BMC Genetics**, v. 9, n. 45, p. 1-12, 2008.
- MELLO, F.; MARTINS, M.F.; COBUCCI, J.A. et al. Análise de diversidade genética do gene da osteopontina em bovinos da raça Girolando. **Revista Brasileira de Zootecnia**, v. 40, n. 11, p. 2374-2377, 2011.
- MELLO, F.; COBUCCI, J.A.; MARTINS, M.F. et al. Association of the polymorphism g.8514C>T in the osteopontin gene (SPP1) with milk yield in the dairy cattle breed Girolando. **Animal Genetics**, v. 43, n. 5, p. 647-647, 2012.

- MEUWISSEN, T. H.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v.157, n.4, p. 1819–1829, 2001.
- PAIVA, D.S.; FONSECA, I.; PINTO, I.S.B. et al. Incidence of bovine leukocyte adhesion deficiency, complex vertebral malformation, and deficiency of uridine-5-monophosphate synthase carriers in Brazilian Girolando cattle. **Genetics and Molecular Research**, v. 12, n. 3, p. 3186-3192, 2013.
- RESENDE, M.D.V.; LOPES, P.S.; SILVA, R.L. et al. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa Florestal Brasileira**, v. 1, n. 56, p. 63–77, 2008.
- SCHNEIDER, G.F.; DEKKER, C. DNA sequencing with nanopores. **Nature Biotechnology**, v.30, n.4, p.326-328, 2012.
- VAN RADEN, P.M. Impact of Genomics on Genetic Improvement. **Advancing Dairy Cattle Genetics: Genomics and Beyond**. 2014
- VAN TASSEL, C.P. The use of SNP chips for selection of dairy cattle. In: Applied Genomics for Sustainable Livestock Breeding Meeting, 1., 2011, Melbourne. **Proceedings...** Melbourne: 2011. (CD-ROM).