

**ARIOSTO ARDILA SILVA**

**MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DE PRODUÇÃO DE  
LEITE NO CROMOSSOMO 6 EM REBANHOS GYR LEITEIRO**

Tese apresentada à  
Universidade Federal de  
Viçosa, como parte das  
exigências do Programa de  
Pós-Graduação em  
Genética e Melhoramento,  
para obtenção do título de  
*Doctor Scientiae*.

**VIÇOSA**  
**BRASIL - MINAS GERAIS**  
**2009**

**ARIOSTO ARDILA SILVA**

**MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DE PRODUÇÃO DE  
LEITE NO CROMOSSOMO 6 EM REBANHOS GIR LEITEIRO**

Tese apresentada à  
Universidade Federal de  
Viçosa, como parte das  
exigências do Programa de  
Pós-Graduação em  
Genética e Melhoramento,  
para obtenção do título de  
*Doctor Scientiae*.

APROVADA: 16 de março de 2009

---

Marco Antonio Machado  
(Co-orientador)

---

Rui Da Silva Verneque  
(Co-orientador)

---

Paulo Luiz Souza Carneiro

---

José Braccini Neto

---

Simone Eliza Facioni Guimarães  
(Orientadora)

## **AGRADECIMENTOS**

À Dr. Simone E. F. Guimarães, pela oportunidade de ser seu orientado, por sua compreensão e confiança.

Ao Dr. Marco Antonio Machado e ao Dr. Rui Da Silva Verneque, pela oportunidade do aprendizado no Programa Nacional de Melhoramento de Gir Leiteiro.

À Dr. Ana Luisa Souza Azevedo, por sua desinteressada colaboração e orientação no experimento.

Aos colegas do laboratório pela ajuda e partilhar, e em especial a Karla Gasparini e Robert Domingues, pela colaboração no experimento.

À Aline por sua desinteressada amizade, colaboração e partilhar em diferentes aspectos do conhecimento e da vida.

À família Lasallista da Colômbia por seu incondicional apoio.

À Universidade Federal de Viçosa, pela oportunidade de realizar os estudos.

À Embrapa gado de leite pela realização do experimento.

## SUMÁRIO

	Página
RESUMO.....	iv
ABSTRACT.....	v
1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	3
2.1 Subespécies <i>Bos taurus</i> e <i>Bos indicus</i> .....	3
2.2 Composição do leite.....	4
2.3 Programa Nacional de Melhoramento Genético do Gir leiteiro no Brasil (PNMGL).....	7
2.4 Metodologias empregadas na obtenção dos valores genéticos.....	9
2.5 Marcadores Moleculares.....	12
2.6 Mapeamento de QTL.....	13
2.7 Delineamento experimental para detecção de QTL.....	16
2.8 Seleção Assistida por Marcadores.....	20
2.9 Identificação de QTL no cromossomo 6 em bovinos de leite..	21
3. MATERIAL E MÉTODO.....	25
3.1 Metodologias para estimação dos valores genéticos.....	25
3.2 Famílias utilizadas para o mapeamento de QTL no rebanho Gir Leiteiro.....	30
3.3 Coleta e extração de DNA.....	31
3.4 Genotipagem com marcadores microssatélites.....	32
3.5 Mapa de ligação.....	35
3.6 Delineamento para detecção de QTL.....	35
4. RESULTADOS E DISCUSSÃO.....	38
4.1 <i>Scoring</i> dos marcadores microssatélites.....	38
4.2 Mapa de ligação.....	43
4.3 Mapeamento de QTL.....	46
5. CONCLUSÕES.....	77
6. RERERÊNCIAS BIBLIOGRÁFICAS.....	78

## RESUMO

SILVA, Ariosto Ardila, D.Sc., Universidade Federal de Viçosa, março de 2009.  
**Mapeamento de QTL para características de produção de leite no cromossomo 6 em rebanhos Gir Leiteiro.** Orientadora: Simone Eliza Facioni Guimarães. Co-Orientadores: Marco Antonio Machado e Rui Da Silva Verneque.

A raça Gir Leiteiro caracterizada por sua rusticidade e adaptação ao trópico é uma boa alternativa para produção de leite e, fundamental na formação da raça Girolando (5/8 Holandesa: 3/8 Gir). O propósito do presente estudo foi mapear QTL sobre o cromossomo 6 para nove características de produção de leite, por meio de um delineamento de filhas. Quatorze famílias de touros Gir Leiteiro com 657 filhas, foram analisadas por meio de 27 marcadores microssatélites. Considerando todas as famílias foram identificados QTL com nível de significância de 1% cromossômico para produção de leite, produção de gordura e porcentagem de gordura. Ao utilizar as famílias 3 e 9, um QTL para produção de leite foi identificado na posição 46 cM, próximo ao marcador BMS2508 e com  $F = 9,64$ ; na família 3 foi mapeado o mesmo QTL para produção de gordura na posição 48 cM próximo ao marcador DIK4382 e com  $F = 16,61$ ; o QTL para porcentagem de gordura foi identificado nas famílias 3 e 4, na posição 56 cM, próximo ao marcador DIK4482, e com  $F = 15,18$ , e para porcentagem de sólidos totais, nas mesmas famílias 3 e 4 foi mapeado um QTL na posição 57 cM, próximo ao marcador DIK4482, com  $F = 7,61$ . Para produção de proteína, produção de lactose e produção de sólidos totais foram mapeados QTL com 5% de significância a nível cromossômico, utilizando as famílias mais significativas. Na região média do BTA6, próximo ao marcador BM143 (53 cM), entre os genes ABCG2 e LAPs, estão localizados seis genes incluindo Osteopontina (OPN), PPARGC1A, FAM13A1 e PKD2, onde, diferentes grupos pesquisaram possíveis associações entre genes nesta região e características de produção de leite em bovinos leiteiros, sendo uma boa alternativa saturar esta região com mais marcadores microssatélites e SNPs, para conferir a presença dos QTL mapeados e, identificar outros QTL ou genes candidatos relacionados com características de produção de leite.

## ABSTRACT

SILVA, Ariosto Ardila, D.Sc., Universidade Federal de Viçosa, March, 2009. **Mapping of QTL for milk production traits on bovine chromosome in Gyr dairy cattle.** Adviser: Simone Eliza Facioni Guimarães. Co-Advisers: Marco Antonio Machado and Rui Da Silva Verneque.

In tropical countries, Gyr (*Bos indicus*) breed is very important due to their heat and parasite tolerance, and essential to girolando cattle formation (5/8 Holstein, 3/8 Gyr). Dairy Gyr is much more adapted to climate and economical traits in Brazil and with crucial qualities for a good milk production on the tropics. The dairy Gyr was the first Zebu breed in Brazil and the world with a program of genetic evaluation since 1985, through of progeny tested sires. This study was done with the objective of identifying QTL on chromosome 6 for nine milk production characteristics through daughters design. Fourteen families of sire Gyr with 657 daughters were analyzed using 27 microsatellite markers. In the analysis across 14 families and within the largest significant families, four-QTL were detects for milk yield, fat yield, fat percentage and total solids yield, at the 1% chromosome-wide. With families three and nine, a QTL was detected for milk yield, in the position 46 cM, near BMS2508 marker and with  $F = 9,64$ ; with family three, a QTL was detected for fat yield, in the position 48 cM, near DIK4382 marker and with  $F = 16,61$ ; the QTL for fat percentage was detected with families three and four, in the position 56 cM, near DIK4482 marker and with  $F = 15,18$ ; and for total solids percentage, a QTL was detected with the families three and four, in the position 57 cM and with  $F = 7,61$ . With the significant families, QTL that affects protein yield, lactose yield and total solids yield, were mapping at the 5% chromosome-wide. The central autosomal region of the chromosome is very problematic for interpretation and localization of data, for possible QTL interference and can present pleiotropic QTL effects. In this study the majority of the QTL were detected in this region. In the middle region of BTA6 next to marker BM143 (53cM) between genes ABCG2 and LAPs are located 6 genes including Osteopontin (OPN), PPARGC1A, FAM13A1 and PKD2. Different groups were investigated the possible association between genes in this region and the milk production characteristics in dairy cattle. The analysis showed the localization of the QTL near this region, which can slow the identification of the candidate genes. The path to solve this problem is

doing genetic mapping with precision using molecular markers SNP (Single Nucleotide Polymorphisms).

## 1. INTRODUÇÃO

O Brasil é detentor de uns dos maiores rebanhos comerciais bovino do mundo, e em sua composição destacam-se as raças zebuínas, introduzidas no país a partir do século XIX, originárias, principalmente da Índia e, em menor escala, do Paquistão. Segundo dados da FAO (Food and Agriculture Organization), o Brasil é o sexto país colocado na classificação mundial de produção de leite bovino, o que equivale a uma produção de mais de 25 bilhões de toneladas de leite por ano, porém, a produção de leite caracteriza-se pela baixa produtividade média, menos de 1.500 litros/vaca/ano (IBGE - Instituto Brasileiro de Geografia e Estatística, 2006).

A vaca Holandesa apresenta maior capacidade de produção de leite no mundo, porém, a dificuldade de adaptação aos trópicos torna difícil e onerosa a produção de leite. Nesse contexto é que a raça Gir mostra a sua grande importância, tanto como raça pura, como nos cruzamentos, contribuindo para a formação do Girolando (5/8 Holandesa: 3/8 Gir). Calor e umidade são fatores reconhecidamente prejudiciais para a reprodução das raças taurinas, e seus efeitos são agravados pela alta produção do leite. A incapacidade do bovino europeu em eliminar o excesso de calor pela deficiência de seu aparelho termorregulador torna difícil sua adaptação aos trópicos, diminuindo sua produção de leite.

Uma alta percentagem do leite produzido no Brasil provém de vacas mestiças leiteiras que apresentam uma alta porcentagem de genes das raças zebuínas. O Brasil tem obtido grandes avanços em ganho genético em seu rebanho leiteiro por meio do melhoramento genético clássico e técnicas reprodutivas como inseminação artificial e mais recentemente, com a transferência de embriões e a fecundação *in vitro*, principalmente para características de fácil mensuração e alta herdabilidade. Características quantitativas complexas, de mensuração tardia, baixa herdabilidade e limitadas pelo sexo têm sido mais difíceis de serem melhoradas com base exclusivamente no emprego do melhoramento convencional. Esses ganhos genéticos poderão ser acelerados mediante o uso dos marcadores moleculares



associados a locos de características quantitativas (QTL), como ferramentas que auxiliam e aprimoram esse tipo de melhoramento.

A maior parte dos estudos de detecção de QTL, associados a características de produção e qualidade no leite do genoma bovino, tem sido realizada em raças leiteiras *Bos taurus*, que apresentam limitações de adaptação e produção à maior parte do clima do Brasil. A pecuária leiteira no Brasil caracteriza-se pela abrangência de animais mestiços e zebuínos, sendo necessário mapearem os locos que afetam as características produtivas em raças leiteiras Zebuínas, como é o caso da raça Gir Leiteiro. O BTA6 é um dos cromossomos mais estudados, sendo que QTL significativos e genes candidatos, associados às características de produção de leite, têm sido mapeados, principalmente, nas raças taurinas. Portanto, espera-se que no mesmo cromossomo da raça Gir Leiteiro sejam encontrados QTL associados às características de produção do leite.

Em bovinos leiteiros, características como produção de leite só se expressa em um sexo, e o intervalo entre gerações é bem amplo, sendo necessários esperar até cinco anos para avaliar um touro; diante disso, a seleção assistida por marcadores torna-se uma ferramenta bem útil para acelerar os ganhos genéticos e reduzir intervalos entre gerações. A identificação de regiões genômicas associadas às características de produção e de qualidade no leite possibilitará uma maior eficiência na seleção para estas características, possibilitando selecionar os futuros touros para o teste de progênie desde menores idades.

O maior objetivo de mapear QTL é encontrar genes/marcadores que possam ser utilizados em programas de melhoramento por meio da seleção assistida por marcadores (MAS) e, especificamente, em gado de leite podem ser empregados para selecionar touros jovens, como futuros candidatos ao teste de progênie, assim, incrementar o diferencial de seleção, reduzir o intervalo de gerações e aumentar o ganho genético. O objetivo do trabalho foi mapear QTL no cromossomo 6 associado às características de produção (de leite, de gordura, de proteína, de lactose e de sólidos totais) e qualidade do leite (percentagens de gordura, proteína, lactose e sólidos totais) na raça bovina Gir leiteiro.

## 2. REVISÃO DE LITERATURA

### 2.1 Subespécies *Bos taurus* e *Bos indicus*

Os bovinos pertencem à família *Bovidae*, subfamília *Bovinae*, incluindo, dentre outros, os gêneros *Bubalus*, *Bos* (boi doméstico) e *Bison*. A comissão Internacional de Nomenclatura Zoológica (ICZN) estabeleceu em 2003 que a denominação mais apropriada fosse *Bos primigenius taurus* e *Bos primigenius indicus*, como duas subespécies. Estudos moleculares e arqueológicos mostram que o processo de domesticação bovina, que se iniciou 10.000 anos atrás, ocorreu em dois pontos diferentes, a partir do *Bos primigenius*, o que explica a grande diversidade genética observada entre raças taurinas e indianas (Edwards *et al.*, 2004; Loftus *et al.*, 1994).

No Brasil, o rebanho bovino foi formado pelo gado de origem européia (taurina), trazido de Portugal e Espanha na época da colonização. Com a importação das raças zebuínas, a maior parte do rebanho foi sofrendo um processo de substituição por meio de cruzamentos absorventes com touros zebus (Santiago, 1986). Isso significa que as vacas taurinas foram cruzadas com touros zebus e, assim também as gerações subsequentes até que os produtos foram se tornando fenotipicamente muito parecidos com o zebu puro.

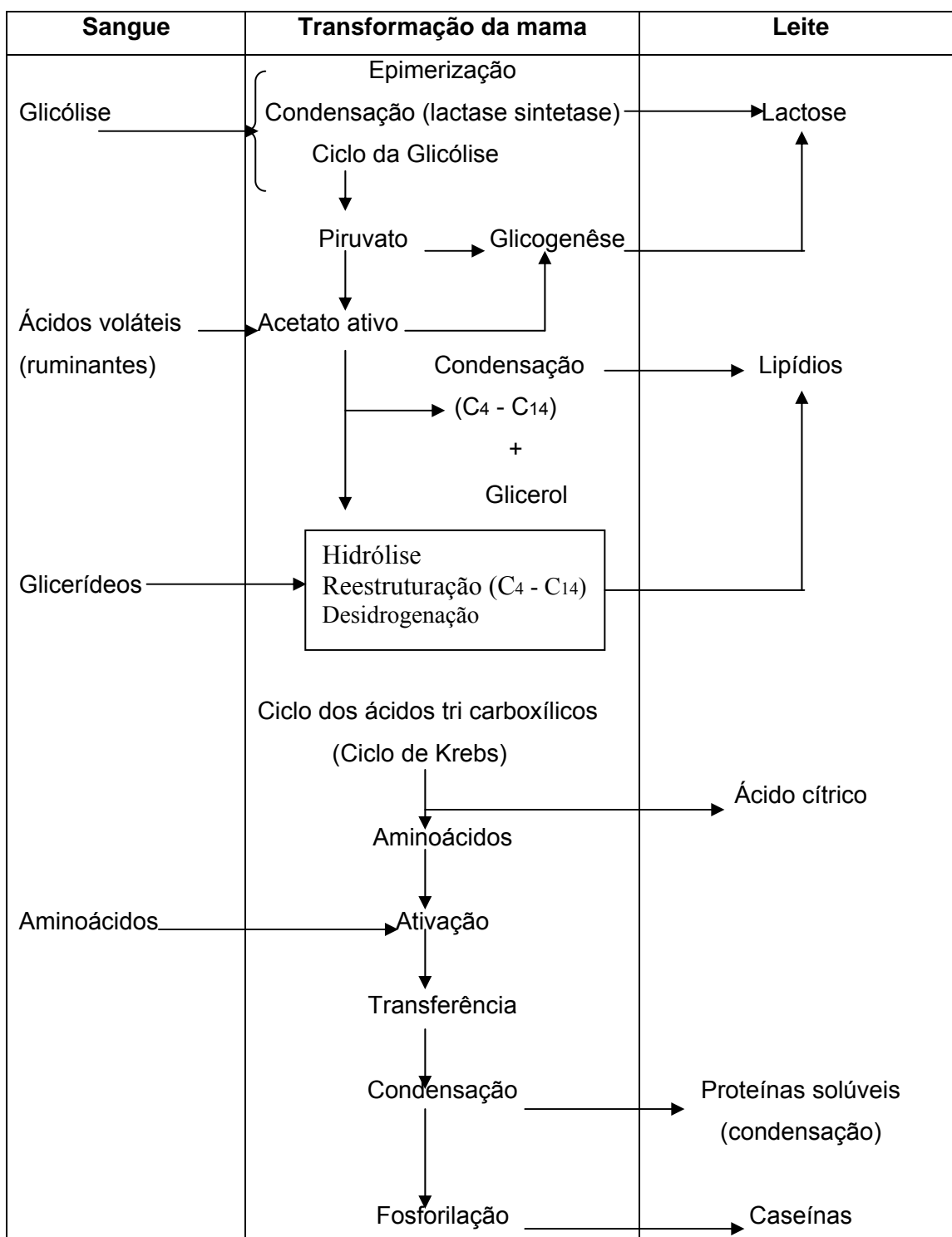
A raça Gir entrou no Brasil provavelmente no ano de 1906, mas as importações mais importantes aconteceram em 1930, 1960 e 1962, e estas tiveram grande contribuição para a formação do pool gênico da raça nos dias atuais. A raça é nativa da Índia, mais especificamente da península de Kathiavar onde foi selecionada para leite (Santos, 2007). No Brasil a raça foi utilizada inicialmente para corte, mas a partir de 1930, iniciou-se a seleção para leite, embora de forma empírica. Nos dias atuais, a raça Gir é utilizada para leite, ainda que exista um grupo de criadores que a utilize também para carne. O Gir leiteiro mostra-se como a raça preferencialmente utilizada em cruzamentos com gado taurino leiteiro, contribuindo com leite, rusticidade, longevidade, docilidade e habilidade materna, apresentando menores infestações de endo e ectoparasitas, e menores incidências de doenças do que

raças do clima temperado, isso traz como consequência um menor uso de carrapaticidas, vermífugos e antibióticos, proporcionando um produto final livre de resíduos, portanto, mais saudável.

## **2.2 Composição do leite**

O leite pode ser definido como um produto da secreção da glândula mamária de mamíferos, e de acordo com o Regulamento Industrial e Sanitária de Produtos de Origem Animal (RIISPOA), artigo 475, denomina-se leite, sem outra especificação, o produto normal, fresco, integral, oriundo da ordenha completa e ininterrupta de vacas sadias. A anatomia do animal e fatores fisiológicos possibilita a capacidade de produção, armazenamento e descida do leite para a cisterna do úbere. No quadro número 1 resume-se a biossíntese dos componentes do leite.

**Quadro 1 - Biossíntese dos componentes do leite**



Fonte: Alias C. H. (1985)

As proteínas resultam parcialmente da síntese de aminoácidos livres e da filtração do plasma sanguíneo. A caseína e a albumina do leite não estão presentes no sangue, sendo sintetizadas a partir dos precursores. Já a globulina do leite é idêntica à do sangue. Os lipídios, os ácidos graxos e o

glicerol que constituem os glicerídeos do leite, procedem, em parte, da corrente sanguínea, porém outra quantidade se sintetiza na mama a partir de moléculas pequenas. No caso dos ruminantes, o acetato e o beta-hidroxibutirato são os precursores mais importantes da síntese. Já a lactose se deriva, principalmente, da glicose sanguínea, e no tecido mamário se une à galactose para ser formada. A água, componente maior no leite (87 a 90%) é retirada diretamente do sangue e serve de veículo para outros constituintes que se encontram em solução e suspensão. Já os minerais do leite, principalmente o cálcio e o fósforo, provêm do sangue por filtração, mas alguns passam por combinações biológicas especiais (Tronco, 2003).

Segundo Madalena (2008), aproximadamente 12% do leite é formado por substâncias sólidas: proteína, gordura, lactose e minerais, sendo a água os 88% restantes. São os sólidos, especialmente as proteínas, que fazem do leite um produto importante, com alto valor nutritivo, e os que servem de matéria-prima para a elaboração dos diferentes produtos lácteos, como queijo, manteiga, iogurtes, sorvetes, etc. A composição média do leite de vaca está especificada no quadro 2.

## Quadro 2 - Composição do leite em percentagens.

<b>Água</b>			87,3
<b>Extrato seco total</b>	Gorduras		3,6
	Extrato seco	Proteínas.....3,3	9,1
	Desengordurado	Lactose.....4,9 Minerais.....0,9	
<b>Vitaminas</b>			-
<b>Enzimas</b>			-
<b>Pigmentos</b>			-
<b>Gases dissolvidos (CO<sub>2</sub>, O<sub>2</sub>, N)</b>			-

Fonte: Tronco (2003).

### 2.3 Programa de Melhoramento do Gir Leiteiro no Brasil

O Gir foi a primeira raça zebuína no Brasil e no mundo a ter um programa de avaliação genética para leite por meio do teste de progênie. O PNMGL, iniciado em 1985 é um trabalho executado pela Embrapa Gado de Leite em parceria com a Associação Brasileira de Criadores de Gir Leiteiro (ABCGIL), e com a participação de diversos órgãos públicos e privados, tais como a Associação Brasileira de Criadores de Zebu (ABCZ), as centrais de processamento de sêmen, CNPq, Fapemig, empresas estaduais de pesquisa agropecuária do Estado de Minas Gerais, criadores de gado Gir Puro e outras fazendas colaboradoras. O objetivo principal do programa é promover o melhoramento genético da raça Gir por meio da identificação e seleção de touros geneticamente superiores para as características de produção (leite, gordura, proteína, lactose e sólidos totais), de conformação e manejo (Verneque *et al.*, 2008).

Segundo Martinez e Verneque (2001), a avaliação de um touro por meio do desempenho de suas progênies constitui-se na maneira mais segura e eficiente de se prever com precisão a capacidade do animal transmitir sua superioridade (ou inferioridade) genética aos seus descendentes. Os resultados do teste de progênie do primeiro grupo de touros da raça Gir foram

publicados em 1993 e, a partir de então, anualmente são publicados resultados de novos grupos de touros. De acordo com estes mesmos autores, os touros depois de avaliados são classificados de acordo com seu valor genético (VG) ou pela capacidade prevista de transmissão (PTA, Predited Transmitting Ability) que é igual à metade do VG. Do anterior, pode-se deduzir que um teste de progênie vem sendo um teste de comparação de touros, e que os melhores classificados devem ser utilizados para produção de progênies da próxima geração.

O sucesso do Programa Nacional de Melhoramento de Gir Leiteiro depende da estrutura, da qualidade e da quantidade dos dados e da metodologia de avaliação genética. Atualmente, tem-se usado nas avaliações genéticas o BLUP (Melhor Previsor Linear Não-Viesado) sob modelo animal com um modelo de repetibilidade. Dentro do PNMGL, Martinez (1989) descreveu objetivamente os principais fatores não-genéticos que podem interferir na estimação do valor genético de um animal: controle leiteiro não-seletivo, acasalamento ao acaso, uso de inseminação artificial, tratamento não-preferencial. Os estágios do teste de progênie no Gir leiteiro no Brasil são apresentados por Verneque *et al.* (1998) como: escolha dos touros a testar, coleta e distribuição de sêmen aos rebanhos colaboradores, número de touros e vacas, controle reprodutivo, nascimentos das progênies, controle produtivo e outras medidas de interesse, avaliação genética e divulgação dos resultados.

No Gir Leiteiro, as avaliações genéticas para as características de produção (leite, gordura, proteína, lactose e sólidos totais), conformação (altura de garupa, perímetro torácico, etc.) e manejo (facilidade de ordenha e temperamento) são realizadas usando-se os procedimentos do modelo animal. De acordo com Verneque *et al.* (2008) o modelo animal, aliado às metodologias adequadas de estimação e predição, representa o que há de mais moderno para calcular as capacidades previstas de transmissão. As avaliações são baseadas nas mensurações do próprio animal (vaca) e nas mensurações de parentes que estão sendo avaliados. As informações do animal propriamente dito, de seus ancestrais, colaterais e suas progênies são incluídas por meio da matriz de parentesco entre os animais avaliados. As informações das famílias das vacas são utilizadas com a inclusão dos registros de produção de todas as fêmeas ancestrais e descendentes (Verneque *et al.*, 2008).

