

**DESCOBRINDO GENES EXPRESSOS NA GLÂNDULA
MAMÁRIA E RELACIONADOS À OCORRÊNCIA E CONTROLE
DA MASTITE BOVINA**

ADILSON FERREIRA DA MOTA

Tese apresentada à Escola Superior de Agricultura
“Luiz de Queiroz”, Universidade de São Paulo, para
obtenção do título de Doutor em Agronomia, Área
de concentração: Ciência Animal e Pastagens.

PIRACICABA
Estado de São Paulo – Brasil
Julho – 2003

**DESCOBRINDO GENES EXPRESSOS NA GLÂNDULA
MAMÁRIA E RELACIONADOS À OCORRÊNCIA E CONTROLE
DA MASTITE BOVINA**

ADILSON FERREIRA DA MOTA

Engenheiro Agrônomo

Orientador: Prof. Dr. **LUIZ LEHMANN COUTINHO**

Co-orientador: Dr. **TAD S. SONSTEGARD**

Tese apresentada à Escola Superior de Agricultura
"Luiz de Queiroz", Universidade de São Paulo, para
obtenção do título de Doutor em Agronomia, Área
de concentração: Ciência Animal e Pastagens.

PIRACICABA
Estado de São Paulo – Brasil
Julho – 2003

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Da Mota, Adilson Ferreira

Descobrimo genes expressos na glândula mamária e relacionados à ocorrência e controle da mastite bovina / Adilson Ferreira Da Mota. - - Piracicaba, 2003.

152 p.

Tese (doutorado) - Escola Superior de Agricultura Luiz de Queiroz, 2003.
Bibliografia.

1. Bovinos 2. Expressão gênica 3. Genética molecular 4. Genomas 5. Glândulas mamárias de animal 6. Mastite animal 7. Resistência genética animal I. Título

CDD 636.20896

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

Dedico à minha esposa Ana Beatriz e aos meus filhos Carolina e João, que – incondicionalmente – me acompanharam e me apoiaram, mesmo em momentos de minhas erradas decisões, ensinando-me o sentido da Fé.

AGRADECIMENTOS

- Ao Dr. Luiz Lehmann Coutinho, professor do Departamento de Zootecnia da ESALQ-USP e Dr. Tad. S. Sonstegard, pesquisador da Unidade de Beltsville, do Serviço de Pesquisa Agrícola, do Departamento da Agricultura dos Estados Unidos da América (BARC-ARS-USDA) pela amizade fraterna, orientação e apoio na realização deste trabalho.
- Ao Dr. Airdem Gonçalves de Assis, coordenador da Embrapa LABEX – EUA e Dra. Terezinha Padilha Bounds, pesquisadora da Embrapa LABEX – EUA pelo apoio para que parte desse trabalho fosse realizado no Departamento da Agricultura dos Estados Unidos da América.
- Aos Drs. Curtis P. Van Tassell, Erin E. Connor, Anthony V. Capuco e Robert J. Wall, pesquisadores do BARC-ARS-USDA pela amizade, colegismo e bom relacionamento vivenciado durante a permanência naquela instituição.
- Ao Ministério da Agricultura, Pecuária e Abastecimento, nas pessoas do Ministro Pratini de Moraes e da chefia do gabinete, Dr. João Cândido Brilhante Neto, bem como à Diretoria da Embrapa, por meio dos Drs. Alberto Duque Portugal e Dante Scollari pelo apoio recebido, por propiciar financiamento das pesquisas, pela concessão de bolsa e pela liberação para realizar parte dos trabalhos no USDA.
- Aos chefes da Embrapa Gado de Leite, Drs. Duarte Vilela e Mario Luiz Martinez pela aprovação das propostas encaminhadas e pela cessão de animais experimentais para a realização das pesquisas.

- A Shirley Weese, Larry Wood, Tina Sphon, Larry Shade e Marsha Campbell, funcionários do BARC-ARS-USDA pelo apoio direto na realização das pesquisas.
- À estagiária Fernanda Luiza de Souza e demais estagiários, aos pós-graduandos, técnicos e funcionários do Laboratório de Biotecnologia Animal da ESALQ-USP pelo bom ambiente de trabalho, amizade e convívio.
- Aos meus pais e minha irmã, meus sogros, meus tios, meus primos, meus sobrinhos e meus amigos pela confiança, apoio e estímulo para realizar este trabalho.
- A todos os professores e funcionários do Departamento de Zootecnia, à Coordenação do curso de pós-graduação em Ciência Animal e Pastagens, aos funcionários da Secretaria de pós-graduação e da Biblioteca Central e a todos que de qualquer maneira tenham contribuído para a realização deste trabalho.

SUMÁRIO

	Página
LISTA DE FIGURAS.....	x
LISTA DE TABELAS.....	xii
RESUMO.....	xiv
SUMMARY.....	xvi
1 INTRODUÇÃO.....	1
2 REVISÃO DE LITERATURA.....	5
2.1 A importância da bovinocultura nacional.....	5
2.2 A história, as características e a importância da raça Gir para o Brasil.....	7
2.3 Emprego da variação genética em pesquisas com animais.....	10
2.4 A mastite bovina e a contagem de células somáticas.....	11
2.5 O Complexo Linfócito Antígeno Bovino.....	13
2.6 O emprego de bibliotecas de cDNA para caracterização e descoberta de genes bovinos expressos na glândula mamária.....	14
3 LOCALIZAÇÃO DE LOCOS GÊNICOS DE CARACTERÍSTICAS QUANTITATIVAS (QTLs) EM GADO DE LEITE. PREMISSAS E ESTRATÉGIAS.....	16
Resumo.....	16
Summary	17
3.1 Introdução.....	17
3.2 Premissas das leis da Genética.....	18
3.2.1 A 1ª Lei da Genética.....	18
3.2.2 A 2ª Lei da Genética.....	19

3.2.3 A 3ª Lei da Genética.....	20
3.3 Marcadores moleculares.....	21
3.3.1. Marcadores microssatélites.....	22
3.4 Populações construídas.....	24
3.4.1 O caso de apenas um loco marcador.....	24
3.4.2 Desenho experimental F2 para um modelo supondo apenas um marcador molecular.....	25
3.4.3 Desenho experimental de retrocruzamento (BC) para modelo com um marcador.....	27
3.4.4 Localização do QTL empregando mais de um marcador molecular flanqueando a região candidata.....	29
3.4.5 Desenho experimental de retrocruzamento (BC) para modelo com mais de um marcador flanqueando a região.....	30
3.4.6 Desenho de filhas ("Daughter Design").....	33
3.4.7 Desenho de netas ("Granddaughter Design").....	34
3.5 O poder para se detectar QTLs.....	36
3.6 Considerações finais.....	37
3.7 Conclusões.....	38
4 DISTRIBUIÇÃO DOS ALELOS DO GENE BOLA-DRB 3 EM BOVINOS DA RAÇA GIR LEITEIRA (<i>Bos indicus</i>).....	39
Resumo.....	39
Summary.....	39
4.1 Introdução.....	40
4.2 Material e Métodos.....	41
4.2.1 Animais e coleta de amostras.....	41
4.2.2 Extração de DNA.....	41
4.2.3 Amplificação do segundo exon do gene BoLA-DRB3.....	42
4.2.4 Eluição dos produtos de PCR e seqüenciamento direto.....	43
4.2.5 Clonagem dos produtos eluídos.....	43
4.2.6 Definição dos alelos.....	44
4.3 Resultados e Discussão.....	44
4.3.1 Comparação das diferentes metodologias de genotipagem.....	44

4.3.2	Frequências alélicas.....	47
4.4	Conclusões.....	50
	Agradecimentos.....	50
5	GENOTIPAGEM DE ALELOS DO GENE BOLA-DRB3 EM BOVINOS DA RAÇA GIR LEITEIRA (<i>BOS INDICUS</i>) POR MEIO DE ELETROFORESE EM GEL DE POLIACRILAMIDA COM GRADIENTE DE TEMPERATURA E SEQÜENCIMENTO.....	51
	Resumo.....	51
	Summary.....	52
5.1	Introdução.....	52
5.2	Material e Métodos.....	54
5.2.1	Extração de DNA e amplificação do gene BoLA-DRB3.....	54
5.2.2	Condições de TGGE.....	54
5.2.3	Amplificação dos alelos separadamente e purificação.....	55
5.2.4	Reações de seqüenciamento.....	55
5.2.5	Designação dos alelos componentes do genótipo do animal.....	55
5.3	Resultados e Discussão.....	56
5.4	Conclusões.....	58
	Agradecimentos.....	58
6	INVESTIGAÇÃO DO POTENCIAL DE DESCOBERTA DE POLIMORFISMO NO EXON 2 DE ALELOS DO GENE BoLA-DRB3 EM VACAS DA RAÇA GIR LEITEIRA (<i>Bos indicus</i>).....	62
	Resumo.....	62
	Summary.....	63
6.1	Introdução.....	63
6.2	Material e Métodos.....	65
6.2.1	Animais experimentais.....	65
6.2.2	Extração de DNA e amplificação do gene BoLA-DRB3.....	65
6.2.3	Clonagem do produto de PCR e seqüenciamento.....	66
6.2.4	Análise das seqüências e genotipagem dos animais.....	66
6.3	Resultados e Discussão.....	67
6.4	Conclusões.....	72

Agradecimentos.....	73
7 CONSTRUÇÃO E CARACTERIZAÇÃO DE BIBLIOTECAS DE cDNA DE GLÂNDULA MAMÁRIA DE BOVINOS DA RAÇA GIR LEITEIRA (<i>Bos indicus</i>) E DA RAÇA HOLANDESA (<i>Bos taurus</i>).....	74
Resumo.....	74
Summary.....	75
7.1 Introdução.....	76
7.2 Material e Métodos.....	78
7.2.1 Coleta de tecidos e desafio com <i>Staphylococcus aureus</i>	78
7.2.2 Preparação do RNA.....	79
7.2.3 Construção das bibliotecas e seqüenciamento.....	79
7.2.4 Análises computacionais de qualidade e agrupamento.....	83
7.3 Resultados e discussão.....	84
7.7 Conclusões.....	93
Agradecimentos.....	93
8 CONCLUSÕES GERAIS.....	94
ANEXOS.....	95
REFERÊNCIAS BIBLIOGRÁFICAS.....	138

LISTA DE FIGURAS

		Página
1	Fotografia de vaca da raça Gir em lactação.....	8
2	Diagrama da segregação independente entre locos sem ligação.....	20
3	Diagrama da segregação entre locos com ligação e frequência de recombinação de 10%.....	20
4	Alelos do microssatélite e sua detecção por PCR e eletroforese.....	23
5	Diagrama do segmento cromossômico de indivíduos da geração F1 com genótipo Qq/Mm para o QTL (Q) e o marcador (M).....	25
6	Diagrama do segmento cromossômico de indivíduos da geração F1 com genótipo MmQqNn para o marcador (M) e marcador (N), flanqueando o QTL (Q).....	30
7	Diagrama do Desenho de Filhas.....	34
8	Diagrama do Desenho de Netas.....	35
9	Eletroforese em géis de agarose 2% com 8 ìl de produto de PCR do gene BoLA-DRB3.....	46
10	Cromatograma da reação de seqüenciamento com a seqüência de nucleotídeos obtida por PCR-DS mostrando múltiplos pontos de superposição de polimorfismo.....	47
11	Gel de poliacrilamida contendo alelos do gene BoLA-DRB3 separados por meio de TGGE usando gradiente de temperatura entre 35°C - 75°C.....	58

12A	Eletroferograma de amostra de RNA total com visualização da qualidade e concentração.....	82
12B	Seleção de perfis para a construção de bibliotecas de cDNA.....	82
12C	Perfil de cDNA recomendável após a recuperação do gel da amplificação por PCR, para preparação de biblioteca.....	82
13	Diagrama esquemático das razões de eliminação das ESTs.....	86
14	Gráfico da redundância observada. Número de ESTs nos elementos únicos de seqüência (UNIQUE).....	87
15	Distribuição da posição relativa dos elementos de seqüência única (UNIQUE) em relação a seqüências completas (<i>full-length open reading frames</i>) de transcritos Humanos.....	91

LISTA DE TABELAS

		Página
1	Número de animais com registro genealógico de nascimento (RGN) de raças zebuínas no Brasil no ano de 2003.....	9
2	Tipos de gametas e suas freqüências derivadas da meiose em uma geração F1.....	25
3	Genótipos, freqüências e valor genotípico esperados na geração F2 obtida do cruzamento entre F1 x F1.....	26
4	Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular.....	26
5	Contrastes ortogonais empregados no cálculo dos efeitos genéticos aditivo (a) e devido à dominância (d) do QTL.....	27
6	Genótipos e freqüências genotípicas no desenho BC.....	28
7	Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular para desenho BC com um marcador.....	28
8	Contrastes ortogonais empregados no cálculo dos efeitos genéticos aditivo (a) e devido à dominância (d) do QTL, para desenho BC com um marcador.....	29
9	Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular para desenho BC com um marcador.....	30

10	Genótipos e frequências genóticas no desenho BC, sendo BC1=(MmQqNn x MMQQNN) e BC2 = (MmQqNn x mmqqnn).....	32
11	Valores esperados das médias da característica de interesse de acordo com as classes dos marcadores moleculares para desenho BC com marcadores flanqueando o QTL.....	33
12	Frequências alélicas do gene BoLA-DRB3 de 28 vacas da raça Gir Leiteira (<i>Bos indicus</i>) genotipadas por seqüenciamento direto de produtos de PCR ou por meio de protocolo de clonagem e seqüenciamento.....	48
13	Identificação do animal, designação dos alelos que integram o genótipo do animal, número de bases que diferem na seqüência, eventos de inserção/deleção, conteúdo de G-C (%), temperatura de fusão dos domínios (°C) e peso molecular (Daltons) para as 28 amostras analisadas. (26 heterozigotos e 2 homozigotos).....	59
14	Frequência dos alelos do gene BoLA-DRB3 em vacas da raça Gir Leiteira (<i>Bos indicus</i>).....	68
15	Seqüências de nucleotídeos do gene BoLA-DRB3 de vacas Gir Leiteira (<i>Bos indicus</i>) dos alelos em que não foi possível obter homologia com alelos depositados nas bases públicas de informações sobre genomas.....	69
16	Alinhamento de nucleotídeos das seqüências do gene BoLA-DRB3 de vacas Gir Leiteira (<i>Bos indicus</i>) dos alelos em que não foi possível obter homologia com alelos depositados nas bases públicas de informações sobre genomas.....	71

17	Seqüências dos <i>primers</i> utilizados na síntese de cDNA e amplificação para a construção das bibliotecas de cDNA.....	80
18	Número de bibliotecas e número de ESTs depositados no GenBank em relação ao status de agrupamento (TC ou ST) e classificação das bibliotecas por raça (Holandesa ou Gir).....	84
19	Análise de agrupamento (<i>Cluster</i>) das bibliotecas BARC-EMBRAPA e resultados das análises realizados pela TIGRrelação ao status de agrupamento (TC ou ST) e classificação das bibliotecas por raça (Holandesa ou Gir).....	90

DESCOBRINDO GENES EXPRESSOS NA GLÂNDULA MAMÁRIA E RELACIONADOS À OCORRÊNCIA E CONTROLE DA MASTITE BOVINA.

Autor: ADILSON FERREIRA DA MOTA
Orientador: Prof. Dr. LUIZ LEHMANN COUTINHO

RESUMO

O Brasil possui o maior rebanho bovino comercial do mundo com sua produção baseada, predominantemente, em animais de origem zebuína (*Bos indicus*), com níveis do potencial genético de produção inferiores aos animais *Bos taurus*. Com a possibilidade de emprego de técnicas de genética molecular, a seleção de indivíduos melhoradores para acasalamento e multiplicação repercutem nos índices de produtividade e produção. A utilização de melhores métodos de estudo da variação genética no nível molecular representa benefícios especialmente nos casos de resistência a doenças, já que essas características apresentam baixa herdabilidade e são afetadas pelo ambiente. O estudo da genética molecular visando o aumento da saúde da glândula mamária de bovinos apresenta a vantagem de poder selecionar animais em idades precoces, antes da produção, para ambos os sexos e, no caso dos machos reprodutores, sem a necessidade de aguardar por informações da saúde da glândula mamária observada somente na descendência dos animais. Para identificar genes com responsabilidade de conferir resistência à mastite bovina, a doença mais importante da bovinocultura, foram realizados diversos experimentos que envolveram um gene do complexo principal de histocompatibilidade (MHC – *Major Histocompatibility Complex*). Os experimentos com o gene BoLA-DRB3, do MHC, permitiram identificar um alelo novo, desenvolver uma técnica mais eficiente de genotipagem de

animais e verificar a distribuição dos alelos em animais da raça Gir Leiteira. Foram também construídas bibliotecas de cDNA utilizando a metodologia de ORESTES e tecido de glândula mamária de animais europeus e zebuínos infectados com *Staphylococcus aureus*. 6.481 sequências expressas (ESTs), obtidas por meio da metodologia de ORESTES, foram depositadas em bases públicas como o GenBank (<http://www.ncbi.nlm.gov/>) e TIGR (<http://www.tigr.org>), sendo 5.950 provenientes de animais *Bos indicus*, o que significou um aumento de vinte vezes no volume de informações disponíveis sobre o transcriptoma de bovinos zebuínos ao tempo deste estudo.

HUNTING EXPRESSED MAMMARY GLAND GENES RELATED TO MASTITIS IN DAIRY CATTLE.

Author: ADILSON FERREIRA DA MOTA
Adviser: Prof. Dr. LUIZ LEHMANN COUTINHO

SUMMARY

Brazil is the largest cattle commercial herd in the world, mainly based in zebu (*Bos indicus*) with lower genetic potential than taurus-derived animals. Molecular genetics brought the possibility of better selection of animals for breeding, which increases production levels. The use of molecular methods for studying genetic variation is especially advantageous for disease resistance because of low herdability and environment effects. Selection by means of molecular genetics in dairy cattle can be done before production, for males and females. To identify genes responsible for conferring resistance to Mastitis, the most important cattle disease, we studied one gene from the MHC (Major Histocompatibility Complex). The experiments with this MHC gene (BoLA-DRB3) identified a new allele in cattle, developed a genotyping technique, and estimate allelic frequencies for Dairy Gir cows (*Bos indicus*). CDNA libraries from mammary gland tissues infected with *Staphylococcus aureus* were also constructed and characterized. 6,481 ESTs (Expressed Sequence Tags) were deposited in GenBank (<http://www.ncbi.nlm.nih.gov/>) and TIGR (<http://www.tigr.org>), being 5,950 from *Bos indicus* cows, which increased by 20-fold the volume of sequence data from the zebu transcriptome, at the time when this study was conducted.

1 INTRODUÇÃO

As pesquisas realizadas com Humanos evoluíram ao longo da História acompanhadas da evolução das pesquisas com bovinos. Ambas compartilharam benefícios e interesses comuns.

Os bovinos vêm acompanhando o Homem ao longo da sua evolução e diversas raças foram geradas mediante a seleção realizada nesses animais. É possível verificar animais mais resistentes a parasitas, adaptados à tração de carga, tolerantes a ambientes tropicais, ou altamente produtivos, tanto para leite quanto para carne. Atualmente, as pesquisas biomédicas, realizadas por meio do emprego de animais como modelo, contribuem para elucidar as interações gênicas que ocorrem em distintas condições, raças e ambientes.

No estágio atual da evolução e da complementaridade entre as pesquisas com Humanos e bovinos e com a introdução da Biologia Molecular na realização de pesquisas genéticas e genômicas, estudos nesta espécie doméstica irá auxiliar a elucidação da evolução Humana e permitirão extrapolar a informação genética para a medicina, bem como para o maior entendimento da Biologia, em benefício do Homem. Estudos sobre o genoma bovino adquire interesse particular nesse caso, já que os bovinos não são roedores ou primatas e as informações sobre seus genes trazem contribuição relevante que permitirá verificar as diferenças que caracterizam os seres humanos e os diferenciam dos demais mamíferos. Todas as espécies animais são resultantes de milhões de anos de mutações e seleção natural. O conhecimento sobre o genoma das espécies permitirá traçar o caminho de volta e entender o caminho da evolução, do Homem e dos animais.

Além de partilhar benefícios e interesses comuns com as pesquisas realizadas em Humanos, as pesquisas com bovinos apresentam o objetivo específico de melhoria das características de saúde do animal e da sua produtividade. Esses aspectos também representam benefícios para o Homem pelo aumento da oferta de alimentos, e também por meio da redução da contaminação de produtos e do meio ambiente, bem como dos custos de aplicação de medicamentos.

As pesquisas com bovinos empregando técnicas de Biologia Molecular já produziram diversos resultados concretos. A genética molecular permitiu desenvolver métodos para selecionar diretamente animais com variantes genéticas para a lacto-globulina (Wilkins & Kuys, 1992) e k-caseína (NG-Kwai-Hang et al., 1990; Zadworny & Kuhnlein, 1990), com repercussões, respectivamente, na produção de leite e qualidade do queijo produzido. Maior produção de leite e proteína (Ron et al., 1994), bem como doenças como a deficiência de adesão leucocitária bovina (BLAD) (Shuster et al., 1992), Mieloencefalopatia degenerativa progressiva (Weaver) (Georges et al., 1993) e Citrulinemia (Dennis et al., 1989) podem ser identificadas previamente e os indivíduos portadores dessas doenças podem ser descartados dos rebanhos. Além disso, é possível identificar animais portadores dos genes de musculatura dupla (GDF-8) (McPherron & Lee, 1997), resistentes à verminose ou com maiores taxas de ganhos de peso. A cada dia novos marcadores são descobertos com efeitos significativos sobre a produção e resistência a doenças em bovinos. Desta forma, recentes avanços da biologia molecular e o refinamento das técnicas de avaliação do DNA dos animais domésticos têm aberto novas perspectivas de impacto nos sistemas de produção. Sobretudo, a Biologia Molecular auxilia na compreensão da relação biológica existente entre os genes e as características fenotípicas de interesse comercial. Proteínas codificadas no genoma estão relacionadas com os processos bioquímicos e fisiológicos envolvidos na produção, na reprodução ou na sanidade de animais domésticos. A identificação dos genes de interesse e a possibilidade de compreensão do seu mecanismo de ação permitem empregá-los pontualmente na seleção de indivíduos portadores de características de interesse comercial. O desafio atual para produzir maiores impactos é determinar os melhores métodos de emprego da variação genética entre animais para incrementar o entendimento a respeito da biologia e identificar aquelas variações que são responsáveis pelas principais alterações fenotípicas.

Dentre os aspectos que são motivos de investigação referente aos bovinos, em especial aos bovinos leiteiros, estão aqueles da saúde e da produção leiteira, relacionados à glândula mamária. As características relativas à produção leiteira apresentam de modo geral maior herdabilidade e ganhos genéticos representativos podem ser obtidos nessas características com o emprego da genética quantitativa clássica. No entanto, a resistência a doenças apresenta baixa herdabilidade (Shook, 1989), sendo reduzidos os ganhos genéticos para essa característica. Em razão disso, vários autores têm realizado estudos empregando técnicas de genética molecular em estudos de saúde da glândula mamária dos bovinos leiteiros (Bernoco & Lewin, 1989; Dietz et al., 1997a; Dietz et al., 1997b; Groenen et al., 1990; Giovambatista et al., 1996; Kelm et al.,

1997; Lewin, 1989; Østergård et al., 1989; Starkenburg et al., 1997) com o objetivo de associar locos de efeito maior com a ocorrência de doenças. O emprego da genética molecular em estudos destinados a oferecer soluções genéticas de aumento da saúde da glândula mamária de bovinos apresentam uma vantagem importante porque os animais podem ser selecionados em idades precoces, antes da produção, para ambos os sexos e, no caso dos machos reprodutores, sem a necessidade de aguardar por informações da saúde da glândula mamária observada somente na descendência dos animais.

Neste contexto, a mastite bovina é o principal problema de saúde relacionado à produção de leite, tanto no mundo quanto no Brasil. Em razão disso, nas últimas décadas diversos estudos vêm sendo realizados na espécie bovina a fim de identificar regiões cromossômicas ou genes candidatos e suas associações com a resistência a esta doença. Esses estudos visam identificar animais produtivos e resistentes à doença, ao mesmo tempo. Entretanto, ainda não há consenso nas informações obtidas. Pesquisas adicionais podem elucidar a complexa interação de processos celulares que ocorre na reação imunológica em resposta a essa doença, e que afeta de maneira significativa à qualidade e os níveis de produção de leite.

Uma das regiões do genoma bovino que tem sido objeto desses estudos é o Complexo Principal de Histocompatibilidade (MHC - *Major Histocompatibility Complex*) que na espécie bovina foi denominado de região BoLA (*Bovine Lymphocyte Antigen*). Sabe-se que estes genes estão envolvidos nos processos celulares relacionados ao sistema imunológico dos animais, sendo altamente polimórficos em diferentes espécies e responsáveis por codificar as proteínas presentes na superfície das células e envolvidas na relação entre antígenos e anticorpos. Em adição, esta região compreende vários genes, mas não há evidências definitivas da ligação entre esses e a resistência à ocorrência de mastite. Estudos independentes e com raças distintas podem contribuir para as confirmações necessárias das associações contraditórias encontradas. Um dos genes dessa região que é altamente expresso e que foi empregado nos estudos mais recentes é o gene BoLA-DRB3. Além deste gene, muitos outros podem estar envolvidos no processo molecular de resistência à mastite. Por isso, construir e caracterizar bibliotecas de cDNA e avaliar as seqüências expressas diferencialmente (ESTs) entre animais suscetíveis e resistentes é uma boa alternativa para identificar novos genes candidatos a estudos de associação com ocorrência de mastite. Especialmente neste momento em que o seqüenciamento automático permite determinar as seqüências de muitos genes com rapidez, esse é um procedimento a ser realizado em pouco tempo e com resultados de impacto na descoberta de genes candidatos.

Nos capítulos apresentados a seguir, o tema: “Descobrir genes expressos na glândula mamária relacionados à ocorrência e controle da mastite bovina” foi escolhido e é apresentado como tese. Buscou-se estudar qual metodologia poderia ser mais bem empregada em bovinos leiteiros para identificar locos ou regiões cromossômicas consideradas candidatas a conter genes que conferem resistência à mastite. Os objetivos também contemplaram o desenvolvimento de metodologias para genotipagem de um gene altamente polimórfico do MHC e verificação de sua associação com a ocorrência da doença. Com a evolução das técnicas de seqüenciamento de DNA e construção de bibliotecas de cDNA, ocorrida durante o período de estudo, objetivou-se gerar seqüências de genes expressos em tecidos de glândula mamária buscando descobrir seqüências novas no genoma bovino com implicações na resistência à mastite.

A relevância do estudo reside no fato de que a descoberta da associação de genes individuais responsáveis por conferir resistência à doença pode levar à utilização desse conhecimento pelo setor produtivo e redução dos índices de ocorrência da mastite, com a conseqüente redução de custos de tratamento dos animais e de depósitos de resíduos de antibióticos, tanto no leite quanto no meio ambiente. Todos os estudos empregaram a raça Gir leiteira (*Bos indicus*), a qual anteriormente não apresentava informações genômicas disponibilizadas publicamente.

2 REVISÃO DA LITERATURA

2.1 A importância da bovinocultura nacional

O Brasil é o terceiro maior produtor de carne bovina do mundo, o terceiro maior exportador mundial e o país onde ocorre o terceiro maior volume de consumo interno mundial (<http://www.agricultura.gov.br/spa/pagespa/index.htm>). A produção de carne bovina brasileira superou 7 milhões de toneladas em 2002, das quais 930 mil toneladas foram exportadas, no mesmo ano. Em 2012, uma década depois, as previsões de produção superam as 9 milhões de toneladas e as exportações daquele ano deverão ter crescido 170% em volume e 250% em valor (ANUALPEC, 2003).

Com relação à produção oficial de leite, o Brasil é o sexto produtor mundial, com cerca de 4% do total produzido no mundo (<http://www.cnpq.embrapa.br/leite/index.php>). Segundo estimativas, o volume de produção superou os 21 bilhões de litros de leite em 2002, sendo que o número de vacas ordenhadas superou os 18,5 milhões de animais. No entanto, a produção de leite no Brasil não é suficiente para atender a demanda de consumo recomendada pelo Ministério da Saúde, que é de 200 litros *per capita*, considerando o produto na forma fluida e na de produtos lácteos. Tomando como base os mais de 175 milhões de brasileiros, um consumo de 600 ml/dia iria demandar 38,3 milhões de litros anuais de leite (<http://www.cnpq.embrapa.br/artigos/LeiteBrasilPrecisa.pdf>).

Segundo as mesmas fontes mencionadas, o Brasil possui o maior rebanho bovino comercial do mundo. Na última década, o efetivo do rebanho do Brasil cresceu em média mais de 15% ao ano, partindo de 152 milhões de animais em 1991 e ultrapassando os 167 milhões em 2003. Com níveis de produtividade da pecuária de corte e leite nacionais semelhantes aos níveis que são observados nos países desenvolvidos, o Brasil poderá atender à demanda de consumo interno de sua população e ainda ser o maior exportador mundial de produtos de origem bovina.

Ainda que os níveis de produtividade dos rebanhos de pecuária bovina brasileiros não sejam elevados, a atividade representa 9% do Produto Interno Bruto (PIB) nacional e as exportações superam os 2 bilhões de dólares anuais. No ano de 2001, mais brasileiros foram empregados nas atividades agrárias do que na construção civil, demonstrando a importância socioeconômica da atividade rural brasileira.

Não obstante a reduzida produtividade da bovinocultura, além do aumento do efetivo, os níveis de produção e produtividade dos rebanhos nacionais vêm aumentando nos últimos anos. A produção bovina tem aumentado em razão do aumento da produtividade e também devido ao aumento das áreas destinadas para a bovinocultura. A introdução de áreas de Cerrado nos sistemas de produção representa uma fonte representativa de aumento da produção agrícola e pecuária. Os Cerrados brasileiros são responsáveis por uma área total de 208 milhões de hectares¹. Atualmente, ainda existem mais de 100 milhões de hectares dos Cerrados a serem integrados aos sistemas de produção brasileiros, representando possibilidade de expansão futura da área destinada para a agropecuária.

O aumento nas características de produção dos animais apresenta vantagens evidentes tanto para produtores quanto para consumidores. A produção animal representa mais da metade da renda proveniente do setor agrícola nos países desenvolvidos e mais de um quarto nos países em desenvolvimento. Estima-se que este padrão se mantenha, pois as projeções de crescimento da atividade de produção animal são de 3,4% nos países desenvolvidos, contra 2,3% de crescimento de produtos de colheita vegetal.

Há projeções também de crescimento na demanda dos produtos de origem bovina nos países do continente asiático. A China cresceu a taxas anuais superiores a 7% nos últimos quatro anos. Outros países daquela região como a Índia, a Coreia do Sul, e a Rússia cresceram a taxas próximas de 5% anuais no período. Como esses países têm mantido a balança comercial favorável, a população tem apresentado elevação do padrão de vida e redução das privações na dieta alimentar, com o aumento no consumo de *commodities*, especialmente grãos e carnes (ANUALPEC, 2003). Como se trata de nações superpopulosas, envolvendo aproximadamente 3 bilhões de habitantes (quase a metade da população do planeta), o seu ingresso no mercado consumidor de proteínas de origem animal justifica previsões de crescimento futuro da demanda internacional por carne bovina e produtos lácteos, em quantidade e qualidade suficientes para suprir as exigências do mercado.

¹ ASSAD, E. (Embrapa Cerrados, Planaltina – DF). **Comunicação pessoal, 2003.**

2.2 A história, as características e a importância da raça Gir para o Brasil

A história da introdução do Gir no Brasil está descrita por Domingues (1973) e Santiago (1975). As primeiras importações de animais da raça Gir provenientes da Índia foram efetuadas por Teófilo de Godoy e Wirmondos M. Borges e iniciaram-se por volta de 1906.

A raça teve origem na floresta de Gyr, na península de Kathiawar, na Índia, onde ainda é criada como raça pura. Os animais dessa raça apresentam na cabeça um perfil ultra convexo peculiar, com fronte larga, lisa e proeminente, de largura e comprimento médios. Os chifres são grossos na base e saem para trás e para baixo com secção elíptica, achatada e cor escura. Os olhos são dispostos bem lateralmente e as orelhas são longas, muito típicas, pendentes e enroladas sobre si mesmas. O dorso e lombo são firmes, largos e horizontais. As costelas são bem arqueadas, cobertas de carne, compridas e separadas. As ancas são largas, a garupa comprida, os membros curtos e musculosos, afastados e aprumados. O úbere é bem desenvolvido, bem irrigado, prolongado para diante e para trás, com quartos simétricos e tetas médias bem distribuídas. As pelagens predominantes são a vermelha, a rosilha e a moura. Os animais apresentam aptidão leiteira e bom rendimento de carne (Figura 1).

Os primeiros trabalhos de seleção foram realizados na década de 30 na Fazenda Experimental de Umbuzeiro, no estado da Paraíba. Em 1948 iniciou-se a seleção também na Fazenda Experimental Getúlio Vargas, em Uberaba, estado de Minas Gerais. A seleção para características produtivas orientadas inicialmente para carne ou leite ocasionou a formação de duas associações de criadores: Associação Brasileira dos Criadores de Gir (ASSOGIR) e Associação Brasileira dos Criadores do Gir Leiteiro (ABCGIL).

A raça Gir apresenta possibilidade de seleção e desenvolvimento, bem como importância para a estrutura genética da população bovina do Brasil. Os programas de seleção e avaliação genética de animais dessa raça vêm identificando animais superiores e despertado o interesse de outros países em adquirir animais, o que possibilita ao Brasil tornar-se exportador de material genético zebuino.



Figura 1 - Fotografia de vaca da raça Gir em lactação

Atualmente, as fazendas atendidas pelo Teste de Progênie da Raça Gir Leiteira realizado pela Embrapa Gado de Leite, em Juiz de Fora – MG, realiza uma avaliação genética das fêmeas em lactação que indicou para o ano de 2002 uma produção média de leite superior a 3.000 kg e uma duração de lactação de 260 dias².

Segundo a Associação Brasileira de Criadores de Zebu – ABCZ, a raça Gir, considerando os animais da raça Gir mocha e da raça Gir, representam 8,8% do total de registro genealógicos de nascimento dos animais zebuínos do Brasil (Tabela 1). Segundo os cálculos da mesma associação, os registros definitivos de animais da raça Gir totalizam cerca de 10% do total de zebuínos brasileiros (<http://www.abcz.org.br>).

A importância desta raça zebuína pode também ser dimensionada quando se avaliam as informações referentes à venda de sêmen no Brasil. Considerando-se o total de sêmen de bovinos leiteiros comercializado no Brasil, a raça Gir situa-se em terceiro lugar com 12,46%, superada apenas pelo Holandês (66,73%) e pelo Jersey (14,60%) (<http://www.cnpqc.embrapa.br/sbma/anais/250.html>). Segundo a mesma fonte, as raças zebuínas, em 2001, comercializaram 1.882.524 doses, representando 43,54% do total de sêmen vendido no Brasil.

² VERNEQUE, R.(Embrapa Gado de Leite, Juiz de Fora – MG). **Comunicação pessoal, 2003.**

Tabela 1 . Número de animais com registro genealógico de nascimento (RGN) de raças zebuínas no Brasil no ano de 2003

Raça	RGN	%
Gir Mocha	36.180	0,58
Gir	516.784	8,29
Guzerá	246.127	3,95
Indubrasil	210.422	3,38
Nelore	4.570.843	73,33
Nelore Mocha	461.996	7,41
Sindi	9.369	0,15
Sindi Mocha	25	0,00
Tabapuã	177.228	2,84
Cangaian	29	0,00
Brahman	4.611	0,07
TOTAL	6.233.614	100,00

Fonte ABCZ – Associação Brasileira de Criadores de Zebu (<http://www.abcz.org.br>)

Além da importância da raça Gir como raça pura e como material genético selecionado para produção leiteira nas condições tropicais do Brasil, a raça também tem sido utilizada em cruzamento com animais da raça Holandesa, produzindo outra raça nacional chamada Girolando, atualmente com mais de 50 mil registros (<http://www.girolando.com.br/noticia.php?id=6>). A raça apresenta versatilidade, rusticidade, vida útil e produção leiteira também elevada.

2.3 Emprego da variação genética em pesquisas com animais

Com o objetivo de aumentar do potencial genético dos animais, com repercussão na produtividade e na produção total, a genética quantitativa clássica utiliza como abordagem a mensuração de variáveis fenotípicas e a coleta de informações a respeito da genealogia dos animais. Análises estatísticas dão indicações dos animais que apresentam genótipos superiores, sem os efeitos de ambiente, do manejo e interações. A genética molecular oferece uma abordagem alternativa, em que os genótipos são medidos diretamente. A análise estatística nesse caso busca identificar como a variação genética está associada com a variação fenotípica observada. Apesar de a alternativa de medir diretamente os genótipos ter sido introduzida há muitos anos, somente após a utilização das ferramentas da genética molecular, essa passou a ser uma alternativa prática e passível de realização em qualquer tipo de organismo vivo.

Correntemente, dentre as metodologias de pesquisa para a utilização de medidas diretas sobre o genótipo dos animais, duas delas merecem especial atenção. Uma das abordagens concebe o conceito e emprego do desequilíbrio de ligação, no qual o conhecimento dos genótipos é empregado como marca (marcador molecular), com endereço localizado e conhecido em determinada região cromossômica. O Capítulo 3 apresenta uma revisão sobre as premissas e estratégias de localização de locos gênicos em gado leiteiro. Nessa situação, a genotipagem de indivíduos, a construção de mapas de ligação, o emprego do conhecimento das medidas fenotípicas e do desequilíbrio de ligação, permitem identificar regiões cromossômicas que apresentam probabilidade estatística de conter os genes responsáveis pelas variações. A associação de marcadores moleculares com características de interesse econômico vem sendo empregada há vários anos. No entanto, antes de empregar um loco de efeito principal como marcador molecular, é necessário estabelecer estimativas precisas do desempenho das características de interesse para os possíveis genótipos (Shook, 1989). Também é necessário testar os efeitos do marcador em diferentes populações e raças, para verificar se os mesmos genótipos apresentam condição de superioridade em diferentes situações. Somente após a verificação de que o marcador encontra-se próximo o suficiente do gene de interesse para que não ocorra recombinação, é que ele poderá ser usado na seleção de animais superiores.

A segunda abordagem concebe genotipar locos de interesse, ou genes, para os quais há indicações substanciadas de envolvimento potencial na ocorrência de fenótipos superiores (genes

candidatos). A genotipagem de polimorfismos nesses genes de interesse e sua associação com os fenótipos constituem alternativas nesse caso. As duas abordagens mencionadas não são mutuamente excludentes e podem ser empregadas em conjunto (e. g., Davies et al., 1994). À medida que as ferramentas de estudos da função e expressão de genes aumentam, os estudos com genes candidatos também são mais freqüentes. No entanto, mesmo quando não há indicações de genes de efeito principal em uma determinada região gênica candidata, o objetivo final dos estudos é alcançado somente quando é descoberto o polimorfismo do gene (ou genes) responsável por determinado fenótipo de interesse.

A identificação de marcadores moleculares ou de polimorfismo genético associados a características de interesse comercial traz contribuição relevante na seleção de indivíduos superiores. Particularmente, no caso dos bovinos leiteiros, a seleção dos machos reprodutores é um processo lento, realizado indiretamente, por intermédio da avaliação genética da progênie. Também há de ser ressaltado que a pecuária leiteira é uma atividade rural importante no contexto sócio-econômico nacional. O leite permanece sendo um alimento nobre e importante no desenvolvimento humano e mesmo as tecnologias mais modernas não puderam produzir um alimento substituto. A correta identificação de reprodutores e matrizes superiores são fundamentais para o aumento dos índices de produção da atividade leiteira. A seleção e o acasalamento desses animais produzem ganhos genéticos permanentes nas gerações futuras. Além de métodos tradicionais de melhoramento animal, é possível empregar estratégias de pesquisa com a utilização de marcadores moleculares ou genes candidatos que podem identificar os locos que contêm os genes de interesse. No caso dos experimentos com marcadores, essas regiões são chamadas de locos de características quantitativas (QTL – *Quantitative Trait Loci*). As estratégias para utilização de marcadores moleculares em bovinos leiteiros representam a possibilidade efetiva de aumentar a precisão da seleção genética de animais leiteiros e será discutida mais profundamente no Capítulo 3 (da Mota et al., 2000).

2.4 A mastite bovina e a contagem de células somáticas

No entanto, há algumas doenças animais que afetam negativamente a produtividade e a produção. A mastite é a principal doença dos rebanhos leiteiros. Os trabalhos de Nonnecke & Harp (1988), Schutz (1993) e Veiga et al. (1994) apresentam informações detalhadas da doença e dos prejuízos por ela causados, tanto no Brasil quanto no mundo.

A doença caracteriza-se por inflamação da glândula mamária da espécie bovina e representa um dos principais prejuízos à bovinocultura leiteira, acarretando redução do volume de produção de leite, bem como a depreciação dos produtos lácteos e podendo promover riscos ao consumo humano.

No Brasil, segundo Veiga et al. (1994), na região sudeste, onde se concentra a maior bacia leiteira do País, cerca de 20% dos quartos mamários são afetados pela doença na forma subclínica, enquanto 1% é afetado clinicamente, apresentando os sintomas da doença. A mastite subclínica é caracterizada pela ocorrência da doença quando não são observadas alterações macroscópicas no leite e quando os sintomas da doença não são percebidos clinicamente. Já a mastite clínica apresenta alterações visíveis no leite com o aparecimento de grumos e mudança acentuada de sabor. O úbere do animal apresenta vermelhidão e há sintomas clínicos da doença com ocorrência de febre. Segundo os mesmos autores, a redução média da produção de leite devido à mastite subclínica é de 35%, enquanto que para a mastite clínica é de 55%, o que resulta em uma perda total diária de 7,5% da produção de leite, equivalentes aproximadamente a um bilhão e cem milhões de litros de leite produzidos a menos por ano.

Os autores mencionam ainda a importância de considerar os riscos à saúde humana pela ingestão de leite contaminado por bactérias potencialmente patogênicas ou contendo resíduos de antibióticos, o que causa resistência ao princípio ativo dos mesmos. A doença também acarreta aumento dos custos com mão-de-obra, com tratamentos dos animais e prejuízos com o descarte de leite.

A mastite é causada por organismos patogênicos que ingressam na glândula mamária através do canal do teto, atingindo e colonizando a cisterna do leite. Os organismos que causam infecções mais frequentes são: *Escherichia coli*, *Staphylococcus aureus*, *Streptococcus agalactiae* e *Streptococcus uberis*. No entanto, a mensuração da incidência de mastite é uma questão controversa e não faz parte, rotineiramente, dos arquivos de dados zootécnicos dos bovinos de leite. Esse fato prejudica a realização de diagnósticos, levantamentos e pesquisas genéticas para identificação e seleção de animais resistentes. A dificuldade em obter dados sobre os índices de ocorrência da doença nos estabelecimentos rurais tem levado à utilização cada vez mais frequente da contagem de células somáticas (CCS) pelas indústrias, cooperativas laticinistas e centros de pesquisa, como medida indireta representativa da ocorrência de mastite. Devido à alta correlação entre a CCS e a ocorrência da mastite, diversas pesquisas com a utilização da CCS como indicador da ocorrência da doença têm sido realizadas (Schutz, 1993).

As células somáticas são células brancas do sangue, leucócitos, que compreendem os linfócitos, macrófagos e neutrófilos, principalmente. A passagem das células do sangue para o leite ocorre em função da reação do organismo à agressão sofrida na glândula mamária. Essas células somáticas têm a importante função de proteger o organismo e migram do sangue para o leite como resultado do processo inflamatório que é desencadeado pela infecção. Como resultado do processo inflamatório, as paredes de vasos sanguíneos se tornam dilatadas e as células secretoras se tornam menos eficientes, ocorrendo morte celular.

A identificação dos genes responsáveis por conferir resistência à ocorrência da mastite nos animais e posterior emprego no melhoramento animal pode constituir uma abordagem efetiva para o incremento genético da resistência a esta doença. Esse processo é especialmente importante no caso de caracteres de baixa herdabilidade (Dietz et al., 1997^a; Shook, 1989; Starkenburg et al., 1997), como a mastite.

2.5 O Complexo Linfócito Antígeno Bovino

Há genes descritos e envolvidos na biologia de resistência a doenças (Dietz et al., 1997a; Kelm et al., 1997; Seyfert et al., 1996) e, dentre esses, há uma família de genes candidatos que fazem parte do Complexo Principal de Histocompatibilidade Bovina (MHC – “Major Histocompatibility Complex”), denominada BoLA (“Bovine Lymphocyte Antigen”) que tem sido empregada para estudos de marcadores moleculares. Estes genes estão envolvidos com os processos fisiológicos e bioquímicos relacionados ao sistema imunológico dos animais. Os genes da família BoLA são altamente polimórficos e codificam proteínas presentes na superfície das células, envolvidas na relação entre antígenos e anticorpos. Os receptores dos linfócitos-T reconhecem fragmentos dos antígenos em associação com as proteínas da superfície das células, permitindo resposta imunológica específica. Portanto, a região BoLA é um alvo potencial na identificação de genes candidatos para resistência a doenças nos estudos da variação molecular. Dentre os produtos dos genes localizados nessa região, há duas classes de antígenos bem definidos. Os antígenos Classe I são encontrados na quase totalidade da superfície das células do organismo, enquanto os antígenos da Classe II são expressos em células especializadas do sistema imunológico, como os linfócitos. Os genes relativos aos antígenos Classe II estão localizados na região BoLA-D, e expressos principalmente nos linfócitos T e linfócitos B, e podem influenciar a produção subsequente de anticorpos (Østergård et al., 1989).

Os primeiros estudos com a região BoLA envolveram o loco BoLA-A, específico para antígenos da Classe I. A partir da introdução das Reações em Cadeia da Polimerase (PCR – “Polimerase Chain Reaction”) (Saiki et al., 1985) e a decorrente amplificação de regiões definidas do genoma, Groenen et al. (1990) sequenciaram regiões do loco BoLA-D que codificam moléculas da Classe II, relacionadas com ação imunológica específica. Desta forma os genes BoLA-DQB e BoLA-DRB passaram a ser candidatos nos estudos de investigação de associação com doenças dos bovinos. O gene BoLA-DRB1 não é expresso e o gene BoLA-DRB2 é pouco expresso (Burke et al., 1991; Russell et al., 1994). Como o gene BoLA-DRB3 é altamente expresso (Burke et al., 1991; Fraser et al., 1996), estudos têm sido realizados (Dietz et al., 1997a; Dietz et al., 1997b; Giovambattista et al., 1996; Kelm et al.; 1997; Starkenburg et al., 1997) com esse gene e sua associação com a ocorrência da mastite ou a contagem de células somáticas, sem que ocorram, no entanto, resultados conclusivos.

2.6 O emprego de bibliotecas de cDNA para caracterização e descoberta de genes bovinos expressos na glândula mamária

Etiquetas de seqüências de genes expressos, ou ESTs (*Expressed Sequence Tags*) são pequenos trechos de DNA, usualmente entre 300 e 500 pares de bases gerados por meio do sequenciamento de genes expressos em determinada célula ou tecido. A idéia central desse procedimento é empregar a seqüência expressa conhecida e, mediante homologia, identificar em que parte do DNA genômico se encontra aquela seqüência. Os estudos de ESTs servem também para caracterizar o padrão de expressão gênica identificando as seqüências expressas em determinada situação. As ESTs se caracterizam pela ausência de *introns*, que são seqüências de DNA intermitentes e que não participam das seqüências que codificam proteínas em um gene.

A identificação de ESTs com base apenas no conhecimento da seqüência do DNA genômico de um organismo superior eucarioto é dificultada pela presença dos *introns* e de regiões intergênicas, que constituem a grande maioria das seqüências do DNA. Uma região codificante do DNA precisa ser convertida, ou transcrita, para que possa se converter em RNA mensageiro (mRNA). É essa molécula de mRNA, que não contém mais os *introns*, que dá origem à proteína, em outro processo denominado tradução. Portanto, a molécula de mRNA é crucial na formação de proteínas. No entanto, essa molécula é instável fora da célula e para não ser degradada necessita ser novamente convertida em uma molécula de dupla-fita através de

enzimas especiais. Essa conversão origina uma molécula de cDNA, DNA complementar, que mantém a seqüência original do mRNA sem a presença dos *introns*, e representa, desta forma as seqüências expressas em células, tecidos e órgãos.