## 2.4 Metodologias utilizadas na obtenção dos valores genéticos

Um número de fatores não-genéticos ou de meio ambiente afeta a produção de leite das vacas e contribui com aproximadamente 75% da variação da produção. Alguns são externos, como região, rebanho e diferenças sazonais de ano para ano. Outros são internos ou de natureza fisiológica, como idade, gestação, lactação, efeitos maternos etc. Muitos destes fatores podem ser medidos e os seus efeitos sobre a produção conhecidos, permitindo o estabelecimento de padrões de variação. Entre os mais importantes destacam-se: duração de lactação, número de ordenhas, idade da vaca, época de parição, período de serviço, período seco etc. Nas avaliações genéticas de vacas e touros é necessário que se considerem todos os fatores de meio ambiente que influenciam o desempenho dos animais, ajustando-se os registros para uma base comum de comparação. Não sendo possível o ajustamento prévio, os efeitos de ambiente devem ser considerados no modelo de avaliação (Teixeira, 2001).

Segundo Ferraz *et al.* (1999), avaliar a qualidade genética de um animal nada mais é do que estimar o seu valor genético aditivo. Jamais se conhecerá o valor que um animal tem como reprodutor, mas, através de metodologias diversas, é possível ter-se uma ideia desse valor. A estimativa do valor genético dos animais depende da herdabilidade do caráter (quanto maior a herdabilidade maior a concordância entre o genótipo e o fenótipo), do número e qualidade das informações, do parentesco entre o animal avaliado e as fontes de informação (quanto mais próximo o parentesco, maior a ênfase que a informação deve ter) e do grau de semelhança fenotípica entre o animal avaliado e as fontes de informação (uma forma de avaliar os efeitos de ambiente que é comum a diferentes fontes de informação).

As metodologias de estimação de valores genéticos dos animais vêm se aperfeiçoando desde as primeiras décadas do século passado, mas apenas a partir dos anos 80 apresentaram avanços maiores e se criaram ferramentas poderosas para a avaliação dos animais, principalmente após o advento dos chamados modelos animais. Nessas metodologias utilizam-se os registros de produção do animal que está sendo avaliado, mas também de todos seus parentes, não importa o grau de parentesco existente. Assim, dados a respeito de primos distantes, avós, filhos, filhas, irmãos e irmãs de um reprodutor são



avaliados segundo complexos modelos estatísticos o que resulta na estimação do seu valor genético.

Os modelos animais, aplicação dos chamados modelos mistos, deduzidos por Henderson nos anos 40 consideram de forma simultânea efeitos fixos (sexo, idade da vaca ao parto, fazenda e grupo de contemporâneos) e aleatórios (como os efeitos diretos dos genes dos animais e os efeitos dos genes da mãe), os efeitos permanentes de ambiente que uma vaca oferece à sua progênie de maneira simultânea; Os efeitos aleatórios dos genes dos animais (direto e materno) são estimados com base nas informações fornecidas pelos registros de produção de todo e qualquer parente do animal. Isso se tornou possível porque na montagem das equações de modelos mistos a matriz de parentesco é adicionada à parte aleatória do sistema, “ligando” os dados de todos os animais que por ventura tenham parentesco. Como os dados estão conectados no momento em que as soluções são obtidas, o valor genético estimado de cada animal é obtido considerando-se as contribuições de cada parente (Sterman e Pereira, 2005).

Na avaliação do mérito genético das vacas para as características de produção, é utilizado o modelo animal com repetibilidade. Medidas repetidas têm aplicação no melhoramento animal por meio da seleção. O parâmetro associado à correlação fenotípica entre medidas repetidas em um mesmo indivíduo está relacionado com a repetibilidade que mede a capacidade dos organismos em repetir a expressão do caráter ao longo de vários períodos de tempo (ou no espaço) no decorrer de suas vidas (Vilela, 2002). Do ponto de vista do melhoramento animal, este parâmetro apresenta importância fundamental na predição de valores genéticos, genotípicos e na inferência sobre o aumento de eficiência seletiva, pelo uso de um determinado número ( $> 1$ ) de medidas por indivíduo (Vilela, 2002).

A repetibilidade no caso características de produção de leite refere-se às correlações fenotípicas entre medidas repetidas no mesmo indivíduo, ou seja, repetibilidade individual, mas no melhoramento animal também é importante repetibilidade maternal. Uma condição essencial para a adoção do modelo de repetibilidade é que as medidas repetidas apresentem correlação genética igual ou próxima a um, fato que confirma o mesmo caráter de uma medida para outra, caso contrário, deve-se adotar um modelo multivariado, o qual trata cada medida como um caráter distinto.

As medidas em diferentes animais são independentes, assim, a covariância concernente é unicamente com medidas no mesmo animal. A estrutura de covariância refere-se a variâncias em tempos individuais e a correlação entre medidas em diferentes tempos no mesmo animal. Há dois aspectos básicos de correlação: duas medidas no mesmo animal são correlacionadas simplesmente porque elas modelam contribuições comuns do animal. Isto é devido à variação entre animais e medidas no mesmo animal, próximo no tempo que são frequentemente mais altamente correlacionadas que medidas afastadas no tempo.

A estrutura fenotípica para três observações de um indivíduo sob este modelo pode ser escrito (Quaas, 1984) como:

$$\text{var} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \sigma_{t1}^2 + \sigma_{pe}^2 + \sigma_g^2 & \sigma_{pe}^2 + \sigma_g^2 & \sigma_{pe}^2 + \sigma_g^2 \\ \sigma_{pe}^2 + \sigma_g^2 & \sigma_{t2}^2 + \sigma_{pe}^2 + \sigma_g^2 & \sigma_{pe}^2 + \sigma_g^2 \\ \sigma_{pe}^2 + \sigma_g^2 & \sigma_{pe}^2 + \sigma_g^2 & \sigma_{t3}^2 + \sigma_{pe}^2 + \sigma_g^2 \end{bmatrix}$$

com  $\sigma_{ti}^2$  = variância do ambiente temporário específica ao registro  $i$ ;  $\sigma_{pe}^2$  = variância devida ao efeito ambiental permanente (variância e covariância são iguais);  $\sigma_g^2$  = variância genotípica. A correlação entre registros de um indivíduo referido, com repetibilidade, é  $(\sigma_g^2 + \sigma_{pe}^2) / (\sigma_y^2)$ . A avaliação genética sob este modelo é concernente não unicamente a valores genéticos preditos, mas também ao efeito ambiental permanente.

De acordo com Verneque (1994), os valores genéticos podem ser melhorados enquanto a sua precisão pelo uso de grandes conjuntos de dados, pela remoção das influências não genéticas, pela incorporação de dados de parentes por meio de modelos mistos apropriados e, especialmente, pelo uso de um método numérico para a estimação dos componentes de covariância e para a precisão do valor genético do animal. A remoção das influências não genéticas foi feita através do uso dos modelos mistos, adotando o modelo animal, que além de permitir realizar os ajustes para os efeitos fixos, tem em conta os dados do próprio animal e de seus parentes através da matriz de parentesco, facilitando, em uma única análise realizar a avaliação do próprio indivíduo e de seus parentes.

Diferentes metodologias têm sido sugeridas para a estimação dos componentes de variância para características múltiplas e dados não balanceados: métodos I, II e III de Henderson (Henderson, 1953); Método de Estimação Quadrática Não-Viesada de Norma Mínima - MINQUE (Minimum Norm Quadratic Unbiased Estimators) (Rao, 1970, 1971a); Método de Estimação Quadrática Não-Viesada de Variância Mínima - MIVQUE (Minimum Variance Quadratic Unbiased Estimation) (Rao, 1971b); Método da Máxima Verossimilhança - ML (Maximum Likelihood) (Hartley e Rao, 1967); Método de Máxima Verossimilhança Restrita - REML (Restricted Maximum Likelihood) (Patterson e Thompson, 1971) e o Método Bayesiano - VEIL (Variance Estimation from Integrated Likelihoods) (Gianola e Foulley, 1990). Baseados no método REML, existem diferentes softwares para o cálculo dos componentes de variância. Os mais importantes são: software DFREML (Derivate Free Restricted Maximum Likelihood) ou Método de Máxima Verossimilhança Restrita utilizando o procedimento Livre de Derivadas; sistema MTC (Multitrait REML Estimation) e o sistema MTDFREML (Multiple Trait Derivate Free Restricted Maximum Likelihood) ou Estimação pelo Método da Máxima Verossimilhança Restrita adotando-se o procedimento Livre de derivadas (DF).

## **2.5 Marcadores Moleculares**

Existe um amplo número de marcadores moleculares, entre os mais comuns, isoenzimas, microssatélites e SNP (Single Nuclotide Polymorphisms). Os marcadores moleculares apresentam algumas vantagens em relação aos marcadores morfológicos, tais como: níveis de polimorfismo, em geral são codominantes e contêm maior quantidade de informação genética por loco (Ferreira e Gratapaglia, 1998).

Os marcadores microssatélites denominados SSR (Simple Sequence Repeats) são constituídos de repetições agrupadas de 1 a 6 pares de base, sendo o dinucleotídeo (TG)<sub>n</sub> a repetição mais abundante em mamíferos (Ellegren, 1993). De acordo com o mapeamento de vários organismos, os microssatélites são distribuídos de forma aleatória e uniforme por todo o genoma (Chin *et al.*, 1996). Estes marcadores moleculares são comuns a todos os organismos eucariotes podendo ser usados em mapeamento de ligação.

Em virtude do alto polimorfismo, as regiões que apresentam microssatélites podem ser facilmente identificadas depois de amplificadas pela PCR e visualização em gel de poliacrilamida ou genotipagem em eletroforese capilar. Estes marcadores moleculares são os marcadores que mais têm contribuído para a elaboração de mapas de ligação na espécie bovina (Kappes *et al.*, 1997), reconstrução de relações filogenéticas, em estudos de genética de populações, teste de paternidade ou em processos de busca de QTL (Darvasi *et al.*, 1993).

Ainda que marcadores microssatélites não identifiquem marcadores ligados diretamente à produção, eles podem ajudar a sua localização de forma indireta, pois os alelos microssatélites segregam junto com os genes ou locos que podem estar associados à produção. Com a estratégia de localização de características quantitativas por meio do uso de microssatélites em procura de polimorfismos, já foi possível a localização de uma série de genes de efeitos maiores (genes que têm um grande efeito na produção ou outras características isoladamente) (Lee, 2002).

Os marcadores SNP são marcadores genéticos que se baseiam na detecção de polimorfismos resultantes da alteração de uma única base no genoma, possuem natureza bi-alélica e são abundantes no genoma, podendo ser encontrados em regiões expressas e não-expressas. No genoma humano foi estimada a presença de um SNP a cada 1000 bases ou menos (Weiner e Hudson, 2002). Do anterior pode-se concluir que, dada sua alta densidade no genoma, adicionada às novas tecnologias de genotipagem, estes marcadores podem ser utilizados no melhoramento assistido por marcadores moleculares, no mapeamento de EST (Expressed Sequence Tags) e na integração de mapas físicos e genéticos.

## **2.6 Mapeamento de QTL**

O mapeamento genômico compreende o mapa de ligação, o mapa físico e o mapa comparativo. Os mapas de ligação são representações estatísticas de marcadores moleculares nos cromossomos, e mostra as distâncias estimadas entre marcadores a partir de frequências de recombinação. Morgan em 1910 foi o primeiro em sugerir que os genes são herdados em conjunto dentro de unidades específicas e, o primeiro mapa de ligação foi elaborado por

seu estudante Stutervant em 1913. O objetivo do mapeamento de QTL é caracterizar genes afetando características e identificar as mutações básicas de variação genética, o que levará a uma importante compreensão da estrutura e função do genoma, sendo um complemento importante ao melhoramento clássico (Olsen *et al.*, 2004).

A detecção de ligação genética entre marcador e QTL depende da existência de desequilíbrio de ligação entre os alelos do loco marcador e do QTL. Se existir equilíbrio de ligação os alelos do loco marcador e dos QTL ocorrerão em combinações equivalentes à distribuição independente, não sendo detectados os QTL. O desequilíbrio produz associação entre o loco marcador e o QTL segregante (Falconer e Mackay, 1996; Lynch e Walsh, 1998). Desequilíbrio de ligação pode ser uma consequência da proximidade física dos genes, ou seja, quando os genes estão fisicamente na mesma molécula de DNA que forma um cromossomo. Em termos gerais, locos que estão fisicamente muito próximos são herdados em bloco, e os distantes, herdados independentemente.

As distâncias de mapa em centiMorgan (cM) ou unidade de mapa, é calculada com base na frequência de recombinação, mas a ocorrência de permutas duplas faz com que essas distâncias não sejam aditivas. Diversas funções de mapeamento transformam essa não aditividade do valor de recombinação, como a de Haldane e Kosambi. A função de Haldane admite independência das permutas nos intervalos adjacentes, enquanto a função de Kosambi considera a interferência. Diferentes metodologias foram deduzidas para a construção de mapas de ligação, como por exemplo: Mapmaker (Lander *et al.*, 1987); Crimap (Green *et al.*, 1990) e GMendel (Liu e Knapp, 1991).

O teste estatístico mais simples utilizado para verificar a presença de ligação entre dois marcadores é o teste  $X^2$ . É um teste qualitativo, pois só comprova a existência ou não de ligação, não sendo possível estimar a frequência de recombinação entre os dois marcadores. Se dois genes segregam individualmente, como previsto na primeira lei de Mendel, porém em conjunto não seguem a lei da independência (segunda lei de Mendel), então eles estão ligados. A independência entre dois genes ou marcadores pode ser verificada pelo teste de  $X^2$  (Schuster e Cruz, 2004). Ligação gênica é a existência de dois ou mais genes localizados num mesmo cromossomo e que possuem menos de 50% de recombinação entre dois marcadores

consecutivos. Quando existe ligação, outros testes precisam ser aplicados, sendo que o procedimento mais aplicado é baseado no método da máxima verossimilhança.

Quando os marcadores estão segregando e, ao mesmo tempo, estão associados aos valores fenotípicos de uma característica de interesse, existe evidência que um gene que está controlando algum componente desta característica está localizado próximo a estes marcadores (Liu, 1998). O princípio básico para a identificação de QTL é que os marcadores moleculares se encontrem ligados aos locos que controlam as características de interesse. Para identificar QTL por ligação, avaliam-se os indivíduos por seu genótipo para o marcador e por seu fenótipo para a característica quantitativa. No caso existir diferenças entre as médias fenotípicas das classes genotípicas estabelecidas para o marcador, pode inferir-se a presença de um QTL ligado a este último.

Existem diferentes métodos para detectar QTL, entre eles, mapeamento por marcas simples, mapeamento por intervalos simples, mapeamento por intervalo composto e mapeamento de múltiplos intervalos. Igualmente, várias metodologias são utilizadas para a detecção de QTL, teste T, análise de variância, regressão linear (teste F) e máxima verossimilhança. Para a análise de QTL pode ser aplicado um teste-t de diferenças entre médias. Alternativamente, podem-se aplicar regressões lineares para cada marcador em relação à(s) característica(s) quantitativa(s), e o grau de significância estatística obtido é indicativo de ligação genética (Grattapaglia e Ferreira, 1998).

No método de máxima verossimilhança, em se tratando da segregação de dois locos em uma população, são possíveis duas situações: locos independentes ou ligação genética. Ao considerar a hipótese de que os genes são independentes e, também, a hipótese alternativa de que os genes estejam ligados, a razão entre as probabilidades associadas a cada hipótese é dada por:  $\Theta = P(\text{genes ligados}) / P(\text{genes independentes})$ . O logaritmo decimal desta razão é conhecido como LOD e, normalmente quando superior a 3, tem-se grande evidência de ligação fatorial:  $\text{LOD} = \text{Log}(\Theta)$  (Schuster e Cruz, 2004).

A metodologia de mapeamento por intervalo foi proposta inicialmente, por Lander e Botstein (1986) que se fundamenta na relação de ligação entre um QTL e os marcadores flanqueadores. Para identificar a associação entre os

marcadores flanqueadores e o QTL, podem ser utilizadas metodologias como modelos lineares, baseados nas diferenças entre as médias dos marcadores, e a metodologia da máxima verossimilhança. Para a análise por intervalo é necessário conhecer a distância entre os marcadores e, a partir dessas informações é investigado todo o intervalo a cada 1 cM, procurando identificar QTL na região. A cada ponto são estimados todos os parâmetros do modelo, seja pela análise de regressão ou pela derivação da função de verossimilhança, gerando duas verossimilhanças, uma baseada na hipótese de presença de QTL, e a outra na ausência de QTL. Os valores obtidos são analisados por meio da razão das duas verossimilhanças gerando um LOD para cada ponto da análise ao longo dos intervalos.

O primeiro mapa de ligação bovino continha 746 marcadores, a maioria microssatélites, com intervalos médios de 5 cM (Barendse *et al.*, 1994). Já no segundo mapa bovino o número de marcadores passou para 1250, e o intervalo médio caiu para 3 cM (Kappes *et al.*, 1997). Um terceiro mapa foi conformado por 2325 marcadores microssatélites com intervalo médio de 1,4 cM (Ihara *et al.*, 2004).

## **2.7 Delineamento Experimental para detecção de QTL**

A detecção e a estimação de QTL individuais são fundamentais para a aplicação da seleção assistida por marcadores (MAS) como ferramenta importante no melhoramento animal. A detecção de QTL depende do tipo de marcadores usados, sua distribuição ao longo do cromossomo, o delineamento e a magnitude do efeito do QTL. O tipo de delineamento experimental afeta o tamanho da amostra necessária para a detecção de QTL (Lynch e Walsh, 1998).

Segundo Walsh (2006), os delineamentos experimentais dividem-se em dois grandes grupos: cruzamento entre linhas endogâmicas e populações exogâmicas (naturais ou segregantes). De acordo com o mesmo autor, a maior diferença entre as análises de QTL usando o cruzamento entre linhas endogâmicas e populações exogâmicas, é que os progenitores do cruzamento entre linhas endogâmicas são geneticamente uniformes, e os de populações exogâmicas são geneticamente variáveis.

Dentro do delineamento para detecção de QTL em grandes populações segregantes em gado de leite encontram-se, principalmente, a análise de meios-irmãs ou delineamento de filhas, e o delineamento de netas. Este último delineamento possui a vantagem de apresentar maior poder por indivíduo genotipado e, logicamente, as amostras de touros para extração de DNA são mais fáceis de coletar (centrais de inseminação), contrário ao delineamento de filhas, em que estas se encontram espalhadas por diferentes rebanhos (Weller *et al.*, 1980).

Entretanto, a estrutura populacional requerida para delineamento de netas (diferentes touros, cada um com um número significativo de filhos testados) é disponível unicamente em grandes populações comerciais. A estrutura populacional para delineamento de filhas (diferentes touros, cada um com um significativo número de filhas) pode ser encontrada em populações de tamanho moderado. Apesar de que o delineamento de filhas apresenta menor poder que o delineamento de netas por indivíduo genotipado, potencialmente, muito mais filhas são disponibilizadas para a análise, compensando o tamanho moderado da população. Entretanto, com múltiplos registros por vaca, a vantagem do delineamento de netas decresce, especialmente para características de alta herdabilidade (Weller *et al.*, 1990).

QTL segregando para características de produção de leite sobre o cromossomo 6 foram encontrados na raça Holandesa U.S. (Georges *et al.*, 1995; Zhang *et al.*, 1998), Holandesa Canadense (Nadesalingam *et al.*, 2001), raça Holandesa (Spelman *et al.*, 1996), Holandesa Alemã (Kuhn *et al.*, 1999), Holandesa Israelense (Lipkin *et al.*, 1998; Ron *et al.*, 2001), Holandesa Chinês (Chen *et al.*, 2006) e na raça Ayrshire (Velkala *et al.*, 1999) por meio de delineamentos de filhas e netas.

O delineamento de filhas tem sido utilizado, principalmente, em gado de leite, pois pelas técnicas reprodutivas, um touro pode ter muitas filhas com dados registrados para diferentes características quantitativas, como quantidade e qualidade do leite. O delineamento de filhas pode ser aplicado para uma única família de meias-irmãs, várias famílias de meias-irmãs e grandes famílias de irmãos completos. Esta aproximação aproveita a estrutura paterna de meio-irmãs, em que os marcadores genotipados são medidos nos touros e em grandes grupos de suas filhas. Quando somente as filhas de um



touro são consideradas, a detecção de um QTL ligado a um marcador pode ser realizada utilizando a seguinte equação:

$$Y_{ijk} = A_i + B_j + e_{ijk} \quad (1),$$

em que,  $Y_{ijk}$  é a medida da característica para o indivíduo  $k$ , do bloco “ $j$ ” e do genótipo  $i$ ;  $A_i$  é o efeito da substituição de alelos recebidos do touro;  $B_j$  é o efeito do “bloco”  $j$  e  $e_{ijk}$  é o efeito residual aleatório associado a cada indivíduo.

O delineamento de filhas foi proposto por Neimann-Soressen e Robertson (1961). Nesse delineamento, as filhas de um touro heterozigoto para o marcador serão genotipadas e mensuradas para as características quantitativas de interesse (Martinez, 2001). Segundo Weller *et al.* (1988), o mapeamento de QTL por meio de marcadores genéticos é baseado na expectativa que, se um indivíduo é heterozigoto em um loco marcador e um QTL ligado, então, exceto recombinação, a progênie recebendo um alelo marcador particular do indivíduo também tenderá a receber o alelo do QTL ligado.

Ao assumir que os dois alelos do QTL estão presentes na população de vacas com frequência  $p$  e  $1-p$ , é possível calcular o valor esperado da característica para cada grupo de progênie que recebeu os alelos alternativos do marcador. Na tabela 1 é apresentado o valor esperado para cada grupo. O contraste entre os dois grupos de progênie é:  $a(1-2r) + d(1-2r)(1-2p)$ , em que,  $a$  e  $d$  são efeitos aditivo e dominante do QTL, e  $r$  é a frequência de recombinação entre os dois locos. Se a frequência dos dois alelos do QTL é igual ( $p=0,5$ ), então o contraste torna-se  $a(1-2r)$ . No caso de ligação completa ( $r=0$ ), o contraste torna-se  $a+d(1-2p)$ , que é a fórmula geral para o efeito da substituição de alelos (Falconer, 1981).

É importante salientar que, embora um QTL esteja segregando na população, um touro, especificamente, pode ser homozigoto para o QTL. Pela razão anterior, a maioria de estudos tem sido baseada na análise de touros heterozigotos. Quando se têm várias famílias, a anterior equação (1) pode conduzir a erros, pois pode ser que vários touros sejam heterozigotos para um marcador ligado a um QTL, mas a relação de ligação para cada touro pode ser diferente. Assim, quando somados os efeitos de todas as progênies nos dois grupos de alelos do marcador, pode não observar nenhum efeito associado aos

alelos deste marcador. O modelo linear apropriado para um delineamento de meios-irmãs com múltiplas famílias é:

$$Y_{ijkl} = S_i + A_{ij} + B_k + e_{ijkl} \quad (2),$$

onde  $S_i$  é o efeito do touro;  $A_{ij}$  é o efeito do alelo  $j$ , pertencente ao touro  $i$ ; e os demais termos como na equação (1) (Martinez, 2001).

Para efeito da análise, o grupo de progênie que recebeu o alelo M1 do touro será comparado com o grupo de progênie que recebeu o alelo M2. Ao manterem-se as pressuposições anteriores e, sendo a distribuição das mães entre os dois grupos ao acaso, então a diferença entre os grupos de progênies para a característica quantitativa será devido ao QTL ligado ao marcador M do touro, assumindo que a origem do alelo do marcador pode ser determinada para as progênies. Os valores genéticos estimados das filhas é uma função das avaliações genéticas de seus pais e mães, em adição a suas próprias produções. Aproximadamente todos os efeitos do touro sobre as avaliações de suas filhas devem ser absorvidos pelo efeito  $S_i$ , enquanto os efeitos das mães sobre as avaliações de suas filhas são incluídos no resíduo (Ron *et al.*, 2001).

**Tabela 1** – Probabilidades dos genótipos no delineamento de filhas e valores esperados da característica quantitativa.