Certamente há muitos outros genes expressos nos tecidos de glândula mamária e que ainda não tiveram suas seqüências determinadas ou tampouco suas funções definidas. Diversos destes genes podem ter sua função biológica associada à rede de ação do sistema imunológico e à resistência à mastite. Uma alternativa para reduzir o volume de seqüenciamento e contribuir para o entendimento biológico do funcionamento do tecido da glândula mamária é analisar apenas as ESTs. Essa tem provado ser uma estratégia eficiente para se obter informações do padrão de expressão gênica. A análise das seqüências ESTs permite identificar novos genes e estabelecer uma coleção daqueles que são expressos em um determinado tecido. Outra potencialidade de um projeto de seqüenciamento de ESTs é permitir a localização de novos marcadores em regiões de interesse do genoma, para os estudos com QTLs (Hatey *et al.*, 1998). Há informação suficiente nas extremidades dos fragmentos de cDNA para localizar essas seqüências nos cromossomos e elaborar mapas de ligação (Adams *et al.*, 1991). Finalmente, a coleção de clones de ESTs pode ser empregada em estudos com unidades de microarranjo de DNA, os *microarrays*, que possibilitam avaliar a expressão de milhares de genes simultaneamente.

Trata-se, portanto de um recurso importante para estudos de expressão gênica e seu emprego neste estudo está detalhado no Capítulo 7.

3 LOCALIZAÇÃO DE LOCOS GÊNICOS DE CARACTERÍSTICAS QUANTITATIVAS (QTLs) EM GADO DE LEITE. PREMISSAS E ESTRATÉGIAS ³

Resumo

Os avanços da biologia molecular têm aberto novas perspectivas de impacto nos sistemas de produção. A identificação dos genes de interesse e a possibilidade de compreensão do seu mecanismo de ação permitem empregá-los como “guias gênicos” utilizados na seleção de indivíduos portadores de características de interesse comercial. A premissa básica desta teoria é que a descendência direta de um ancestral diplóide e heterozigoto para um determinado loco gênico de característica quantitativa (QTL - *Quantitative Trait Locus*) de interesse pode ser dividida em duas classes com níveis de produção diferentes, cada uma delas recebendo um dos alelos parentais. O objetivo deste texto é apresentar uma revisão, na forma de introdução básica compreensível, e uma discussão a respeito de algumas alternativas dentre as diversas metodologias empregadas na localização de QTLs em Gado Leiteiro. Dezenas de milhares de regiões cromossômicas são candidatas a integrar QTLs de interesse econômico. A localização destas regiões cromossômicas e a possibilidade de empregá-las no melhoramento animal por meio da seleção assistida por marcadores moleculares (MAS – *Marker Assisted Selection*) representam uma contribuição efetiva que pode ser oferecida aos programas de melhoramento, aos produtores e aos consumidores em geral.

³ DA MOTA, A. F.; MARTINEZ, M. L.; COUTINHO, L. L. Localização de locos gênicos de características quantitativas (QTLs) em gado de leite. Premissas e estratégias. **Multiciência**, v.4, n.1, pp. 18-37, 2000. ARTIGO PUBLICADO.

Summary

Recent advances in molecular biology are opening new insights to impact production systems. The identification of major genes and the understanding of their biological action will allow their use as “gene markers” to select bulls and cows with commercial traits of interest. The basic premise of this theory is that a diploid heterozygote ancestral has two classes of descendants. Each half progeny receiving one of the two different production level alleles. The objective of this manuscript is to offer a basic and comprehensive review on the subject and a discussion through the methodologies used in Dairy Cattle for QTL mapping. Tens of thousands of different chromosome regions are putative to contain QTLs. Mapping these regions and using them as molecular markers in animal breeding by “marker assisted selection” (MAS) will effectively contribute to breeding programs, farmers and consumers.

3.1 Introdução

Os recentes avanços da biologia molecular e o refinamento das técnicas de avaliação do DNA dos animais domésticos têm aberto novas perspectivas de impacto nos sistemas de produção. Sobretudo, essa tecnologia permite compreender a relação biológica existente entre os genes e as características de interesse comercial. Mais especificamente, ocorre que as enzimas e proteínas codificadas no genoma estão relacionadas com os processos bioquímicos e fisiológicos envolvidos na produção, na reprodução e na sanidade dos animais. A identificação dos genes de interesse e a possibilidade de compreensão do seu mecanismo de ação permitem empregá-los como “guias gênicos”, que podem, dentre outras utilidades, ser empregados na seleção de indivíduos portadores de características de interesse comercial.

Os modelos aplicados em Genética Quantitativa baseiam-se em evidências de que as características métricas de interesse econômico são, em geral, expressas por muitos locos gênicos, denominados Locos de Características Quantitativas (QTL – “Quantitative Trait Loci”). As diferenças nos fenótipos dos animais são resultantes, portanto, da ação individual de vários genes, bem como do efeito das suas interações, na forma de entropia e pleiotropia, e também da ação dos fatores ambientais.

As características quantitativas dos organismos animais apresentam variação contínua nas progênes de ancestrais específicos. O processo de identificação dos genes responsáveis pela variação inclui análises dessas características e o emprego de ferramentas biométricas para

localização da região cromossômica (QTL) para a qual os alelos de um ou mais genes de função ainda desconhecida afetam a característica de interesse. A premissa básica desta teoria é que a descendência direta de um ancestral diplóide e heterozigoto para um determinado QTL de interesse pode ser dividida em duas classes com níveis de produção diferentes, cada uma delas recebendo um dos alelos parentais. Caso os alelos representem efeitos genéticos significativos sobre a característica de interesse, é estatisticamente provável que as duas classes de descendentes também apresentem diferença fenotípica na característica quantitativa.

Apesar da ampla discussão sobre o assunto, a localização de QTLs pressupõe conhecimentos em áreas distintas. O objetivo deste texto é apresentar uma revisão, na forma de introdução básica compreensível, e uma discussão a respeito de algumas alternativas dentre as diversas metodologias empregadas na localização de QTLs em Gado Leiteiro. Espera-se contribuir trazendo o assunto a um número maior de pessoas, já que é indiscutível a importância atual de realizar a localização de QTLs nas espécies domésticas. Dezenas de milhares de regiões cromossômicas são candidatas a integrar QTLs de interesse econômico. A localização destas regiões cromossômicas e a possibilidade de emprego no melhoramento animal por meio da seleção assistida por marcadores moleculares (MAS – “Marker Assisted Selection”) representam uma contribuição efetiva que pode ser oferecida aos programas de melhoramento, aos produtores e aos consumidores em geral.

3.2 Premissas das leis da Genética

As leis da Genética já estavam consagradas quando a Biotecnologia tornou-se realidade. No entanto, mesmo considerando que estas leis foram formuladas sem a comprovação da existência dos genes, mas somente baseadas em inferências, elas permanecem válidas. A genética Mendeliana foi concebida somente por meio da expressão dos genes na progênie para características visuais, e suas frequências estabelecidas com base nos fenótipos dos ancestrais. Gregor Mendel (1823-1884) formulou duas das três leis que servem de base aos estudos da genética de organismos sexuados.

3.2.1 A 1ª Lei da Genética

Em termos atuais, a primeira lei afirma que cada indivíduo carrega duas cópias de cada gene e que somente uma dessas duas cópias é transmitida para cada filho. O filho, portanto, irá

receber uma cópia completa de cada gene, recebida de cada progenitor. Assim, recebendo uma cópia do pai e uma da mãe, o filho terá restituído seu genótipo com duas cópias de cada gene.

Indivíduos com duas cópias de cada gene são chamados de diplóides.

A primeira lei foi formulada por Mendel e é confirmada quando os organismos diplóides produzem células germinativas com apenas uma cópia de cada gene - espermatozóides e óvulos. Durante o processo de divisão celular em que essa transformação ocorre - meiose - as duas cópias de cada gene irão separar-se, ou "segregar", uma da outra. Esse evento produziu o nome da primeira lei de Mendel: "Lei da Segregação".

A segregação pode ser observada na progênie de indivíduos heterozigotos, com dois alelos distintos. Metade da descendência direta receberá um dos alelos, enquanto a outra metade receberá o outro alelo, com a mesma probabilidade.

A primeira lei de Mendel descreve os acontecimentos referentes à herança gênica quando apenas um gene é considerado. A sua segunda lei foi formulada para descrever o que ocorre quando dois genes diferentes são transmitidos por herança.

3.2.2 A 2ª Lei da Genética

A segunda lei da Genética, também formulada por Mendel, em termos atuais, afirma que a segregação dos alelos em cada loco não exerce influência sobre a segregação em outro loco. Ou seja, as segregações que ocorrem em dois locos são independentes. Assim, a segunda lei de Mendel ou a "Lei da Segregação Independente" estabelece que locos independentes produzem segregação também independente. Indivíduos heterozigotos para dois locos, supondo loco A e loco B, produzem gametas com frequência igual à multiplicação da probabilidade de ocorrência dos gametas para cada loco. No caso em que o indivíduo apresenta genótipo A/a (loco A) e B/b (loco B), a chance de que um gameta seja A (loco A) é de 50% e a chance de que seja B (loco B) também é de 50%. Portanto, a chance de que um gameta seja AB (loco A e loco B) será de $0,50 \times 0,50 = 0,25$, ou 25%. A Figura 2 apresenta uma ilustração da segregação independente.

Sabe-se que a segunda lei de Mendel somente é válida no caso de genes que não estão ligados no mesmo cromossomo. Quando dois genes com dois alelos têm ligação cromossômica, as frequências, para cada uma das possíveis combinações dos alelos, serão desviadas de 25%. Duas combinações de alelos irão representar os arranjos de ligação dos cromossomos parentais e estas serão, cada uma, transmitida a uma frequência superior a 25%. As duas classes restantes representam os arranjos recombinantes e irão ocorrer a uma frequência inferior a 25%. No caso

extremo de ligação absoluta, somente as duas classes parentais serão transmitidas, cada uma à frequência de 50%.

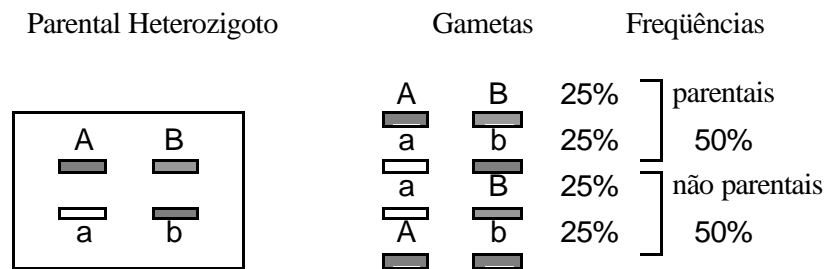


Figura 2 - Diagrama da segregação independente entre locos sem ligação

3.2.3 A 3ª Lei da Genética

Para os casos de genes em ligação, as fases de associação ou repulsão são consequência da localização dos genes no mesmo cromossomo. Alelos em fase de associação estão localizados na metade homóloga do cromossomo parental, ao contrário daqueles em fase de repulsão.

A Figura 3 apresenta um diagrama das frequências dos gametas produzidos por um indivíduo heterozigoto para os locos A e B, com uma frequência de recombinação de 10%.

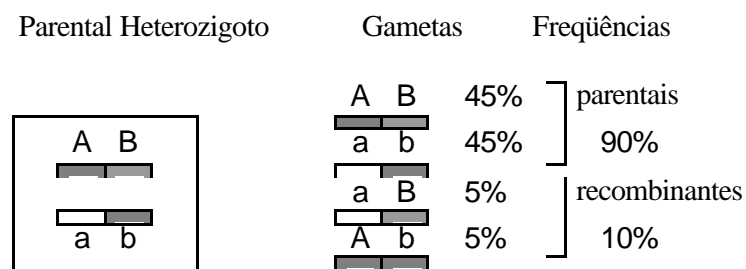


Figura 3 - Diagrama da segregação entre locos com ligação e frequência de recombinação de 10%

Na ilustração apresentada na Figura 3, um indivíduo heterozigoto cujo genótipo pode ser descrito como sendo AB/ab produz 90% de seus gametas da mesma forma apresentada no seu genoma e 10% dos gametas produzidos apresentam uma forma não-parental, ou recombinante.

Isso é consequência da ocorrência do fenômeno de “crossing-over” durante a divisão celular, meiose, que deu origem aos gametas. Quanto menor a distância entre os genes ligados no mesmo cromossomo, como mostrado na Figura 3, menores são as chances de que os eventos de “crossing-over” possam ocorrer entre os dois genes em ligação. Portanto, pela genotipagem da progênie de indivíduos cruzados e pela determinação da frequência de recombinação, torna-se possível estimar a distância genética entre os genes ligados.

As distâncias genéticas são medidas em centimorgans (cM), sendo 1 cM definido como a distância entre dois locos cuja recombinação ocorre em 1% dos indivíduos da progênie. Importante salientar que mesmo sendo a recombinação proporcional à quantidade de DNA que separa dois locos no mesmo cromossomo, quando esta distância for grande, a frequência de recombinação irá aproximar-se de 50%, tornando os resultados similares aos observados entre locos sem ligação.

Há ainda que se considerar a situação em que ocorre mais de um evento de “crossing-over” entre dois genes durante a meiose (recombinação dupla). No entanto, à medida que a distância genética é reduzida, a frequência de ocorrência de mais de um evento de “crossing-over” é reduzida exponencialmente. Para dois genes distantes 10 cM entre si, as chances de ocorrência de dois eventos de “crossing-over” entre eles é de apenas 1%. Assim, neste texto, com a finalidade de simplificação, não serão consideradas as ocorrências de recombinações duplas entre o marcador genético e o QTL.

As três leis da genética apresentadas resumem a compreensão mínima necessária de como se emprega o conhecimento do desequilíbrio da ligação e da frequência de recombinação dos alelos para utilizar marcadores genéticos na localização de QTLs de interesse. O raciocínio é equivalente para o emprego de marcadores moleculares polimórficos e exemplos serão apresentados a seguir para os desenhos experimentais especialmente construídos para localizar novos QTLs em Gado Leiteiro.

Outra questão importante dentro do contexto de localização de QTL, e que deverá ser alvo de escolha durante o processo, diz respeito aos marcadores empregados.

3.3 Marcadores moleculares

Como a localização e identificação de um QTL não são conhecidas *a priori*, não há como discriminar os genótipos do QTL dos indivíduos parentais ou de sua progênie. Assim, são empregadas marcas com localização genômica conhecida, marcadores moleculares, estrategicamente dispostos ao longo do genoma. A hipótese empregada para a localização de um

QTL é de que estas marcas, ou marcadores moleculares polimórficos, estarão próximas e, portanto, suficientemente ligadas ao QTL, de forma que, na maioria da progênie de um indivíduo heterozigoto para o marcador, será possível associar as variações na característica quantitativa de interesse com o genótipo do marcador, que será indicativo do genótipo do próprio QTL. A associação entre os genótipos do QTL e do marcador não ocorrerá quando tiver ocorrido “crossing-over” na meiose da formação de gametas da geração anterior, indicativo do desequilíbrio da ligação. A frequência de ocorrência de recombinação causada por “crossing-over” dos cromossomos dos indivíduos parentais permitirá estabelecer a estimativa da distância genética entre o marcador e o QTL. Como a localização do marcador é conhecida, será possível conhecer a estimativa da localização do QTL.

3.3.1. Marcadores microssatélites

Sem a existência da reação de PCR, a maioria das regiões microssatélites não teria utilidade como marcadores genéticos (Silver, 1995). As regiões microssatélites são marcadores moleculares especiais que consistem de pequenas seqüências com um a quatro nucleotídeos, repetidas em “tandem” ao longo do genoma. São mais frequentes nos genomas de organismos eucarióticos e são bem distribuídas ao acaso. Os marcadores microssatélites constituem a classe de marcadores que apresenta o maior grau de polimorfismo entre os marcadores disponíveis (Ferreira & Grattapaglia, 1996). Regiões do DNA podem ser amplificadas por PCR com “primers” específicos, complementares a regiões flanqueadoras às seqüências repetidas da região microssatélite. Esses fragmentos amplificados apresentam polimorfismo extenso, resultante de diferentes números de seqüências de nucleotídeos repetidas. Cada fragmento amplificado de tamanho diferente representa um alelo diferente do mesmo loco. A Figura 4 apresenta um diagrama representando os possíveis alelos de um microssatélite e como os genótipos podem ser detectados. A diferença existente entre o comprimento das seqüências na região microssatélite produz fragmentos amplificados com diferentes mobilidades, o que permite identificar padrões ou bandas por meio da eletroforese em gel de poliacrilamida.

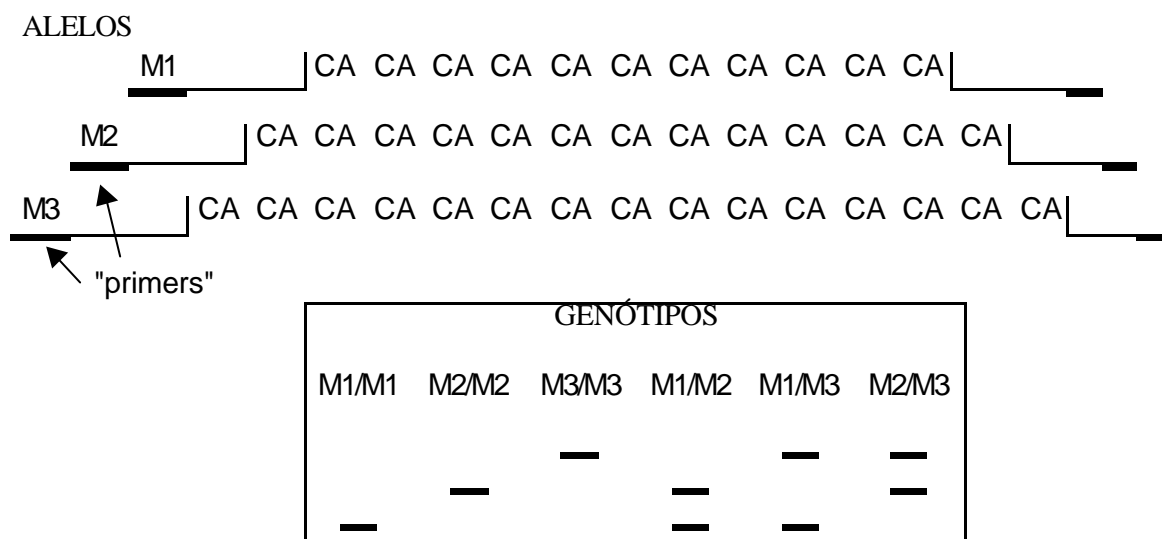


Figura 4 - Alelos do microsatélite e sua detecção por PCR e eletroforese. Três diferentes alelos compostos da repetição dos nucleotídeos "CA". Os sítios de anelamento dos "primers" específicos empregados na amplificação estão representados junto à identificação dos alelos. Os padrões do gel representam possíveis genótipos com combinações dos alelos indicados

Considerando a estimativa de que nos genomas das espécies de animais domésticos podem estar disponíveis de 50.000 a 150.000 regiões microsatélites (Hillel et al., 1992), a facilidade em obter a genotipagem de animais pela reação de PCR e também o elevado conteúdo de informação de polimorfismo, qualquer população onde ocorre segregação pode ser alvo de estudo de ligação ou de mapeamento genético empregando microsatélites.

Segundo Ferreira & Grattapaglia (1996) a limitação básica que existe para aplicação mais ampla destes estudos empregando microsatélites recai na necessidade de trabalho técnico especializado e equipamentos laboratoriais sofisticados. Segundo os mesmos autores, no entanto, esta técnica será cada vez mais utilizada em um número cada vez maior de espécies.

Baseado no exposto, considerando que é possível empregar populações construídas nos estudos para detectar e mapear QTL de interesse por meio da análise de ligação e através do emprego de marcadores moleculares microsatélites, resta estabelecer o desenho experimental a ser utilizado na construção da população de interesse e a metodologia necessária para estimar a localização precisa do QTL.

3.4 Populações construídas

Há diversos métodos de detecção de QTL, como: Distribuição Multimodal em seleção divergente; Distribuição Bimodal em retrocruzamento com seleção; Distribuição não-normal; Heterogeneidade da variância dentro de famílias; e Semelhança entre a progênie e os pais (Falconer & MacKay, 1996). No entanto, a abordagem com maior poder em detectar um gene de efeito principal que afeta uma característica quantitativa é realizada pela Análise Complexa da Segregação, que será aqui abordada em maior profundidade. Uma das soluções para permitir a análise da segregação é a construção de populações formadas a partir do cruzamento entre duas diferentes linhagens consanguíneas ou entre raças puras (P1 x P2), gerando uma produção de indivíduos – teoricamente - heterozigotos na progênie (F1). A partir da geração F1, há duas opções de novos cruzamentos e progênies (Knapp et al., 1990; Soller & Beckmann, 1983). A primeira delas é o desenho experimental de retrocruzamento (BC), quando a geração F1 será cruzada novamente com os parentais (P1 e P2, sendo BC1 e BC2, respectivamente). A segunda delas é a produção de uma geração F2, obtida pelo cruzamento entre indivíduos oriundos da geração F1 (F1 X F1).

3.4.1 O caso de apenas um loco marcador

Além da escolha do desenho experimental, podem ser utilizados diferentes modelos genéticos (Arús & Moreno-González, 1993). Para exemplificação, aqui será suposta a situação em que se emprega apenas um marcador molecular microssatélite (M) e que este possui apenas dois alelos ("M" e "m"). Será considerado que o QTL de interesse possui apenas dois alelos ("Q" e "q"), supondo também uma frequência de recombinação (r), os indivíduos de F1 obtidos do cruzamento entre P1 e P2 teriam o diagrama de seus cromossomos conforme apresentado na Figura 5, a seguir.

Os tipos de gametas em F1 e as frequências de ocorrência são apresentados na Tabela 2.

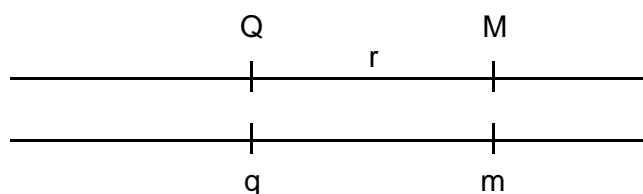


Figura 5 - Diagrama do segmento cromossômico de indivíduos da geração F1 com genótipo Qq/Mm para o QTL (Q) e o marcador (M), respectivamente, quando "r" é a frequência de recombinação

Tabela 2 . Tipos de gametas e suas frequências derivadas da meiose em uma geração F1

Gametas	Frequências esperadas
QM	$(1 - r) / 2$
qM	$r / 2$
Qm	$r / 2$
qm	$(1 - r) / 2$

3.4.2 Desenho experimental F2 para um modelo supondo apenas um marcador molecular

Supondo um desenho experimental de F2, havendo o cruzamento entre si dos gametas descritos na Tabela 2, seguindo Soller & Beckmann (1983) e admitindo que +a, d e -a sejam os efeitos genotípicos para os genótipos QQ, Qq e qq, respectivamente, os genótipos da geração F2 produzida, bem como suas frequências e valor genotípico são apresentados na Tabela 3.

Tabela 3 . Genótipos, frequências e valor genotípico esperados na geração F2 obtida do cruzamento entre F1 x F1

Genótipo	frequências	valor genotípico
QQMM	$(1 - r)^2 / 4$	+ a
QQMm	$r (1 - r) / 2$	+ a
QQmm	$r^2 / 4$	+ a
QqMM	$r (1 - r) / 2$	d
QqMm	$(1 - 2r + 2r^2) / 2$	d
Qqmm	$r (1 - r) / 2$	d
qqMM	$r^2 / 4$	- a
qqMm	$r (1 - r) / 2$	- a
qqmm	$(1 - r)^2 / 4$	- a

O valor da estimativa da taxa de recombinação, indicativa da distância genética existente entre o marcador e o QTL, é obtido por cálculo algébrico, agrupando-se os indivíduos em três classes de marcadores, a saber: MM, Mm e mm. No entanto, esta estimativa só tem sentido quando houver diferença significativa entre as médias observadas para a característica fenotípica de interesse para as classes de marcadores mencionadas. As médias esperadas para as três classes de marcadores, obtidas pela combinação entre as frequências e os valores genotípicos apresentados na Tabela 3, estão na Tabela 4.