Alelo paterno do marcador	Alelo paterno do QTL	Probabilidade do alelo paterno do QTL	Alelo materno do QTL	Probabilidade do alelo materno do QTL	Valor da característica	Valor esperado da característica para o genótipo do marcador
M1	Q1	1-r	Q1	p	a	$a(p-r)+d(1-r-p+2rp)$
	Q2	R	Q2	1-p	d	
M2	Q2	1-r	Q1	p	d	$a(r+p-1)+d(p+r-2rp)$
	Q1	R	Q2	1-p	-a	

Fonte: Weller e Ron (2000)

## 2.8 Seleção Assistida por Marcadores

O melhoramento genético clássico em gado de leite utiliza o teste de touros, realizado por meio da avaliação de suas progênes. Para isso os touros jovens são selecionados e utilizados mediante inseminação artificial, constituindo-se, assim, a próxima geração. De acordo com Verneque *et al.* (1998), o processo leva no mínimo cinco anos para se obter um touro geneticamente testado, com ganho genético para produção de leite por ano de 1%. Esse período de tempo pode ser reduzido segundo Machado e Martinez, (2001), porque depois do nascimento dos animais podem ser genotipados, e com a identificação de marcadores ligados às características de interesse econômico, touros jovens podem ser selecionados, mantendo-se assim, aqueles animais com alto potencial genético, enquanto os que apresentarem baixo potencial genético podem ser descartados, evitando custos necessários para a manutenção do animal por um período de tempo determinado.

A diferença da seleção baseada no fenótipo, a MAS possui as vantagens de não sofrer efeitos do meio ambiente, não requer o abate do

animal e não ser limitada a um sexo. Segundo Machado *et al.* (2001), a MAS pode ser realizada após do nascimento do animal ou até no mesmo embrião, reduzindo o intervalo entre gerações e aumentando a intensidade de seleção. Por meio da MAS, também é possível a identificação e introgressão dos alelos das raças melhoradas em raças nativas adaptadas, aumentando as características de produção (Beckmann e Soller, 1987).

Em programas de avaliação genética por meio do teste de progênie a aplicação mais prática da MAS consiste em utilizar a informação da associação dos marcadores moleculares com QTL, para pré-selecionar touros jovens para o teste de progênie. A MAS para bovinos de leite tem sido avaliada em diferentes estudos, mostrando incremento nos ganhos genéticos (Spelman e Garrick, 1992). Dois esquemas da MAS têm sido avaliados: um baseado na seleção dentro de famílias e outro utilizando informações dos marcadores moleculares no BLUP. Dentro de famílias, a seleção envolve decisões baseadas em informações dos valores genéticos estimados e as provenientes dos QTL. A MAS baseada no BLUP compreende modelos mistos incorporando as informações dos QTL, e as decisões são feitas sobre os valores genéticos estimados que combinem componentes de QTL e os poligenes (Brascamp *et al.*, 1993).

A Seleção Assistida por Marcadores Moleculares fundamenta-se no princípio de que, se um gene ou grupo de genes está associado a um marcador molecular identificável, então ao selecionar para este marcador, a seleção será mais eficaz que ao selecionar para uma característica. Segundo Dudley (1994), a MAS compreende duas etapas básicas: identificação de associações entre alelos do marcador e alelos do QTL, e o uso destas associações para o desenvolvimento de populações melhoradas.

## **2.9 Identificação de QTL no cromossomo 6 em bovinos de leite**

Para produção e composição do leite as pesquisas foram feitas, inicialmente para genes candidatos, especialmente genes para proteínas lácteas, as quais apresentam efeitos apenas modestos sobre produtividade. O primeiro experimento utilizando marcadores distribuídos por todo o genoma, com um grande número de touros e utilizando delineamento de filhas (Grand-Daughter Design) na raça Holandesa foi conduzido nos Estados Unidos por

Georges *et al.* (1995) que identificaram evidências de QTL para características produtivas nos cromossomos 1, 6, 9, 10 e 20.

Khatkar *et al.* (2004) utilizando a metodologia de Meta-análise encontraram QTL praticamente em todos os 29 autossomos do genoma bovino leiteiro, mas os cromossomos 3, 6, 9, 14, 20 e 23 foram reportados como possuidores de QTL com efeitos pleiotrópicos sobre características múltiplas de produção de leite. De maneira similar Smaragdov (2006), em um estudo de mapeamento genético de locos responsáveis para características de produção em gado de leite, encontrou que o número de QTL e as posições estavam especialmente nos cromossomos 1, 3, 6, 14, 20 e 23.

O cromossomo 6 é considerado como um cromossomo importante para o estudo de características de produção de leite associadas a QTL, pelo número de QTL detectados e pelos genes candidatos que ali se encontram. Freyer *et al.* (2003) pesquisando sobre QTL pleiotrópicos no cromossomo 6 afetando características de produção de leite, encontraram dois QTL significativos nas posições 49 e 64 cM, QTL sugestivos na região 68 cM para produção de gordura, e na região cromossômica 71 cM para produção de proteína; com análise multivariada encontraram um QTL pleiotrópico na região 68 cM, afetando produção de gordura e produção de proteína. Olsen *et al.* (2004), mediante a análise de desequilíbrio e equilíbrio de ligação, refinaram uma posição onde previamente havia sido encontrado um QTL afetando características de produção de leite em um intervalo de 7,5 cM, flanqueado pelos marcadores BMS2508 (9,73 cM) e FBN12 (17,27 cM), sendo que os resultados mais significativos encontrados foram percentagem de gordura e proteína.

Chen *et al.* (2006), utilizando 14 marcadores microssatélites numa região com cobertura de 63,5 cM no cromossomo 6 bovino, em uma população da raça Holandesa Chinês, mediante delineamento de filhas por meio de 26 famílias, com as 2 famílias mais significativas e com um modelo de QTL fixado, encontraram um QTL afetando produção de gordura, perto do marcador BMS470 (32 cM), com um nível de significância 5% genômico; com o modelo de 2 QTL fixados encontraram dois QTL afetando as três características de produção: leite, proteína e gordura.

Segundo Smaragdov (2006) em uma revisão de estudos representativos de detecção de QTL no cromossomo 6 do genoma de bovinos leiteiros, foram

encontrados QTL para produção de leite nas posições 7, 28, 39, 40, 44, 47, 48, 49, 53, 56 e 64 cM; para produção de proteína QTL encontrados nas posições 7, 24, 35, 47, 48, 49, 57, 59 e 67 cM, respectivamente; para percentagem de proteína foram encontrados QTL nas posições 7, 35, 36, 43, 46 47, 48, 49, 50 e 58 cM; para produção de gordura, segundo este estudo, estiveram presentes nas posições 5, 35, 45, 49, 51, 52, 57, 58 e 62 cM; e para percentagem de gordura os QTL encontrados estiveram nas posições 5, 35, 40, 46, 47, 49, 51 e 58 cM.

De acordo com a metodologia de meta-análise, Khatkar *et al.* (2004) encontraram QTL no BTA6 altamente significativos, para produção de leite nas posições 50 e 87 cM, produção de proteína um QTL na posição 52 cM, percentagem de proteína apresentou um QTL na posição 49 cM, produção de gordura um QTL na posição 51 cM e para percentagem de gordura dois QTL nas posições 48 e 113 cM, respectivamente.

Os seguintes QTL significativos foram mapeados também no cromossomo 6: George *et al.* (1995) encontraram um QTL na posição 40 cM para produção de leite; Ron *et al.* (2001) mapearam um QTL para produção de leite na posição 47 cM; Spelman *et al.* (1996) um QTL para percentagem de proteína a 13 cM; Ron *et al.* (2001) encontraram dois QTL para percentagem de proteína nas posições 53 e 56 cM; Ashwell *et al.* (2002) um QTL para porcentagem de proteína na posição 58 cM e Boichard *et al.* (2003) também para porcentagem de proteína na posição 98 cM; para produção de gordura encontraram QTL significativos, Rodriguez *et al.* (2002) e Kuhn *et al.* (1999) nas posições 1 e 53 cM, respectivamente; para porcentagem de gordura dois QTL mapeados significativamente por Ashwell *et al.* (2002 e 2004) nas posições 49 e 58 cM e Ron *et al.* (2001) nas posições 50 e 54 cM.

Além do número considerável de QTL encontrados no cromossomo 6, foram identificados genes candidatos importantes associados às características de produção de leite. Os genes da caseína e da albumina estão localizados no cromossomo 6. O gene da caseína têm efeitos sobre produção de leite (Velmala *et al.*, 1995; Ikonen *et al.*, 2001; Prinzenberg *et al.*, 2003). A kappa-caseína é uma das proteínas coaguláveis do leite, e em taurinos a presença do alelo B da kappa-caseína tem sido associada a uma coagulação mais eficiente, obtendo como consequência aumento no rendimento na produção do queijo, e associado a um aumento da quantidade e concentração de proteína no leite.

Mais recentemente, diferentes genes candidatos no BTA6 (*Bos taurus* autosome 6) como FAM13A1, PPARGC1A e OPN, foram reportados com efeitos potenciais sobre características de produção de leite (Cohen *et al.*, 2004; Weikard *et al.*, 2005; Schnabel *et al.*, 2005). O gene FAM13A1 está relacionado com produção de proteína, e o gene OPN (osteopontina) está relacionado com produção de leite, porcentagem de gordura e proteína.

O BTA6 é um dos cromossomos em que maior número de QTL têm sido mapeados e relacionados com características de produção de leite. No momento existem mais de 100 QTL identificados ([http://www.animalgenome.org/cgi\\_bin/QTLdb/brows](http://www.animalgenome.org/cgi_bin/QTLdb/brows)) no cromossomo 6 e, em sua maioria para características de produção de leite (Olsen *et al.*, 2004). Este cromossomo apresenta mais de 111 milhões de pares de bases, e foram identificados 420 genes codificadores de proteínas, 36 pseudogenes, 98.411 SNP, e 56 genes que codificam diferentes tipos de RNAs ([http://www.ensembl.org/Bos\\_taurus/index.html](http://www.ensembl.org/Bos_taurus/index.html)).

### 3 MATERIAL E MÉTODOS

#### 3.1 Valores genéticos para características de produção de leite

As características de produção geralmente possuem herdabilidade médias, enquanto as características de porcentagem no leite apresentam alta herdabilidade. A genética dá à vaca a capacidade de produzir leite, enquanto o ambiente fornece as condições para ela produzir. A produção de leite é, portanto, a soma dos efeitos do meio, da genética da vaca e da possível interação do meio com a genética (Everett *et al.*, 1976). Na tabela 2 são apresentados valores de herdabilidade e correlações entre as características de produção do leite para a raça Gir Leiteiro

**Tabela 2** - Estimativas de herdabilidade ( $h^2$ ) para características de produção e suas respectivas correlações genéticas ( $r_G$ ) entre produção de leite e produções de gordura, proteína e sólidos totais na raça Gir Leiteiro.

Características	$h^2$	$r_G$
Produção de leite	0,30	1,00
Produção de gordura	0,31	0,96
Produção de proteína	0,34	0,91
Produção de lactose	0,20	0,99
Produção de sólidos totais	0,31	0,96

**Fonte.** Sumário Brasileiro de touros Gir (Verneque, 2008).

Os valores fenotípicos das características de produção do leite, proteína, gordura, lactose e sólidos totais, e porcentagem de proteína, gordura, lactose e sólidos totais foram obtidos do Arquivo Zootécnico Nacional (AZN) do ano 2008. No controle leiteiro, mensalmente um técnico credenciado realizou a pesagem da produção leiteira de cada uma das matrizes dos diferentes rebanhos pertencentes ao programa. O intervalo entre as pesagens é no máximo de 75 dias, sendo que a primeira pesagem de leite não é inferior a 5



dias e nem superior a 45 dias da parição. As informações obtidas no controle leiteiro possibilitaram a realização das avaliações genéticas das matrizes, dos animais jovens e dos touros.

Com o parto das filhas dos touros em teste, iniciou-se o controle produtivo dessas filhas e também das suas companheiras de rebanho. Foi obtida mensalmente a produção de leite em duas ordenhas diárias e uma amostra de leite por matriz foi enviada para o laboratório de controle de qualidade do leite na Embrapa Gado de Leite onde foram obtidas as porcentagens de gordura, proteína, lactose e sólidos totais. Para serem obtidas as produções de gordura, proteína, lactose e sólidos totais, a produção de leite foi multiplicada pelas respectivas porcentagens de gordura, proteína, lactose ou sólidos totais, para um determinado controle, obtendo-se ao final uma curva de lactação.

Uma lactação completa é definida como a produção de leite iniciada a partir do quinto dia após do parto, mas não superior a 45 dias, e continuando pelo menos 50 dias após o início da lactação. O termo curva de lactação refere-se a uma representação gráfica da variação da produção diária de leite de uma vaca em função do tempo de lactação. A avaliação genética das vacas para produção de leite foi baseada na produção acumulada na lactação com um período padrão de 305 dias, cujo total foi estimado a partir das produções medidas nos dias dos controles leiteiros, geralmente realizados em intervalos de 30 dias.

O método TIM (Test Interval Method) de Everett e Carter (1968) foi empregado para calcular as produções por lactação a partir dos controles parciais. O método TIM ajusta a produção no começo e no final da curva de lactação. Os valores obtidos por Everett e Carter (1968), após de simulações de 403 lactações de vacas Holandesas, para ser empregados na equação de fatores de correção para o primeiro controle leiteiro de "x" dias de duração e por idade da matriz, estão na tabela 3.

**Tabela 3** - Equação para fatores de correção para o primeiro período de controle de “X” dias de duração por idade da vaca.

$$y = \mu + b_1 X^a + b_2 X + b_3 (1.0/X^2)$$

Idade (anos)	$\mu$	$b_1$	$b_2$	$b_3$	a	$R^2$
2	0,8553	$0,0000 \times 10^0$	$0,1933 \times 10^{-2}$	$0,1619 \times 10^0$	0,00	0,983
3	0,8995	$0,2041 \times 10^{-2}$	$-0,2478 \times 10^{-2}$	$0,1090 \times 10^0$	1,19	0,997
4	0,9103	$0,9057 \times 10^{-2}$	$-0,1232 \times 10^{-1}$	$0,1081 \times 10^0$	1,10	0,991
5	0,9018	$0,2561 \times 10^{-2}$	$-0,4018 \times 10^{-2}$	$0,1204 \times 10^0$	1,19	0,998
6+	0,8950	$0,2793 \times 10^{-2}$	$-0,4305 \times 10^{-2}$	$0,1192 \times 10^0$	1,19	0,994

Fonte: Everett e Carter (1968).

A equação para fatores de correção para o período de controle de “X” dias de duração no último controle é:  $y = \mu + b_1 X^a + b_2 X$ . Após ajuste das produções do primeiro e do último controle na curva de lactação foi obtida a média dentre cada intervalo de dois controles consecutivos da curva de controle e multiplicada pelo período de tempo, sendo somadas essas produções parciais para obter a produção total por lactação. Na tabela 4 encontram-se os valores obtidos para ser empregados na equação para fatores de correção para um período de controle de “X” dias de duração dado o último dia do controle.

**Tabela 4** - Equação para fatores de correção para um período de controle de “X” dias de duração dado o último dia do controle.

$$y = \mu + b_1X^a + b_2X$$

Dia do último controle	$\mu$	$b_1$	$b_2$	a	$R^2$
4	1,0064	$-0,2350 \times 10^{-6}$	$0,1814 \times 10^{-2}$	2,60	0,995
5	1,0057	$-0,1214 \times 10^{-6}$	$0,1636 \times 10^{-2}$	2,71	0,993
6	1,0018	$-0,2068 \times 10^{-6}$	$0,1543 \times 10^{-2}$	2,60	0,991
7	1,0006	$-0,6780 \times 10^{-7}$	$0,1392 \times 10^{-2}$	2,80	0,991
8	0,9994	$-0,6176 \times 10^{-7}$	$0,1302 \times 10^{-2}$	2,81	0,989
:	:	:	:	:	:
40	0,9975	$-0,1253 \times 10^{-4}$	$0,4536 \times 10^{-3}$	1,71	0,595

Fonte: Everett e Carter (1968).

As PTAs para produção de leite foram calculadas em análise unicamente para essa produção. As análises das características de produção de gordura, proteína, lactose e sólidos totais foram realizadas considerando duas características por vez, em que a produção de leite sempre foi a âncora delas. Isso torna a confiabilidade das estimativas das PTAs maior, se existe correlação genética diferente de zero entre as características. A base genética foi estimada como a média das PTAs de todas as filhas do touro em teste nascidas no ano 2000, e calculada para todas as características. O grupo de manejo foi definido considerando-se o rebanho, ano de parto e a estação de parição da vaca.

O modelo estatístico empregado ( $y = Xb + Za + Wpe + e$ ) na avaliação genética dos animais considera como efeitos fixos, rebanho-ano de parto, estação de parição da vaca, composição genética da filha do touro e a idade da vaca ao parto e os efeitos aleatórios foram os erros e o efeito do animal (vaca, pai e mãe). Para avaliação genética das vacas foi incluído o efeito aleatório de meio permanente por possuir medidas repetidas no mesmo animal.

A média da produção de leite até 305 dias de lactação considerada na base de dados foi de  $2.783 \pm 1.292$  kg, produção de gordura  $112 \pm 55$  kg, produção de proteína  $90 \pm 44$  kg, e sólidos totais  $337 \pm 171$  kg. A duração

média da lactação foi de  $286 \pm 77$  dias e idade média ao primeiro parto foi de  $43,2 \pm 7,4$  meses. A base genética de produção de leite considerada para a avaliação foi a média do valor genético das filhas dos touros nascidas no ano 2000, cujo valor foi de 190 kg. Isso significa que, a cada valor genético de cada animal avaliado foram deduzidos 190 kg, sendo que a média do valor genético da produção do leite, dos animais nascidos no ano 2000, foi movida para 0 (zero). De igual forma se fez para as outras características: gordura (7,81 kg), proteína (6,72 kg) e sólidos totais (26,18 kg).

O modelo de repetibilidade usualmente é da forma:

$$y = Xb + Za + Wpe + e,$$

em que  $y$  = vetor de observações;  $b$  = vetor de efeitos fixos;  $a$  = vetor de efeito animal, aleatório;  $pe$  = vetor de efeitos ambientais permanentes e efeitos genéticos não-aditivos;  $e$  = vetor de efeitos residuais aleatórios;  $X, Z, W$ , são matrizes de incidência de efeitos fixos, animal e de ambiente permanente, respectivamente. O vetor "a" unicamente inclui efeito animal aleatório aditivo, conseqüentemente, efeitos genéticos não-aditivos estão incluídos no termo  $pe$ . É assumido também que, efeito de ambiente permanente e efeito residual, estão independentemente distribuídos com medias zero e variâncias  $\sigma_{pe}^2$  e  $\sigma_e^2$ , respectivamente. Assim,

$$\text{var}(pe) = I\sigma_{pe}^2; \text{var}(e) = I\sigma_e^2 = R; \text{var}(a) = A\sigma_a^2 \text{ e } \text{var}(y) = ZAZ'\sigma_a^2 + WI\sigma_{pe}^2 W' + R$$

Equações do modelo misto para o melhor estimador linear não-viesado (BLUE) de funções estimáveis de  $b$ , e para a melhor predição linear não-viesado (BLUP) de  $a$  e  $pe$ , são:

$$\begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{pe} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z & X'R^{-1}W \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1}I/\sigma_a^2 & Z'R^{-1}W \\ W'R^{-1}X & W'R^{-1}Z & W'R^{-1}W + I(1/\sigma_{pe}^2) \end{bmatrix}^{-1} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \\ W'R^{-1}y \end{bmatrix}$$

Entretanto, a equação de modelos mistos (MME) com fatoração de  $R$  :

$$\begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{pe} \end{bmatrix} = \begin{bmatrix} X'X & X'Z & X'W \\ Z'X & Z'Z + A^{-1}\alpha_1 & Z'W \\ W'X & W'Z & W'W + I\alpha_2 \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \\ W'y \end{bmatrix}$$

em que,  $\alpha_1 = \sigma_e^2/\sigma_a^2$  e  $\alpha_2 = \sigma_e^2/\sigma_{pe}^2$

O software MTDFREML (Multiple Trait Derivate Free Restricted Maximum Likelihood) ou Estimação pelo Método da Máxima Verossimilhança Restrita adotando-se o procedimento Livre de derivadas (DF). Este último foi utilizado para o cálculo dos componentes de variâncias e os valores genéticos no presente estudo. O MTDFREML foi desenvolvido por Boldman *et al.* (1993). O sistema é conformado por três programas, descritos na tabela 5.

**Tabela 5** - Programas do sistema MTDFREML

Programa	Propósito
MTDFNRM	Programa principal para o cálculo de $A^{-1}$ e recodificação para número do animal, pai e mãe (ou pai, pai do pai e avô materna).
MTDFPREP	Lê o arquivo do pedigree, recodifica animais e níveis de efeitos fixos e fatores aleatórios e calcula estatísticas descritivas de variáveis de modelo.
MTDFRUN	Utiliza os dados preparados pelos programas anteriores e calcula componentes de (co) variância pelo máximo logaritmo da função de verossimilhança, com um algoritmo DFREML.

Fonte: Verneque (1994).

O procedimento faz reordenamento e fatoração das equações de modelos mistos multivariados e operação de matrizes esparsas, por meio do conjunto de rotinas SPARSPAK e, no processo de procura do mínimo de  $-2\text{Log}(L_2)$ , adota o método simplex (politopo). Para sua convergência com método simplex, o processo se encerra quando o programa atinge o número máximo de interações informado, ou quando a variância do simplex for:  $\text{Var}(-2\text{Log}(L_2)) < 10^{-9}$ . O processo de estimação deve ser repetido (tanto se utilizando o método simplex ou o método Powell), até que a diferença de valores em  $-2\text{Log}(L_2)$ , entre duas rodadas completas, seja inferior a 0.002 unidades (Verneque, 1994). Os resultados são apresentados nos arquivos de saída MTDF76 (variâncias e covariâncias) e MTDF78 (valores genéticos).

### 3.2 Famílias utilizadas para o mapeamento de QTL

Para o desenvolvimento deste trabalho foram utilizadas 657 vacas Gir pertencentes ao controle leiteiro, filhas de 14 touros Gir, dos quais 12 foram positivos para o teste de progênie e os outros 2 foram pais de touros positivos para o mesmo teste. O critério utilizado para o número de filhas por touro baseou-se na existência de amostras de DNA para as vacas filhas por touro, sendo vinte o número mínimo de filhas para ser considerada uma família. Na tabela 6 são apresentados famílias, número de filhas por touro e os touros classificados pela PTA para leite.

**Tabela 6.** Famílias de meios-irmãs, número de filhas por touro e classificação pela PTA para leite.

Famílias	Código do touro	Nº Filhas	Classificação (PTA)
1	001	127	16
2	129	94	10
3	224	68	31
4	293	62	26
5	356	59	-
6	416	46	90
7	463	34	33
8	498	28	65
9	527	28	23
10	556	24	-
11	581	23	69
12	605	22	70
13	628	22	1
14	651	20	38

### 3.3 Coleta e extração de DNA

Algumas amostras de DNA correspondentes a vacas e touros foram obtidas do Banco de DNA de Gado de Leite da Embrapa. As outras amostras necessárias para a constituição de famílias e o uso do delineamento de filhas foram obtidas de amostras de sangue (vacas) e sêmen (touros), coletadas em diferentes rebanhos do Brasil, que fazem parte do PNMGL. Para a extração do DNA seguiu-se o protocolo modificado de Sambrook e Russel (2001), coletando-se dois tubos vacutainer de 10 mL ( com 0,5 mL de EDTA a 0,5 M) de sangue, por via endovenosa. As amostras foram centrifugadas, sendo retirado o anel de células brancas, aproximadamente 500 µL de leucócitos por

cada amostra foram transferidos para um microtubo de 2 mL. As células vermelhas do sangue foram desintegradas com tampão de lise em quantidade apropriada e o resíduo celular pelitizado por centrifugação. A desintegração foi repetida até que o precipitado branco foi obtido. As células brancas do sangue foram recolocadas em suspensão e incubadas em tampão de digestão (10 mM Tris-HCL pH 7,6, 10 mM EDTA pH 8,0, 50 mM NaCl, 0,02% SDS, com 40 µL de proteinase K por amostra) a 50 °C até o dia seguinte. As proteínas foram extraídas duas vezes em fenol: clorofórmio:isoamil (25:24:1), duas vezes em clorofórmio:isoamil (24:1), e então precipitadas em etanol gelado. O DNA foi solubilizado em tampão TE (10 mM Tris-HCL, 1 mM EDTA) e congelado até o momento do uso.

A extração de DNA a partir de sêmen utilizou duas palhetas de sêmen em um microtubo de 2mL, centrifugando a 14000 r.p.m. por 30 segundos e descartando o sobrenadante. O restante do protocolo foi semelhante ao usado na extração do DNA a partir de sangue. Para quantificação do DNA foi usado o equipamento Nanodrop e a qualidade foi conferida pela razão de absorbância 260/280 nm.

### **3.4 Genotipagem com marcadores microssatélites**

Para a genotipagem do cromossomo 6 foram utilizados marcadores microssatélites, escolhidos do mapa consenso e disponibilizado pelo MARC/USDA (Meat Animal Research Center / United States Department of Agriculture - <http://www.marc.usda.gov/genome/genome.html>). Na escolha dos marcadores levaram-se em conta sua posição, número de alelos, e um mínimo de 50% de heterozigidade. Foram selecionados 27 marcadores microssatélites ao longo do cromossomo 6, com distância média entre marcadores de 4,9 cM, que foram utilizados para genotipar 14 touros e 657 vacas, em um total de 671 animais.

Para as reações de PCR, foram testadas três concentrações de MgCl<sub>2</sub> (1,5 mM; 2,0 mM e 2,5 mM). Também foram testadas cinco temperaturas de anelamento (TA) para cada *primer*: 50, 52, 54, 56 e 58 °C. As reações constaram de 45 ng de DNA nuclear; 02 mM de dNTPs; 20 mM tris pH 8,3; 50 mM KCl, 0,1 µM de cada *primer* e a concentração de MgCl<sub>2</sub> foi definida para cada marcador. Os *primers forward* foram marcados na extremidade 5' com os

seguintes tipos de fluorescência: FAM, HEX e TAMRA. As amplificações foram feitas em termociclador, com as seguintes etapas: 94 °C (1 minuto), TA (1 minuto), 72 °C (1 minuto) por 35 ciclos, e uma extensão geral de 45 minutos a 72 °C.

Os produtos das reações de PCR foram precipitados utilizando 3 µL de acetato de amônio (7,5 M), 30 µL de etanol (90%), 30 µL de etanol (70%). Após da centrifugação das placas, o DNA foi resuspendido em 20 µL de água. Os produtos das amplificações foram genotipados por eletroforese capilar no equipamento MegaBACE 1000 (GE Healthcare, NYSE, Germany). Os genótipos foram analisados no programa Fragment Profiler e os dados exportados para uma planilha no Excel. Foram definidas as condições ótimas de TA e MgCl<sub>2</sub> para cada marcador (Tabela 7).



**Tabela 7.** Posição, nome do marcador, Fluoróforo para marcar o *primer forward*, concentração de cloreto de magnésio (MgCl<sub>2</sub> mM), temperatura de ligação do mapa do rebanho Gir Leiteiro do PNMGL (TL-P °C) da Embrapa.

Posição (cM)	Marcador	Fluoróforo	MgCl <sub>2</sub>	TL-PNMGL (°C)
0,00	ILST093	FAM	1,5	52
9,02	DI4408	HEX	1,5	54
15,3	DIK5285	FAM	1,5	56
20,1	DIK4498	TAMRA	1,5	52
29,3	MNB66	FAM	1,5	54
35,39	BM1329	HEX	1,5	56
38,16	DIK1058	HEX	1,5	54
43,93	BMS2508	FAM	1,5	52
50,09	DIK4382	TAMRA	1,5	56
54,5	DIK4482	HEX	1,5	54
60,2	MNB-208	FAM	1,5	52
63,86	BM4322	FAM	1,5	54
67,4	BMS470	TAMRA	1,5	54
71,5	DIK3026	HEX	1,5	50
75,27	DIK2294	FAM	1,5	58
81,96	DIK4867	TAMRA	1,5	56
87,26	ILSTS035	FAM	1,5	54
90,5	DIK4574	FAM	1,5	52
93,8	BMS5021	TAMRA	1,5	54
96,98	AFR227	TAMRA	1,5	56
101,4	DIK2174	FAM	1,5	54
107,12	DIK4827	HEX	1,5	56
109,9	DIK2995	HEX	1,5	52
115,32	DIK1182	HEX	1,5	56
121,49	DIK2690	TAMRA	1,5	56
127,49	BM2320	TAMRA	1,5	56
130,78	DIK4992	HEX	2,0	54

As informações de frequência de alelos ( $p_i$ ), heterozigosidade esperada ( $H_e$ ) e observada ( $H_o$ ), e conteúdo de informação de polimorfismo (PIC) dos marcadores microssatélites, foram obtidas utilizando o programa CERVUS versão 2.0 (Marshall *et al.*, 1998), com as seguintes equações:

$$PIC = 1 - \sum_i p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i p_j^2, \quad p_i = K_i/N$$

$$H_e = 1 - \sum_{i=1}^n p_i^2 = 2 \sum_{i=1}^n \sum_{j=i+1}^{n-1} p_i p_j$$

em que:  $K_i$  é o número de observações do alelo  $i$  ( $i=1, \dots, m$ ),  $N$  é o número total de informações ( $2n$  alelos) e  $p_i$  e  $p_j$  são as informações de frequência de alelos. A  $H_o$  é a razão entre o número de heterozigotos e o número total de indivíduos.

### 3.5 Mapas de Ligação

O mapa consenso foi gerado levando-se em conta diferentes trabalhos realizados com as raças Angus, Australian Friesian, Boran, Brahman, Brangus, Charolês, Gelbvieh, Gir, Hereford, Holandês, Indubrasil, Nelore, N'Dame, Normando, Piemontesa, Sahiwal e Simental (Bishop *et al.*, 1994; Barendse *et al.*, 1994).

O mapa de ligação para o cromossomo 6 foi construído com o programa Crimap (Green *et al.*, 1990). Este determina a frequência de recombinação e a melhor ordem para os marcadores no grupo de ligação. A ordem dos marcadores é realizada tendo em conta os valores maiores do  $\text{Log}_{10}$  da máxima verossimilhança (LOD). O mapa de ligação foi determinado pelo método de máxima verossimilhança e, a função de mapeamento empregada foi a de Kosambi, para transformar unidades de mapa de recombinação em centiMorgans (cM).

### 3.6 Delineamento para detecção de QTL

No presente estudo foi utilizado o delineamento de meias-irmãs ou delineamento de filhas, proposto inicialmente por Neimann-Soressen e Robertson (1961). As análises preliminares de QTL para todos os marcadores foram realizadas pelo seguinte modelo:

$$BV_{ijkl} = S_{ij} + M_{ijk} + e_{ijk}$$

em que:  $BV_{ijkl}$  é o EBV (Estimated Breeding Values) para a característica  $i$  da vaca  $l$ , que recebeu o alelo paterno  $k$ ;  $S_{ij}$  é o efeito do touro  $j$  sobre a característica  $i$ ;  $M_{ijk}$  é o efeito do alelo paterno  $k$  do touro  $j$  sobre a característica  $i$ , e  $e_{ijk}$  é o resíduo aleatório associado com cada registro. Um

efeito de alelo paterno significativo é indicativo de um QTL ligado ao marcador molecular segregando (Ron *et al.*, 2004).

Produção de leite, gordura, proteína, lactose e sólidos totais, e porcentagens para essas mesmas características, sobre 305 dias foram analisadas pelo modelo animal de repetibilidade, e as avaliações foram realizadas tendo em conta os registros até novembro de 2008. Médias, desvios padrões, valores máximos e mínimos das avaliações genéticas das 657 vacas genotipadas para as nove características analisadas são apresentadas na tabela 8.

As análises para identificação de QTL no cromossomo 6 foram realizadas em mapeamento por intervalo, com um delineamento de meios-irmãs, utilizando o software GridQTL (<http://www.gridqtl.org.uk>) (Seaton *et al.*, 2006). As análises consistiram, primeiramente, na determinação da probabilidade de Identidade-*Por-Descendência* (IBD) para localização específica das regiões cromossômicas referente aos dados de marcadores múltiplos e, logo, o modelo estatístico adequou-se às observações e aos coeficientes de IBD (Seaton *et al.*, 2006). A segunda etapa da análise dos dados consistiu na regressão dos dados fenotípicos e os coeficientes que foram calculadas previamente.

Nos resultados das informações, primeiro, a informação genótipica do marcador é resumida em termos do número de marcadores, alelos por marcador e loco informativo por touro. O marcador informativo por touro é calculado pela contagem do número de progênie para a qual o alelo herdado do touro pode ser determinado sem ambigüidades e, segundo, os resultados são revelados para a escolha da análise do QTL. O sinal do efeito de substituição é arbitrário, porque a escolha do alelo 1 ou alelo 2 do touro é arbitrária. O teste estatístico é um teste F-padrão com número de graus de liberdade no numerador, igual ao número de touros que são informativos em uma dada localização do cromossomo (Knott *et al.*, 1996). Nos valores genéticos estimados (EBV) os pesos são proporcionais à confiabilidade do EBV. Gridqtl utiliza esses pesos para computar em um modelo equivalente sobre o qual os cálculos adicionais estão baseados.

Os limiares de significância a nível cromossômico (1% e 5%) foram calculados no programa GridQTL a partir de testes de permutação (10.000) (Churchill e Doerge, 1994). O intervalo de confiança (IC 95%) foi obtido usando

a aproximação de Qui-quadrado ( $X^2$ ), conforme é descrito por Pérez-Enciso *et al*, (2000). A proporção da variação fenotípica explicada pelo QTL foi calculada de acordo com Knott *et al*, (1996):

$$h_{QTL}^2 = 4x[1 - (MSEf)/(MSEr)],$$

em que: MSE f é quadrado médio do modelo completo (presença de QTL), MSE r é quadrado médio do erro do modelo reduzido (ausência de QTL).

**Tabela 8** - Média, desvio padrão, mínima e máxima dos valores genéticos estimados para nove características de produção de leite (657 vacas).

Característica	Média	Desvio padrão	Mínimo	Máximo
Produção de leite, kg	682,083	401,645	-330,8	1720,8
Produção de gordura, kg	22,29	16,047	-15,4	84,6
Produção de proteína, kg	12,626	11,572	-12,3	56,2
Produção de lactose, kg	15,407	13,166	-12,2	61,5
Produção de sólidos t, kg	45,837	33,008	-44,3	182,1
Percentagem de gordura, %	-0,0388	0,1006	-0,315	0,329
Percentagem de proteína, %	-0,0119	0,04961	-0,189	0,139
Percentagem de lactose, %	0,026	0,0677	-0,190	0,264
Percentagem de sólidos, %	0,095	0,1923	-0,456	0,864

## 4 RESULTADOS E DISCUSSÃO

### 4.1 *Scoring* dos marcadores microssatélites

Foram realizadas 18.117 genotipagens correspondentes aos 671 animais analisados, dos quais, 14 foram touros e 657 filhas, por 27 marcadores microssatélites. As vacas que não herdaram alelos paternos para mais de 5 locos não foram consideradas filhas e foram descartadas para as análises. Ron *et al.* (1996) aplicaram uma probabilidade mínima de 90% de paternidade sobre os marcadores para que uma vaca fosse considerada filha de um determinado touro.

Foi possível identificar 286 alelos para os 27 locos analisados (Tabela 9), dando uma média de 10,6 alelos por marcador. Os locos DIK4574 e DIK2995 apresentaram o menor número de alelos (6), e o loco com maior número de alelos (21) correspondeu ao BM2320. O mapa consenso apresentou 243 alelos na população estudada, o que corresponde a uma média de 9 alelos por marcador. Embora, para a formação do mapa consenso fossem utilizadas maiores populações e diferentes raças que a raça do presente estudo, esta última apresentou maior diversidade, o que pode ser explicado por uma menor intensidade de seleção nas raças zebuínas. Somente 5 marcadores (18,51%) do mapa consenso apresentaram maior número de alelos que a população Gir; 4 marcadores (14,81%) apresentaram o mesmo número de alelos, e 18 marcadores (66,66%) da população em estudo apresentaram maior número de alelos que no mapa consenso.

Os marcadores DIK5285, DIK4574 e DIK4992 apresentaram igual tamanho mínimo em pares de base para a população Gir e o mapa de referência do MARC. O loco ILSTS093 apresentou igual tamanho máximo em pares de base para os dois mapas. Na população Gir o marcador DIK3026 apresentou um tamanho mínimo em pares de base, muito maior que o mesmo marcador no mapa consenso (75 pb). Já o marcador DIK4382 do mapa MARC apresentou um tamanho mínimo em pares de base maior que o respectivo marcador na população Gir (161). Para o tamanho máximo em pares de base,

o marcador DIK5285 na população estudada apresentou 44 pares de base a mais que o mapa consenso. Os restantes 20 locos apresentaram diferenças quanto ao tamanho mínimo e/o máximo de pares de base para as duas populações, em média de 5 pares de base. Pode-se concluir que foram identificados novos alelos dentro da população Gir em estudo, reforçando a tese de que as raças zebuínas precisam de mais estudos moleculares, do contrário, essa diversidade genômica observada nas raças zebuínas pode ser comprometida e reduzida sua confiabilidade na utilização de marcadores moleculares para a seleção assistida, e, em estudos de teste de paternidade.

**Tabela 9** - Locos, número de alelos, tamanho mínimo e máximo em pares de base (pb), nos mapas MARC e o obtido do PNMGL.

Loco	MARC			PNMGL		
	N° Ale.	Min (pb)	Max (pb)	N° Ale.	Min (pb)	Max (pb)
ILSTS093	21	179	202	9	172	202
DIK4408	9	182	201	12	185	203
DIK5285	7	185	201	17	185	245
DIK4498	8	202	218	8	204	220
MNB-66	11	182	202	12	181	206
BM1329	9	137	161	11	145	169
DIK1058	7	131	157	8	134	162
BMS2508	9	87	111	8	86	115
DIK4382	10	188	380	14	349	387
DIK4482	7	179	199	8	173	197
MNB-208	7	131	153	9	133	155
BM4322	6	165	187	12	149	185
BMS470	9	59	85	9	64	87
DIK3026	9	338	364	9	263	363
DIK2294	10	189	217	17	195	214
DIK4867	7	217	232	8	213	228
ILSTS035	19	208	266	18	214	274
DIK4574	10	191	215	6	191	221
BMS5021	8	176	190	11	161	191
AFR227	11	96	120	11	80	116
DIK2174	5	228	242	8	234	250
DIK4827	7	189	210	10	177	205
DIK2995	5	207	221	6	205	217
DIK1182	14	298	333	9	301	336
DIK2690	4	213	221	8	205	223
BM2320	10	128	152	21	122	158
DIK4992	4	201	207	7	201	213

Foram calculadas heterozigidade observada ( $H_o$ ) e esperada ( $H_e$ ), e o conteúdo de informação polimórfica (PIC) dos 27 marcadores microssatélites (Tabela 10). Estas análises tiveram como objetivo verificar o grau de

polimorfismo entre todos os 27 marcadores e dentro de cada marcador. A média de  $H_o$  na população Gir Leiteiro foi de 0,67 e no mapa MARC de 0,65. O marcador que apresentou maior  $H_o$  dentro da população Gir Leiteiro foi DIK2294 (0,848), e o menor foi o loco MBS5021 (0,348). Segundo Ott (1992), um loco é considerado polimórfico se  $H_o$  foi superior a 0,1 e altamente polimórfico se  $H_o$  foi superior a 0,7. Na população estudada 11 locos foram altamente polimórficos e 16 foram polimórficos. A média do conteúdo de informação polimórfica (PIC) na população Gir foi de 0,6472, considerada altamente polimórfica. Segundo Botstein *et al.* (1980) um PIC superior a 0,5 é considerado altamente polimórfico, entre 0,25 e 0,5 é moderadamente polimórfico, e menor que 0,25 é considerado pouco polimórfico. O marcador que apresentou maior PIC foi DIK2294 (0,848) e o menor PIC foi DIK4574 (0,372). O PIC é um índice que serve, também, para determinar o grau de polimorfismo de um loco.

Em populações em que o genótipo dos parentais é variável, as famílias a serem genotipadas podem ser escolhidas baseando-se na heterozigidade dos marcadores, pois quando um marcador apresenta grande número de alelos, os valores do PIC são semelhantes a  $H_e$  e, a medida que o número de alelos aumenta, o PIC e a heterozigidade aumentam (Liu, 1998).

Os marcadores BM1329, BMS5021 e BM2320 apresentaram frequências de alelos nulos de 0,07, 0,24 e 0,16, respectivamente. Segundo Summer *et al.* (1997), a frequência de alelos nulos que segrega em cada loco é calculada por meio de um algoritmo iterativo baseado na diferença entre a frequência observada e a frequência esperada de homozigotos. O alelo nulo pode ser qualquer alelo que não é identificado pela genotipagem que acontece por causa de mutações no lugar de ligação em um ou nos dois *primers*. Os locos com frequência de alelos nulos maiores ou igual a 0,05 são considerados com frequências altas.

Após os processos de amplificação e genotipagem, foi calculada uma média de 4,1% de falhas em toda a população estudada. Esta porcentagem está dentro do normal, pois algumas amostras de DNA mais velhas comprometeram os processos de PCR e de genotipagem automática. O marcador DIK3026 apresentou a maior porcentagem de falhas (11,24%), e o loco ILSTS0936, o menor número de falhas (0,62%).



Pode-se concluir que aqueles marcadores com alta heterozigosidade e conteúdo de informação polimórfica podem ser selecionados para estudos que impliquem detecção de QTL, genealogias, identidade genética e teste de paternidade para a raça Gir Leiteiro e/ou para outras raças zebuínas.

**Tabela 10** - Locos, heterozigosidade observada ( $H_o$ ), heterozigosidade esperada ( $H_e$ ), conteúdo de informação de polimorfismo (PIC) e percentagens de falhas (%F) para os 27 marcadores microssatélites analisados na população Gir Leiteiro do PNMGL.

<b>Locos</b>	<b><math>H_o</math></b>	<b><math>H_e</math></b>	<b>PIC</b>	<b>% F</b>
ILSTS093	0,765	0,703	0,675	3,08
DIK4408	0,796	0,800	0,776	3,23
DIK5285	0,785	0,772	0,745	2,46
DIK4498	0,461	0,469	0,390	2,92
MNB-66	0,863	0,843	0,822	2,46
BM1329	0,552	0,640	0,602	4,006
DIK1058	0,748	0,740	0,705	4,93
BMS2508	0,684	0,745	0,703	2,62
DIK4382	0,840	0,852	0,834	8,16
DIK4482	0,605	0,581	0,503	2,62
MNB-208	0,815	0,773	0,740	2,16
BM4322	0,519	0,555	0,479	5,55
BMS470	0,658	0,668	0,610	0,77
DIK3026	0,605	0,591	0,560	11,24
DIK2294	0,848	0,888	0,877	2,31
DIK4867	0,667	0,674	0,623	0,61
ILSTS035	0,820	0,779	0,752	3,23
DIK4574	0,493	0,465	0,372	0,924
BMS5021	0,348	0,560	0,505	3,235
AFR227	0,635	0,668	0,605	1,54
DIK2174	0,689	0,682	0,635	2,31
DIK4827	0,705	0,718	0,665	2,77
DIK2995	0,579	0,649	0,578	6,78
DIK1182	0,639	0,718	0,694	4,77
DIK2690	0,718	0,678	0,619	4,93
BM2320	0,632	0,870	0,856	8,01
DIK4992	0,651	0,610	0,549	2,31

#### 4.2 Mapa de ligação

As distâncias de mapa entre os 27 marcadores analisados foram computadas com a opção “fixed” do Crimap (Green *et al.*, 1990). Para a organização dos dados, uma família de meios-irmãs foi conformada pela filha, pai e a mãe. Na maioria dos touros, foram encontradas filhas partilhando à mesma mãe, estas foram consideradas como famílias de irmãs completas. Na tabela 11 encontram-se as famílias de meios-irmãs e irmãs completas para cada touro e o total de famílias por touro.

**Tabela 11** – Famílias, código de Touro, famílias de meios-irmãs e irmãs completas para a realização do mapa de ligação.

Famílias	Código do touro	Famílias de meios-irmãs	Famílias de irmãs completas	Total
1	001	88	13	101
2	129	53	12	65
3	224	57	5	62
4	293	39	10	49
5	356	39	9	48
6	416	32	6	38
7	463	21	5	26
8	498	22	3	25
9	527	22	2	24
10	556	15	4	19
11	581	14	4	18
12	605	18	2	20
13	628	22	0	22
14	651	20	0	20
	Total	462	75	537

Segundo Lynch e Walsh (1998) um mapa de ligação está em função do número de marcadores, dos animais utilizados e do tipo de delineamento em estudo. Para gerar o mapa do cromossomo 6 dentro do lineamento de filhas, foi montada uma planilha em que os dados foram organizados pelas famílias com as informações de pais e filhas (Tabela 12). Esta planilha continha 537 famílias para o mapa gerado da raça Gir Leiteiro do PNMGL e 19 marcadores microssatélites, sendo retirados os marcadores com maior distorção.

Na acurácia do mapa de ligação, o número de indivíduos está relacionado, uma vez que, quanto maior o número de meioses informativas, maior a probabilidade de detectar recombinação entre pequenas distâncias, característica em populações comerciais. O número médio de meioses informativas encontradas dentro da população Gir Leiteiro (167,9) foi menor que a média do mapa MARC (467,3). Isto pode ser explicado pelo número de indivíduos e o delineamento utilizado. O marcador ILSTS0936 foi o único que apresentou maior número de meioses informativas (164) para o mapa da população Gir. A ordem dos marcadores não foi alterada após a formação dos novos mapas. Para todas as análises de mapeamento foi utilizado o mapa consenso, porque tem sido o mapa comumente aceitado e porque mostrou que as diferenças em frequências de recombinação estimadas não apresentaram viés no teste para QTL ou efeitos de QTL estimados (Haley e Knott, 1992).

**Tabela 12** - Comparação entre o mapa consenso MARC/USDA, e o mapa estabelecido da população Gir Leiteiro, referente à posição (Pos), distância (Dis) e meioses informativas (M.I.) para locos no cromossomo 6.

Locos	Mapa MARC			Mapa PNMGL		
	Pos.	Dis.	M.I.	Pos.	Dis.	M.I.
ILSTS093	0,00	-	2397	0,00	-	198
DIK5285	15,362	15,3	299	15,2	15,2	207
DIK4498	20,131	4,8	170	22,0	6,8	130
MNB-66	29,374	9,243	301	33,6	11,6	218
DIK1058	38,169	8,9	320	40,2	6,6	171
BMS2508	43,936	5,767	1872	48,3	8,1	238
DIK4382	50,097	6,161	342	54,8	6,5	185
DIK4482	54,516	4,419	318	59,7	4,9	136
MNB-208	60,218	5,702	249	66,7	7,0	153
DIK3026	71,561	11,3	260	76,0	9,4	153
DIK2294	75,27	3,709	281	83,0	7,0	272
DIK4867	81,961	6,695	132	91,7	8,6	164
ILSTS035	87,265	5,304	625	96,9	5,2	181
DIK4574	90,505	3,24	287	98,9	2,0	79
DIK4827	107,12	16,6	274	112,3	13,4	199
DIK2995	109,92	2,8	184	114,4	2,1	158
DIK1182	115,32	5,4	298	120,7	6,3	179
DIK4492	130,781	15,5	241	134,5	13,8	169

### 4.3 Mapeamento de QTL

Na tabela 13 encontram-se os touros e as filhas genotipadas por marcador, e o número de touros heterozigotos por marcador que foram utilizados para as análises de mapeamento de QTL. O loco que apresentou maior número de filhas genotipadas foi o DIK4574 com 614 filhas, correspondente a 93%. Já os marcadores com menor número de filhas genotipadas foram DIK4382 e DIK3026 com 527 (80%) e 532 (80%), cada um para as respectivas análises. Todos os 14 touros foram heterozigotos para o marcador BMS2508 que, além de sua alta heterozigosidade esperada e PIC (Tabela 10) o converte em um forte candidato para testes de identidade e

paternidade nas raças zebuínas por seu alto poder de discriminação individual. A metade dos touros foram heterozigotos para os marcadores DIK4867 e DIK4992, sendo os marcadores que apresentaram o menor número de touros heterozigotos. O touro 651 apresentou o maior número de marcadores heterozigotos, 25 em total, enquanto que o touro 628 foi heterozigoto para 16 marcadores microssatélites.