Tabela 4 . Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular.

Classe do marcador	média esperada da característica
MM	$\mu + a (1 - 2r) + 2rd (1 - r)$
Mm	$\mu + d (1 - 2r + 2r^2)$
mm	$\mu + a (2r - 1) + 2rd (1 - r)$

Contrastes ortogonais realizados entre as médias observadas nos indivíduos classificados para o marcador permitem estimativas dos efeitos do QTL, referentes aos efeitos aditivos e de dominância. Considerando [MM] como sendo a média da característica de interesse observada nos indivíduos classificados como apresentando o genótipo MM para o marcador. Considerando, de forma análoga, a média [Mm] para os indivíduos da classe Mm e a média [mm] para os da classe mm, os contrastes são apresentados na Tabela 5.

Tabela 5 . Contrastes ortogonais empregados no cálculo dos efeitos genéticos aditivo (a) e devido à dominância (d) do QTL

Contrastes Ortogonais

$$\begin{aligned} ([MM] - [mm]) / 2 &= a (1 - 2r) \\ [Mm] - (([MM] - [mm]) / 2) &= d (1 - 2r)^2 \\ (2[Mm] - [MM] - [mm]) / ([MM] - [mm]) &= (d / a) (1 - 2r) \end{aligned}$$

Para a situação em que se emprega um desenho experimental de retrocruzamento (BC), por meio do retrocruzamento com P1 e P2, respectivamente BC1 e BC2, com um marcador molecular, há alterações nas freqüências dos genótipos, nos valores genotípicos esperados e nas médias esperadas para o marcador, bem como para os contrastes.

3.4.3 Desenho experimental de retrocruzamento (BC) para modelo com um marcador

Considerando um retrocruzamento realizado com o cruzamento entre uma geração F1 (QqMm) e as linhas ou raças parentais P1 (QQMM) e P2 (qqmm), as freqüências genotípicas que seriam observadas estão dispostas na Tabela 6.

A separação dos indivíduos em classes de genótipo para os marcadores, de acordo com o retrocruzamento, BC1 ou BC2, estão apresentadas na Tabela 7, onde μ_{BC1} e μ_{BC2} representam as médias dos valores genotípicos para os demais genes que contribuem para a característica de interesse, respectivamente para BC1 e BC2. Os valores estimados para contrastes ortogonais estão apresentados na Tabela 8.

Segundo Arús & Moreno-Gonzáles (1993), o método empregando apenas um marcador molecular apresenta algumas desvantagens, dentre elas destaca-se que pode ocorrer que os

valores de "a" e "d" sejam subestimados e quando o valor de "r" for maior, esse viés também será maior.

Tabela 6 . Genótipos e frequências genotípicas no desenho BC, sendo BC1= (QqMm x QQMM) e BC2 = (QqMm x qqmm)

Genótipo	Frequências Genotípicas	
	BC1	BC2
QQMM	$(1 - r) / 2$	
QQMm	$r / 2$	
QqMM	$r / 2$	
QqMm	$(1 - r) / 2$	
QqMm		$(1 - r) / 2$
Qqmm		$r / 2$
QqMm		$r / 2$
Qqmm		$(1 - r) / 2$

Tabela 7 . Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular para desenho BC com um marcador

Classe do marcador	média esperada da característica
MM (BC1)	$\mu_{BC1} + a (1 - r) + r d$
Mm(BC1)	$\mu_{BC1} + d (1 - r) + r a$
Mm(BC2)	$\mu_{BC2} + d (1 - r) - r a$
mm(BC2)	$\mu_{BC2} + a (r - 1) + r d$

Tabela 8 . Contrastes ortogonais empregados no cálculo dos efeitos genéticos aditivo (a) e devido à dominância (d) do QTL, para desenho BC com um marcador

Contrastes Ortogonais
$([MM(BC1)] - [Mm(BC1)] + [Mm(BC2)] - [mm(BC2)]) / 2 = a (1 - 2r)$
$([Mm(BC1)] + [Mm(BC2)] - [MM(BC1)] - [mm(BC2)]) / 2 = d (1 - 2r)$

Outra desvantagem do modelo com apenas um marcador é que apesar de haver indicações da distância entre o marcador e o QTL, sua localização não permanece estabelecida, podendo este estar a jusante ou a montante do marcador. Baseado nas desvantagens e limitações do método utilizando apenas um marcador, Lander & Botstein (1989) e Knapp et al. (1990) sugeriram o emprego de marcadores flanqueando a região candidata ao QTL.

3.4.4 Localização do QTL empregando mais de um marcador molecular flanqueando a região candidata

Com a utilização de marcadores microsátélites estrategicamente dispostos ao longo do genoma, diversos estudos têm sido desenvolvidos com maior densidade na localização de QTLs (Georges, 1995).

O método consiste em cruzar duas linhagens ou raças (P1 x P2), de genótipos MMNN (P1) e mmnn (P2) para os marcadores M e N localizados no mesmo cromossomo e com relativa proximidade. Existindo QTL (Q) localizado entre os marcadores (Figura 6), o genótipo da geração F1 está apresentado na Tabela 9, em que r_1 é a frequência de recombinação entre o marcador M e o QTL e r_2 a frequência de recombinação entre o marcador N e o QTL.

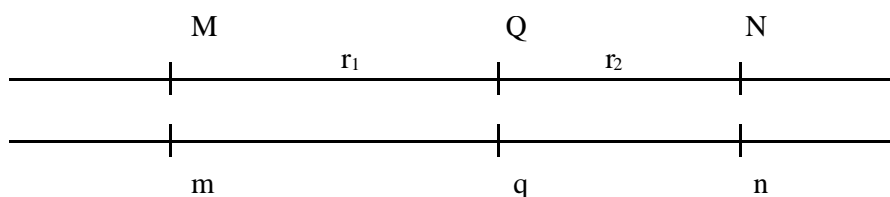


Figura 6 - Diagrama do segmento cromossômico de indivíduos da geração F1 com genótipo MmQqNn para o marcador (M) e marcador (N), flanqueando o QTL (Q), quando r_1 e r_2 são as frequências de recombinação entre o marcador M e o QTL e a frequência de recombinação entre o marcador N e o QTL, respectivamente

Tabela 9. Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular para desenho BC com um marcador

Classe do marcador	média esperada da característica
MM (BC1)	$\mu_{BC1} + a (1 - r) + r d$
Mm(BC1)	$\mu_{BC1} + d (1 - r) + r a$
Mm(BC2)	$\mu_{BC2} + d (1 - r) - r a$
mm(BC2)	$\mu_{BC2} + a (r - 1) + r d$

3.4.5 Desenho experimental de retrocruzamento (BC) para modelo com mais de um marcador flanqueando a região

Considerando a alternativa de empregar retrocruzamento com os parentais P1 e P2, as frequências genóticas e os genótipos para BC1 (F1 x P1) e BC2 (F1 x P2) são apresentados na Tabela 10. Conforme mencionado, as tabelas consideram desprezíveis as chances de ocorrência de mais de um evento de "crossing-over" entre os marcadores (ou seja, $r_1 \times r_2 = 0$).

A Tabela 11 permite visualizar as classes dos marcadores detectáveis nos indivíduos e as médias esperadas para a característica de interesse, sendo "a" e "d" efeitos genéticos aditivo e devido à dominância do QTL, respectivamente.

Considerando a distância entre os marcadores M e N como sendo r, de forma que na ausência de duplas recombinações, $r = r_1 + r_2$, de acordo com Arús & Moreno-González (1993), teremos que a proporção r_1/r será dada por: $r_1/r = (\theta_1 - \theta_2 + \theta_3 - \theta_4) / 4 (\theta_1 - \theta_4) + (\theta_7 - \theta_8 + \theta_5 - \theta_6) / 4 (\theta_5 - \theta_8)$; e os valores de "a" e "d", calculados de acordo com as expressões: $a = (\theta_1 - \theta_4 + \theta_5 - \theta_8) / 2$; $d = (\theta_5 - \theta_8 - \theta_1 + \theta_4) / 2$. Com base nas expressões apresentadas, pode-se realizar a estimativa de localização do QTL, bem como seu valor genético aditivo e de dominância.

O modelo de F2 encontra-se descrito em Carbonell et al. (1992), com marcadores flanqueando a região do QTL, desenvolvido atendendo o mesmo raciocínio empregado para o desenho BC. Vários outros autores publicaram trabalhos considerando os desenhos mencionados e indicando testes apropriados (Haley & Knott, 1992; Lander & Botstein, 1989; Lebowitz et al., 1987; Soller & Genizi, 1978; Visscher & Haley, 1996; Weller, 1986). Segundo Falconer & MacKay (1996), o desenho de F2 teria vantagens sobre o desenho BC, já que aquele necessita de menor número de informações para que os testes de detecção do QTL atinjam poder estatístico equivalente.

Há também de ser considerada a possibilidade de ser realizada a genotipagem seletiva dos animais produzidos a partir do cruzamento de linhagens ou raças. A genotipagem seletiva foi proposta por Lebowitz et al. (1987) e constitui uma alternativa interessante na redução de custos mediante a genotipagem dos indivíduos distribuídos nos extremos da curva de distribuição dos valores da característica quantitativa avaliada. Darvasi & Soller (1992) demonstraram que a genotipagem seletiva pode reduzir o número de genotipagens para um mesmo poder do teste. Bovenhuis & Spelman (1996) encontraram resultados semelhantes. Segundo Hillel et al. (1992), empregando-se 5% da população total com a genotipagem de indivíduos distantes ao menos dois desvios-padrão da média, essas informações contribuem com 28% do total da informação de ligação disponível.

Tabela 10. Genótipos e frequências genóticas no desenho BC, sendo BC1= (MmQqNn x MMQQNN) e BC2 = (MmQqNn x mmqqnn). r1 e r2 são as frequências de recombinação entre o loco M e o QTL (Q) e a frequência de recombinação entre os locos N e Q, respectivamente

Genótipo	Frequências Genóticas	
	BC1	BC2
MMQQNN	$(1 - r_1 - r_2) / 2$	
MMQQNn	$r_2 / 2$	
MMQqNn	$r_1 / 2$	
MmQQNN	$r_1 / 2$	
MmQqNN	$r_2 / 2$	
MmQqNn	$(1 - r_1 - r_2) / 2$	
MmQqNn		$(1 - r_1 - r_2) / 2$
MmQqnn		$r_2 / 2$
Mmqqnn		$r_1 / 2$
mmQqNn		$r_1 / 2$
mmqQnN		$r_2 / 2$
mmqqnn		$(1 - r_1 - r_2) / 2$

Tabela 11 . Valores esperados das médias da característica de interesse de acordo com as classes dos marcadores moleculares para desenho BC com marcadores flanqueando o QTL.

Retrocruzamento	Classe do marcador	média esperada da característica
BC1	MMNN	$\theta_1 = \mu_{BC1} + a$
	MMNn	$\theta_2 = \mu_{BC1} + d r_1 + a r_2$
	MmNN	$\theta_3 = \mu_{BC1} + d r_2 + a r_1$
	MmNn	$\theta_4 = \mu_{BC1} + d$
BC2	MmNn	$\theta_5 = \mu_{BC2} + d$
	Mmnn	$\theta_6 = \mu_{BC2} + d r_2 - a r_1$
	mmNn	$\theta_7 = \mu_{BC2} + d r_1 - a r_2$
	mmnn	$\theta_8 = \mu_{BC2} - a$

3.4.6 Desenho de filhas (*Daughter Design*)

Especialmente em gado de leite, os desenhos experimentais são fundamentais. Os touros não expressam a característica de interesse dos animais, que é a lactação. Há também uso freqüente de inseminação artificial e a conseqüente multiplicação em muitos rebanhos de parte do genoma dos touros em níveis elevados.

A teoria do "Desenho de Filhas" consiste na pressuposição de existência de um touro heterozigoto para um determinado marcador e que possui um grande número de descendentes. É possível, portanto, testar a presença de ligação entre marcador e QTL na sua progênie por meio da diferença entre as médias dos valores da característica de interesse nas filhas, sendo que estas podem ter recebido um ou outro dos alelos do marcador do touro. Na maioria dos casos, é interessante incluir maior número de touros no estudo, para reduzir as chances de que o touro seja heterozigoto para o marcador e homozigoto para o QTL. Uma análise da variância entre os

subgrupos de filhas que receberam distintos alelos permite avaliar a significância da diferença entre as médias dos fenótipos observados. O diagrama da Figura 7 ilustra o Desenho de Filhas. No Desenho de Filhas, todos os animais são genotipados para o marcador e são coletadas informações da produção da característica de interesse das filhas.

Segundo Soller & Genizi (1978) e Weller et al. (1990) para se detectar um QTL de efeito médio de substituição de alelos inferior a $\frac{1}{2}$ desvio-padrão será necessário genotipar e acompanhar a produção de milhares de filhas de diversos touros.

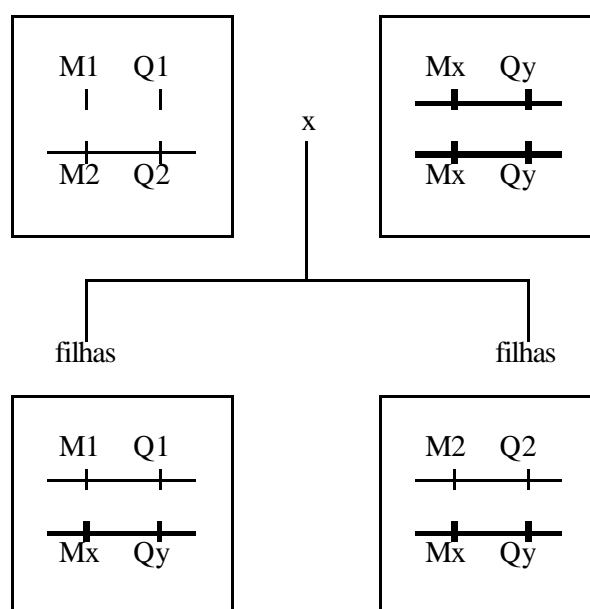


Figura 7 - Diagrama do Desenho de Filhas. M1 e M2 são alelos do marcador molecular do touro. Mx alelos dos marcadores das matrizes. Q1 e Q2 alelos do QTL do touro. Qy alelos do QTL das matrizes. — Cromossomo do touro; — cromossomo das matrizes

3.4.7 Desenho de netas (*Granddaughter Design*)

Neste desenho, os touros possuidores de filhos que também são utilizados nos programas de inseminação artificial, são genotipados em busca de marcadores heterozigotos. Os seus filhos também são genotipados e são coletadas informações da característica fenotípica de interesse somente nas netas. Segundo Weller et al. (1990), o número de genotipagens necessárias no

Desenho de Netas é inferior ao necessário no Desenho de Filhas, mantendo-se o mesmo poder do teste. Os autores também consideram vantajoso utilizar apenas sêmen dos touros para a genotipagem dos animais, sem a necessidade de coleta de sangue das matrizes para essa finalidade.

Na Figura 8 está ilustrado o diagrama do Desenho de Netas.

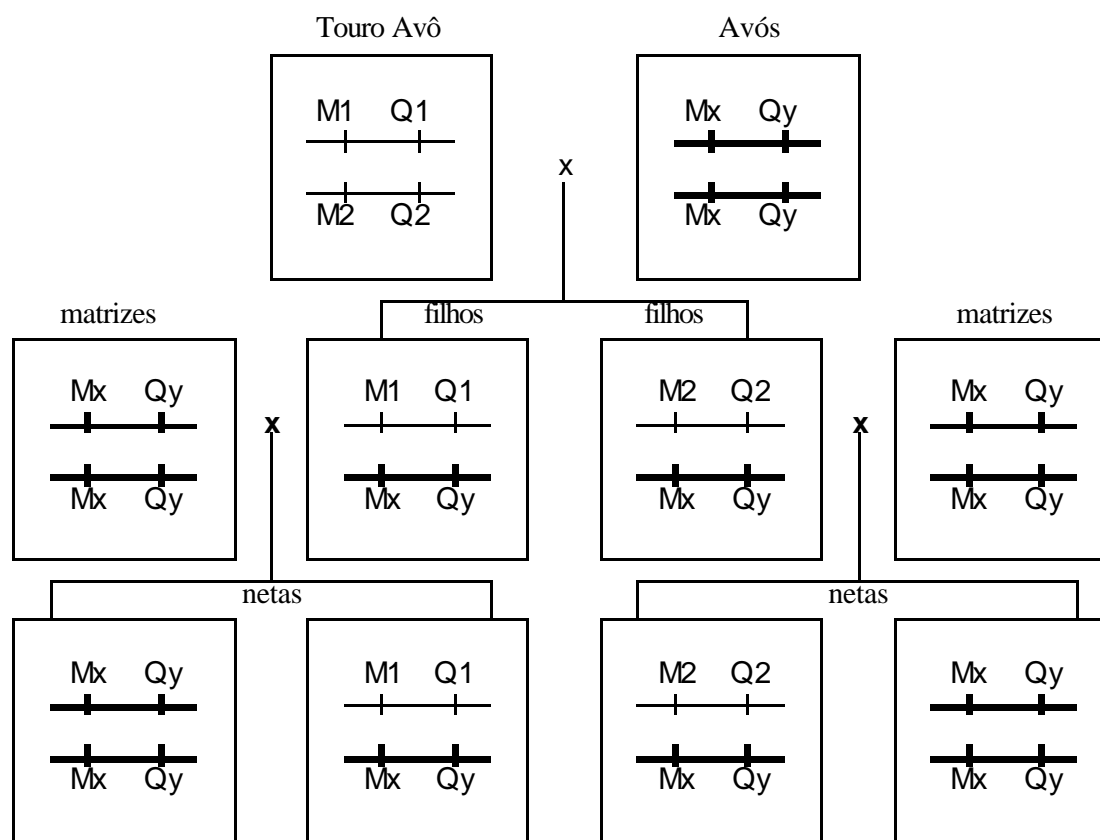


Figura 8 - Diagrama do Desenho de Netas. M1 e M2 são alelos do marcador molecular do touro. Mx alelos dos marcadores das matrizes. Q1 e Q2 alelos do QTL do touro e dos filhos. Qy alelos do QTL das matrizes. — Cromossomo do touro; — cromossomo das matrizes

O Desenho de Netas utiliza registros genealógicos e, segundo Georges et al. (1995), com a redução da necessidade de genotipagens dos animais, para características de baixa

herdabilidade, esse desenho experimental proporciona a redução de aproximadamente 3,5 vezes no tempo destinado para genotipagem em relação ao "Desenho de Filhas".

Segundo Weller et al. (1990), uma questão importante reside em conhecer a estrutura da população para julgar o desenho mais apropriado para análise. O Desenho de Filhas requer entre 5 e 20 touros com 200 a 2.000 filhas cada um para sua realização com razoável poder de detecção. Segundo este autor, a aplicação do Desenho de Filhas seria interessante em rebanhos numerosos e com uso intensivo da inseminação artificial, com distâncias reduzidas entre os locais de criação e com reduzido número de touros provados usados na inseminação a cada ano. O Desenho de Netas seria mais apropriado em rebanhos nos quais os touros classificados como "elite" possuíssem filhos também empregados nos programas de inseminação artificial e preferencialmente pudesse ser obtido sêmen desses animais, em uma única coleta, em centrais de inseminação artificial. Esse desenho também seria mais vantajoso nos casos em que há um grande número de matrizes cujas informações zootécnicas estejam reunidas e organizadas. Com relação aos métodos de análise do QTL, podem ser realizadas análises pelo método da análise de variância (Weller et al., 1990), pelo método da máxima verossimilhança (Weller, 1986), ou pelo método do Qui-quadrado (Geldermann, 1975).

3.5 O poder para se detectar QTLs

Uma das preocupações recentes entre os autores com relação aos distintos desenhos e modelos experimentais empregados para a localização de QTLs refere-se ao poder do teste para detectar a ligação deste com o marcador. Nos estudos com vários marcadores moleculares distribuídos ao longo do genoma, tem sido empregado com frequência o método da máxima verossimilhança. A principal diferença em relação ao método dos quadrados mínimos é que a máxima verossimilhança independe da distribuição dos dados. As estimativas de máxima verossimilhança (MLE - "Maximum Likelihood Estimation") são obtidas em análises realizadas por computadores e alguns programas foram desenvolvidos por grupos de pesquisa e estão disponíveis mediante solicitação. Usualmente as análises destas estimativas geram gráficos de "LOD scores" (LOD, para "log odds") que apresentam a verossimilhança da ligação entre dois locos geneticamente próximos um do outro. O "LOD score" é o logaritmo na base dez de uma razão, na qual o numerador corresponde às chances de que os dados provenham de um QTL, enquanto o denominador corresponde às chances de que não exista nenhum QTL ligado aos

marcadores para a característica em estudo. Valores de "LOD scores" que evitam a detecção de um QTL falso-positivo com 95% de confiabilidade quando são testados 60 marcadores distribuídos em uma região do genoma com 1.200 cM foram estimados como sendo de 2,4 (Lander & Botstein, 1989). O gráfico é construído calculando-se os valores de "LOD scores" para posições definidas dentro dos intervalos entre os marcadores.

A possibilidade de que sejam detectados QTLs falsos-positivos tem levado vários autores a realizarem estudos para estabelecer o poder dos testes estatísticos relativos a diferentes números de animais utilizados na análise (Martinez, 1998; Moody et al., 1997; Simpson, 1989; Soller & Genezi, 1978; Weller et al., 1990).

3.6 Considerações finais

Uma abordagem tradicional em genética quantitativa é avaliar os fenótipos dos animais e dados a respeito de sua genealogia e inferir sobre a base genética que causa a variação fenotípica observada. A abordagem alternativa trazida pela genética molecular inclui a genotipagem direta dos animais nesse contexto, mas também a avaliação de como a localização dos genes e a elaboração de mapas genéticos podem auxiliar a explicar a variação fenotípica.

A identificação de genes de efeito principal que são responsáveis por características métricas comercialmente importantes em animais domésticos irá requerer a identificação de locos candidatos. A localização pode ser realizada em populações abertas ou em populações construídas para este fim, no entanto uma das dificuldades nas metodologias para se detectar QTLs envolve o tamanho dos experimentos, mesmo para detectar efeitos genéticos principais. Outra complicação é que ainda que animais apresentem o mesmo genótipo para os principais genes relacionados com determinada característica quantitativa, verifica-se variação nos níveis de produção devido a fatores não-genéticos ou ambientais. Há também que se considerar que nos animais de grande porte os intervalos entre cada geração são grandes e assim, as populações construídas com grandes animais trazem a necessidade de experimentos com múltiplas gerações, o que pode ser atingido somente após muitos anos de estudo.

Dentre as principais possibilidades de utilização do conhecimento sobre o genoma dos animais domésticos pode-se destacar o uso da seleção de animais assistida por marcadores moleculares, especialmente quando as características de interesse apresentarem baixa herdabilidade ou correlações genéticas indesejáveis. Também se pode destacar a possibilidade de seleção precoce de animais destinados à reprodução. Outro importante aspecto a mencionar

refere-se ao isolamento, clonagem e manipulação direta de genes que controlem características quantitativas desejáveis. A engenharia genética de animais como forma de alterar suas características produtivas pode multiplicar muitas vezes os níveis de produção. Tudo isso depende inicialmente da localização correta e eficiente dos QTLs e posteriormente da identificação dos genes específicos responsáveis pelo fenótipo de interesse.

Não obstante às dificuldades, experimentos dessa natureza vêm sendo conduzidos, pois não há como declinar de tarefa científica tão fundamental quanto a de contribuir na construção dos mapas genéticos das espécies comerciais mais importantes. As perspectivas existentes na vasta literatura sobre o assunto são de que as próximas décadas irão se beneficiar muito das informações básicas a respeito da estrutura e organização dos genomas. Inúmeras possibilidades advindas desse conhecimento e da exploração eficiente da variabilidade genética estão previstas.

3.7 Conclusões

- O emprego de marcadores moleculares permite identificar regiões cromossômicas que contêm gene(s) responsáveis pela ocorrência de mastite.
- Populações construídas de gado de leite com desenhos experimentais pré-estabelecidos são os melhores recursos para essa finalidade, já que é nesses casos que se verifica o maior desequilíbrio de ligação.
- O estudo de populações construídas com o estudo do fenótipo e marcadores moleculares constitui uma alternativa inteligente de utilizar a variação genética e selecionar indivíduos superiores, com possibilidades de impacto nos sistemas de produção.

4 DISTRIBUIÇÃO DOS ALELOS DO GENE BOLA-DRB3 EM BOVINOS DA RAÇA GIR LEITEIRA (*Bos indicus*)⁴

* A seqüência de nucleotídeos reportada foi submetida ao GenBank e recebeu número de acesso: AF209710.

Resumo

Bovinos da raça Gir Leiteira caracterizam-se como uma raça bem adaptada às condições tropicais e para a qual não há registro de informações moleculares a respeito do complexo principal de histocompatibilidade (MHC – “Major Histocompatibility Complex”). O segundo “exon” do gene BoLA-DRB3 foi amplificado por meio da reação em cadeia da polimerase (PCR- Polymerase Chain Reaction) em amostras de DNA de 28 vacas da raça Gir Leiteira. Duas metodologias de genotipagem foram empregadas: Seqüenciamento direto dos produtos de PCR (PCR-DS); e clonagem seguida de seqüenciamento dos fragmentos amplificados do gene BoLA-DRB3 (CLO-S). Dezesete alelos foram identificados, inclusive alguns somente recentemente descritos em zebuínos. As freqüências alélicas variaram entre 2 e 18 %. Os alelos mais freqüentes foram *3601 (18%), *2201 (14%) e *2101 (11%).

Summary

Brazilian dairy Gir (*Bos indicus*) cattle are a tropical, well-adapted breed, for which no information on the major histocompatibility complex (MHC) is currently available. The second exon of the bovine lymphocyte antigen (BoLA-DRB3) was amplified by polymerase chain reaction (PCR) of DNA samples from 28 Brazilian dairy Gir cows. Two experimental genotyping approaches were used: direct sequencing of PCR gene products (PCR-DS) and

⁴ DA MOTA, A. F.; GABRIEL, J. E; MARTINEZ, M. L.; COUTINHO, L. L. Distribution of bovine lymphocyte antigen (BoLA -DRB3) alleles in Brazilian dairy Gir cattle (*Bos indicus*) **European Journal of Immunogenetics**, v. 29, p. 223-227. 2002. ARTIGO PUBLICADO.

sequencing of cloned PCR fragments (CLO-S). Results demonstrate the viability of both typing approaches. PCR-DS allowed typing of 39% of the animals while the remainders were genotyped by CLO-S. Seventeen BoLA-DRB3 alleles were assigned, including some that were only recently described for zebu cattle. Allelic frequencies ranged from 0.02 to 0.18. The most frequent alleles were *3601 (frequency = 0.18), *2201 (0.14) and *2101 (0.11).

4.1 Introdução

No Brasil, os bovinos da raça Gir foram inicialmente importados da Índia no início do século XX. Esses animais apresentam características como corpo compacto, barbela desenvolvida, cabeça convexa e pelagem rosilha. Segundo Santiago (1987), em 1938 a seleção para fêmeas destinadas à produção leiteira foi iniciada na Estação Experimental de Umbuzeiro, no estado brasileiro da Paraíba. Nos dias atuais, a raça apresenta animais bem adaptados às condições tropicais do Brasil e o desempenho leiteiro de vacas puras situa-se ao redor de 2.700 kg de leite por lactação ajustada para 305 dias (<http://www.cnpq.embrapa.br/gir>).

A mastite representa a doença mais comum a afetar os rebanhos leiteiros do Brasil (Veiga et al., 1994). A inflamação da glândula mamária reduz a produção de leite e os preços dos produtos, bem como causam riscos ao consumo humano. A identificação de bovinos resistentes à mastite vem sendo associada por diversos autores ao genótipo dos animais para os genes do Complexo Principal de Histocompatibilidade (MHC) (Dietz et al., 1997b). Há vários estudos na literatura, nos quais a associação genética entre o gene BoLA-DRB3 e a ocorrência de mastite vêm sendo investigadas (Dietz et al., 1997a,b; Gilliespie et al., 1999; Giovambattista et al., 1996; Kelm et al., 1997; Sharif et al., 1998; Starkenburg et al., 1997). Nos bovinos da raça Gir Leiteira, os problemas decorrentes da ocorrência de mastite também representam uma importante fonte de redução de produção e lucro na pecuária.

O produto do gene BoLA-DRB3 é uma proteína de ligação, presente na superfície celular e relacionada com a formação do complexo antígeno-anticorpo, sendo portanto uma importante molécula associada com a resposta imunológica específica. Até hoje, 107 diferentes alelos foram identificados no gene BoLA-DRB3 (www.projects.roslin.ac.uk), sendo esse extenso polimorfismo responsável pela dificuldade em genotipar corretamente os animais. Os primeiros estudos empregaram a metodologia de PCR-RFLP (RFLP – "Restriction Fragment Length Polymorphism") com o emprego de enzimas de restrição e digestão de fragmentos de tamanho variável em função das variações das seqüências dos alelos - polimorfismo (Gelhaus et al., 1995;

Sitte et al., 1995; Van Eijk et al., 1992). Essa metodologia está bem caracterizada, no entanto, como consequência do extenso polimorfismo observado para esse gene, diversos alelos não podem ser discriminados desta forma. A metodologia de PCR-RFLP permite discriminar apenas 56 alelos diferentes dentre os 107 alelos descobertos até o momento. Como não há informações sobre o gene BoLA-DRB3 em animais da raça Gir Leiteira, o propósito do presente trabalho é identificar os alelos presentes nesses animais e verificar a sua distribuição para essa raça tropical zebuína. Para atingir esse objetivo, duas metodologias experimentais serão empregadas para genotipar os animais: seqüenciamento direto dos produtos de PCR (PCR-DS); e clonagem seguida de seqüenciamento dos fragmentos amplificados do gene BoLA-DRB3 (CLO-S).

4.2 Material e Métodos

4.2.1 Animais e coleta de amostras

Vinte oito vacas da raça Gir Leiteira (*Bos indicus*) pertencentes à Embrapa Gado de Leite, localizada em Juiz de Fora, MG, foram utilizadas. Amostras de sangue de 5 ml foram coletadas nos animais e conservadas em 0,5% de EDTA, sendo congeladas a -20°C para análise posterior.

4.2.2 Extração de DNA

O protocolo de extração consistiu de adição de 7 ml de tampão de lise (0,32 M sacarose, 12 mM Tris-HCl pH 7,5, 5 mM MgCl₂ e 1% Triton 100) a 5 ml de sangue. As amostras foram a seguir centrifugadas a 2.000 g por 20 minutos. O sobrenadante foi em seguida descartado e mais 3 ml de tampão de lise foram adicionados às amostras que, após serem ressuspensas, foram centrifugadas a 2.000 g durante três minutos, novamente. Esse passo foi repetido até que um pellet claro fosse obtido. O DNA foi extraído da amostra mediante a suspensão da amostra em 900 μl de uma solução contendo 450 μl de tampão de proteinase K (NaCl 0,375 M ; EDTA 0,12 M, pH 8,0), 45 μl de proteinase K (20 mg/ml) e 20 μl de SDS 20% (*Sodium dodecyl sulfate*). As amostras foram a seguir incubadas por 16 horas a 55°C e na manhã seguinte a proteína foi precipitada mediante a adição de 500 μl de NaCl 5M. Após centrifugação a 2.000 g durante 25

minutos, o sobrenadante ($\approx 1050 \mu\text{l}$) foi distribuído em três tubos "eppendorf" de $1500 \mu\text{l}$ ($350 \mu\text{l}/\text{cada}$) e $1000 \mu\text{l}$ de etanol absoluto foram adicionados a cada tubo. O DNA das amostras foi precipitado após centrifugação a $10.000 g$ por 15 minutos. O sobrenadante foi descartado e as amostras foram lavadas em etanol 70%. Ao final, as amostras foram secas a 37°C por uma hora e cada *pellet* de DNA foi ressuscitado em $100 \mu\text{l}$ de água milli-Q (Millipore Corporation, Bedford, MA).

4.2.3 Amplificação do segundo exon do gene BoLA-DRB3

Oligonucleotídeos iniciadores (*primers*) publicados por Van Eijk et al. (1992) foram utilizados para amplificação de uma região genômica do segundo exon do gene BoLA-DRB3. Os *primers* HL030 (5'- ATCCTCTCTCTGCAGCACATTTCC - 3') e HL031 (5'- TTTAATTCGCGCTCACCTCGCCGCT - 3') foram utilizados em uma primeira reação de amplificação. Essa primeira reação foi conduzida empregando 100 ng de DNA ($5 \mu\text{l}$) em um volume total de $25 \mu\text{l}$ contendo 1 x tampão de PCR, 10 mM dNTP mix, 50 mM MgCl, $0,5 \text{ mM}$ de cada *primer* e uma unidade da enzima *Taq* DNA polimerase (Gibco BRL, New York, NY). As condições térmicas de amplificação para essa primeira parte da PCR contiveram uma etapa inicial de desnaturação de três minutos a 94°C seguida de dez ciclos de dez segundos a 94°C , 30 segundos a 60°C e 30 segundos a 72°C , permanecendo uma etapa de extensão final de cinco minutos a 72°C . Após a primeira etapa de amplificação, uma segunda reação de PCR semi-aninhada foi conduzida com $10 \mu\text{l}$ da solução da primeira reação de PCR distribuídas em cinco novos tubos ($2 \mu\text{l}/\text{tubo}$) com concentrações de reagentes conforme descrito para a primeira etapa de PCR, exceto que nesta etapa foram utilizados os *primers* HL030 e HL032 (5'- TCGCCGCTGCACAGTGAACTCTC - 3'). O *primer* HL032 possui uma seqüência interna em relação à seqüência amplificada na primeira etapa de PCR, possuindo uma superposição de oito bases com o *primer* HL031 (sublinhada no texto acima). As condições térmicas nesta segunda etapa da reação e PCR foram 25 ciclos com 40 segundos a 94°C e 30 segundos a 65°C , seguidos de cinco minutos a 72°C . As reações foram realizadas em placas de 96 amostras (Perkin-Elmer Applied Biosystems, Foster City, CA).

4.2.4 Eluição dos produtos de PCR e seqüenciamento direto

Brevemente, a eletroforese das amostras amplificadas foi conduzida em gel de agarose 2% (Gibco BRL, New York, NY) aplicando-se 100 μ l das reações de PCR e combinando-se quatro tubos. Fatias de gel com as bandas específicas dos produtos do gene BoLA-DRB3 foram eluídas e purificadas utilizando as recomendações do fabricante do produto Sephaglas BandPrep Kit (Amersham-Pharmacia Biotech Inc., Piscataway, NJ). Dois μ l do produto purificado com 200-500 ng de amostra de DNA amplificado foram submetidos à reação de seqüenciamento em um volume total de 15 μ l contendo 2 μ l de ABI Prism Dye Terminator Cycle Sequencing Core Kit com *AmpliTaq* DNA polimerase (Perkin-Elmer Applied Biosystems), 1 μ l de *primer* 5 μ M, sendo HL030 ou HL032 (consistente com a reação de seqüenciamento sendo direta ou reversa) e 2 μ l de tampão de seqüenciamento (200 mM Tris-HCl, pH 9,0, e 5 mM MgCl₂). As reações foram realizadas de acordo com o seguinte perfil térmico: 96 °C durante 3 minutos, 25 ciclos de 96 °C durante dez segundos, 50 °C durante cinco segundos e, ao final, uma extensão de quatro minutos a 60 °C. As amostras, após a reação, foram precipitadas em isopropanol 90%, lavadas em etanol 75% e secas à temperatura ambiente. A seguir, as amostras foram ressuspendidas em 4 μ l de formamida deionizada : EDTA 50 mM, pH 8,0 (5:1) e 2 μ l foram aplicados em um seqüenciador automático ABI 377 (Perkin-Elmer Applied Biosystems) e submetidos à eletroforese durante sete horas.

4.2.5 Clonagem dos produtos eluídos

As amostras que revelaram a possibilidade de identificação de novos alelos, bem como aquelas para as quais a genotipagem direta dos produtos de PCR mostrou-se não-confiável, foram submetidas à clonagem. Esta foi realizada no sítio *Sma* I do vetor pUC18 utilizando o protocolo de ligação SureClone Ligation Kit (Amersham-Pharmacia Biotech Inc.) e transformadas em células competentes JM109 da bactéria *Escherichia coli* (Hanahan, 1983). Clones com cópias dos alelos do gene BoLA-DRB3 foram seqüenciados mediante as mesmas condições acima descritas, exceção feita para o uso de *primers* diretos ou reversos universais M13 (Perkin-Elmer Applied Biosystems).

4.2.6 Definição dos alelos

A designação dos alelos que compunham o genótipo de cada animal foi realizada inicialmente após exame das seqüências obtidas com o seqüenciamento direto (PCR-DS) com os *primers* direto e reverso. As seqüências consenso das duas leituras obtidas por meio de exames dos cromatogramas foram comparadas com as seqüências de alelos previamente identificados utilizando o software Sequencher versão 3.0 (Gene Code Corporation, Ann Arbor, MI). Com o objetivo de dirimir dúvidas e ambigüidades em amostras de indivíduos heterozigotos, somente os alelos que produziram combinações possíveis de leituras dos alelos identificados anteriormente foram consideradas. A genotipagem foi definida somente nos casos em que não havia dúvida com relação à leitura das bases nos cromatogramas e a combinação de alelos previamente identificados era precisa, em todas as posições. Nos casos em que a genotipagem por PCR-DS não foi possível, as amostras foram clonadas e seqüenciadas (CLO-S).

4.3 Resultados e discussão

4.3.1 Comparação das diferentes metodologias de genotipagem

Foram testadas duas metodologias de seqüenciamento de produtos de PCR localizados no segundo exon do gene BoLA-DRB3. Na primeira, os produtos eram seqüenciados diretamente (PCR-DS) após a amplificação; e na segunda, os produtos amplificados eram clonados antes de serem seqüenciados (CLO-S). Quando os produtos da amplificação foram seqüenciados diretamente (PCR-DS) sem a eluição dos fragmentos em géis de agarose, a correta identificação dos alelos não foi possível. Os resultados mostraram que, mesmo sendo a reação de PCR, uma técnica consagrada, rápida e efetiva de amplificação de fragmentos específicos de DNA, o seqüenciamento de produtos de PCR exige produtos isolados e purificados para que a reação seja bem-sucedida com a leitura correta das bases da região do fragmento de interesse. Quando os produtos de PCR foram eluídos e purificados em géis de agarose, os *primers* restantes das reações anteriores foram retirados e a nova reação de seqüenciamento foi realizada na presença específica do *primer* incluído para a realização da reação de seqüenciamento,

produzindo uma leitura única. Alterações na qualidade das amostras após a purificação podem ser visualizadas na Figura 9.

A genotipagem foi possível com o emprego de PCR-DS, no entanto, as múltiplas combinações das seqüências provenientes de dois alelos sendo seqüenciados em uma única reação, quando os animais eram heterozigotos, resultaram em elevado consumo de tempo para análise dos resultados. Ainda, ao final, a identificação dos alelos em alguns casos permaneceu ambígua e a amostra teve de ser clonada e seqüenciada para que a genotipagem fosse realizada com sucesso.

O seqüenciamento dos clones produziu seqüências de melhor qualidade e a possibilidade de imediata genotipagem (Figura 10). Todos os animais tiveram ao menos duas seqüências provenientes de PCR-DS, uma empregando o *primer* direto e outra com o *primer* reverso. Dentre as 28 vacas analisadas, 11 animais (39,3%) foram corretamente genotipados por meio de PCR-DS. Duas vacas dentre essas 11 necessitaram de quatro reações de seqüenciamento para permitir uma genotipagem confiável. Material genético das restantes 17 vacas foi clonado para que a genotipagem do animal fosse possível. CLO-S também foi realizado nas 11 vacas genotipadas por meio de PCR-DS para conferir a genotipagem que havia sido realizada nesses animais com essa metodologia.

PCR-DS apresentou vantagens e desvantagens. Primeiramente, as seqüências são obtidas rapidamente, pois não há necessidade de clonagem dos fragmentos de PCR dos animais. Outra vantagem apresentada pela técnica é que apenas 64 reações de seqüenciamento foram necessárias para genotipar 11 animais. No entanto, mesmo quando amostras purificadas e isoladas de géis de agarose foram submetidas a reações de seqüenciamento, análises das seqüências foram extremamente demoradas devido ao elevado nível de polimorfismo e, ainda assim, em muitos casos a clonagem das amostras mostrou-se necessária para a correta genotipagem dos animais.

CLO-S foi uma metodologia mais trabalhosa, sendo necessárias 297 reações de seqüenciamento para genotipar os restantes 17 animais. Quando os produtos de PCR dos animais foram clonados, diversas seqüências de boa qualidade foram necessárias para formar um consenso confiável da seqüência, evitando erros de incorporação de nucleotídeos ou de edição da enzima *Taq* DNA polimerase. A incorporação errada de nucleotídeos e a observação de erros da enzima *Taq* DNA polimerase está discutida na literatura (Gyllestein & Allen, 1993) e foi observada na realização deste trabalho (dados não apresentados).

Para sobrepor as limitações associadas com cada uma das duas metodologias, PCR-DS e CLO-S, uma possível solução discutida na literatura indica a possibilidade da separação dos alelos antes da reação de seqüenciamento. A utilização de SSCP (*Single strand conformation polymorphism*) (Mok et al., 1993) ou DGGE (*Denaturing gradient gel electrophoresis*) (Aldriege et al., 1998) permitiram separação de alelos. Os alelos do exon 2 do gene BoLA-DRB3 apresentam extenso polimorfismo e as metodologias mencionadas podem contribuir na realização da separação de alelos antes da realização da reação de seqüenciamento.

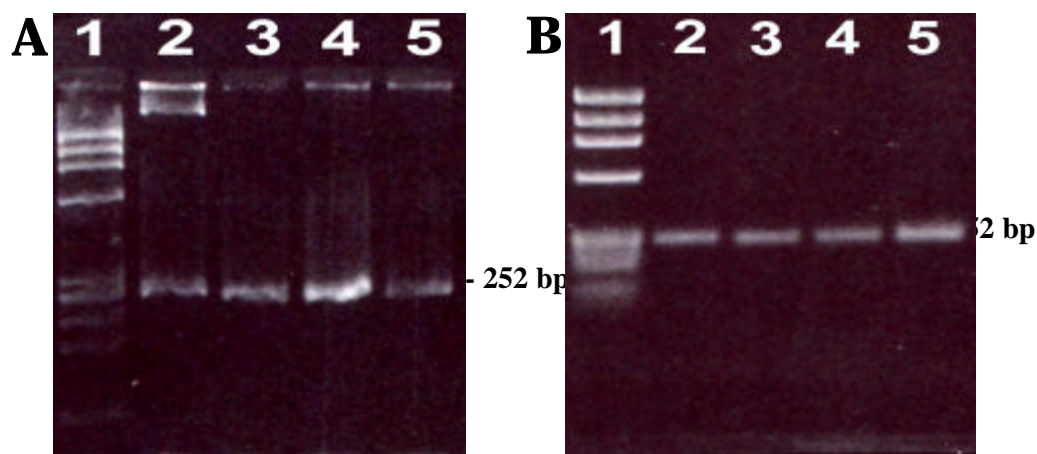


Figura 9 - (a) Eletroforese em géis de agarose 2% com 8 μ l de produto de PCR do gene BoLA-DRB3 (252 bp) após duas etapas de reação de PCR semi-aninhadas com *primers* específicos. Ruídos no sinal das bandas e amostras que apresentaram excesso de DNA genômico (linha 2); formação de rastro (linha 4); e quantidades de amplificação variável dos produtos (linhas 3, 4 e 5) estão representadas. (b) As mesmas amostras e volumes estão apresentados após a eluição das bandas de PCR do gene BoLA-DRB3 e purificação. (linha 1 = λ X174 *Hae* III, 1,5 μ g).

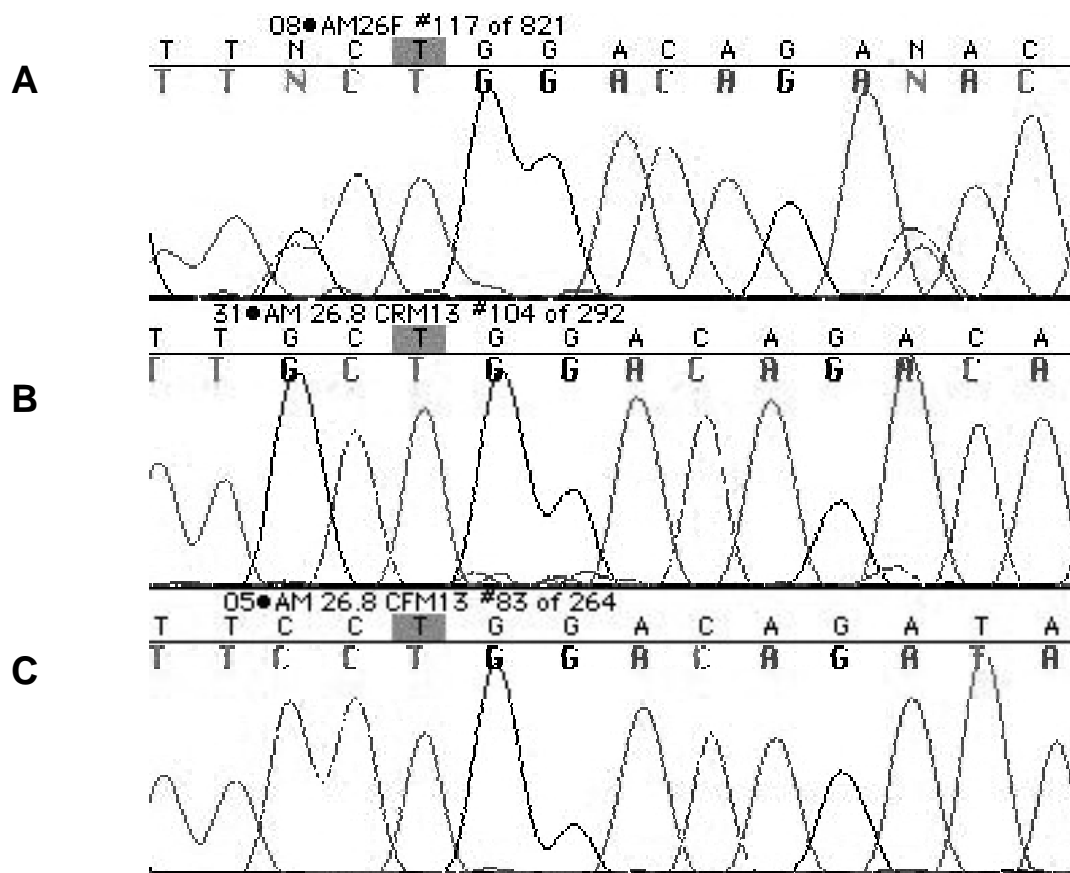


Figura 10 - (a) Cromatograma da reação de seqüenciamento com a seqüência de nucleotídeos obtida por PCR-DS mostrando múltiplos pontos de superposição de polimorfismo. (N). (b) As mesmas amostras obtidas por meio de CLO-S mostrando a seqüência de um dos alelos separadamente. (c) Cromatograma mostrando a seqüência de nucleotídeos do outro alelo. Seqüência de nucleotídeos do alelo BoLA-DRB3*1703 depositado no GenBank (accession number AF209710): 5' -tggagtattc tacggcgag tgcatttct tcaacgggac cgagcgggtg cggtacctgg acagatactt ctataatgga gaagagtacg cgcgcttca cagcgactgg ggcgagtcc gggcgggtgac cgagctgggg cggccggacg ccaagtactg gaacagccag aaggagatcc tggagcggga gcgggcttat gtggacacgt actgcagaca caactacggg gtcggtgaga gtttactgt gc-3'.

4.3.2 Freqüências alélicas

Freqüências alélicas do gene BoLA-DRB3 em matrizes da raça Gir Leiteira (*Bos indicus*) variaram entre 0,02 e 0,18 (Tabela 12). Os alelos menos freqüentes foram *0201, *1001, *14011 e *1602, cada um deles com a freqüência de 0,0179. Entretanto, os alelos mais freqüentes foram *3601 (0,1786), *2201 (0,1429) e *2101 (9,1071), correspondendo por mais de

43% das frequências alélicas. A sequência de nucleotídeos do alelo BoLA-DRB3*1703 foi depositada no GenBank (<http://www.ncbi.nlm.nih.gov>) com 252 bp (accession number AF209710) (Figura 10), o qual apresenta 3 bp a mais na extremidade 3'- do que a última submissão realizada recentemente para esse alelo por outro estudo independente em animais da raça Brahma (Maillard et al., 1999).