**Tabela 13** - Locos, filhas genotipadas, número de touros homozigotos e heterozigotos por marcador.

<b>Locos</b>	<b>Filhas genotipadas (%)</b>	<b>Código dos touros homozigotos</b>	<b>Total de touros heterozigotos</b>
ILSTS093	588 (89%)	129-224-527	11
DIK4408	586 (89%)	416-527-556	11
DIK5285	573 (87%)	556	13
DIK4498	599 (91%)	129-224-356-416-527-581	8
MNB-66	575 (87%)	224-356	12
BM1329	560 (85%)	356-416-463-498-556-628	8
DIK1058	579 (88%)	1-605	12
BMS2508	554 (84%)	-	14
DIK4382	527 (80%)	556	13
DIK4482	602 (91%)	1-293-463-556	10
MNB-208	591 (89%)	1-129-356	11
BM4322	576 (87%)	129-224-356-498-581-628	8
BMS470	608 (92%)	463-498-581-628	10
DIK3026	532 (80%)	1-581-628	11
DIK2294	577 (87%)	605	13
DIK4867	566 (86%)	224-356-463-498-527-556-605	7
ILSTS035	573 (87%)	1-356-527-628	10
DIK4574	614 (93%)	129-293-356-463-556-581	8
BMS5021	539 (82%)	1-224-581-628	10
AFR227	598 (91%)	293-498-527-605	10
DIK2174	599 (91%)	1-527	12
DIK4827	585 (89%)	1	13
DIK2995	544 (82%)	293-527	12
DIK1182	539 (82%)	463-556-605-628	10
DIK2690	564 (85%)	527-581-628	11
BM2320	433 (65%)	498	13
DIK4992	599 (91%)	224-356-463-498-581-605-628	7

De acordo com as leis de Mendel, cada pai e progênie devem apresentar ao menos um alelo comum. Discrepâncias entre genótipos de uma vaca e seu suposto pai (touro) podem ser devido a mutações, erros na

genotipagem ou erros nos registros de paternidade (Ron *et al.*, 2004). As fêmeas que não apresentaram pelo menos um alelo do touro para mais de cinco locos (> 5) foram descartadas para a realização das diferentes análises. Um total de 35 vacas (5,32%) não foi considerado filhas por incorreta paternidade, estando de acordo com Ron *et al.* (1996) e Ron *et al.* (2001), que aplicando delineamento de filhas consideraram uma paternidade mínima de 90%. Para as famílias 498 e 628, todas as fêmeas foram filhas, portanto, assumidas nas análises. A família 651 apresentou o maior número de filhas descartadas, 4 de 20, o que é equivalente a 20% (Tabela 14).

**Tabela 14** - Número de fêmeas não-filhas nas 14 famílias da população Gir.

Famílias	Código do touro	Não-filhas (%)	Total
1	1	6 (4,7%)	127
2	129	4 (4,3%)	94
3	224	5 (7,4%)	68
4	293	2 (3,2%)	62
5	356	2 (3,4%)	59
6	416	3 (6,5%)	46
7	463	2 (5,9%)	34
8	498	0 (0,0%)	28
9	527	3 (10,7%)	28
10	556	2 (8,4%)	24
11	581	1 (4,4%)	23
12	605	1 (8,7%)	23
13	628	0 (0,0%)	22
14	651	4 (20%)	20

O número total de fêmeas genotipadas por família e o número de filhas informativas por cada marcador são detalhadas na tabela 15. Filhas informativas são aquelas que apresentam genótipos diferentes a seu pai (Ron *et al.*, 1996). Segundo Martinez (1989), no delineamento de filhas não é possível determinar inequivocamente a origem do alelo marcador para todas as progênes, ao contrário do que acontece com linhagens consanguíneas. Se uma filha tem o mesmo genótipo do pai, ela pode ter recebido qualquer um dos



alelos de qualquer um dos pais e, em geral, as informações dessas progênies não são utilizadas nas análises.

Na tabela 15 também se encontram os touros que foram homozigotos para um determinado marcador. Os touros 356 (família 5) e 628 (família 13) apresentaram o maior número de locos homozigotos (10), o que é equivalente a 37,03%. Já o touro 651 (família 14) só foi homozigoto para 2 marcadores microssatélites, o que é equivale a 7,4%. Isto é importante porque se vários touros heterozigotos estão segregando para um QTL específico em uma determinada região cromossômica e há touros que são homozigotos para os marcadores que estão flanqueando essa posição do QTL, as famílias desses touros homozigotos não serão informativas, o que não contribui para a detecção do QTL em particular.

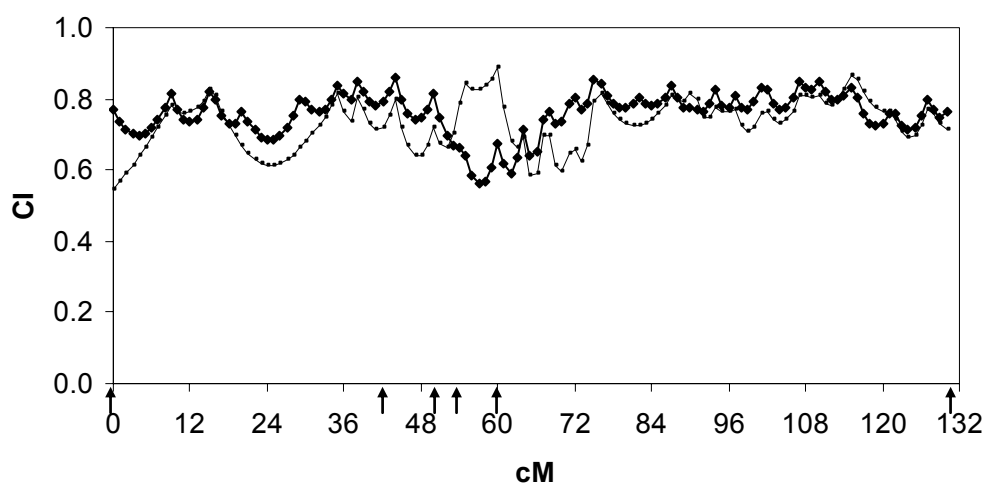
Importante sublinhar que na família 7 (touro 463) só 9 filhas de 37 foram informativas para o marcador DIK4498 (20,13 cM), o que equivale a 26,47%. Isso significa que aproximadamente 74% das filhas genotipadas não foram informativas para esse marcador específico, tanto por apresentarem genótipos ambíguos (mesmo genótipo do touro), descarte como possível filha, falhas no processo de amplificação ou genotipagem ou por não apresentar nenhum alelo do touro.

**Tabela 15** - Famílias, número de filhas genotipadas por família e filhas informativas por loco (H = touro homocigoto).

Famílias	Nº Filhas													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Locos	127	94	68	62	59	46	34	28	28	24	23	22	22	20
ILSTS093	103	H	H	55	48	37	22	21	H	22	15	16	17	18
DIK4408	105	67	62	60	48	H	24	28	H	H	18	22	20	18
DIK5285	96	69	59	51	54	44	30	26	27	H	23	21	21	16
DIK4498	79	H	H	45	H	H	9	14	H	20	H	21	19	13
MNB-66	121	83	H	46	H	43	22	27	24	24	16	12	16	18
BM1329	94	83	66	56	H	H	H	H	21	H	22	17	H	15
DIK1058	H	81	49	57	33	42	32	20	18	24	19	H	18	H
BMS2508	97	80	60	62	49	44	28	23	26	20	14	18	16	16
DIK4382	113	77	64	62	55	46	30	22	24	H	12	18	20	17
DIK4482	H	60	46	H	30	30	H	19	16	H	10	16	9	11
MNB-208	H	H	65	58	H	33	29	19	27	18	22	17	16	17
BM4322	85	67	H	41	H	36	23	H	24	8	H	13	H	11
BMS470	95	68	49	51	36	36	H	H	22	14	H	17	H	15
DIK3026	H	64	41	61	43	43	29	21	26	15	H	18	H	18
DIK2294	110	87	60	48	44	44	30	23	23	24	18	H	19	18
DIK4867	117	64	H	64	H	28	H	H	H	H	22	H	14	14
ILSTS035	H	82	48	62	H	42	34	23	H	20	22	14	H	20
DIK4574	73	H	26	H	H	23	H	8	16	H	H	11	11	16
BMS5021	H	89	H	60	45	45	34	25	23	12	H	14	H	20
AFR227	94	73	44	H	22	22	20	H	H	14	14	H	12	19
DIK2174	H	78	49	49	35	35	29	20	H	15	23	19	19	12
DIK4827	H	72	57	52	41	41	28	25	19	18	16	16	19	19
DIK2995	86	59	42	53	H	46	19	18	H	17	16	18	21	20
DIK1182	120	82	54	48	45	45	H	20	27	H	21	H	H	18
DIK2690	84	H	48	54	42	42	29	22	H	15	H	17	H	16
BM2320	126	85	62	60	45	45	H	27	27	23	22	21	H	H
DIK4992	77	81	H	51	H	37	H	H	15	12	H	H	H	15

O conteúdo de informação (CI) para a produção de leite está representado na figura 1. O CI foi calculado da variância das probabilidades

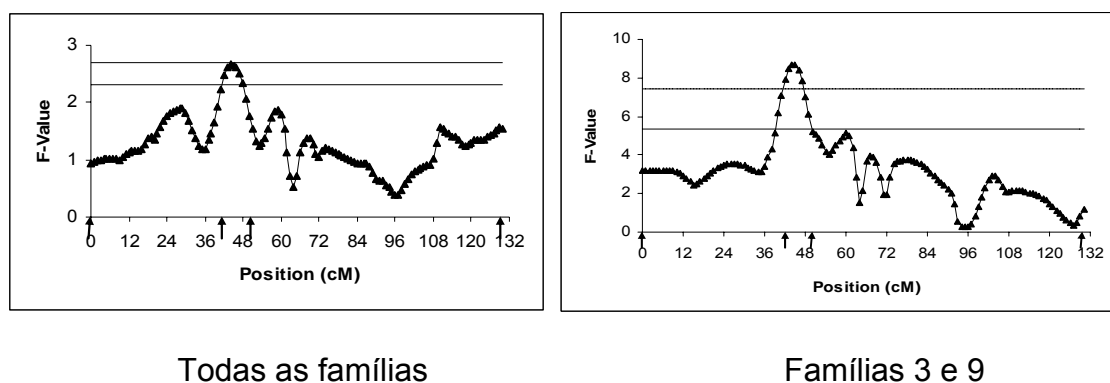
condicionais de herdar uma região cromossômica a cada cM como uma proporção da variância quando o descendente verdadeiro é conhecido. O CI através de todas as famílias ao longo do cromossomo 6 foi maior que nas famílias 3 e 9 juntas, exceto na região compreendida entre os marcadores DIK4482 e MNB-208.



**Figura 1** - Conteúdo de informação (CI) considerando todas as famílias (▲) e nas famílias 3 e 9 (—) para produção de leite. As setas indicam a posição dos marcadores BMS2508 (43,93 cM) e DIK4382 (50,09 cM).

Após a análise, considerando todas as famílias para cada característica de produção de leite, foram identificadas as famílias mais significativas de acordo com os valores absolutos de “t”, que são as significâncias dos efeitos dentro de cada família do alelo marcador paterno. As famílias que apresentaram um nível de significância igual ou maior a 5% para um QTL específico foram consideradas significativas e levadas em conta para a detecção de QTL ao longo do cromossomo 6. Chen *et al.* (2006) identificaram QTL afetando características de produção de leite sobre BTA6 em uma população Holandesa com um delineamento de filhas utilizando 26 famílias e 2.260 fêmeas, para isso, escolheram as famílias mais significativas para cada característica; Ron *et al.* (2001) através da análise de QTL sobre o cromossomo 6 em uma população Holandesa por delineamento de filhas, empregando 9 famílias com 2.978 filhas, selecionaram as famílias mais significativas para a identificação de QTL.

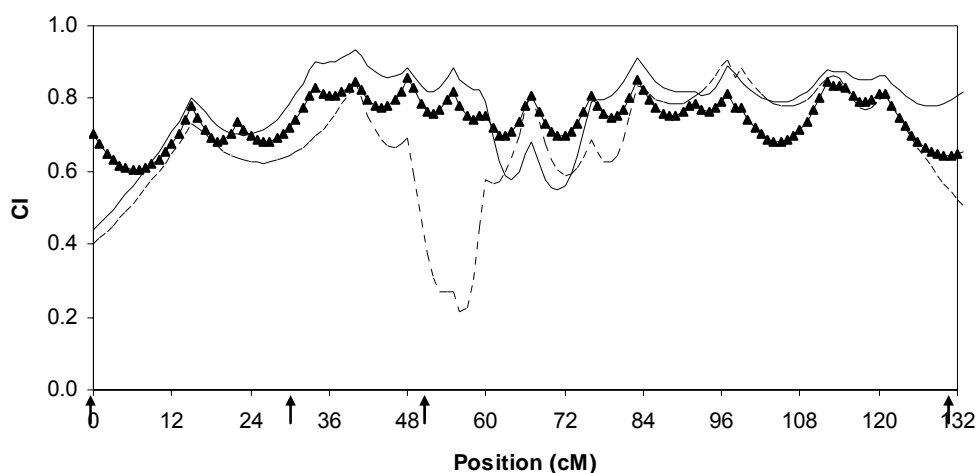
Considerando todas as famílias para a característica produção de leite, foi identificado um QTL (Figura 2) com  $P < 0,01$  e  $F = 2,66$  na posição 44 cM, próximo ao marcador BMS2508 (43,93 cM). Entre os marcadores BMS2508 (43,93 cM) e DIK4382 (50,09 cM) os seguintes estudos observaram QTL para produção de leite: George *et al.* (1995) detectaram um QTL na posição 47 cM; Freyer *et al.* (2002) identificaram um QTL na posição 49 cM; Ron *et al.* (2001) detectaram um QTL na posição 47 cM; Cohen *et al.* (2002) encontraram um QTL na posição 49 cM e Khatkar *et al.* (2004), identificaram um QTL na posição 50 cM. O marcador BMS2508 apresentou  $H_o=0,68$  e  $PIC=0,75$  (Tabela 10), além disso, todos os 14 touros foram heterozigotos para esse marcador (Tabela 12); e o marcador DIK4382 apresentou uma  $H_o=0,84$  e  $PIC=0,85$ , sendo que 13 touros foram heterozigotos para o mencionado marcador. Com as famílias 3 e 9, para a mesma característica, foi identificado o mesmo QTL na posição 46 cM e com  $F=9,64$ .



**Figura 2** - Valor de F para produção de leite considerando todas as famílias e as famílias 3 e 9. As setas indicam a posição dos marcadores BMS2508 e DIK4382, respectivamente. — 5% e ---- 1% de significância a nível cromossômico.

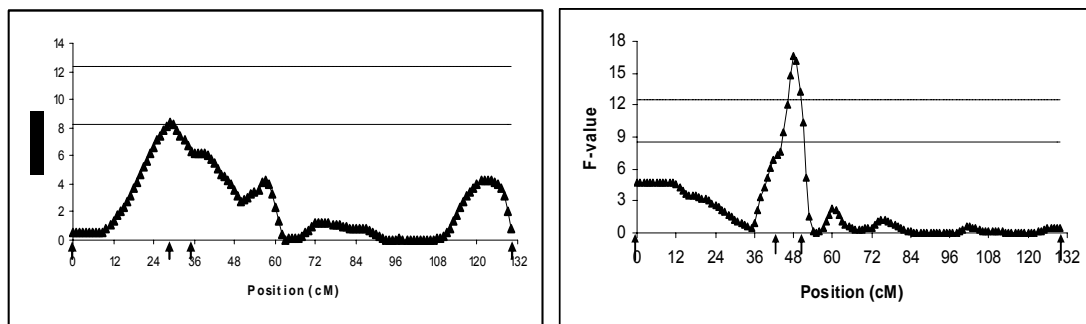
Na figura 3 encontra-se o conteúdo de informação para a produção de gordura. Ao longo do cromossomo 6 a família 2 apresentou maior CI com relação a todas as famílias e à família 3, exceto nas regiões cromossômicas entre os locos ILSTS093 (0 cM) e DIK4408 (9,028 cM), o mesmo ocorreu entre os marcadores DIK4382 (50,09 cM) e MNB-208 (60,21) em que o CI para todas as famílias foi maior. A família 3 apresentou um CI menor (0,3) com relação às

outras famílias na região cromossômica compreendida entre os marcadores DIK4382 (50,09 cM) e MNB-208 (60,21 cM).



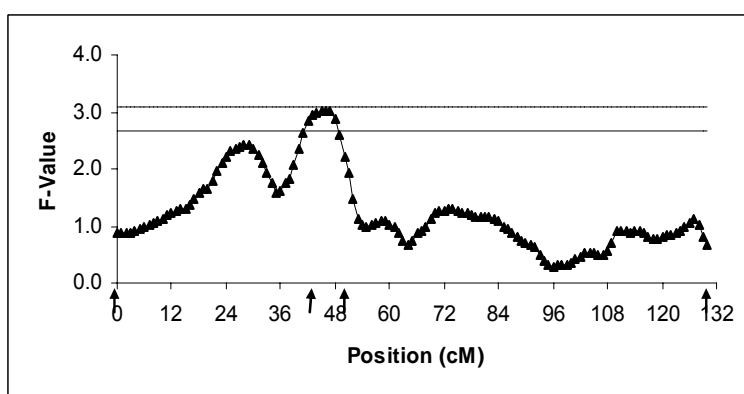
**Figura 3** - Conteúdo de informação (CI) consideradas todas as famílias ( $\blacktriangle$ ), família 2 (—) e família 3 (---) para produção de gordura. As setas indicam a posição dos marcadores MNB-66 (29,37 cM) e DIK4382 (50,09 cM).

Ao considerar todas as famílias, um QTL para produção de gordura foi identificado com  $P < 0,01$  e  $F = 3,03$  (Figura 4) na posição cromossômica 46 cM. Próximo ao marcador BMS2508 (43,93 cM) na família 3, foi mapeado o mesmo QTL na posição 48 cM com  $P < 0,01$  e  $F = 16,61$ , sendo a única família que apresentou um nível de significância igual ou maior a 5%. Nesta mesma região para produção de gordura, Cohen *et al.* (2004) identificaram um QTL na posição 49 cM e Ron *et al.* (2001) na posição 45 cM. Para esta mesma característica foi identificada a presença de outro QTL com  $P < 0,05$  e  $F = 8,32$  na posição 29 cM na família 2, próximo ao marcador MNB-66 (29,3 cM). Zhang *et al.* (1998) encontraram um QTL na posição 24 cM utilizando uma população de 14 famílias da raça Holandesa Americana com um delineamento de netas.



Família 2

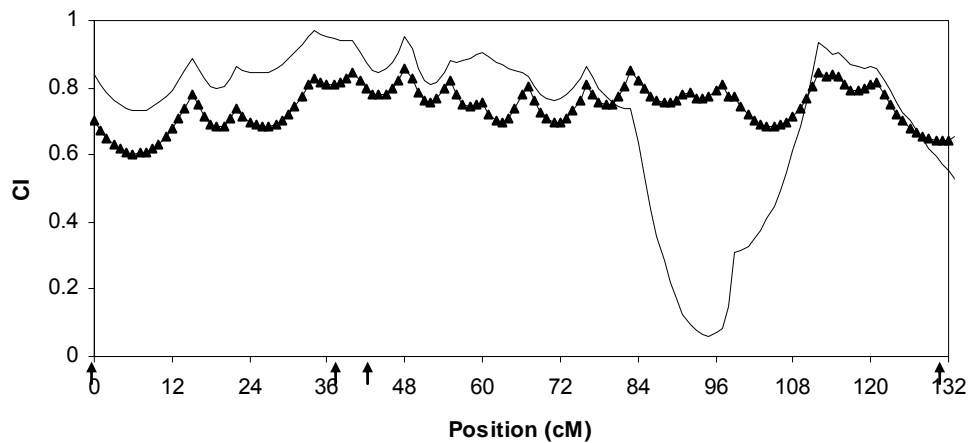
Família 3



Todas as famílias

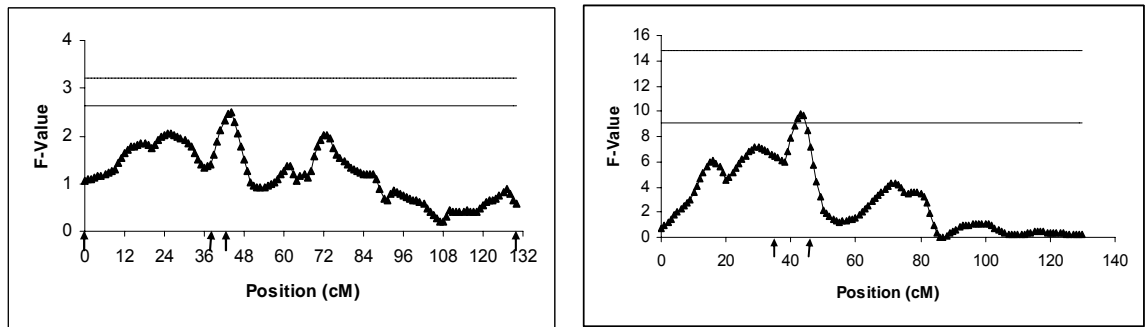
**Figura 4** - Valor de F para a característica produção de gordura considerando todas as famílias e nas famílias 2 e 3 juntas. As setas indicam a posição dos marcadores MNB66 (29,37 cM) e BM1329 (35,39 cM) para a família 2, e, BMS2508 (43,93 cM) e DIK4382 (50,09 cM) para todas as famílias e para a família 3. — 5% e ----- 1% de significância a nível cromossômico.

Na figura 5 está representado o conteúdo de informação considerando todas as famílias e a família 8 para produção de proteína. O CI na família 8 foi maior entre os marcadores ILSTS093 (0,00 cM) e DIK4867 (81,9 cM), mas apresentou uma diminuição drástica na região cromossômica entre os marcadores DIK4867 e DIK4827 (107,12 cM). Em geral, a curva de CI com todas as famílias foi mais estável ao longo do cromossomo que na família 8.



**Figura 5** - Conteúdo de informação (CI) considerando todas as famílias (▲) e a família 8 para produção de proteína (—). As setas indicam a posição dos marcadores DIK1058 (38,17 cM) e BMS2508 (43,93 cM).

Ao Considerar todas as famílias para produção de proteína não foi identificado nenhum QTL com  $P < 0,05$ . A família 8 apresentou maior valor de  $t$  para produção de proteína, sendo a única a dar o nível de significância exigido para o mapeamento de QTL. Nessa família foi identificado um QTL com  $P < 0,05$  e  $F = 9,74$ , próximo ao marcador BMS2508 (43,93 cM) e na posição 43 cM (Figura 6). Na região compreendida entre os marcadores BMS2508 (43,93 cM) e DIK4382 (50,09 cM) para produção de proteína, Bennwitz *et al.* (2003) identificaram um QTL na posição 43 cM; Olsen *et al.* (2004) encontraram um QTL na região 46 cM; Viitala *et al.* (2003) identificaram um QTL na posição 47 cM; Ashwell *et al.* (2004) encontraram um QTL na posição 48 cM e Cohen *et al.* (2002) observaram um QTL na posição 49 cM.



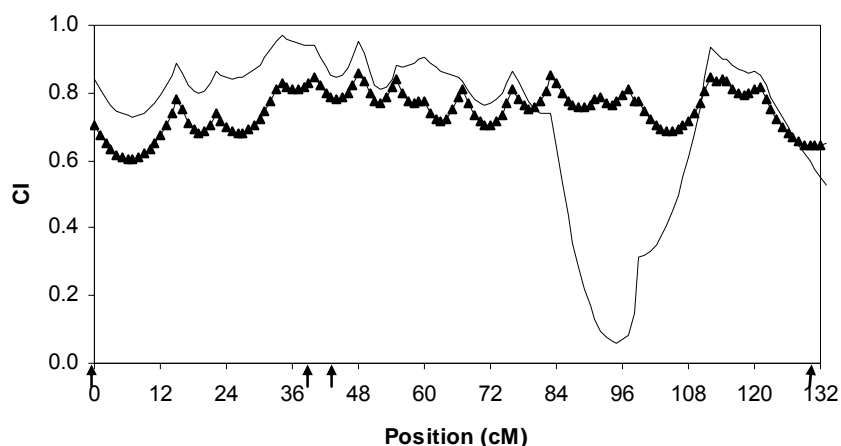
Todas as famílias

Família 8

**Figura 6** - Valor de F para a característica produção de proteína considerando todas as famílias e a família 8. As setas indicam a posição dos marcadores DIK1058 (38,16 cM), BMS2508 (43,93 cM) e DIK4382 (50,09 cM). — 5% e ---- 1% de significância a nível cromossômico.

O conteúdo de informação para produção de lactose ao longo do cromossomo 6 apresentou diferenças significativas na região compreendida entre os marcadores DIK4867 (81,96 cM) e DIK4827 (107,12 cM), sendo maior ao considerar todas as famílias (Figura 7). Assim como para a produção de proteína, o conteúdo de informação na família 8 foi maior na região compreendida entre os marcadores ILSTS093 (0,0 cM) e DIK4867 (81,96 cM). Os QTL identificados na família 8 para esta característica estavam localizados na região onde o conteúdo de informação foi maior em relação a todas as famílias.



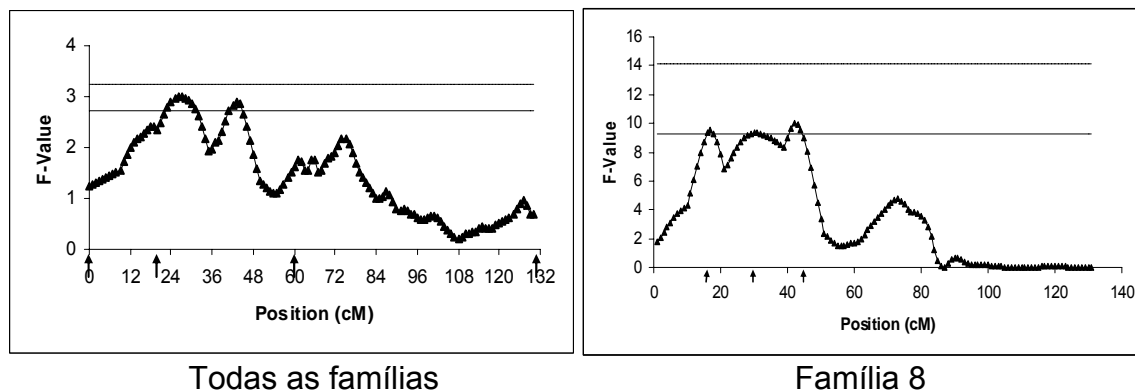


**Figura 7** - Conteúdo de informação (CI) considerando todas as famílias (▲) e na família 8 para produção de lactose (—). As setas indicam a posição dos marcadores DIK1058 (38,16 cM) e BMS2508 (43,93 cM).

Ao considerar todas as famílias foram identificados dois QTL (Figura 8): Um QTL com  $P < 0,05$  e  $F = 2,93$  na posição 29 cM próximo ao marcador MNB-66 (29,37 cM) e outro na posição 43 cM próximo ao marcador BMS2508 (43,93 cM) com  $P < 0,05$  e  $F = 2,88$ . Com a família 8, a figura anterior identifica a existência de três QTL para produção de lactose: Um QTL com  $P < 0,05$  e  $F = 9,52$ , na posição 17 cM, próximo ao marcador DIK5285 (15,36 cM), um segundo QTL com  $P < 0,05$  e  $F = 9,29$ , na posição 29 cM, próximo ao marcador MNB-66 (29,37 cM), e o terceiro QTL com  $P < 0,05$  e  $F = 9,96$  na posição 43 cM.

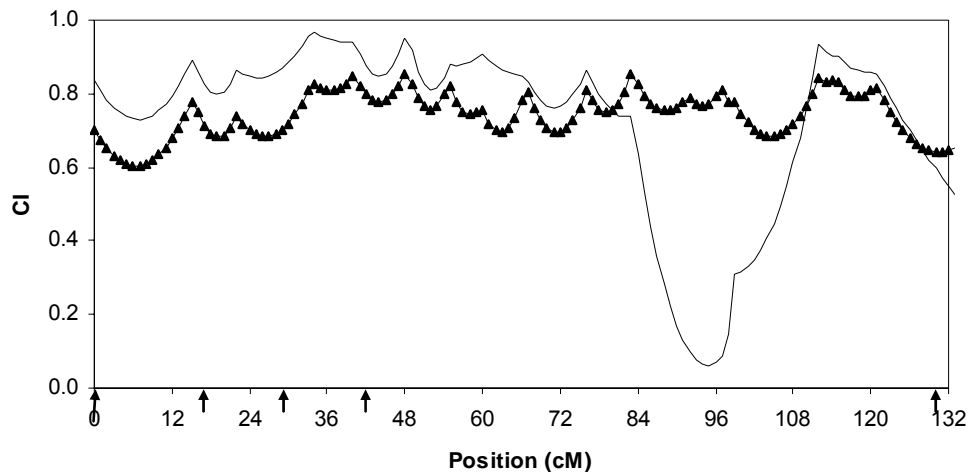
Segundo Schuster e Cruz (2004) se houver mais de um QTL no mesmo grupo de ligação, o teste estatístico na posição que está sendo avaliado é afetado por todos os QTL, a posição e os efeitos estimados dos QTL identificados podem ser viesados e o poder do teste reduzido. Além disto, pode ocorrer o aparecimento de QTL fantasmas, ou seja, a detecção de um QTL com efeito significativo dentro de um intervalo em que o mesmo não acontece pela presença de outros QTL em intervalos adjacentes. Para avaliar a presença de três, dois ou um QTL, realizou-se uma nova análise no software GridQTL, que permite a identificação de dois QTL em um mesmo grupo de ligação. Em todas as famílias foram identificados dois QTL nas posições 29 e 43 cM, e o valor de  $F$  encontrado quando se comparou 2 QTL vs 0 QTL foi de 2,62, porém

o F encontrado quando se comparou 2 QTL vs 1 QTL foi de 2,17. Utilizando a mesma opção do GridQTL para identificar dois QTL no mesmo grupo de ligação com a família 8, foi conferido o QTL na posição 43 cM.



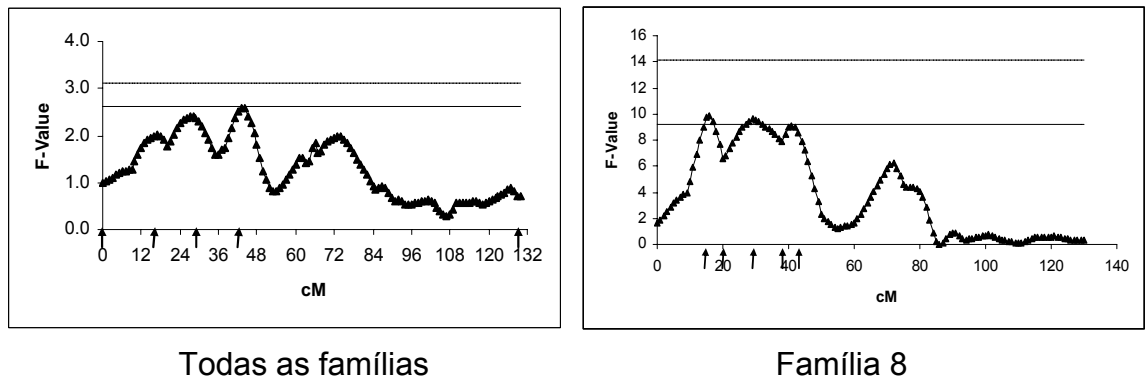
**Figura 8** - Valor de F para a característica produção de lactose considerando todas as famílias e a família 8. As setas indicam a posição dos marcadores DIK4498 (20,13 cM), MNB-66 (29,37 cM) e BMS2508 (43,93 cM). — 5% e ---- 1% de significância a nível cromossômico.

O conteúdo de informação para a característica de produção de sólidos totais encontra-se na figura 9. A família 8 apresentou maior conteúdo de informação entre os marcadores ILSTS093 (0,0 cM) e DIK4867 (81,96 cM), mas na região compreendida entre os marcadores DIK4867 e DIK1182 (115,32 cM) apresentou uma diminuição considerável. Na região marcada pelos locos DIK1182 e DIK2690 (121,5 cM), novamente o conteúdo de informação na família 8 foi maior que em todas as famílias, mas na parte final do cromossomo foi inferior.



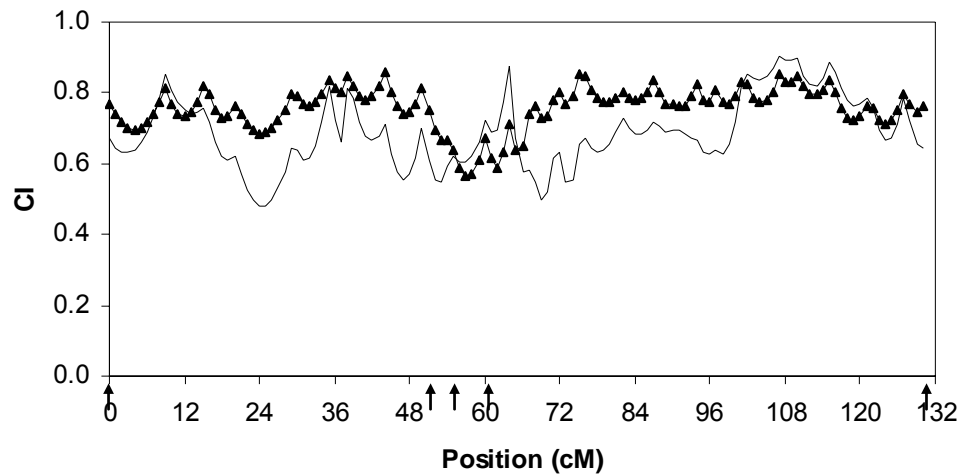
**Figura 9** - Conteúdo de informação considerando todas as famílias (▲) e a família 8 para produção de sólidos totais (—). As setas indicam a posição dos marcadores DIK5285 (15,36 cM), MNB-66 (29,37 cM) e BMS2508 (43,93 cM).

Considerando todas as famílias um QTL foi detectado para produção de sólidos totais, na posição 43 cM, com  $P < 0,05$  e  $F = 2,59$  (Figura 10). O QTL foi encontrado entre os marcadores DIK1058 (38,16 cM) e BMS2508 (43,93 cM). A família 8 foi a única família que apresentou o nível de significância de  $P < 0,05$  para mapeamento de QTL, e nela foram identificados três possíveis QTL, nas posições 17, 30 e 42 cM, respectivamente. Com a opção para identificar dois QTL no mesmo grupo de ligação, por meio do software GridQTL, foi conferido o QTL na posição 42 cM, próximo ao marcador BMS2508 (43,93 cM). Os sólidos totais ou extrato seco total compreende todos os componentes do leite, exceto a água. Sua importância reside no fato de que alguns laticínios pagam ou bonificam de acordo ao conteúdo de extrato seco, não entanto, poucos trabalhos existem na literatura identificando QTL relacionados com esta característica.



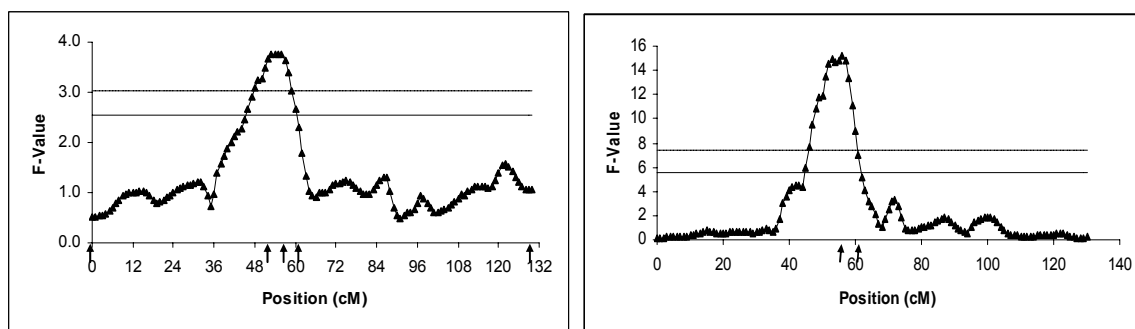
**Figura 10** - Valor de F para produção de sólidos totais considerando todas as famílias e a família 8. As setas indicam a posição dos marcadores DIK5285 (15,36 cM), MNB-66 (29,37 cM) e BMS2508 (43,93 cM). — 5% e ---- 1% de significância a nível cromossômico.

O conteúdo de informação para percentagem de gordura está representado na figura 11. Considerando todas as famílias seu conteúdo de informação foi maior que nas famílias 3 e 4, ao longo do cromossomo, exceto para as regiões cromossômicas entre os marcadores DIK4482 (54,51 cM) e BM4322 (63,86 cM), e no final do cromossomo entre os marcadores DIK2174 (101,4 cM) e DIK2690 (121,5 cM). Em termos gerais, o conteúdo de informação para todas as famílias foi mais estável ao longo do cromossomo que na família 8, cuja curva apresentou mais variações.



**Figura 11** - Conteúdo de informação considerando todas as famílias (▲) e as famílias 3 e 4 (—) para percentagem de gordura. As setas indicam a posição dos marcadores DIK4382 (50,09 cM), DIK4482 (54,51 cM) e MNB-208 (60,21 cM).

Para a porcentagem de gordura considerando todas as famílias foi identificado um QTL com  $P < 0,01$  e  $F = 3,76$ , na posição 53 cM (Figura 12), próximo ao marcador DIK4482 (54,51 cM). Com as famílias 3 e 4, as mais significativas, foi identificado o QTL, na posição 56 cM, próximo ao marcador DIK4482, com  $P < 0,01$  e  $F = 15,18$ . Nesta região, Ron *et al.* (2001) detectaram dois QTL nas posições 50 e 54 cM, Georges *et al.* (1995) identificaram um QTL na posição 51 cM, Boichard *et al.* (1997) identificaram um QTL na posição 58 cM, e Ashwell *et al.* (2002) mapearam um QTL na posição 58 cM.

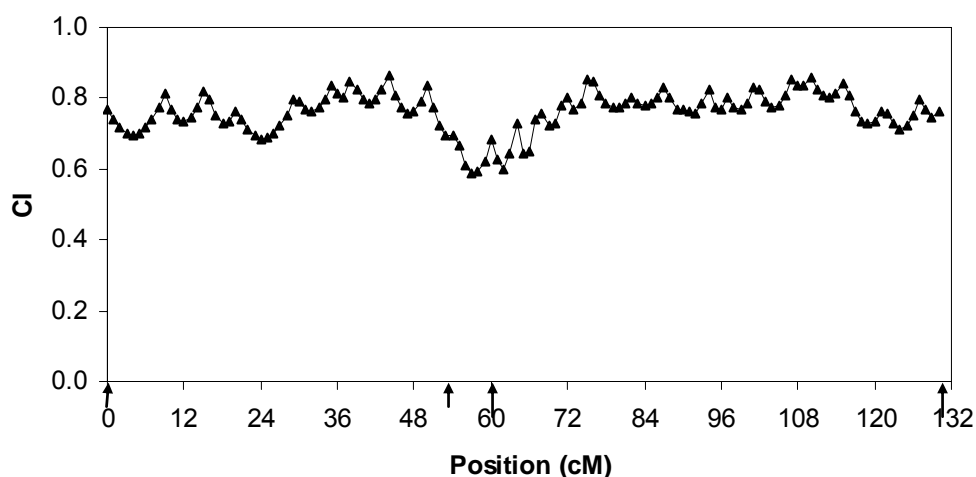


Todas as famílias

Famílias 3 e 4

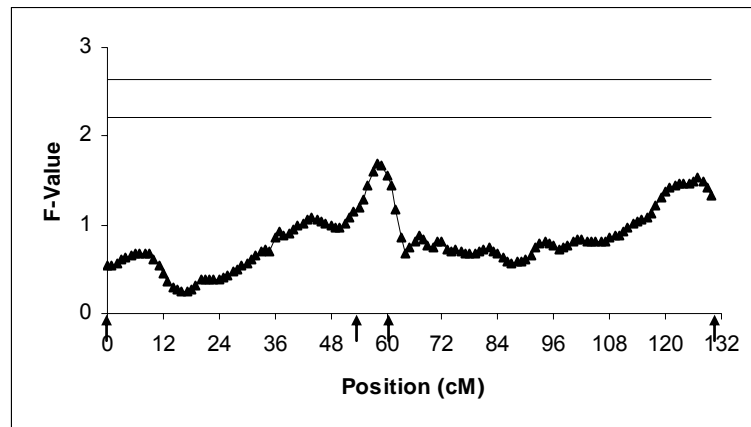
**Figura 12** - Valor de F para porcentagem de gordura considerando todas as famílias e as famílias 3 e 4. As setas indicam a posição dos marcadores DIK4382 (50,09 cM), DIK4482 (54,51 cM) e MNB-208 (60,21 cM). — 5% e ---- 1% de significância a nível cromossômico.

Na figura 13 encontra-se unicamente o conteúdo de informação que considerou todas as famílias para a característica da porcentagem de proteína, pois em nenhuma família identificou-se QTL com nível de significância de  $P < 0,05$ . Em geral, o conteúdo de informação ao longo do cromossomo foi maior que 0,7, exceto na região compreendida entre os marcadores MNB-208 (60,21 cM) e BMS470 (67,4 cM), onde o conteúdo diminuiu para 0,6.



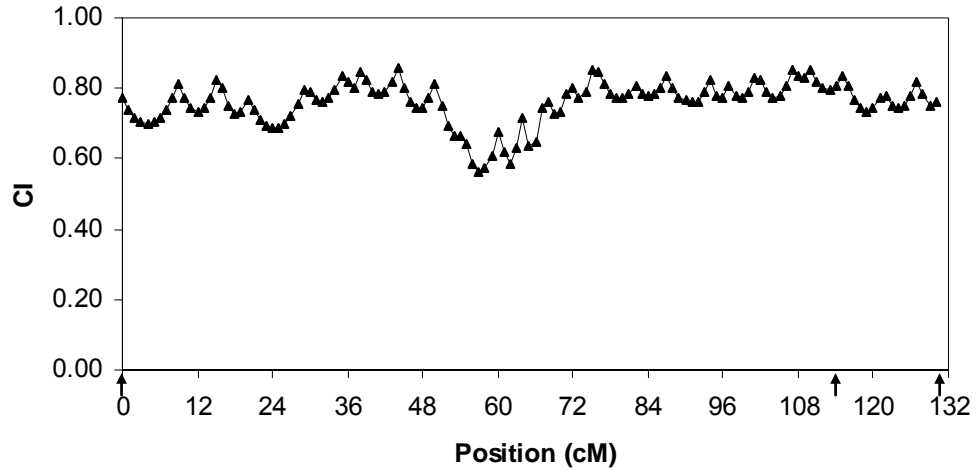
**Figura 13** - Conteúdo de informação considerando todas as famílias para porcentagem de proteína ( $\blacktriangle$ ). As setas indicam a posição dos marcadores ILST093 (0,0 cM), DIK4482 (54,51 cM), MNB-208 (60,21 cM) e DIK4492 (130,78 cM).

Ao considerar todas as famílias para porcentagem de proteína, um QTL (Figura 14) que não alcançou o nível de significância de  $P < 0,05$  e  $F = 1,68$  foi identificado na posição 58 cM. Com as famílias 2, 3, 7 e 13, o mesmo QTL foi mapeado, com  $P < 0,05$  e  $F = 4,57$ , na posição 58 cM, próximo ao marcador MNB-208 (60,2 cM). Nessa região Ron *et al.* (2001) detectaram um QTL na posição 56 cM, próximo ao marcador BM143 (53,72 cM); Olsen *et al.* (2002) identificaram um QTL na posição 58 cM; Freyer *et al.* (2002) detectaram um QTL na posição 58 cM e Ashwell *et al.* (2002) encontraram um QTL na mesma região (58 cM), próximo ao marcador BMS518 (58,96 cM).



**Figura 14** - Valor de F para porcentagem de proteína considerando todas as famílias. As setas indicam a posição dos marcadores DIK4482 (54,51 cM) e MNB-208 (60,21 cM). — 5% e ---- 1% de significância a nível cromossômico.

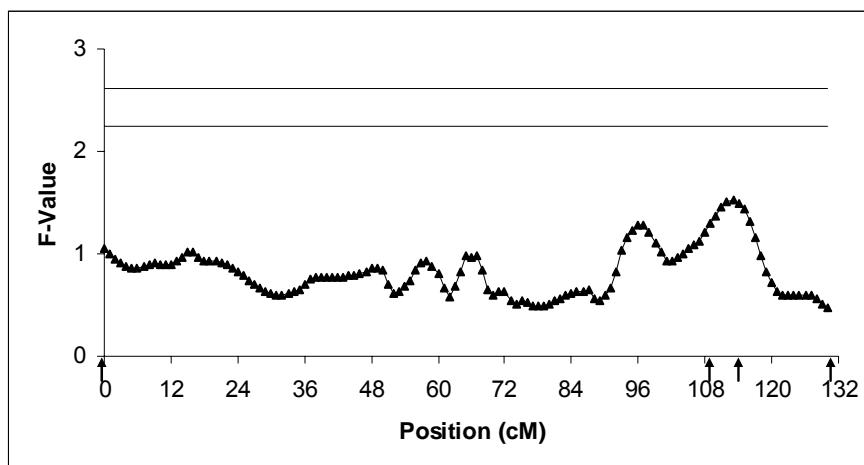
Igual à porcentagem de proteína, na característica porcentagem de lactose, não foi identificada QTL com  $P < 0,05$  em nenhuma família. A curva do conteúdo de informação considerando todas as famílias (Figura 15), ao longo do cromossomo 6, apresentou valores acima de 0,7, exceto na região compreendida entre os marcadores DIK4482 (54,51 M) e DIK3026 (71,56 cM), cujo conteúdo de informação foi inferior a 0,7.



**Figura 15** - Conteúdo de informação considerando todas as famílias para percentagem de lactose ( $\blacktriangle$ ). As setas indicam a posição dos marcadores ILST093 (0,0 cM), DIK1182 (115,32 cM) e DIK4992 (130,78 cM).

Não foram identificados QTL para percentagem de lactose com  $P < 0,05$ , ao considerar todas as famílias (Figura 16), nem ao fazer a análise em cada família, embora ao utilizar todas as famílias o maior pico foi localizado na posição 113 cM com  $F = 1,51$ . Com as famílias 4, 6, 7 e 8 um QTL com  $P < 0,01$  e  $F = 5,29$  foi identificado na posição 113 cM, próximo ao marcador DIK1182 (115,32 cM). Para produção de lactose a literatura não apresenta mapeamento ou identificação de QTL.

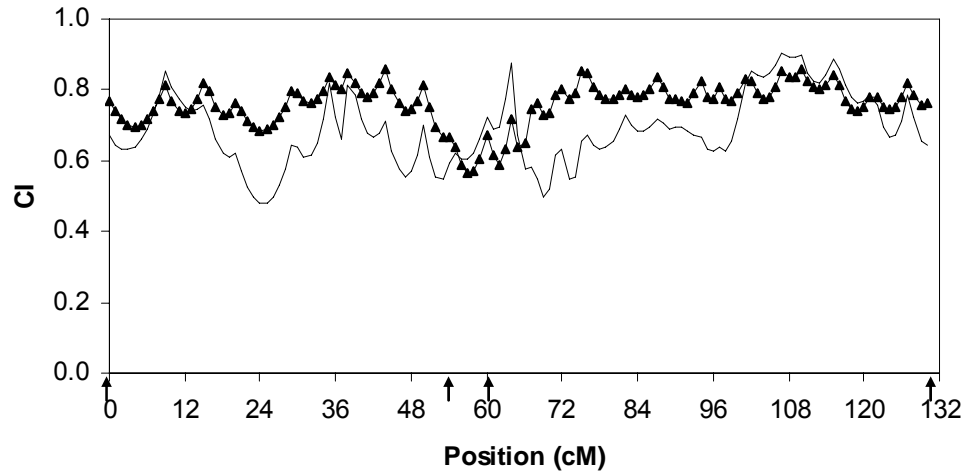




Todas as famílias

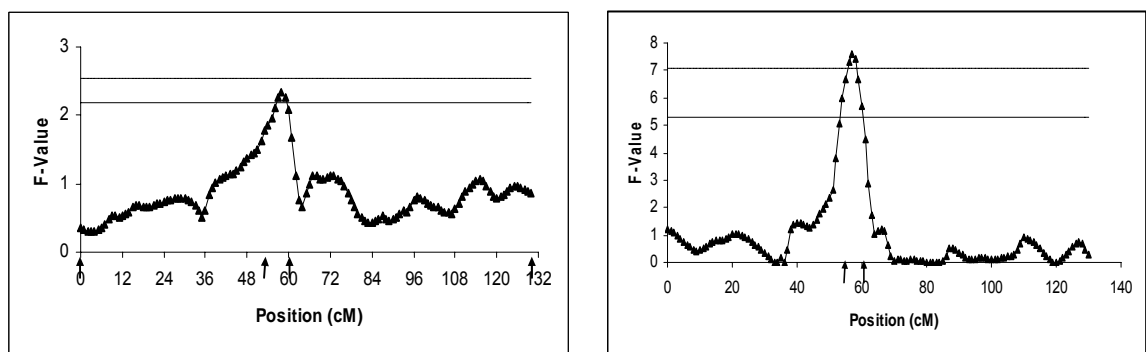
**Figura 16** - Valor de F para porcentagem de lactose considerando todas as famílias. As setas indicam a posição dos marcadores DIK2995 (109,92 cM) e DIK1182 (115,32 cM) no cromossomo 6. — 5% e ---- 1% de significância a nível cromossômico.