Tabela 12 . Frequências alélicas do gene BoLA-DRB3 de 28 vacas da raça Gir Leiteira (*Bos indicus*) genotipadas por seqüenciamento direto de produtos de PCR ou por meio de protocolo de clonagem e seqüenciamento

Alelos do gene BoLA-DRB3	Número de cópias	%
*0201	1	1,79
*1001	1	1,79
*14011	1	1,79
*1602	1	1,79
*25011	1	1,79
*4101	1	1,79
*1703	2	3,57
*2704	2	3,57
*3001	2	3,57
*2802	3	5,36
*3101	3	5,36
*1801	4	7,14
*2801	4	7,14
*3301	4	7,14
*2101	6	10,71
*2201	8	14,29
*3601	10	17,86

Outros estudos investigaram as frequências alélicas do gene BoLA-DRB3. Kelm et al. (1997) genotiparam 131 vacas Holandesas empregando a técnica de PCR-RFLP e utilizando nomenclatura descrita por Van Eijk et al. (1992) e Gelhaus et al. (1995). Esses autores

observaram que os alelos DRB3.2*8 e *11 foram os mais freqüentes (21.31 e 17.62%, respectivamente). Sharif et al. (1998) descreveram freqüências alélicas de 20.1% para o alelo DRB3.2*8 e 14.9% para *11, em animais da raça Holandesa, enquanto para a raça Jersey o alelo *8 estava presente na freqüência de 1.7% e o alelo *11 não foi observado. Gilliespie et al. (1999), em um estudo com matrizes da raça Jersey, encontrou freqüências alélicas de 11,3 % para o alelo DRB3.2*8, enquanto o alelo *11 não foi observado.

O padrão observado para o alelo DRB3.2*8, obtido por PCR-RFLP corresponde ao do alelo DRB3.2*1201 e DRB3.2*1202 (<http://www.projects.roslin.ac.uk/bola/drftab.html>). No presente estudo, utilizando animais da raça Gir Leiteira criados no Brasil, os alelos *1201 e *1202 não foram observados. O alelo DRB3.2*11, o qual corresponde aos alelos DRB3.2*0901 e 0902, também não foi observado neste estudo. Em contraste, o alelo DRB3.2*2201 (DRB3.2*06 considerando PCR-RFLP), o qual apresentou uma freqüência de 0,003 no estudo de Giovambattista et al. (1996) com gado Crioulo da Argentina, foi o segundo mais freqüente no presente estudo (0,1429). Em adição, os alelos *3601, *2201, *2101, *1801 e *2801 (alelos por PCR-RFLP: *20, *06, *35, *18 e *31, respectivamente), responsáveis por mais de 60% das freqüências alélicas no presente estudo, não foram observados em animais da raça Holandesa nos Estados Unidos (Kelm et al., 1997) e representaram somente 2% das freqüências alélicas em animais da raça Holandesa no Canadá (Sharif et al., 1998).

Um dos fatores de incerteza sobre a distribuição de alelos do gene BoLA-DRB3.2, especialmente em animais zebuínos, deve-se à falta de estudos com animais dessa origem em regiões tropicais. Esse é o primeiro estudo com o gene BoLA-DRB3 em animais da raça Gir Leiteira, realizado no Brasil. A variação dos resultados em relação aos já existentes na literatura indica que as freqüências alélicas do gene BoLA-DRB3 devem, ao menos até certo ponto, depender da raça do animal, como consequência da população fundadora e da pressão de seleção realizada ao longo dos anos.

No momento, quase todos os alelos previamente identificados e depositados no GenBank tiveram sua origem em animais de raças de origem européia. Mesmo considerando a população reduzida sobre a qual o presente estudo foi realizado, foi possível verificar a existência de 17 alelos, alguns dos quais apenas recentemente descritos para animais de origem zebuína (Maillard et al., 1999). Esses resultados permitem concluir que novos alelos, ainda não identificados, restam ser descobertos em estudos futuros em animais de raças tropicais de origem zebuína.

Desta forma, foi aqui demonstrada a possibilidade de genotipagem empregando PCR-DS e CLO-S. O procedimento de CLO-S foi utilizado gerando resultados confiáveis de genotipagem. Em outros locos que apresentem menor polimorfismo, a técnica de PCR-DS poderá ser empregada preferencialmente, produzindo resultados com maior desempenho. Essa técnica é substancialmente menos dependente de trabalho, podendo também gerar genotipagem correta dos animais. Em situações nas quais a genotipagem for ambígua, a clonagem e seqüenciamento isolado dos alelos podem gerar uma genotipagem definitiva.

4.4 Conclusões

- O gene BoLA-DRB3 é altamente polimórfico.
- As técnicas correntes de genotipagem desse gene, por meio de PCR-RFLP não permitem discriminar todos os alelos conhecidos.
- Genotipar os animais através da clonagem-sequenciamento é um método trabalhoso e dispendioso, não aconselhável para estudos com grande número de animais.
- O sequenciamento direto permitiu genotipagem dos animais, no entanto essa seria uma técnica mais apropriada para genes que apresentam menor polimorfismo.

Agradecimentos

A. F. da Mota agradece à Embrapa por recebimento de bolsa de doutorado. L. L. Coutinho e M. L. Martinez agradecem ao CNPq por bolsa de produtividade em pesquisa. J. E. Gabriel agradece à FAPESP por bolsa de doutoramento.

5 GENOTIPAGEM DE ALELOS DO GENE BOLA-DRB3 EM BOVINOS DA RAÇA GIR LEITEIRA (*BOS INDICUS*) POR MEIO DE ELETROFORESE EM GEL DE POLIACRILAMIDA COM GRADIENTE DE TEMPERATURA E SEQÜENCIMENTO ⁵

Resumo

BoLA-DRB3 é um gene integrante do Complexo Principal de Histocompatibilidade (MHC – Major Histocompatibility Complex) em bovinos. O produto do gene é uma proteína relacionada com a formação do complexo antígeno-anticorpo associada com a resposta imunológica específica. A tarefa de genotipar corretamente os alelos do gene BoLA-DRB3 é complexa em razão do extenso polimorfismo apresentado nestes locos. As técnicas correntes para a designação dos alelos empregam a digestão de fragmentos amplificados com enzimas de restrição (PCR-RFLP) (PCR – Polymerase Chain Reaction) (RFLP – Restriction Fragment Length Polymorphism), o seqüenciamento direto de produtos de PCR (PCR-DS), a clonagem seguida de seqüenciamento dos alelos (CLO-S) ou a realização de géis com gradiente de desnaturação para separação dos alelos previamente ao seqüenciamento (DGGE – Denaturing Gradient Gel Electrophoresis). Essas técnicas apresentam desvantagens como: consumo elevado de tempo e mão-de-obra, não são capazes de discriminar todos os alelos já identificados ou não são facilmente reproduzíveis. O objetivo do presente trabalho foi desenvolver a técnica de genotipagem do gene BoLA-DRB3 por meio da eletroforese em gel com gradiente de temperatura (TGGE – *Temperature Gradient Gel Electrophoresis*). Amostras de DNA de 28 vacas da raça Gir Leiteira (*Bos indicus*) foram amplificadas com primers específicos para o gene BoLA-DRB3 e submetidas a TGGE. Os alelos foram novamente amplificados, purificados e seqüenciados. A separação de alelos foi possível em 21 dos 26 indivíduos heterozigotos (81%). Os resultados indicaram que duas leituras de seqüenciamento (direto e reverso) são suficientes

⁵ DA MOTA, A. F.; MARTINEZ, M. L.; COUTINHO, L. L. *European Journal of Immunogenetics*, no prelo. ACEITO PARA PUBLICAÇÃO.

para uma genotipagem precisa dos alelos do gene BoLA-DRB3. A separação dos alelos desse gene mediante TGGE representa a possibilidade de genotipagem confiável com técnica de alto desempenho (*high-throughput*), crítica para permitir a realização de estudos de associação de alelos com características de sanidade dos animais.

Summary

BoLA-DRB3 is a gene of the Major Histocompatibility Complex (MHC) in cattle. The product of the BoLA-DRB3 is a protein related to formation of an antigen-antibody complex associated with immune response. Genotyping BoLA-DRB3 gene is relatively complex due to the extensive polymorphism associated with this locus. Current techniques for assignment of genotypes are restriction fragment length polymorphism of DNA amplified fragments (PCR-RFLP), direct sequencing of PCR products, cloning-sequencing, and denaturant-gradient gel electrophoresis. These techniques are either time-consuming, do not discriminate all possible alleles, or are not readily reproducible. The objective of this study was to genotype the gene BoLA-DRB3 using of temperature-gradient gel electrophoresis (TGGE) to separate alleles before sequencing. PCR reactions made out of 28 DNA samples from Gir Dairy cows (*Bos indicus* Brazilian breed) were submitted to TGGE. New PCR products were generated from separated alleles, purified, and sequenced. Allele separation was possible in 21 out of 26 heterozygote samples (81%). Results indicate that two sequence reads (forward and reverse) were sufficient to accurate genotype BoLA-DRB3 alleles. Separation of alleles by TGGE provides a high-throughput reliable typing of BoLA-DRB3 alleles, critical to performing disease association studies in cattle.

5.1 Introdução

O produto do gene BoLA (Bovine Lymphocyte Antigen) é uma proteína de ligação envolvida na formação do complexo antígeno-anticorpo que tem papel essencial na especificidade da resposta imunológica. O gene BoLA-DRB3 é altamente expresso e atualmente, mais de 100 alelos diferentes do segundo exon do gene BoLA-DRB3 estão identificados e depositados no GenBank (<http://www.ncbi.nlm.nih.gov/>). A genotipagem dos genes da família BoLA é relativamente complexa devido ao elevado polimorfismo observado nesses genes. Um paralelo pode ser observado com o gene ortólogo HLA-DRB1 expresso em Humanos onde mais

de 290 alelos (<http://www.anthonynolan.org.uk/HIG/lists/class2list.html>) já foram identificados. A diversidade dos genes DRB3 bovinos é responsável por aumentar a dificuldade em obter-se genotipagem confiável dos animais em estudo, especialmente para associação com doenças quando informações de *pedigree* não estão disponíveis.

Estudos desenvolvidos com bovinos buscaram associar a presença de alelos específicos do gene BoLA-DRB3 com a resistência à mastite (Dietz et al., 1997; Gelhaus et al., 1995; Gilliespie et al., 1999; Kelm et al., 1997; Van Eijk et al., 1992; Sharif et al., 1998) e levaram a diferentes conclusões. Nos estudos mais recentes, a genotipagem foi realizada por meio de amplificação e obtenção de um padrão de eletroforese após digestão com enzimas de restrição (PCR-RFLP). Essa metodologia não é capaz de determinar precisamente todas as diferenças conhecidas entre os alelos identificados e esta pode ser a causa de diferentes conclusões nos estudos mencionados. Mais modernamente, com a possibilidade de seqüenciamento automático realizado em diversos laboratórios, o seqüenciamento direto de produto de PCR ou a clonagem seguida de seqüenciamento dos alelos têm sido citados na literatura (da Mota et al., 2002). No entanto, a clonagem seguida de seqüenciamento apresenta a desvantagem de ser uma técnica trabalhosa e dispendiosa, sendo necessárias várias seqüências para obtenção do genótipo do animal. O seqüenciamento direto de produtos de PCR amplifica e apresenta a seqüência dos dois alelos conjuntamente, não sendo possível em todos os casos discriminar as diferenças observadas pelo polimorfismo em múltiplas posições, especialmente no caso de deleções e inserções.

Um método mais robusto para genotipagem emprega gel de poliacrilamida com gradiente de desnaturação (DGGE) e permite separar os alelos durante a eletroforese, anteriormente à reação de seqüenciamento (Aldridge, 1998). A separação dos alelos ocorre devido à diferença no padrão de eletroforese do alelo causada por mudanças na estrutura secundária deste. A desvantagem apresentada pela técnica reside na dificuldade em reproduzir o gradiente de desnaturação do gel, e, conseqüentemente, reproduzir o padrão de eletroforese.

O presente estudo objetivou desenvolver a técnica de genotipagem do gene BoLA-DRB3 por meio da eletroforese em gel com gradiente de temperatura (TGGE – Temperature Gradient Gel Electrophoresis) seguida de reação de seqüenciamento. Para tanto foram utilizadas amostras de DNA provenientes de animais da raça Gir Leiteira (*Bos indicus*).

5.2 Material e Métodos

5.2.1 Extração de DNA e amplificação do gene BoLA-DRB3

Cinco ml de amostras de sangue foram coletados de 28 vacas da raça Gir Leiteira criadas na Embrapa Gado de Leite, Juiz de Fora, MG. A extração de DNA foi procedida conforme publicado anteriormente (da Mota et al., 2002), o mesmo tendo ocorrido com a amplificação do segundo exon do gene BoLA-DRB3, exceção somente para o número de ciclos da segunda etapa de PCR que foi realizada em 35 ciclos.

5.2.2 Condições de TGGE

A eletroforese foi conduzida em um sistema de TGGE com o gel (9 cm x 9 cm) disposto horizontalmente (Biometra, Göttingen, Germany). O estabelecimento do gradiente da temperatura para condições ótimas de eletroforese foi estabelecido com base em eletroforese perpendicular e pelo cálculo de perfis de desnaturação térmica e mobilidade dos fragmentos de dupla-fita de DNA (Steger, 1994). As condições de eletroforese foram realizadas em gel com 8% de poliacrilamida (peso/volume) (29:1 acrilamida:bis-acrilamida) em tampão 0.2 x Na-TAE pH 8.4 (Acetato de Sódio 1 M, EDTA 10 mM, TRIS 400 mM). Um gradiente térmico variando entre 35°C e 75°C foi empregado nos fragmentos previamente amplificados. Aproximadamente 3 µg de amostra (1 µg/11) foram misturados na proporção de 2:1 com o tampão de carregamento (0.2 x Na-TAE pH 8.4, 0.1% Triton-X 100, 0.01% Bromophenol Blue dye, 0.01% Xylene Cyanol dye). As amostras foram submetidas a uma pré-corrída paralela durante 15 minutos a 100 V, sendo imediatamente submetidas a 70 minutos de corrida a 250 V. Os géis foram corados em solução de brometo de etídio (0,5 µg/ml) e visualizados através de luz ultra-violeta.

A seguir realizou-se a excisão de tiras do gel de poliacrilamida contendo os alelos separados para eluição. Pequenas bandas do gel com os alelos separados foram transferidas para tubos "eppendorf" com 50 µl de solução de eluição para géis de poliacrilamida (TRIS-HCl 10 mM pH 8,5, KCl 50 mM, MgCl₂ 1.5 mM, 0.1% Triton X-100) e foram incubadas à temperatura ambiente durante 16 horas para eluição do DNA.

5.2.3 Amplificação dos alelos separadamente e purificação

Dois μl (50 ng) da solução de eluição para géis de poliacrilamida contendo as bandas dos alelos separados durante a eletroforese serviram de molde para a reamplificação dos alelos por PCR. As mesmas condições de PCR descritas acima foram utilizadas. Após a reação, a solução foi purificada utilizando o GFXTM Purification Kit (Amersham Pharmacia Biotech, Piscataway, NJ), seguindo as recomendações do fabricante.

5.2.4 Reações de seqüenciamento

Após a purificação, dois μl of de produtos purificados com aproximadamente 200-500 ng de DNA foram submetidos à reação de seqüenciamento em um volume total de 15 μl com 2 μl de ABI Prism Dye Terminator Cycle Sequencing Core Kit com AmpliTaq DNA Polymerase (Perkin-Elmer-Applied Biosystems, Foster City, CA), 1 μl de *primer* HL030 3.2 μM ou HL032 (consistente com o emprego de primer reverso ou direto) e 2 μl de tampão de seqüenciamento. As condições térmicas das reações foram realizadas de acordo com as instruções do fabricante. As amostras foram precipitadas em isopropanol 90%, lavadas em etanol 70% e secas à temperatura ambiente. Cada uma das amostras dos alelos, após a reação de seqüenciamento, purificação e precipitação, foram ressuspensas em 4 μl de formamida deionizada:EDTA 50 mM pH 8.0 (5 : 1), e 2 μl foram carregados em um equipamento de seqüenciamento automático ABI 377 Automated Sequencer (Perkin-Elmer-Applied Biosystems, Foster City, CA) e submetidas à eletroforese durante sete horas.

5.2.5 Designação dos alelos componentes do genótipo do animal

Alelos separados por TGGE, eluídos e amplificados, foram submetidos à purificação e posteriormente à reação de seqüenciamento. A designação do genótipo foi realizada examinando as seqüências direta e reversa dos dois alelos do animal. As seqüências das duas reações e os cromatogramas foram comparados com todos os alelos identificados e depositados no GenBank. Foi empregado o software SequencherTM assignment software (Gene Code Corporation, Ann

Arbor, MI). Todas as amostras foram amplificadas independentemente, clonadas e seqüenciadas para confirmação dos genótipos estabelecidos por meio de TGGE.

5.3 Resultados e Discussão

Foi desenvolvida no presente estudo uma metodologia de genotipagem do segundo exon do gene BoLA-DRB3. Essa metodologia produz incrementos na acurácia e no desempenho da genotipagem. Todas as amostras foram amplificadas por meio de PCR e submetidas a condições otimizadas de eletroforese em géis de poliacrilamida com gradiente de temperatura (TGGE). Os melhores resultados foram alcançados com o tempo de eletroforese de 70 minutos, voltagem de 250 V e gradiente de temperatura variando entre 35°C e 75°C (Figura 11).

Havia 26 animais heterozigotos entre as 28 vacas estudadas. Quando as amostras foram submetidas à eletroforese nas condições acima descritas, foi possível separar os alelos das amostras em 81% dos casos (21 animais). Em alguns animais não foi possível observar a separação dos alelos, mesmo sendo heterozigotos (Figura 11A, linhas 1 e 6). TGGE mostrou-se uma metodologia consistente com eletroforese de géis reproduzíveis. Na Figura 11A e 11B a mesma amostra, localizada na linha 4, apresenta o mesmo padrão eletroforético em dois géis independentes.

Teoricamente, cada alelo deve apresentar um padrão de eletroforese em gel de poliacrilamida. Isso é devido ao fato de que fragmentos de fitas duplas de DNA (dsDNA - Double-strand DNA), mesmo quando apresentam o mesmo comprimento medido em pares de bases, apresentam estruturas secundárias com conformação distinta em função das diferenças nas seqüências causadas por mutações. O aumento na temperatura, ocorrido durante a eletroforese em um equipamento com gradiente de temperatura, causa a desnaturação das fitas duplas das moléculas de dsDNA em regiões específicas, nos domínios de fusão, que são alterados devido a alterações nas seqüências dos fragmentos de dsDNA. Dessa forma, a desnaturação e modificação das estruturas secundárias em temperaturas e momentos distintos durante a eletroforese prevêm modificação na mobilidade eletroforética dentre os alelos, e separação dos mesmos.

Devido à presença de moléculas híbridas formadas por uma fita simples de um dos alelos que se liga à fita complementar do outro alelo - heteroduplexes – e ao diferente padrão eletroforético dessas moléculas em resposta à composição de bases, é possível visualizar quatro bandas em uma amostra de um indivíduo heterozigoto (Figura 11A, linha 3). As duas bandas

mais inferiores representam moléculas homoduplexes, enquanto as outras duas correspondem a moléculas heteroduplexes.

Em algumas amostras foi possível observar a completa separação dos alelos homoduplexes (Figura 11A, linhas 2 e 5, e Figura 11B, linhas 1, 2, 3, 4 e 6) que foram posteriormente eluídos e seqüenciados. Na Tabela 13 é possível visualizar a comparação de algumas características dos diferentes alelos. Alguns dos alelos que não puderam sofrer separação por meio de TGGE apresentaram o mesmo número de pares de bases, domínios de fusão ocorrendo à temperatura semelhante e conteúdos equivalentes de bases G-C (i. e., ID=7 e ID=27). Entretanto, alelos que apresentaram diferenças mais expressivas no conteúdo de bases G-C, nas seqüências de nucleotídeos e no peso molecular não foram separadas (i. e., ID= 6 e ID=28). É possível observar também que alelos com eventos de inserção ou deleção foram separados em alguns casos (ID= 4, 21, 22 e 24); no entanto as amostras ID=10 e ID=12, com a inserção e deleção de 3 nucleotídeo, não foram separadas.

É possível também verificar na Tabela 13 que alelos pouco distintos em relação à seqüência dos nucleotídeos foram separados por TGGE (ID= 5 e ID=13). Portanto, os resultados sugerem que não é possível inferir se um fragmento amplificado nas condições descritas será separado por meio de TGGE, baseado nas informações do peso molecular, temperatura de fusão dos domínios, número de pares de bases ou conteúdo de G-C na seqüência amplificada.

Quando há alterações nas bases de nucleotídeos ao longo das seqüências, diferentes estruturas secundárias ocorrem e ocasionam diferenças nas formas das fitas duplas de DNA. As estruturas secundárias de DNA apresentam grutas e proeminências. Uma explicação a respeito da relação entre a seqüência e a modificação nas estruturas secundárias está na alteração do ângulo de interação entre as bases nitrogenadas, resultando em maiores ou menores grutas ou diferença nos diâmetros das duplas hélices. Simulações podem ser utilizadas para determinar estruturas secundárias que podem ser resultantes de alterações nas seqüências de dsDNA (<http://www.biophys.uni-duesseldorf.de/local/POLAND/poland.html>). Alguns alelos foram testados empregando esta metodologia, no entanto não foi possível chegar à conclusão definitiva, pois alguns alelos apresentaram diferenças nas estruturas secundárias simuladas, mas não foram separados por TGGE nas condições descritas (dados não apresentados).

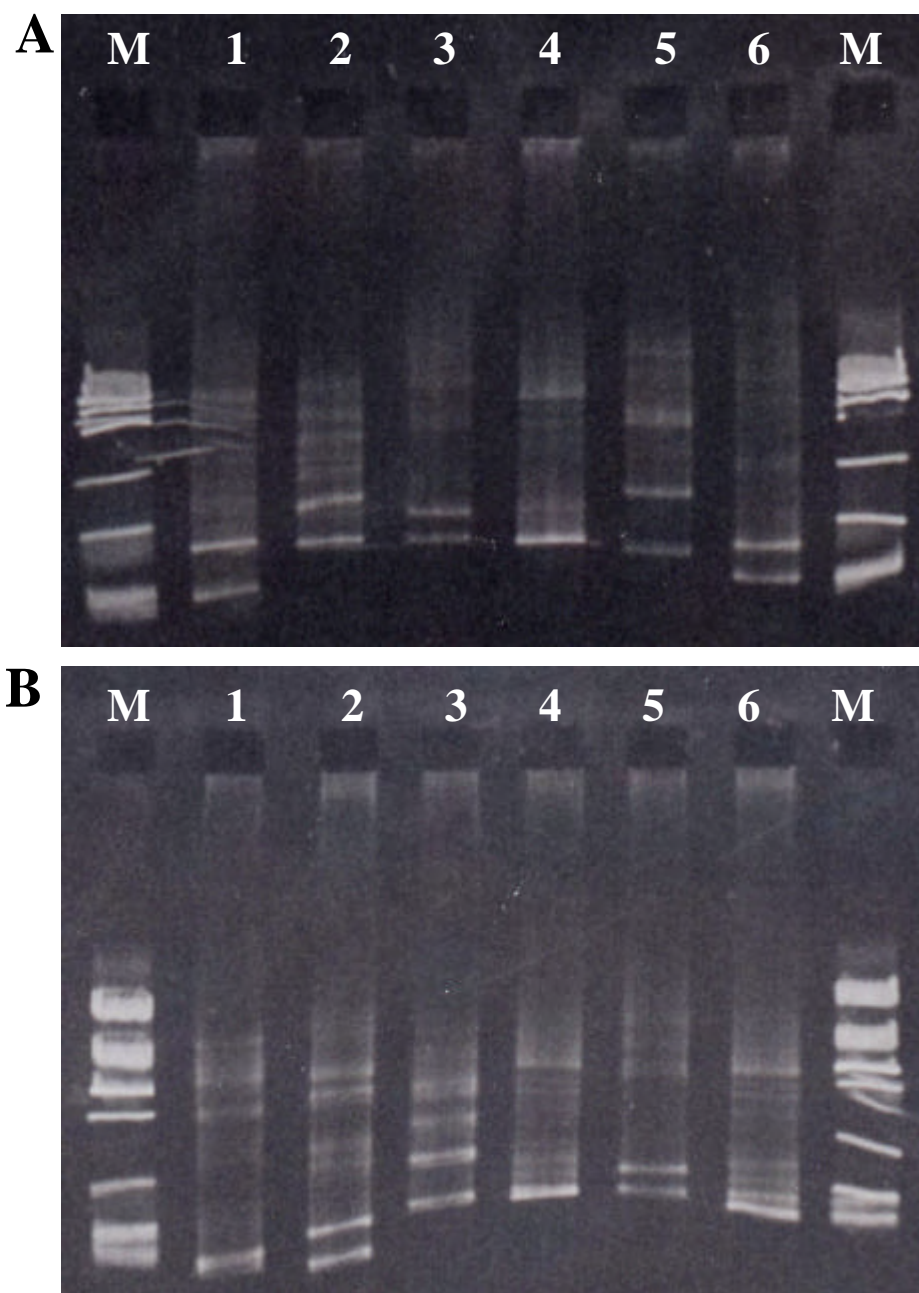


Figura 11. Gel de poliacrilamida contendo alelos do gene BoLA-DRB3 separados por meio de TGGE usando gradiente de temperatura entre 35°C - 75°C. As amostras foram submetidas à eletroforese durante 70 minutos a 250 V. A) Gel mostrando a separação dos alelos em cinco amostras de indivíduos heterozigotos (linhas 1, 2, 3, 5, 6). B) Gel mostrando a mesma amostra de um indivíduo homozigoto do gel A (Linha 4) e algumas amostras de indivíduos heterozigotos que não foram separadas (linhas 1, 6). Linhas designadas como “M” contêm ÔX174 *Hae*III, 1.5 ìg.