O conteúdo de informação para porcentagem de sólidos totais, considerando todas as famílias, encontra-se na figura 17. Em termos gerais, a curva de conteúdo de informação considerando todas as famílias foi mais estável ao longo do cromossomo 6. Para porcentagem de sólidos totais, as famílias mais significativas foram 3 e 4, cuja curva de conteúdo de informação só foi superior nas regiões compreendidas pelos marcadores DIK4482 (54,51 cM) e MNB-208 (60,21 cM), onde posteriormente foi encontrado um QTL significativo, igual ao da região flanqueada pelos marcadores DIK2174 (101,4 cM) e DIK4827 (107,12 cM).



**Figura 17** - Conteúdo de informação (CI) considerando todas as famílias (▲) e nas famílias 3 e 4 para porcentagem de sólidos totais. As setas indicam a posição dos marcadores DIK4482 (54,51 cM) e MNB-208 (60,21 cM).

Considerando todas as famílias para porcentagem de sólidos totais, um QTL foi identificado com  $P < 0,05$  e  $F = 2,34$  na posição 58 cM, próximo ao marcador MNB-208 (60,21 cM). Ao empregar as famílias 3 e 4, as mais significativas, o mesmo QTL foi identificado com  $P < 0,01$  e  $F = 7,61$ , mas na posição 58 cM, próximo ao marcador MNB-208.



Todas as famílias

Famílias 3 e 4

**Figura 18** - Valor de F para porcentagem de sólidos totais nas famílias 3 e 4. As setas indicam a posição dos marcadores DIK4482 (54,51 cM) e MNB-208 (60,21 cM) no cromossomo 6. — 5% e ---- 1% de significância a nível cromossômico.

A localização dos QTL com significância  $P < 0,01$  cromossômico encontram-se na tabela 16, o mesmo que a análise ANOVA para as famílias mais significativas. Para produção de leite considerando todas as famílias foi identificado um QTL ( $P < 0,01$ ), com  $F = 2,66$  na posição 44 cM, com as famílias 3 e 9 foi obtido o mesmo QTL na posição 46 cM, mas com um  $F$  maior (9,64). Para produção de gordura considerando todas as famílias foi identificado um QTL na posição 46 cM com  $P < 0,01$  e  $F = 3,03$ , e com a família 3, o mesmo QTL foi mapeado na posição 48 cM, mas com  $F = 16,61$ . Para a percentagem de gordura conta todas as famílias, um QTL foi identificado na posição 53 cM, com  $P < 0,02$  e  $F = 3,76$  e, com as famílias 3 e 4 o mesmo QTL foi identificado na posição 56 cM, mas com  $F = 15,18$ . Para porcentagem de lactose considerando todas as famílias, um QTL foi mapeado na posição 58 cM, com  $P < 0,01$  e  $F = 2,34$  e, com as famílias 3 e 4 o mesmo QTL foi identificado na posição 57 cM, mas com  $F = 7,61$ .

A família 3 foi significativa para os QTL identificados com  $P < 0,01$  cromossômico. Da mesma forma as famílias 3 e 4 foram significativas na detecção de QTL para porcentagem de gordura e sólidos totais. Os QTL para produção de leite e gordura estavam localizados na mesma região cromossômica, entre os marcadores BMS2508 e DIK4382. O marcador DIK4482 foi o loco mais próximo para porcentagem de gordura e sólidos totais. Produção e porcentagem de gordura apresentaram os maiores valores de  $F$ , 16,61 e 15,18, respectivamente.

**Tabela 16** - Posição dos QTL com seus marcadores próximos, valor F, limiar com significância de 1% cromossômico, famílias significativas após análise considerando todas as famílias, efeito de substituição do alelo do QTL para cada família, desvio-padrão e valor absoluto de t.

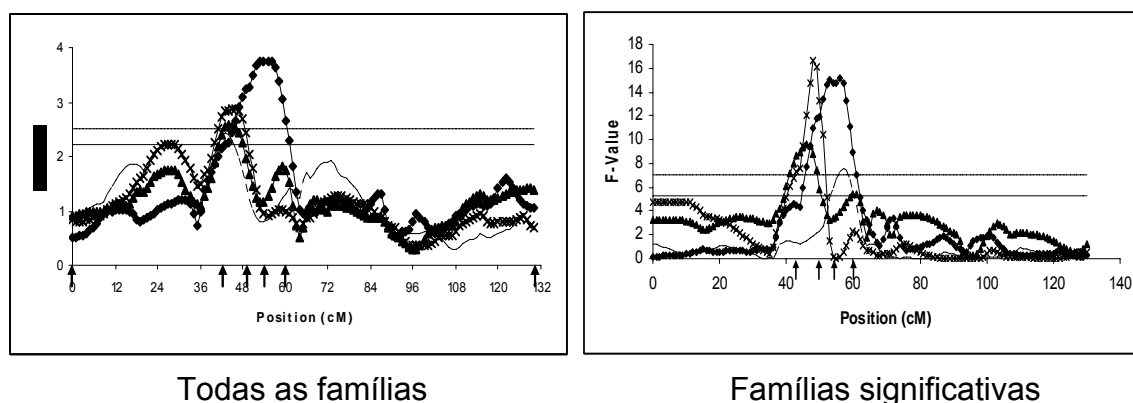
Característica	Posição (cM)	Marcador próximo	Valor F <sup>a</sup>	Limiar	Família	Efeito <sup>b</sup>	DP <sup>c</sup>	ABS(t) <sup>d</sup>
PL	46	BMS2508	9,64	7,95	3(224)	385,51	120,23	3,206
					9(527)	-400,52	133,49	3,00
PG	48	DIK4382	16,61	12,25	3(224)	19,56	4,801	4,075
%G	56	DIK4482	15,18	7,425	3(224)	0,066	0,0196	3,365
					4(293)	0,168	0,0387	4,361
%S	57	DIK4482	7,61	7,075	3(224)	-0,14	0,053	2,612
					4(293)	0,29	0,101	2,869

PL = Produção de leite; PG = Produção de gordura; %G = Porcentagem de gordura; %S = Porcentagem de sólidos totais. <sup>a</sup> Efeito de significância do alelo marcador paterno sobre os respectivos touros heterozigotos. <sup>b</sup> Efeito de substituição. <sup>c</sup> Desvio padrão. <sup>d</sup> Efeito de significância dentro de famílias do alelo marcador paterno.

A família 3 foi significativa para os QTL identificados com  $P < 0,01$ . Da mesma maneira, as famílias 3 e 4 foram significativas na detecção de QTL para porcentagem de gordura e sólidos totais. Os QTL para produção de leite e gordura estavam localizados na mesma região cromossômica, entre os marcadores BMS2508 e DIK4382. O marcador DIK4482 foi o loco mais próximo para porcentagem de gordura e sólidos totais. Produção e porcentagem de gordura apresentaram os maiores valores de F, 16,61 e 15,18, respectivamente (Figura 19).

Como os QTL identificados para produção de gordura e leite, porcentagem de gordura e lactose estiveram nas mesmas regiões cromossômicas, isso pode ser explicado pela correlação entre duas características, que medem a tendência de variarem numa mesma direção (correlação positiva) ou em direção oposta (correlação negativa). A correlação

negativa entre a produção de leite e porcentagem de gordura, por exemplo, dificulta a seleção de vacas para o aumento simultâneo da produção de leite e alta porcentagem de gordura. Ao selecionar vacas para alta porcentagem de gordura e sólidos totais no leite, se teria animais com menor produção de leite. A correlação próxima a zero indica que duas características não são relacionadas, ou seja, são independentes.



**Figura 19** - Valor de F para as características porcentagem de gordura ( $\diamond$ ), produção de gordura (x), produção de leite ( $\blacktriangle$ ) e porcentagem de sólidos totais (----) com todas as famílias e as mais significativas. As setas indicam a posição dos marcadores BMS2508 (43,93 cM), DIK4382 (50,09 cM), DIK4482 (54,51 cM) e MNB-208 (60,21 cM). — 5% e ----1% cromossômico.

Na tabela 17 encontram-se as posições dos QTL identificados com limiar de 5% de significância para as características de produção de proteína, lactose e sólidos totais. Para produção de proteína o QTL na posição 44 cM e com  $F = 2,51$  ao considerar todas as famílias, não alcançou o nível de  $P < 0,05$  cromossômico, mas com a família 9 o mesmo QTL foi localizado na posição 43 cM, com  $P < 0,05$  e  $F = 9,74$ . Para produção de lactose um QTL considerando todas as famílias foi mapeado na posição 26 cM, com  $P < 0,05$  e  $F = 3$ , e ao considerar a família 8, outro QTL foi localizado na posição 41 cM, com  $P < 0,05$  e  $F = 10$ . Para a característica produção de sólidos totais, ao considerar todas as famílias, um QTL foi mapeado na posição 43 cM, com  $P < 0,05$  e  $F = 2,59$ , e ao ter em conta só a família 8, outro QTL foi identificado na posição 16 cM, com  $P < 0,05$  e  $F = 9,88$ .

Somente a família 8 foi significativa para a identificação de QTL com  $P < 0,05$ , apresentando valores de F muito similares: Produção de lactose apresentou o maior valor de  $F = 10$ , seguido de produção de sólidos totais (9,88) e por último produção de lactose (9,74).

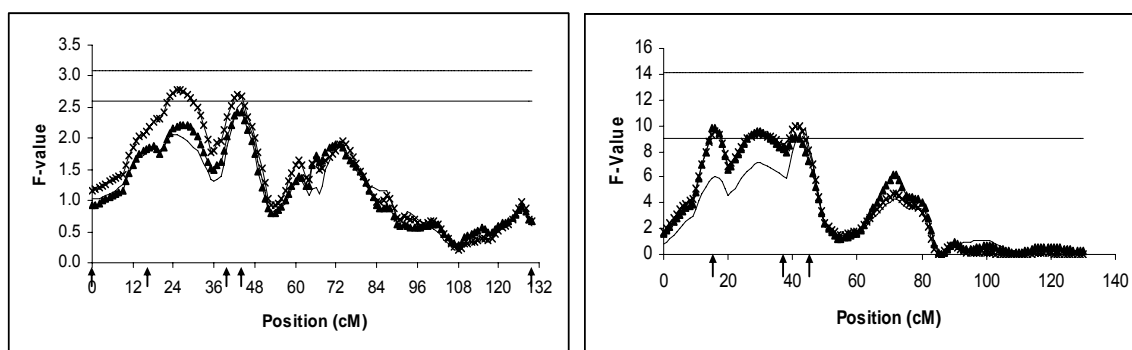
**Tabela 17** - Posição dos QTL com seus marcadores próximos, valor F, limiar (5%), famílias significativas após a análise considerando todas as famílias, efeito de substituição do alelo do QTL para cada família, desvio-padrão e valor absoluto de t.

Característica	Posição (cM)	Marcador próximo	Valor F <sup>a</sup>	Limiar	Família	Efeito <sup>b</sup>	DP <sup>c</sup>	ABS(t) <sup>d</sup>
PP	43	BMS2508	9,74	9,049	8(498)	-12,62	4,042	3,12
PLac	41	BMS2508	10	9,21	8(498)	-15,32	4,84	3,16
PS	16	DIK5285	9,88	9,24	8(498)	-42,16	13,41	3,14

PP = Produção de proteína; PLac = Produção de lactose; PS = Produção de sólidos totais.

<sup>a</sup> Efeito de significância do alelo marcador paterno computado sobre os respectivos touros heterozigotos. <sup>b</sup> Efeito de substituição. <sup>c</sup> Desvio padrão. <sup>d</sup> Efeito de significância dentro de família do alelo marcador paterno.

Os QTL para produção de proteína e lactose ficaram localizados na mesma região cromossômica, flanqueada pelos marcadores DIK1058 (38,16 cM) e BMS2508 (43,93 cM). O QTL para produção de sólidos totais estava localizado na posição 16 cM, embora outro QTL no mesmo grupo de ligação fosse localizado na posição 41 cM, tanto na família 8 como considerando todas as famílias. Ao se considerar outro QTL para produção de sólidos totais na posição 41 cM, estes três QTL estariam na mesma região cromossômica, próximos ao mesmo marcador BMS2508, e novamente essa proximidade na localização pode ser explicada pela correlação positiva entre elas.



Todas as famílias

Família 8

**Figura 20** - Valor de F para produção de proteína (—), produção de lactose (x) e produção de sólidos totais (▲) com todas as famílias e a família 8. As setas indicam a posição dos marcadores DIK5285 (15,36 cM), DIK1058 (38,16 cM) e BMS2508 (43,93 cM). — 5% e --- 1% de significância a nível cromossômico.

Na tabela 18 encontra-se o resumo dos QTL identificados sobre o cromossomo 6 com nível de  $P < 0,01$  cromossômico. Os QTL para as características de produção de leite e gordura encontram-se entre os marcadores BMS2508 (43,93 cM) e DIK4382 (50,09 cM), o mesmo que os QTL para porcentagem de gordura e sólidos totais, entre os marcadores DIK4482 (54,51 cM) e MNB-208 (60,21 cM). Isto pode ser explicado pela correlação positiva que existe entre as características de produção do leite. A principal causa de correlação genética é a pleiotropia, que é a propriedade de um gene afetar dois ou mais caracteres, de tal maneira que se um gene for segregado causará variação simultânea na característica que ele afeta.

Na mesma tabela estão as proporções das variações fenotípicas explicadas por cada QTL. Para produção de leite (PL) o QTL com  $P < 0,01$  e  $F = 9,64$  identificado explicou 0,62 da variação fenotípica. O QTL ( $P < 0,01$ ) com  $F = 16,61$  para produção de gordura (PG) explicou a maior variação fenotípica (0,75). Para a característica porcentagem de gordura (%G) o QTL ( $P < 0,01$ ) e  $F = 15,18$  identificado, explicou 0,70 da variação fenotípica. O QTL ( $P < 0,01$ ) e  $F = 7,61$  para porcentagem de sólidos totais (%S), explicou a menor variação fenotípica dos QTL identificados (0,37). Os valores para os QTL identificados com nível de  $P < 0,01$ , e que explicaram a variação fenotípica, ficaram dentro

do intervalo (0,2-0,6) para ser considerados como significativos, de acordo com os estudos integrais de Khaktar *et al.* (2004) e Smaragdov (2006).

**Tabela 18** - Resumo das análises de associação para os QTL que foram identificados no cromossomo 6 com nível de  $P < 0,01$  cromossômico, indicando o valor F, posição em que foi detectado o QTL, intervalo de confiança (I.C) e a porcentagem de variação fenotípica explicada pelo QTL encontrado.

Característica	Posição (cM)	Valor de F	I.C- 95% (cM)	$h^2_{QTL}$
PL	46	9,64	41-102	0,62(0,14)
PG	48	16,61	5 - 75	0,75 (0,17)
%G	56	15,18	48 - 59,5	0,70 (0,12)
%S	57	7,61	2 - 112	0,37 (0,16)

$h^2_{QTL}$  é a variação fenotípica explicada pelo QTL encontrado; entre parênteses a variação fenotípica explicada pelo QTL quando foram utilizadas todas as famílias.

Dadas as correlações altas e positivas entre as características de produção, e a localização dos QTL para a maioria destas características na mesma região cromossômica, pode-se estar frente a um dilema: é um único QTL pleiotrópico sobre esta região ou vários QTL particulares? Ron *et al.* (2001) apresentaram evidências para um único QTL afetando produção de leite, produção de proteína, produção de gordura, porcentagem de proteína e porcentagem de gordura no BTA6, na posição 48 cM em diferentes famílias. Freyer *et al.* (2003) reportaram um QTL no BTA6 na posição 58 cM para produção de proteína e porcentagem de gordura. Lipkin *et al.* (2002) estimaram no genoma de bovinos leiteiros aproximadamente 40 QTL com efeitos pleiotrópicos sobre produção de leite, produção e porcentagem de proteína.

Na expressão de QTL pleiotrópicos, podem ser indiretamente estabelecidos pelas correlações positivas e negativas que existem entre diferentes características de produção e conteúdo no leite. Por exemplo, QTL afetando a produção de leite com aumento (diminuição) nas características de porcentagem no leite, provavelmente afeta fatores osmóticos, tais como



lactose, íons de cálcio ou receptores que regulam o transporte molecular através da membrana celular. Esses fatores podem incrementar (diminuir) a pressão osmótica, a qual garante um incremento (diminuição) do conteúdo de água em leite e, dessa maneira, aumenta (diminui) a porcentagem de gordura ou proteína no leite.

Dada a significância e quantidade de QTL e genes candidatos encontrados e relacionados com as características do leite, Khaktar *et al.* (2004) e Samaragdov (2006), com metodologias como meta-análise, realizaram estudos globais no genoma e em cada cromossomo, sendo o BTA6 um dos mais representativos. De acordo com estes estudos no cromossomo 6 existem três regiões cromossômicas, nas quais podem ser encontrados QTL: uma primeira região está localizada na parte pericentromérica do BTA6 onde foram identificados QTL para %G e %P na posição  $7 \pm 4$  cM; em uma segunda região foram detectados QTL para PL, PP e %P no intervalo 28 - 35 cM e, na terceira região, bem definida, Ron *et al.* (2001) identificaram QTL pleiotrópicos para diferentes características no leite na raça Holandesa ( $50 \pm 4$  cM) e Olsen *et al.* (2004) na raça Norway ( $46 \pm 4$  cM).

As análises com dados totais confirmam que esta região ancora fortes QTL para PL (40 cM); PL, PP, PG, %P e %G (48 cM), e QTL para PL, PP, PG, %P e %G (58-62 cM). Assim, de acordo com os anteriores estudos o BTA6 carrega 6 QTL pleiotrópicos para PL, PP, PG, %P e %G nas regiões 7, 30-35, 48, 58-60, 75-78 e 90-100 cM. As regiões estimadas com a metodologia de meta-análise para características de produção e conteúdo no leite encontraram QTL altamente significativos: para PL ( $49, 5 \pm 5$  cM;  $86,7 \pm 8$  cM); PP ( $52 \pm 7$  cM); %P ( $49, 4 \pm 1,8$  cM;  $91,1 \pm 7,6$  cM); PG ( $51 \pm 6$  cM); %G ( $48,1 \pm 2,8$  cM;  $113 \pm 15$  cM).

O aumento da população, a procura de marcadores informativos heterozigotos para todos os touros participantes no delineamento de filhas, ou saturando ainda mais algumas regiões sobre o cromossomo 6 poderá aumentar o valor de F e diminuir o nível de significância, no caso de QTL mapeados com nível de significância  $P < 0,05$ , ou para QTL que não alcançaram este nível de significância, como é o caso das características de porcentagem de proteína e lactose

. A existência de dois ou mais QTL dentro de um mesmo grupo de ligação, com efeitos contrários, pode identificar incorretamente a presença de

QTL. Com o desenvolvimento dos mapas comparativos da espécie humana, camundongo e bovino (Itoh *et al.*, 2005) pode ser procurados segmentos de sintonia entre humanos e bovinos, por exemplo, para identificar aquelas regiões em que os QTL foram localizados e procurar possíveis genes candidatos dentro dessas regiões.

Diferentes estudos de identificação de QTL relacionados com características de produção no leite mostram que a localização de QTL identificados é aproximada, sendo uma alternativa fazer um mapeamento de precisão, utilizando marcadores SNP. Um primeiro passo já foi dado por Blott *et al.* (2003) e Olsen *et al.* (2004) mapeando genes candidatos com SNP.

Um QTL de efeito maior associado com produção de proteína e porcentagem de gordura foi localizado em um intervalo de 4 cM na posição 55 cM próximo ao marcador BM143 (53,724 cM). Nessa região o gene candidato FAM13A1 pertencente à família FAM13A que faz parte dos genes que codificam proteínas da matriz extracelular (ECM), foi determinado por mapeamento comparativo (Cohen *et al.*, 2004) com regiões ortólogas no cromossomo 4 em humanos, e relacionado com proteínas da matriz extracelular, envolvido na formação do osso e estruturas lóbulo-alveolar da glândula mamária. Um QTL altamente significativo para porcentagem de gordura na posição 56 cM próximo ao marcador DIK4482 (54,51 cM) foi identificado no presente trabalho, e um QTL para porcentagem de sólidos totais na posição 57 cM.

Em gado leiteiro Norueguês, Olsen *et al.* (2004) saturaram a região compreendida entre os marcadores BMS2508 (43,93 cM) e FBN12 (51,43 cM), encontrando um QTL altamente significativo, causando uma redução em %G e %P, como também um incremento menor na PL. Nessa mesma região identificaram um gene candidato, policistina 2 (PKD2), que codifica para proteínas de membrana, envolvendo interações célula-célula/matriz, interagindo com o genes PKD1, participando na homeostase intracelular de cálcio e em outros sinais de tradução (Nauli *et al.*, 2003). Diferenças na expressão de PKD2 podem alterar o conteúdo de água no leite e, assim reduzir a porcentagem de proteína e gordura no leite. Nessa região foram identificados QTL com  $P < 0,01$  para produção de leite e gordura, na população Gir Leiteiro.

Weikard *et al.* (2005) reportaram uma associação significativa entre um SNP no íntron 9 com o gene PPARGC1A (Peroxisome Proliferator Activated

Receptor Gamma, Coativador 1, Alpha), também conhecido como PGC1 $\alpha$ , próximo ao marcador BM143 e que tem um papel importante no metabolismo de energia, gordura e glicose. Segundo os pesquisadores, este gene está organizado em 13 éxons compreendendo 6.261 pb expressados em diferentes tecidos. Eles determinaram que o cDNA do PPARGC1A e sequências de proteínas mostram similaridade substancial de 92-95% com os respectivos genes ortólogos em humanos, rato e camundongo.

Leonard *et al.* (2005) e Schnabel *et al.* (2005) confirmaram uma associação entre o gene da osteopontina (OPN) e percentagem de proteína em uma população de Holandesa Americana. O gene de OPN é um forte candidato funcional para QTL. OPN é uma glicoproteína que tem funções na mediação das interações de matriz celular e sinalamento celular, completamente ligado com integrinas e receptores CD44 e expressado em diferentes tecidos, sendo necessário para o normal desenvolvimento da glândula mamária e lactação em roedores. Em um modelo de camundongo transgênico foi possível evidenciar o papel do OPN no desenvolvimento da glândula mamária e na lactação. Snabel *et al.* (2005) apresentaram evidências do OPN bovino com efeitos significativos sobre PL, PG e PP.

Outro gene importante que se encontra nesta região crítica é ABCG2 (ATP Binding Cassette Sub Family G Member 2), que afeta principalmente %P e %G. Jonker *et al.* (2005) demonstraram que o ABCG2 é responsável pela secreção ativa dos substratos clínico e toxicologicamente importantes no leite de camundongo. O ABCG2 em *Bos taurus* está associado com o incremento de produção de gordura, porcentagem de gordura e proteína, como também uma diminuição na produção do leite e proteína.

## 5. CONCLUSÕES

- Foram identificados QTL com nível de significância  $P < 0,01$  para produção de leite, produção de gordura, porcentagem de gordura e porcentagem de sólidos totais. Com nível de  $P < 0,05$  foram mapeados QTL para produção de proteína, produção de lactose e produção de sólidos totais.
- Embora exista uma grande diversidade genética observada entre raças taurinas e indianas, os QTL encontrados no cromossomo 6 da raça Gir relacionados com características de produção de leite são conservados, o que prova indiretamente que alelos de genes candidatos relacionados com essas características de produção no mesmo cromossomo foram fixados.
- Para as características de produção de leite, gordura, proteína e sólidos totais foram identificados QTL na mesma região cromossômica. QTL para porcentagem de gordura e sólidos totais foram identificados na mesma região cromossômica, demonstrando a alta correlação entre estas características.
- Os QTL para porcentagem de gordura identificado na posição 56 cM explicou 0,70 da variação fenotípica, e para produção de gordura na posição 48 cM explicou 0,75 da variação fenotípica.
- Os QTL mapeados foram encontrados em uma região considerada crítica pela presença de QTL pleitrópicos e genes candidatos como ABCG2, LAP3, HCAP-6, PKD2, FAM13A1, PPARGC1A e OPN.
- Foram identificados novos marcadores microssatélites com alta  $H_o$  e PIC, para ser empregados em estudos de identificação de QTL, identidade genética e teste de paternidade, principalmente nas raças zebuínas que requerem ser mais exploradas com a genética molecular.

## 6. REFERENCIAS BIBLIOGRAFICAS

- ASHWELL, M. S.; HEYEN, D. W.; SONSTEGARD, T. S.; VAN, C. P., and TASSELL, Y. Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. **Journal Dairy Science**, 87: 468-475, 2004.
- ASHWELL, M. S.; VAN TASSELL, C. P.; SONSTEGARD, T. S. A genome scan to identify quantitative trait loci affecting economically important traits in a US Holstein population. **Journal of Dairy Science**, 84:2535-2542, 2002.
- BARENDSE, W.; ARMITAGE, S. M.; CLAYTON, D.; LI, L.; NEIBERGS, H.; NAN Z.; GROSSE, M.; CREIGHTON, P.; MCCARTHY, F.; RON, M.; SOLLER, M.; FRIES, R.; MACGRAW, R. A.; MOORE, S. S.; TEALE, A.; GEORGES, M.; WOMACK, J.E.; HETZEL, D.J.S. A preliminary map of the bovine genome. **Nature Genetics** 6, 227-235, 1994.
- BECKMANN, J. S., and SOLLER, M. Molecular markers in animal genetics improvement. **Biotechnology** 5, 573-576, 1987.
- BENNEWITZ, J.; REINSCH, N.; GROHS, C. Combined Analysis of data two grand daughter designs: A simple strategy for QTL confirmation and increasing experimental power in dairy cattle. **Genet. Sel. Evol.**, v. 35, p. 319-338, 2003.
- BISHOP, M. D.; KAPPES, S. M.; KEELE, J. W.; STONE, R.T.; SUNDEN, S. L. F.; HAWKINS, G. A.; SOLINAS TOLDO, S.; FRIES, R.; CROSZ, M. D.; YOO, J.; BEATTIE, C. W. A genetic linkage map for cattle. **Genetics** 136, 619-625, 1994.
- BLOTT, S.; KIM, J. J.; MOISIO, S. Molecular dissection of a QTL: A phenylalanine-to-Tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effects on milk yield composition. **Genetic**, v. 163, p. 253-266, 2003.
- BOICHARD, D., and BISHOP, M. P. detection of QTL influencing milk production and mastitis resistance with a granddaughter design in Holstein cattle, paper presented at 48<sup>th</sup> annual meeting of European Association for Animal Production, Vienna, Austria, 1997.
- BOICHARD, D.; GROHS, C.; BOURGEORS, F.; CERQUEIRA, F.; FAUGERAS, R.; NEAU, A.; RUPP, R.; AMIGUES, Y.; BOSCHER, M.; LEVÉZIEL, H. Detection of genes influencing economic traits in free French dairy cattle breeds. **Genet. Sel. Evol.** 35, 77-101, 2003.
- BOLDMAN, K. G.; KRIESE, L. A.; VAN VLECK, L. D.; KACHMAN, S. D. **A Manual for use MTDFFEML**; a set of programs to obtain estimates of variances and

covariances (DRAFT). Lincoln, Department of Agriculture/Agricultural Research Service, 1993. 120p.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DARVIS, R. W. Construction of genetic linkage map in man using restriction fragment length polymorphisms. **The American Journal of Human Genetics** 32, 314-331, 1980.

BRASCAMP, E. W.; VAN ARENDONK, J. A. M.; GROEN, A. F. Economic appraisal of the utilization of genetic markers in dairy cattle breeding. *Journal of Dairy Science*, v. 76, n° 4, 1204-1213, 1993.

CHEN, H. Y.; ZHANG, Q.; YIN, C. C.; WANG, C. K.; GONG, W. J.; MEI, G. Detection of Quantitative Trait Loci Affecting Milk Production Traits on Bovine Chromosome 6 in a Chinese Holstein Population by the Daughter Design. **Journal Animal Science**, 89: 782-790, 2006.

CHIN, E. C. L.; SENIOR, M.L.; SHU, H.; SMITH, J. C. S. Maize simple repetitive DNA sequences: abundance and allele variation. **Genome** 39, 866-873, 1996.

CHURCHILL, G. A.; DOERGE, R. W. Empirical threshold values for quantitative trait mapping. **Genetics** 138, 963-971, 1994.

COHEN, M.; REICHESTEIN, M.; EVERTS-VAN DER WIND, A.; HEON-LEE, J.; SHANI, M.; LEWIN, H. A.; WELLER, J. I.; RON, M.; SEROUSSI, E. Cloning and characterization of FAM13A1 - A gene near a milk protein QTL on BTA6: Evidence for population-wide linkage disequilibrium in Israeli Holsteins. *Genomics* 84, 374-383, 2004.

DARVASI A.; WEINREB A.; MINKE, V. *et al.* Detecting marker-QTL linkage and estimating QTL gene effect and map location using saturated genetic map. **Genetics**, v. 134, n. 3, pp. 943-51, 1993.

DUDLEY, J. W. Linkage disequilibrium in crosses between Illinois maize strains divergently selected for protein percentage. **Theoretical and Applied Genetics**. 87, 1016-1020, 1994.

EDWARDS, C. J.; MACHUGH, D. E.; DOBNEY, K. M.; MARTIN, L.; RUSSEL, N.; HORWITZ, L. K.; MCINTOSH, S. K.; MACDONALD, K. C.; HELMER, D.; TRESSET, A.; VIGNE, J. D.; BRADLEY, D. G. Ancient DNA Analysis of cattle remains: limits and prospect. **Journal of Archaeological Science**, v. 31, p. 695-710, 2004.

ELLEGREN, H. Genome Analysis with microsatellites markers. Department of Animal Breeding and Genetics. **Swedish University of Agriculture Science**, 1993.

ESTADÍSTICAS In: Anualpec 2004. **Anuário de pecuária Brasileira**. São Paulo: FNP - Consultoria e Agroinformática, p. 63, 2004.

EVERETT, R.W. and CARTER, H.W. Accuracy of test interval method of calculating dairy herd improvement association records. **Journal Dairy Science**, v. 51, n° 12, p. 1931LSTS093941, 1968.

- EVERETT, R. W.; KEOWN, J. F.; CLAPP, E. E.; Relationships among type, production and stability in Holstein cattle. **Journal of Dairy Science**, Champaign, v. 59, p. 1505-1510, 1976.
- FALCONER, D.S. **Introduction to quantitative genetics**. 2 ed. New York: Logman, 1981.
- FALCONER, D. S.; MACKAY, T. F. C. **Introduction to quantitative genetics**. Harlow: Longman Group, 464p., 1996.
- FERRAZ, J. B. S.; ELER, J. P.; GOLDEN, B. I. Análise genética do composto Montana tropical. **Ver. Brs. Reprod. Animal**, v. 23, p. 111-113, 1999.
- FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3ed. Brasília: Embrapa, CENARGEN, 220p, 1998.
- FREYER, G.; KUHN, C.; WEIKARD, R. Multiple QTL on chromosome six in dairy cattle affecting yield and content traits. **Journal Animal Breeding Genetic**, v. 119, p. 69-82, 2002.
- FREYER, G.; SORENSEN, P.; KÜHN, C.; WEIKARD, R.; HOESCHELET, I. Search for Pleiotropic on chromosome BTA6 affecting yield traits of milk production. **Journal Dairy Science**, v. 86, n. 3, 2003.
- GEORGES, M., D.; NIELSEN, M.; MACKINNON, A.; MISHRA, R.; OKIMOTO, A. T.; PASQUINO, L. S.; SARGEANT, A.; SORENSEN, M. R.; STEELE, X.; ZHAO, J.; WOMACK, E., and HOESCHELE, I. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. **Genetics**, 139:907-920, 1995.
- GIANOLA, D., and FOULLEY, J. L. Variance estimation from integrated likelihood (VEIL). **Génétique, Sélection, Évolution**, Paris, 22: 403-417, 1990.
- GREEN, P.; FALLS, K.; CROOKS, S. **CRIMAP Documentation** version 2.4 St. Louis: Washington University School of Medicine, 1990.
- HALEY, C. S. and KNOTT, S. A. A simple regression method for mapping quantitative loci in line crosses using flanking markers. **Heredity** 69, 315-324, 1992.
- HARTLEY, H.O.; RAO, J. N. K. Maximum-likelihood estimation for the mixed analysis of variance model. **Biometrika**, London, 54 (1): 93 -108, 1967.
- HENDERSON, C.R. Estimation of variance and covariance components. **Biometrics**, Washington, 9: 22-52, June 1953.
- IHARA, N.; TAKASUGA, A.; MIZOSHITA, K.; TAKEDA, H.; SUGIMOTO, M.; MIZOGUCHI, Y. A Comprehensive Genetic Map of the Cattle Genome based on 3802 Microsatellites. **Genome Research**, p. 1987-1998, 2004.

- IKONEN, T.; BOVENHUIS, H.; OJALA, M.; RUOTTINEN, O.; GEORGES, M. Association between casein haplotypes and first lactation milk production traits in Finnish Ayrshire cows, **Journal Dairy Sciences**, 84: 507-514, 2001.
- ITOH, T.; WATANAB, T.; IHARA, N.; MARIANI, P.; BEATTIE, C. W. A Comprehensive Radiation Hybrid Map of the Bovine Genome Comprising 5593 loci. **Genomics**, 85: 413-419, 2005.
- JONKER, J. W.; MERINO, G.; MUSTERS, S.; HERWAARDEN, E.; BOLSCHER, E. WAGENAAR, E.; MESMAN, E.; DALE, T. C.; SCHINKEL, H. The breast cancer resistance protein BCRP (ABCG2) concentrates drugs and carcinogenic xenotoxins into milk. **Nat. Med.** 11: 127, 129, 2005.
- KAPPES, S. M.; KEELE, J. W.; STONE, R.T. *et al.* A second-generation linkage map of the bovine genome. **Genome Research**, v. 7, pp. 36-38, 1997.
- KHATKAR, M. S.; THOMSON, P. C.; TAMMEN, I.; RAADSMA, H. W. Quantitative trait loci mapping in dairy cattle: review and meta-analysis, **Genetics. Selection. Evolution**. 36: 163-190, 2004.
- KNOTT, S. A.; ELSEN, J. M.; HALEY, C. S. Methods for multiple-marker mapping of quantitative trait loci in half sib populations. **Theor App Genet**, 93: 71-80, 1996.
- KÜHN, C.; FREYER, G.; WEIKARD, R.; GOLDAMMER, T., and SCHWERIN, M. Detection of QTL for milk production traits in cattle by application of a specifically developed marker map of BTA6. **Animal Genetics**, 30:333-340, 1999.
- LANDER, E. S.; BOLSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics** 121, 185-199, 1989.
- LEDIC, I. L.; FERNANDES, A. R. **Cartilha para avaliação morfológica de animais**. Uberaba, MG: ABCGIL, 24p, 2005.
- LEE, H.; DEKKERS, J. C.; SOLLER, M. *et al.* Application of the false discovery rate to quantitative trait loci interval mapping with multiple traits. **Genetics**, v. 161, n. 2, pp. 905-914, 2002.
- LEONARD, S; KHATIB, H.; SCHUTZKUS, V.; CHANG, Y. M.; MALTECCA, C. Effects of the osteopontin gene variants on milk production traits in dairy cattle. **Journal Dairy Science**, 88: 4083-4086, 2005.
- LIPKIN, E.; GROSMAN, G.; TCHOURZYNA, E. QTL affecting milk protein percent also affect milk yield and protein yield. **Proc. 7<sup>th</sup> World Congress. On Genetics Applied to Livestock Production**, Montpellier, France, 2002, p. 55-58.
- LIPKIN, E.; MOSIG, M.O.; DARVASI, A.; EZRA, E.; SHALOM, A. *et al.* In: Quantitative traits locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: analysis of milk protein percentage. **Genetics** 149: 1557-1567, 1998.



- LIU, B. H. *Statistical genomics: Linkage, mapping, and QTL analyses*: **CRC Press**. 611p. 1998.
- LIU, B. H.; KNAPP, S. J. A reference manual for G-Mendel, 1:0. **Statistical software for Mendelian segregation and linkage analysis**. Oregon State, Uni. Corvallis, 1991
- LOFTUS, R. T.; MACHUGH, D. E.; BRADLEY, D. G. *et al.* Evidence for two independent domestications of the cattle. **Proceedings of National Academy of Science**, v. 91, p. 2757-2761, 1994.
- LYNCH, M.; WALSH, B. **Genetics and Analysis of Quantitative Traits**. Sunderland, Massachusetts, U.S.A., 1998.
- MACHADO, M.A. e MARTINEZ, M.L. Acelerando o Melhoramento com o Mapeamento do genoma Bovino. **Informe Agropecuário** 22, 98-104, 2001.
- MADALENA, F. E. Estratégias de uso de recursos genéticos visando melhorar a qualidade do leite e derivados. **VII Simpósio Brasileiro de Melhoramento Animal. São Carlos, SP**, 10 e 11 de julho de 2008.
- MARSHALL, T. C.; SLATE, J.; KRUIK, L. E. B.; PEMBERTON, J. M.; Statistical confidence for likelihood-based paternity inference in natural population. **Molecular Ecology** 7, 639-655, 1998.
- MARTINEZ, M. L. Avaliação genética de touros puros e mestiços. **In: Simpósio sobre Produção Animal. Anais, Piracicaba, FEALQ**, p. 77-106, 1989.
- MARTINEZ, M.L.; VERNEQUE, R.S. Programa Nacional de Melhoramento Genético. **In: Produção e rusticidade. Gir leiteiro, a solução para os trópicos. Revista Balde Branco**, São Paulo, SP, p. 8-10, 2001.
- NADESALINGAM, J.; PLANTE, Y.; GIBSON, P. Detection of QTL for milk production on Chromosome 1 and 6 of Holstein cattle. **Mammalian Genome** 12:27-31, 2001.
- NAULI, S. M.; ALENGHAT, F.J.; LUO, Y.; WILLIAMS, E.; VASSILEY, P. Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. **Nat. Genet.** 33: 129-137, 2003.
- NEIMANN-SØRENSEN, A., and ROBERTSON, A. **Acta Agri. Scand.** 11: 163-196, 1965.
- OLSEN, H. G.; GOMEZ-RAYA, L.; VAGE, D. I. Genome scan for QTL affecting milk production in Norwegian dairy cattle. **Journal Dairy Science**, v. 85, p. 3124-3130, 2002.
- OLSEN, H. G.; LIEN, S.; SVENDSEN, M.; NILSEN, H.; ROSETH, A.; AASLAND, OPSAL M.; MEUWISSEN, T. H. G. Fine mapping of milk production QTL on BTA6 by combined linkage and linkage disequilibrium analysis. **Journal of Dairy Science** 87, 690-698, 2004.

- OLSEN, H.G.; LIEN, S.; SVENDSEN, H.; NILSEN, H.; ROSETH, A.; OPSAL, M. A.; MEUWISSEN, T. H. OLSEN, H. G.; LIEN, S.; GAUTIER, M.; NILSEN, H.; ROSETH, A.; BERG, P. R.; SUNDSAASEN, K.; SVENDSEN, M.; MEUWISSEN, T. E. Mapping of a milk production quantitative trait locus to a 420-kb Region on Bovine Chromosome 6. **Genetics**, 169: 275-283, 2005.
- OTT, J. Strategies for characterization highly polymorphic markers in human gene mapping. **American Journal of Human Genetics**, 51: 283-290, 1992.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, London, 58: 545-554, 1971.
- PÉREZ-ENCISO, M.; CLOP, A.; NOGUERA, J. L.; OVILO, C.; COLL, A.; FOLCH, J. M.; BABOT, D.; ESTANY, J.; OLIVER, M. A.; DIAZ, I.; SANCHEZ, A. A QTL on pig chromosome 4 affects fatty acid metabolism: Evidence from an Iberian by Landrace intercross. **Journal of Animal Science** 78, 2525-2531, 2000.
- PRINZENBERG, E. M.; WEIMANN, C.; BRANDT, H.; BENNEWITZ, J.; KALM, E.; SCHWERIN, M.; ARHARDT, G. Polymorphism of the bovine CSN1S1 promoter: Linkage mapping, intragenic haplotypes, and effects on milk production traits. **Journal dairy Sci.** 86: 269DIK4408705, 2003.
- QUASS, R. L. **Linear Prediction in BLUP school handbook: use of mixed models for prediction and estimation of (Co)variance components.** Animal Genetics, and Breeding Unit, University of New England, New South Wales, Australia, 76p, 1984.
- RAO, C.R. Estimation of variance and covariance components MINQUE theory, **Journal of Multivariate Analysis**, San Diego 1: 257-275, 1971a.
- RAO, C.R. Minimum variance quadratic unbiased estimation of variance components. **Journal of multivariate analysis**, San Diego, 1: 445-456, 1971b.
- RODRIGUEZ-ZAS, S. L.; SOUTHEY, B. R.; HEYEN, D. W.; LEWIN, H. A. Interval and composite interval mapping of somatic cell score, yield and components of milk in dairy cattle. **Journal Dairy Science**, v. 85, 3081-3091, 2002.
- RON, M.; BLANK, Y.; BAND, M.; EZRA, E.; WELLER, J. I. Misidentification rate in Israeli Dairy cattle population and its implications for genetic improvement. **Journal Dairy Science**, 79: 676-681, 1996.
- RON, M.; KLIGER, D.; FELDMESSER, E.; SEROUSSI, E.; EZRA, E. and WELLER, J. I. Multiple quantitative trait locus analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design. **Genetics**, 159:727-735, 2001.
- SAMBROOK and RUSSELL (2001). **Molecular Cloning: A laboratory Manual**, third edition.
- SANTIAGO, A. A. O Zebu na Índia, no Brasil e no mundo. **Instituto Campineiro de Ensino Agrícola**, p. 745, 1986.

- SANTOS, R. O Gir e o leite na pecuária fundamental. Uberaba, MG. Editora **Agropecuária Tropical**, p. 456, 2007.
- SCHNABEL, R. D.; KIM, J. J.; ASHWELL, M. S.; SONSTEGARD, T. S.; VAN TASELL, C. P.; CONNOR, E. E.; TAYLOR, J. F. Fine-mapping milk production quantitative trait loci on BTA: Analysis the bovine osteopontin gene. **Proc. Natl. Acad. Sci. USA**, 2005.
- SEATON, G.; HERNANDEZ, J.; GRUNCHEC, J. A.; WHITE, I.; ALLEN, J.; DE KONING, D. J.; WEI, W.; BERRY, D.; HALEY, C.; KNOTT, S. GridQTL: A grid portal for QTL mapping of compute intensive datasets. **Proc. 8<sup>th</sup> World Congress on Genetics Applied to Livestock Production, Belo Horizonte, MG**, Brasil, 2006.
- SHUSTER, I.; CRUZ, C. D. **Estatística Genômica** - aplicada a populações derivadas de cruzamentos controlados. Viçosa: UFV, 2004, 568 p.
- SMARAGDOV, M. G. Genetic Mapping of Loci Responsible for Milk Production Traits in Dairy Cattle. **Russian Journal of Genetics**, v. 42, n. 1, pp. 1-12, 2006.
- SCHNABEL, R. D.; KIM, J. J.; ASHWELL, M. S.; SONSTEGARD, T. S.; VAN TASELL, C. P.; CONNOR, E. E.; TAYLOR, F. F. Fine-mapping milk production quantitative trait loci on BTA6: Analysis of osteopontin gene. **Proc. Natl. Acad. Sci. USA** 102: 6896-6001, 2003.
- SPELMAN, R. J.; COPPIETERS, W.; KARIM, L.; ARENDONK, J. A. and BOVENHUIS, H. Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population. **Genetics**, 144:1799-1808, 1996.
- TEIXEIRA, N. M. Fatores não genéticos que afetam a produção do leite. In: **Melhoramento Genético de Bovinos de Leite**. Juiz de Fora, MG. Embrapa, p. 103, 2001.
- TRONCO, V. M. **Manual para inspeção da qualidade do leite**. 2 ed. - Santa Maria: Ed. da UFSM, p. 192, 2003.
- VAN der WERF, J.H.J.; BÔER, I. J. M. Estimation of genetic variance from selected base populations. **Proc. Joint. Annu. Meeting of ADSA and ASAS**, Lexington, USA, 1989.
- VELMALA, R. J.; VIKKI, K.; TANILA, A. Casein haplotypes and their association with milk production traits in Finnish Ayrshire cattle. **Animal Genetic**, 26, 419-425, 1995.
- VELMALA, R. J.; VIKKI, H. J.; ELO, K. T.; KONING, D. J.; MÄKI-TANILA, A. V. A search of quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle. **Animal Genetic**. 30: 13-26, 1999.
- VERENEQUE, R. S. **Procedimentos numéricos e estimação de componentes de variância em análise multivariada pelo método de máxima verossimilhança restrita-modelos mistos aplicados ao melhoramento**

**animal**, Dissertação (tese doutorado). Faculdade de Agronomia, Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP, 1994.

VERNEQUE, R. S.; TEODORO, R. L.; MARTINES, L. M.; **Melhoramento Genético das raças Gir e Guzerá pelo Teste de Progênie**. Juiz de Fora, MG: Embrapa-CNPGL, 1998. 28p.

VERNEQUE, R. S.; PEIXOTO, M.G.; FILHO, A.E.; MACHADO, M.A.; BARBOSA, M.V.; FERNANDES, A.R.; e MACHADO, C.H. **Programa Nacional de Melhoramento do Gir Leiteiro - Sumario Brasileiro de Touros - Resultado do Teste de Progênie**. Embrapa Gado de Leite, Juiz de Fora, MG, 2008.

VIITELA, S. M.; SCHULMAN, N. F.; KONING, D. J. QTL affecting milk production traits in Finnish Ayrshire Dairy Cattle. **Journal Dairy Science**, v. 86, p. 1828-1836, 2003.

WALSH, F. 2006. **Strengthening Family Resilience**. Guilford Press, New York. p. 372

WEIKARD, R. C.; KÜHN, C.; GOLDAMER, T.; FREYER, G.; SCHWERIN, M. The bovine PPARGC1A gene: Molecular characterization and association of a SNP with variation of milk synthesis. **Physiology Genomic**. 21:1-13, 2005.

WEINER, M. P.; HUDSON, T, J. Introduction to SNPs: Discovery of markers for disease. *BioTechniques* 32, S4-S13, 2002.

WELLER, J. L; SOLLER, M; BRODY, T. Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* x *L. pimpinellifolium*) by means of genetic markers. **Genetics** 118: 329-339, 1988.

WELLER, J. I.; KASHI, Y.; SOLLER, M, Power of “ daughter” and “granddaughter” designs for genetic mapping of quantitative traits in dairy cattle using genetic markers. **Journal of Dairy Science**, v. 73, p. 2525-2537, 1990.

WELLER, J. I., and RON, M. Estimation and detection of QTL in dairy cattle. **XXXVII Reunião Annual da SBZ, Viçosa, MG**. 2000.

ZHANG, Q.; BOICHARD, D.; HOESCHELE, I.; ERNST, C.; EGGEN, A. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbreed pedigree. **Genetics** 149; 159-173, 1998.