Tabela 13. Identificação do animal, designação dos alelos que integram o genótipo do animal, número de bases que diferem na sequência, eventos de inserção/deleção, conteúdo de G-C (%), temperatura de fusão dos domínios (°C) e peso molecular (Daltons) para as 28 amostras analisadas. (26 heterozigotos e 2 homozigotos)

Animal ID	BoLA-DRB3 Alelos	Dif. nas Seq. (bp)	Ins./ Del. (bp)	G-C (%)	Temp. fusão (°C)	Peso Mol. (Daltons)
1	*14011	26	-	61	87	78252
	*25011			61	87	77774
2	*2101	20	-	60	87	77924
	*3601			60	87	77497
3	*2802	25	-	61	87	73494
	*3601			60	87	77497
4	*4101	17	3	60	87	76551
	*2101			60	87	77924
5	*2801	16	-	61	87	77861
	*2201			62	87	77948
6**	*1801	21	-	62	88	77759
	2802			61	87	73494
7**	*1703	23	-	60	87	77494
	*2801			61	87	77861
8	*1801	22	-	62	88	77759
	*4201			61	87	77629
9	*1001	18	-	61	87	77774
	*2704			60	87	77969
10/12**	*3301	21	3	60	87	77157
	*3601			60	87	77497
11	*1602	26	-	58	86	77482
	*3001			60	87	77879
13	*2704	13	-	60	87	77969
	*3101			61	87	78234
14	*2201	30	-	62	87	77948
	*3601			60	87	77497
15	*2101	20	-	60	87	77924
	*3601			60	87	77497
16	*4201	23	-	61	87	77629
	*3601			60	87	77497
17	*1801	26	-	62	88	77759
	*2201			62	87	77948
18	*3001	28	-	60	87	77879
	*2201			62	87	77948
19	*2801	16	-	61	87	77861
	*2201			62	87	77948
21	*0201	23	3	62	88	76867
	*2201			62	87	77948
22	*1703	18	3	60	87	77494
	*3301			60	87	77157
23	*2801	16	-	61	87	77861
	*2101			60	87	77924
24	*3301	25	3	60	87	77157
	*3101			61	87	78234
26	*2101	13	1	60	87	77924
	*2802			61	87	73494
27**	*3101	23	-	61	87	78234
	*3601			60	87	77497
28**	*1801	19	-	62	88	77759
	*2101			60	87	77924

**Amostras que não puderam ser separadas por TGGE.

A estrutura específica de uma combinação de alelos não é previsível antes da identificação e genotipagem dos alelos. Alelos que se apresentaram semelhantes não puderam ser separados, ainda que as condições de eletroforese tenham sido otimizadas para permitir a melhor separação para o gene BoLA-DRB3. Nestes casos se torna necessário adicionar uma terminação rica em conteúdo G-C à extremidade da molécula durante a reação de PCR. Essa metodologia já foi apresentada na literatura com resultados favoráveis em casos em que a separação não era possível (Kappes et al., 1995). Outra possibilidade de genotipagem confiável nesses casos é o emprego da metodologia de donagem e seqüenciamento do fragmento amplificado (da Mota et al., 2002).

5.4 Conclusões

- TGGE é uma técnica rápida e confiável de genotipar alelos do gene BoLA-DRB3.
- As amostras puderam ser separadas em 81% dos casos. Os resultados indicam que duas seqüências (reversa e direta) são suficientes para a correta identificação das bases no exon 2 do gene BoLA-DRB3.
- A separação dos alelos por meio de TGGE e posterior seqüenciamento reduzem as despesas de genotipagem e o tempo envolvido nessa atividade. A técnica é consistente e com alto desempenho, apresentando géis reproduzíveis e permitindo a realização de estudos de associação de alelos com características de sanidade dos animais.
- É possível a descoberta de novos alelos, especialmente importante nos estudos com animais de raças zebuínas, ainda não contemplados em números mais expressivos em estudos dessa natureza. A técnica pode ser ainda utilizada com a incorporação de fragmentos ricos em conteúdo de G-C, denominados grampos, cuja utilização após otimização das condições pode levar à separação de alelos em maior percentual.

Agradecimentos

Adilson F. da Mota agradece a Embrapa por ter recebido bolsa de doutoramento. Luiz Lehmann Coutinho e Mario Luiz Martinez agradecem ao CNPq por recebimento de bolsa de produtividade em pesquisa. Os autores agradecem o Dr. Tad S. Sonstegard por revisar criticamente o manuscrito. Todos os experimentos com animais estão de acordo com a legislação brasileira e de acordo com as normas de conduta da Embrapa para experimentos com animais.

6 INVESTIGAÇÃO DO POTENCIAL DE DESCOBERTA DE POLIMORFISMO NO EXON 2 DE ALELOS DO GENE BoLA-DRB3 EM VACAS DA RAÇA GIR LEITEIRA (*Bos indicus*)

Resumo

A mastite é o principal problema sanitário que ocasiona prejuízos na pecuária leiteira. Há um gene do Complexo Principal de Histocompatibilidade (MHC – *Major Histocompatibility Complex*) em bovinos, chamado BoLA-DRB3 que está relacionado com a formação do complexo antígeno-anticorpo e associado com a resposta imunológica específica. Os estudos realizados para verificar a associação de alelos desse gene com a ocorrência de mastite ou contagem de células somáticas apresentam resultados inconsistentes, sendo que a maioria dos alelos foi caracterizada em bovinos de raças européias. A utilização de bovinos zebuínos pode trazer contribuição ao conhecimento com a descoberta de novos alelos. Nesse sentido foram empregadas quarenta amostras de DNA de vacas da raça Gir Leiteira (*Bos indicus*) estratificadas nos extremos da população, considerando as contagens de células somáticas durante lactações anteriores. Um *pool* de amostras dos animais com contagens superiores foi amplificado na região do exon 2 do gene BoLA-DRB3. O mesmo foi realizado para animais com contagens reduzidas de células somáticas. O produto da amplificação foi clonado e foram seqüenciados 29 clones. Os resultados permitiram identificar 18 alelos do gene BoLA-DRB3, sendo que 11 alelos foram considerados potencialmente como novos, não descritos anteriormente. Sete alelos confirmaram seqüências previamente publicadas, representando uma frequência de 58 % no genótipo dos animais estudados. Os alelos mais frequentes foram: *0902 (21 %), *3601 (10 %) e *2801 (10 %). O alinhamento das seqüências obtidas revelou as regiões de polimorfismo. Ficou evidente que animais da raça Gir Leiteira (*Bos indicus*) devem ser utilizados em estudos futuros destinados a identificação de novos alelos do gene BoLA-DRB3 e em estudos de associação com a ocorrência de mastite.

Summary

Mastitis is a major sanitary issue reducing profit in dairy industry. There is a gene named BoLA-DRB3 from the bovine MHC (*Major Histocompatibility Complex*) related to antigen complex formation that is associated with immune response, which has been used in association studies with somatic cells count and mastitis occurrence in cattle. These studies have reached different results and most of the alleles represented *Bos taurus* cattle breeds. Use of *Bos indicus* animals will bring new contribution to these studies and possible discovery of new alleles. Thus, a pool of samples from Dairy Gir cows (*Bos indicus*) stratified for somatic cells count was amplified in exon 2 of BoLA-DRB3. Further cloning and sequencing of 29 clones revealed 18 different alleles from gene BoLA-DRB3. Eleven alleles were not previously identified. Most frequent alleles were: *0902 (21 %), *3601 (10 %), and *2801 (10 %). Sequence alignment showed positions of the polymorphism. Dairy Gir (*Bos indicus*) is an important breed for future studies conducted to discover new BoLA-DRB3 alleles in association with mastitis occurrence or somatics cell count.

6.1 Introdução

As proteínas codificadas pelos genes da região MHC (*Major Histocompatibility Complex*) são altamente polimórficas e têm um importante papel na resposta imunológica efetiva frente a infecções causadas por organismos patogênicos. Dentre as doenças dos bovinos, a mastite é o principal fator sanitário causador de prejuízo ao setor leiteiro, devido à diminuição da qualidade e quantidade do leite produzido pela matriz infectada (Fox et al., 1991; Harmon, 1994; McCarthy et al., 1988; Sordilho et al., 1997). A mastite é caracterizada pela inflamação de um ou mais quartos do úbere da matriz. Segundo Harmon (1994), a mastite clínica é caracterizada por alterações inflamatórias visíveis no úbere (dor à palpação, vermelhidão e endurecimento), sendo que o leite pode apresentar-se aguado ou com grumos, podendo também se transformar em uma secreção espessa e purulenta. Na forma clínica aguda da doença, há o aumento da temperatura retal, letargia, anorexia e até mesmo a morte do animal.

Em estudo realizado por Harmon (1994), o autor salienta que a inflamação causa também o aumento da contagem de células somáticas (CCS). O mesmo autor salienta que toxinas produzidas por bactérias danificam o tecido responsável pela produção de leite ocasionando a diminuição da quantidade de leite produzido. Os bloqueios de pequenos ductos e

os danos ocasionados às células epiteliais podem resultar na cicatrização e formação de tecido fibroso, conseqüentemente, na redução da produção láctea na porção atingida da glândula.

A mastite possui etiologia múltipla, podendo ser causada por ordenha mal realizada, fatores anatômicos defeituosos na glândula mamária do animal, como diâmetro reduzido do canal do teto, agravados pela presença de bactérias que colonizam e causam a infecção no animal (Schukken et al., 1997; Smith et al., 1985). Além de infectar as matrizes, estes microorganismos se multiplicam no leite e liberam substâncias que podem causar quadros de intoxicação alimentar em consumidores (Marcus & Kehrlí, 1994). Segundo Veiga et al. (1994), os riscos à saúde humana pela ingestão do leite são provocados pela contaminação por bactérias patogênicas, pelo conteúdo de resíduos de antibióticos, ou ainda por substâncias que são produzidas pelas bactérias.

Marcus & Kerli (1994) salientam que as células somáticas presentes no leite são constituídas de vários tipos de células, principalmente neutrófilos, macrófagos, linfócitos, eosinófilos, e células do epitélio da glândula mamária. O número elevado de células somáticas presentes no leite é decorrente da resposta do organismo animal à infecção.

Há estudos na literatura que têm investigado a associação genética de um dos genes da região MHC com a resistência à mastite (Dietz et al., 1997a; Dietz et al., 1997b; Gilliespie et al., 1999; Giovambatista et al., 1996; Kelm et al., 1997; Starkenburg et al., 1997; Sharif et al., 1998). Esse gene, denominado BoLA (*Bovine Leucocyte Antigen*) codifica uma proteína de superfície celular que está envolvida na formação do complexo antígeno-anticorpo, relacionado ao sistema imunológico dos animais. O gene BoLA-DRB3 é altamente polimórfico, e está localizado no cromossomo 23 dos bovinos (Gilliespie et al., 1999). Há atualmente 105 alelos desse gene depositados nas bases públicas de informações moleculares sobre os genes BoLA (<http://www.projects.roslin.ac.uk/bola/drbtabs.html>). A maior parte dos alelos descobertos nesse gene foram provenientes de estudos realizados em animais de raças européias. Em Humanos, os genes da região DR- β estão localizados no cromossomo 6 e também apresentam extenso polimorfismo (http://www.nmdpresearch.org/HLA/hla_educational_resources.html). Estudos em zebuínos (*Bos indicus*) podem contribuir com a descoberta de polimorfismo ainda não descrito nesse gene, podendo este estar associado à maior capacidade de resistência aos patógenos causadores da mastite.

Os estudos independentes realizados para verificar a associação do gene BoLA-DRB3 com a ocorrência de mastite não chegaram a um consenso. Dietz et al. (1997a) reportou que o alelo 16 estava associado com animais que apresentaram um aumento significativo na CCS, e

Kelm et al. (1997) também encontrou uma associação positiva entre o alelo 16 e o aumento da CCS; entretanto; no estudo de Starkeburg et al. (1997) foi relatado que o mesmo alelo 16 estava associado com uma significativa diminuição da CCS. Em adição, Dietz et al. (1997a) indica que o alelo 22 estava relacionado com a diminuição da CCS, durante a segunda lactação, mas no estudo de Starkeburg et al. (1997), foi encontrada a associação do alelo 22 com a elevação da CCS, durante a primeira lactação.

Uma das justificativas para a ausência de consistência nos resultados reside na metodologia empregada na maioria dos estudos realizados (PCR-RFLP), que não permite discriminar entre os diversos alelos existentes. Outro fato é a possibilidade que alelos ainda não descritos estejam relacionados com a contagem de células somáticas e com a resistência à mastite. Há, portanto, necessidade de realizar mais estudos nesse loco a fim de identificar alelos novos desse gene, especialmente em zebuínos. Pesquisas adicionais podem conduzir à descoberta de novos alelos e elucidar a interação de processos celulares que ocorrem na reação imunológica em resposta à mastite, que afeta de maneira significativa a qualidade e os níveis da produção de leite.

6.2 Material e Métodos

6.2.1 Animais experimentais

Foram utilizadas 40 vacas da raça Gir Leiteira (*Bos indicus*) integrantes de duas fazendas particulares que participam do Teste de Progênie da Embrapa Gado de Leite, com lactações completas ocorridas em três anos. Estes 40 animais foram selecionados nos extremos da população, considerando os *quartis* superiores e inferiores da distribuição da contagem de células somáticas. Do total, vinte animais apresentaram alta contagem de células somáticas, e os restantes apresentaram baixa contagem.

6.2.2 Extração de DNA e amplificação do gene BoLA-DRB3

Cinco ml de sangue foram coletados dos animais, sendo que a extração de DNA foi procedida conforme publicado anteriormente (da Mota et al., 2002), o mesmo tendo ocorrido

com a amplificação do segundo exon do gene BoLA-DRB3, exceção somente para o número de ciclos da segunda etapa de PCR que foi realizada em 35 ciclos.

6.2.3 Clonagem do produto de PCR e seqüenciamento

Após a reação de PCR, foram reunidos em um único tubo *epENDORF* 10 µl equimolarmente distribuídos de amostras amplificadas dos animais que apresentaram alta CCS e em outro tubo foram reunidos 10 µl das amostras amplificadas dos animais que apresentaram baixa CCS. A seguir, as amostras foram purificadas e empregadas separadamente na clonagem no vetor *pCR4-Blunt-TOPO* usando o kit *Zero Blunt® TOPO® PCR Cloning Kit* (Invitrogen), sendo a transformação realizada em células competentes *E. coli (DH5a)*. O protocolo de seqüenciamento foi realizado segundo especificações do fabricante do kit *Big Dye Terminator Cycle Sequencing Ready* (Perkin-Elmer Applied Biosystems). Após a reação de seqüenciamento, as amostras foram purificadas e aplicadas no seqüenciador automático ABI 3100 (*Applied Biosystems*).

6.2.4 Análise das seqüências e genotipagem dos animais

As seqüências obtidas dos clones foram analisadas com o auxílio de programas computacionais específicos, sendo retiradas as seqüências incompletas e as que apresentavam qualidade inferior a 18 (Phred-Phrap-Consed). Também foram retiradas as seqüências com homologia com o vetor utilizado. Os alelos foram identificados por meio da análise do alinhamento das seqüências com os alelos anteriormente identificados e disponíveis no GenBank (<http://www.ncbi.nlm.nih.gov/>), com o emprego do programa Clustal-W (<http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html>). A identificação de cada alelo foi realizada apenas quando a seqüência consenso tinha homologia em todas as posições com o alelo publicado previamente.

6.3 Resultados e Discussão

Foram seqüenciados 29 clones, permitindo a identificação de 18 alelos do gene BoLA-DRB3 (Tabela 14). Dentre estes, foram identificados 11 alelos considerados potencialmente como novos, não descritos anteriormente.

Sete alelos confirmaram seqüências previamente publicadas, representando uma freqüência alélica de 58 % de alelos já conhecidos nos animais estudados. Os demais alelos (42 %) representam seqüências ainda não publicadas. Os alelos mais freqüentes foram: *0902 (21 %), *3601 (10 %) e *2801 (10 %). Em estudo anterior (da Mota et al., 2002) os alelos *3601 e *2801 já haviam sido identificados em vacas da raça Gir Leiteira, com freqüências, respectivamente, de 18 e sete por cento. No entanto, o alelo *0902 ainda não havia sido encontrado em vacas Gir, demonstrando a variação passível de ser encontrada com a utilização de animais provenientes de outras raças e origens. As seqüências identificadas nos alelos novos estão apresentadas a seguir na Tabela 15. A Tabela 16 permite visualizar o alinhamento e as diferenças entre as seqüências apresentadas na Tabela 15.

Tabela 14 . Freqüência dos alelos do gene BoLA-DRB3 em vacas da raça Gir Leiteira (*Bos indicus*)

Alelo BoLA-DRB3	Número de cópias
*0902 (*11) [†]	6
*3601 (*20)	3
*2801 (*31)	3
*4101 (*29)	2
*1201 (*8)	1
*4301 (*54)	1
*1501 (*16)	1
H04 [‡]	2
A01 [‡]	1
A02 [‡]	1
A05 [‡]	1
B01 [‡]	1
C01 [‡]	1
C02 [‡]	1
D02 [‡]	1
F01 [‡]	1
F02 [‡]	1
F08 [‡]	1

[†] Entre parênteses, a nomenclatura segundo PCR-RFLP

[‡] Alelo sem homólogo nas bases públicas de informações sobre genoma.

Tabela 15 . Sequências de nucleotídeos do gene BoLA-DRB3 de vacas Gir Leiteira (*Bos indicus*) dos alelos em que não foi possível obter homologia com alelos depositados nas bases públicas de informações sobre genomas

A01	TGGAGTATTATAAGGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCG GTTCTGGACAGACTTCTATAATGGAGAAAAGTTCGTGCGCTTCGACAGC GACTGGGGCGAGTCCGGGCGGTGACCGAGCTGGGGCGGCAGGACGCCGAG CAGTGGAACAGCCAGAAGGACACCCTGGAGCGGGAGCGGGCCTATGTGGAC ACGTA CTGCAGACACA ACTACGGGGTTCGGTGAGAGTTTCACTGAA
A02	TGGAGTATTCTAAGAGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCG GTTCTGGAGAGATCCTTCTATAATGGAGAAGAGAACGTGCGCTTCGACAGC GACTGGGGCGAGTACCGGGCGGTGACCGAGCTAGGGCGGCCGGACGCCGAG TACTGGAACAGCCAGAAGGACTTCCTGGAGCAGAGGGCGGGCCGCGGTGGAC ACGTA CTGCAGACACA ACTACGGGGTTCGGTGAGAGTTTCACTGTG
A05	TGGAGTATTCTAAGAGAGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGATACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGC GACTGGGGCGAGTACCGGGCGGTGACCGAGCTAGGGCGGCCGGACGCCGAG TACTGGAACAGCCAGAAGGAGATCCTGGAGCGGGCGCGGGCCGCGGTGGAC ACGTA CTGCAGACACA ACTACGGGGTTCGGTGAGAGTTTCACTGTG
B01	TGCAGTATCATAAGGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGACTTCTATAATGGAGAAGAGTACGTGCGCTTCGACAGC GACTGGAGCGAGTACCGGGCGGTGACCGAGCTAGGGCGGCCGGACGCCGAG TACTGAAACAGCCAGAAGGAGATCCTGAAGCGGCCGCGGCCCGCGGTGAAC ACGTA CTGCAGACACA ACTACGGGGTTCGGTGAGAGTTTCACTGTG
C01	TGCAGTATCATAAGGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGC GACTGGGGCGAGTCCGGGCGGTGACCGAGCTGGGGCGGCCGGCCGCCGAG TACTGGAACAGCCAGAAGGACTTCCTGGAGGAGAGGGCGGGCCGAGGTGGAC ACGTA CTGCAGACACA ACTACGGGGTTCGGTGAGAGTTTCACTGTG
C02	TGGAGTATTATAAGGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCG GTTCTGGACAGACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGC GACTGGGGCGAGTCCGGGCGGTGACCGAGCTGGGGCGGCAGGACGCCGAG CAGTGGAACAGCCAGAAGGACACCCTGGAGCGGGAGCGGGCCTATGTGGAC ACGTA CTGCAGACACA ACTACGGGGTTCGGTGAGAGTTTCACTGAA

Tabela 15 . Sequências de nucleotídeos do gene BoLA-DRB3 de vacas Gir Leiteira (*Bos indicus*) dos alelos em que não foi possível obter homologia com alelos depositados nas bases públicas de informações sobre genomas

D02	TGCAGTATCATAAGGGCGAGTGTCATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGACTTCTATAATGGAGAAAAGTGCGTGCCTTCGACAGC GACTGGGACGAGTTCCGGGCGGTGACCGAGCTGGGGCGGCCGGCCCGGAG TACTGGAACAGCCAGAAGGACTTCCTGGAGGAGAGGGCGGGCCGAGGTGGAC ACGTACTGCAGACACAACACTACGGGGGCGTGGAGAGTTTCACTGTG
F01	TGGAGTATTCTAAGAGCGAGTGTCATTTCTTCAACGGGACCGAGCGGGTGCG GTTCTGACAGATACTTCCATAATGGAGAAGAGTACGTGCCTTCGACAGC GACTGGGGCGAGTACCGGGCGGTGACCGAGCTGGGGCAGCGGGTTCGCCGAG TACTGCAACAGCCAGAAGGACTTCCTGGAGCGGGCGCGGGCCCGGTTGGAC ACGTACTGCAGACACAACACTACGGGGTTCGGTGAGAGTTTCACTGTG
F02	TGCAGTATCATAAGGGCGAGTGTCATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGACTTCTATAATGGAGAAGAGTACGTGCCTTCGACAGC GACTGGAGCGAGTACCGGGCGGTGACCGAGCTAGGGCGGCCGGACGCCGAG TACTGGAACAGCCAGGAGATCCTGGAGCGGGCGCGGGCCCGGTTGGACACG TACTGCAGACACAACACTACGGGGTTCGGTGAGAGTTTCACTGTG
F08	TGCAGTATCATAAGGGCGAGTGTCATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGACTTCTATAATGGAGAAGAGTTCGTGCCTTCGACAGC GACTGGGGCGAGTACCGGGCGGTGACCGAGCTAGGGCGGCCGGTTCGCCGAG CAGTTGAACGGCCAGAAGGACACCCTGGAGCGGGAGCGGGCCTATGTGGAC ACGTACTGCAGACACAACACTACGGGGTTCGGTGAGAGTTTCACTGTG
H04	TGGAGTATTGTAAGAGAGAGTGTCATTTCTTCAACGGGACCGAGCGGGTGCG GTTGCTGGACAGATACTTCTATAATGGAGAAGAGTACGTGCCTTCGACAGC GACTGGAGCGAGTACCGGGCGGTGACCGAGCTAGGGCGGCCGGACGCCGAG TACTGGAACAGCCAGGAGATCCTGGAGCGGGCGCGGGCCCGGTTGGACACG TACTGCAGACACAACACTACGGGGTTCGGTGAGAGTTTCACTGTG

Tabela 16 . Alinhamento de nucleotídeos das seqüências do gene BoLA-DRB3 de vacas Gir Leiteira (*Bos indicus*) dos alelos em que não foi possível obter homologia com alelos depositados nas bases públicas de informações sobre genomas

Ident.*	Seqüência de nucleotídeos	
A02	TGGAGTATTCTAAGAGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
F01	TGGAGTATTCTAAGAGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
A05	TGGAGTATTCTAAGAGAGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
H04	TGGAGTATTCTAAGAGAGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
B01	TGCAGTATCATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
F02	TGCAGTATCATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
C01	TGCAGTATCATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
D02	TGCAGTATCATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
A01	TGGAGTATTATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
C02	TGGAGTATTATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
F08	TGCAGTATCATAAAGGCGAGTGTCAATTTCTTCAACGGGACCGAGCGGGTGCGGTTCCTGG	60
	** ***** * *****	
A02	AGAGATCCTTCTATAATGGAGAAGAGAACGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
F01	ACAGATACTTCCATAATGGAGAAGAGTACGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
A05	ACAGATACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
H04	ACAGATACTTCTATAATGGAGAAGAGTACGTGCGCTTCGACAGCGACTGGAGCGAGTACC	120
B01	ACAGACACTTCTATAATGGAGAAGAGTACGTGCGCTTCGACAGCGACTGGAGCGAGTACC	120
F02	ACAGACACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
C01	ACAGACACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
D02	ACAGACACTTCTATAATGGAGAAAAGTTCGTGCGCTTCGACAGCGACTGGGACGAGTACC	120
A01	ACAGACACTTCTATAATGGAGAAAAGTTCGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
C02	ACAGACACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
F08	ACAGACACTTCTATAATGGAGAAGAGTTCGTGCGCTTCGACAGCGACTGGGGCGAGTACC	120
	* *** ***** ** ***** ***** **	
A02	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACTTCC	180
F01	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACTTCC	180
A05	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGAGATCC	180
H04	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCA--GGAGATCC	177
B01	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGAGATCC	180
F02	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCA--GGAGATCC	177
C01	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACTTCC	180
D02	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACTTCC	180
A01	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACACCC	180
C02	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACACCC	180
F08	GGGCGGTGACCGAGCTAGGGCGGCAGCCGAGTACTGGAAACAGCCAGAAGGACACCC	180
	***** ***** ** * ***** * * *** ***** **	
A02	TGGAGCAGAGGCGGGCCGCGGTGGACACGTA CTGCAGACACA ACTACGGGGTCGGTGAGA	240
F01	TGGAGCAGAGGCGGGCCGCGGTGGACACGTA CTGCAGACACA ACTACGGGGTCGGTGAGA	240
A05	TGGAGCAGAGGCGGGCCGCGGTGGACACGTA CTGCAGACACA ACTACGGGGTCGGTGAGA	240
H04	TGGAGCAGAGGCGGGCCGCGGTGGACACGTA CTGCAGACACA ACTACGGGGTCGGTGAGA	237
B01	TGAAGCGCCGCGGGCCGCGGTGGACACGTA CTGCAGACACA ACTACGGGGTCGGTGA	240
F02	TGGAGCAGAGGCGGGCCGCGGTGGACACGTA CTGCAGACACA ACTACGGGGTCGGTGAGA	237

Tabela 16 . Alinhamento de nucleotídeos das seqüências do gene BoLA-DRB3 de vacas Gir Leiteira (*Bos indicus*) dos alelos em que não foi possível obter homologia com alelos depositados nas bases públicas de informações sobre genomas

C01	TGGAGGAGAGGCGGGCCGAGGTGGACACGTA	CTGCAGACACA	ACTACGGGGGTGTGGAGA	240
D02	TGGAGGAGAGGCGGGCCGAGGTGGACACGTA	CTGCAGACACA	ACTACGGGGGCGTGGAGA	240
A01	TGGAGCGGGAGCGGGCCTATGTGGACACGTA	CTGCAGACACA	ACTACGGGGTCGGTGAGA	240
C02	TGGAGCGGGAGCGGGCCTATGTGGACACGTA	CTGCAGACACA	ACTACGGGGTCGGTGAGA	240
F08	TGGAGCGGGAGCGGGCCTATGTGGACACGTA	CTGCAGACACA	ACTACGGGGTCGTTGAGA	240
	** ** *	**** **	*** *****	* ** *
A02	GTTTCACTGTG			251
F01	GTTTCACTGTG			251
A05	GTTTCACTGTG			251
H04	GTTTCACTGTG			248
B01	GTTTCACTGTG			251
F02	GTTTCACTGTG			248
C01	GTTTCACTGTG			251
D02	GTTTCACTGTG			251
A01	GTTTCACTGAA			251
C02	GTTTCACTGAA			251
F08	GTTTCACTGTG			251

* A ordem das seqüências foi estabelecida em função da similaridade entre as mesmas.

As Tabelas 15 e 16 permitem observar que as seqüências apresentaram extenso polimorfismo. As seqüências foram obtidas de seqüenciamento único (*single-pass*) e foram submetidas à análise de qualidade. O processo eliminou as seqüências que apresentaram qualidade inferior a 20, no critério dos softwares Phred-Phrap-Consed, bem como das seqüências dos vetores. Cumpre mencionar que mesmo com os critérios estabelecidos, as seqüências podem ter sofrido erro introduzido pela enzima *Taq* DNA polimerase. Esse fato está previsto na literatura com uma freqüência de 1:10.000 bases (Gyllestein & Allen, 1993), e mesmo entre novos alelos depositados por outros autores no gene BoLA-DRB3, há alelos que foram retirados após dupla checagem e repetição das ampliações (e.g., alelo *0102, <http://www.projects.roslin.ac.uk/bola/drftab.html>).

6.4 Conclusões

- O emprego da técnica de clonagem e seqüenciamento de um *pool* de amostras estratificadas nos extremos da população permitiram identificar alelos

potencialmente novos do gene BoLA-DRB3 em vacas da raça Gir Leiteira em 42 % das amostras.

- Há alelos não descritos e estudos de associação desses alelos com a ocorrência de mastite ou com a contagem de células somáticas podem auxiliar a elucidar a resistência a esta doença.
- A alta percentagem de alelos novos justifica a realização de outros estudos desse gene, com a genotipagem individual utilizando as técnicas desenvolvidas como o TGGE, apresentando como resultado a descoberta de alelos novos do gene BoLA-DRB3 em animais *Bos indicus*.

Agradecimentos

Os autores agradecem Fernanda Luiza de Souza e Lílian Giotto Zaros pela condução dos experimentos. Adilson F. da Mota agradece a Embrapa por ter recebido bolsa de doutoramento. Luiz Lehmann Coutinho e Mario Luiz Martinez agradecem ao CNPq por recebimento de bolsa de produtividade em pesquisa. Todos os experimentos envolvendo animais estão de acordo com a legislação brasileira e de acordo com as normas éticas, de bem estar e conforto da Embrapa para experimentos com animais.

7 CONSTRUÇÃO E CARACTERIZAÇÃO DE BIBLIOTECAS DE cDNA DE GLÂNDULA MAMÁRIA DE BOVINOS DA RAÇA GIR LEITEIRA (*Bos indicus*) E DA RAÇA HOLANDESA (*Bos taurus*) ⁶

* As seqüências de nucleotídeos reportadas foram submetidas à base de dados do GenBank e receberam os números de acesso: CA998517 até CA998586, CB040527 até CB042862 e CB059218 até CB063292.

Resumo

Atualmente, há aproximadamente uma ordem de magnitude menos informações a respeito de seqüências de nucleotídeos para o transcriptoma bovino do que a quantidade disponível para os transcriptomas do camundongo e do Homem. Além disso, há uma marcante ausência de informação do genoma das raças zebuínas (*Bos indicus*), as quais dominam a produção de carne e leite bovina nos mercados tropicais. Portanto, gerar mais etiquetas de seqüência expressas (EST – *Expressed Sequence Tags*) derivadas de gado zebuino é importante para caracterizar a diversidade genética existente entre as raças, referente à saúde e produção dos animais, bem como auxiliar na correta anotação, identificação e classificação de genes bovinos. Essa informação será essencial na realização de estudos de genômica funcional. O presente estudo empregou uma modificação na estratégia de produção de EST, conhecida como ORESTES (*Open Reading Frame – ORF EST*). Foram construídas e caracterizadas 33 bibliotecas de cDNA usando a metodologia de ORESTES. Foram depositadas 6.481 ESTs em bases públicas de informações sobre genomas empregando tecidos de glândula mamária de uma novilha da raça Holandesa (*Bos taurus*) e de vacas da raça Gir Leiteira (*Bos indicus*) infectadas durante o período seco. O alinhamento e agrupamento das 6.481 EST geraram 1.700 elementos únicos de seqüência, e dentre esses, 1.157 não produziram alinhamentos com as seqüências depositadas no *Bos taurus* gene index (BtGI) (ver 8.0), considerando limite de $e=1,0 \times 10^{10}$. No entanto, 326

⁶ Artigo enviado para publicação no periódico Animal Genetics.

desses 1.157 elementos únicos de seqüência foram alinhados significativamente com ESTs de camundongos e do Homem, demonstrando que essas bibliotecas podem ser empregadas eficientemente para descobrir novas seqüências gênicas em bovinos. Uma análise de comparação da posição do alinhamento das ESTs bovinas geradas com genes Humanos seqüenciados em toda sua extensão (*full-length*) demonstrou que a maioria dos alinhamentos ocorreu na parte central dos transcritos, o que sugere estar relacionado com a relativa alta taxa de descoberta de seqüências do presente estudo.

Summary

Currently, there is approximately an order of magnitude less sequence information in public databases than that available for the transcriptomes of human and mouse. Furthermore, there is a complete lack of sequence information from *Bos indicus* (zebu) breeds, which dominate the large cattle markets of tropical production areas. Generating expression sequence tags (EST) from zebu cattle is important to better characterize bovine genetic diversity related to health and production and to more thoroughly identify, annotate, and classify bovine genes. This information will be essential in functional genomic studies. We used a modification on the strategy to produce EST, named ORESTES (Open Reading Frame –ORF EST). We constructed and characterized 33 cDNA libraries using ORESTES methodology. We deposited 6,481 EST in public databases using healthy and infected involuted mammary gland tissues from Holstein and mature Brazilian Dairy Gir cows (zebu). Alignment and assembly of the 6,481 EST generated 1,700 unique sequence elements, and of these, 1,157 did not significantly match sequence assemblies available in the *Bos taurus* gene index (BtGI-TIGR), with a threshold of $e=10 \times 10^{-10}$. However, a total of 326 of these 1,157 sequence elements did comparatively align with mouse and human EST demonstrating that these libraries can be used to efficiently discover novel sequences. Furthermore, comparison of the alignment position of all bovine sequence elements matching human full-length reference cDNA sequences suggested that the majority of arbitrarily priming events for cDNA synthesis occurred towards the central portion of the transcripts, which may be correlated to the relatively high rate of novel sequence discovery.

7.1 Introdução

Bibliotecas de cDNA são coleções de moléculas em fita dupla de DNA complementar (cDNA), sintetizado *in vitro* a partir do RNA mensageiro (mRNA), inseridas em plasmídios e transformadas em bactérias competentes, de maneira a permitir armazenagem e propagação.

Um dos principais objetivos dos estudos em genômica realizados nos mamíferos é caracterizar o padrão de expressão gênica que corresponde a eventos fisiológicos importantes relacionados com a produção e a saúde dos animais. Um recurso necessário para atender esse objetivo é obter seqüências dos clones das bibliotecas de cDNA que representam a expressão do mRNA que estava ocorrendo em determinada célula ou tecido submetido a determinada circunstância de estudo. Cada uma dessas seqüências é chamada de Etiqueta de Seqüência Expressa (*EST – Expressed Sequence Tag*). Atualmente, há mais de 5.000.000 ESTs disponíveis para Humanos e 3.600.000 ESTs para camundongos. As coleções de ESTs para essas espécies continuam sendo ampliadas em larga escala e mais bibliotecas representando estádios de saúde ou de desenvolvimento seguem sendo amostradas para identificar genes e verificar as vias metabólicas nas quais os mesmos estão envolvidos. Esse volume de informações representa ao menos uma ordem de magnitude mais informação existente sobre o genoma dessas espécies do que sobre o genoma bovino.

Em bovinos, seqüências derivadas dos transcritos têm sido organizadas em um *Bos taurus* Gene Index (BtGI, <http://www.tigr.org/tdb/tgi/btgi/>) para prover informações úteis no desenho ou interpretação dos experimentos de genômica funcional. No entanto, mais ESTs são necessárias do que as atuais 30 0.000 ESTs para que seja possível identificar, anotar e classificar os genes específicos da espécie bovina. Maior número de informações sobre os transcritos será fundamental para interpretar resultados gerados através de experimentos de genômica funcional, especialmente devido à completa ausência de informação sobre as seqüências gênicas dos zebuínos (*Bos indicus*). Essas raças dominam o mercado de produção bovina nas regiões tropicais e a geração de ESTs provenientes desses animais é importante para construir ferramentas de estudos genômicos que podem ser utilizadas para caracterizar a diversidade genética entre os animais de origem européia (*Bos taurus*) e os de origem zebuína (*Bos indicus*), em relação à saúde e produção.

Alem do fato de que mais ESTs são necessárias e de que a quase totalidade fora obtida de animais *Bos taurus*, a maioria das ESTs foi produzida com amostras de múltiplos tecidos (Smith *et al.*, 2001). Desta forma, as ESTs bovinas, na maioria, foram produzidas com o objetivo

de manter uma elevada taxa de descoberta de novas ESTs e não com o objetivo de amostrar tecidos específicos ou situações fisiológicas de interesse. Por isso, não há um volume representativo de ESTs provenientes de tecidos da glândula mamária. Há somente 23.200 ESTs obtidas do parênquima da glândula mamária em diversas situações fisiológicas (Sonstegard et al., 2002). Não há, portanto, informações suficientes e representativas sobre os transcritos expressos de glândula mamária e suas relações com as características de produção e de saúde da glândula do animal. Nos estudos em Humanos, as informações não estão relacionadas à produção ou sanidade. As ESTs de tecido mamário em Humanos estão relacionadas, na maioria, à presença de adenocarcinoma, sendo provenientes de tecidos obtidos de tumores (<http://www.ncbi.nlm.nih.gov/UniGene/lbrowse.cgi?ORG=Hs>).

Outra característica importante das ESTs de bovinos recai sobre a metodologia de construção das bibliotecas de cDNA empregadas nesses estudos. As bibliotecas mais importantes e representativas geraram seqüências nas extremidades 5' - dos transcritos (Smith et al., 2001; Sonstegard et al., 2002). Outros estudos produziram seqüências a partir da extremidade 3' - dos transcritos (Takasuga et al., 2002). Para o caso de mensagens curtas, as seqüências obtidas das duas extremidades dos mesmos genes podem ser alinhadas empregando programas de computador desenvolvidos para esse objetivo através da superposição de seqüências idênticas. No entanto, para mensagens longas não é possível obter as informações sobre a região central dos genes, o que produz diferentes resultados no agrupamento das informações sobre os mesmos genes.

No presente estudo, foi empregada uma modificação na estratégia de produzir ESTs denominada ORESTES (*Open Reading Frame – ORF EST*), para a geração de uma série de bibliotecas de cDNA chamadas BARC-EMBRAPA. A metodologia de ORESTES, desenvolvida por Dias Neto et al. (2000) e De Souza et al. (2000), visa gerar seqüências aleatoriamente ao longo dos transcritos, ao invés de iniciar a produção de ESTs nas extremidades desses (Camargo et al., 2001), e não havia sido anteriormente empregada em bovinos. A metodologia de construção de bibliotecas de cDNA por ORESTES foi desenvolvida para otimizar e maximizar a descoberta de novos genes bovinos e minimizar os custos de produção de ESTs.

O presente estudo reporta o depósito de 6.481 ESTs, derivadas da glândula mamária de bovinos, em bases públicas de informações sobre genomas. A maioria dessas está localizada na região central dos genes e representam uma contribuição significativa às informações disponíveis sobre expressão gênica bovina, especialmente em *Bos indicus* e referentes exclusivamente à glândula mamária.

7.2 Material e Métodos

7.2.1 Coleta de tecidos e desafio com *Staphylococcus aureus*

Tecidos da glândula mamária de uma novilha da raça Holandesa foram coletados no abatedouro localizado no centro de pesquisas *Beltsville Agricultural Research Center* (BARC-ARS-USDA) para construir as bibliotecas de cDNA derivadas de *Bos taurus*. Os tecidos para as bibliotecas de *Bos indicus* foram importados por permissão do *United States Veterinary Permit for Importation and Transportation of Controlled Materials and Organisms and Vectors* (Permit# 48425), vindos do Brasil. Esses tecidos foram coletados no Centro da Embrapa Gado de Leite, em Juiz de Fora - MG, durante o abate de quatro vacas adultas da raça Gir Leiteira que estavam no período seco entre lactações. Esses animais foram previamente desafiados através de uma inoculação intramamária de solução contendo cepas de *Staphylococcus aureus*. A estirpe 284 de *Staphylococcus aureus* foi previamente isolada de um caso clínico de infecção e mantida congelada a -20°C em leite magro contendo 10% de glicerol. Para a preparação das doses, foi seguido o protocolo descrito por Schukken et al. (1999). Após determinar o número de unidades formadoras de colônia (CFU – *Colony Forming Units*), a cultura foi diluída em PBS para padronização da dose de inoculação entre os tetos dos animais e cada dose conteve 300 cfu/ml contidos em um volume de 5,0 ml. Antes das inoculações, secreções de cada quarto mamário foram coletadas e enviadas para exame de cultura bacteriana. Apenas as vacas que apresentaram resultados negativos prévios para o exame microbiológico foram inoculadas e tiveram amostras de tecidos coletadas. Antes das inoculações, os animais tiveram seus tetos depositados por imersão em solução de iodophor 1% e limpos com etanol 70% embebido em toalha de papel. As infusões se deram através da introdução de uma seringa estéril através do canal do teto, seguida de uma breve massagem dirigindo a solução inoculante na direção da cisterna do leite. Os animais foram abatidos no abatedouro local, depois de decorridas 72 horas da inoculação, de acordo com os protocolos éticos, de bem estar e conforto dos animais aprovados pela Embrapa. Tecidos da glândula mamária foram então coletados e imediatamente submersos em RNALater (Ambion, Austin, Texas, USA), sendo posteriormente congelados à temperatura de -20°C até o momento da extração do RNA total. As bibliotecas foram construídas utilizando amostras de um único teto ou quarto de um animal ou amostras múltiplas resultantes da reunião de amostras individuais do parênquima da glândula mamária.

7.2.2 Preparação do RNA

As amostras de RNA foram isoladas e purificadas em um protocolo de três etapas que compreenderam a extração com *Qiagen RNeasy Mini Kit* (Qiagen, Hilden, Germany), seguida da remoção de traços de DNA genômico com o tratamento de *Dnase I* utilizando *Rnase-free Dnase Set* (Qiagen, Hilden, Germany), e posteriormente a purificação do mRNA com *MicroPoly(A) Purist* (Ambion, Abingdon, Oxon, UK).

Todas as amostras foram submetidas à avaliação de concentração e qualidade empregando o equipamento *Agilent 2100 Bioanalyser* (Agilent, Palo Alto, California, USA). A integridade das amostras foi avaliada através do *RNA 6000 Nano Assay* (Agilent, Palo Alto, California, USA) empregando as recomendações do fabricante. Procedeu-se um controle de qualidade rigoroso para avaliar a qualidade das amostras de RNA e somente as que apresentaram qualidade superior nos eletroferogramas (Figura 12A) e uma relação 28S:18S superior a 1,9 foram utilizadas nas etapas seguintes do protocolo.

7.2.3 Construção das bibliotecas e seqüenciamento

As bibliotecas foram construídas utilizando modificações nos protocolos descritos por Dias Neto et al. (2000), De Souza et al. (2000) e Camargo et al. (2001). Amostras de mRNA purificadas (± 30 ng) foram aquecidas a 65 °C durante 10 minutos e submetidos à síntese de cDNA mediante transcrição reversa utilizando *SuperScript II First-strand Synthesis System for RT-PCR* (Invitrogen, Carlsbad, California, USA). As reações de *first-strand* foram realizadas a 37 °C durante 60 minutos na presença de 200 unidades de transcriptase reversa, 1 x RT Buffer, 10 μ M dNTPs, 10 mM dithiothreitol (DTT) e 10 pmoles de *primers*, em um volume final de 10 μ l. Inicialmente, durante a otimização inicial da metodologia de ORESTES, *primers* ao redor de 23 pares de bases que amplificam genes da família dos receptores do estrógeno foram empregados. Para aumentar a diversidade dos produtos gerados, nas bibliotecas de *Bos indicus*, foram empregados *primers* escolhidos aleatoriamente com teores de Guanina e Citosina ao redor de 50-75% e com 16 pares de bases. A seqüência dos *primers* utilizados esta apresentada na Tabela 17. Para a amplificação de cDNA, um décimo da reação de *first-strand* de cDNA foi submetido à amplificação por PCR na presença do mesmo *primer* utilizado na primeira reação, conforme descrito por Camargo et al. (2001) e De Souza et al. (2000). A estimativa da

complexidade relativa potencial dos perfis gerados por PCR foi verificada com a visualização por meio de eletroforese em gel de agarose 1% (p/v) (Figuras 12B e 12C). Produtos com peso molecular entre 500 e 1500 bp e apresentando diversidade potencial das mensagens (Figura 12C) sofreram excisão do gel e foram submetidos à purificação e tiveram as extremidades coesivas tratadas com *Klenow fragment of T4 DNA Polymerase I* (Amersham, Piscataway, New Jersey, EUA). Após a eluição e tratamento das extremidades com a enzima *Klenow*, as moléculas com as extremidades tornadas abruptas e com o peso molecular situado entre os parâmetros mencionados foram clonadas em vetores e transformadas em bactérias competentes (*Escherichia coli*) (Invitrogen, Carlsbad, CA, EUA). Colônias positivas resultantes da transformação foram coletadas individualmente e alocadas em placas de 384 amostras contendo 0,2 ml de meio líquido Luria Bertani contendo 100 µg/ml de ampicilina e crescidas por 16 horas.

Tabela 17. Sequências dos *primers* utilizados na síntese de cDNA e amplificação para a construção das bibliotecas de cDNA

<i>Primer</i>	bp	Sequência (5' 3')
Erâ – 2A	23	CCCTGTGACCAAAGACTTGTGTC
Erâ – 2B	23	GACTCACCTGCTGAATGCTGTGA
Erá – 1A	20	TGTGTGGAAGGCATGGTGGGA
Erá – 1B	21	GGCACCACGTTCTTGCACTTC
Err1 – 1A	23	GGCATGGCATAACAGCTTCTCAGG
Err1 – 1B	23	AGTACAGCTGTCCGGCCTCCAAC
AM1	16	GGTGCCACTGGACTCG
AM2	16	CTGGACTCGTGCGTCA
AM3	16	CACTGTGCGACTCGTC
AM4	16	CGTCGTACTIONGACGGA
AM5	16	CGCGAGTCGCGCACTT
AM6	16	CAGCACTGTGTGCGCGT

A produção de ESTs, antes do arquivamento das colônias em estoques com glicerol, foi realizada mediante a coleta de uma alíquota de 2 µl do meio crescido submetidos a reação de PCR conforme protocolos descritos por Hoffman e colaboradores (<http://microarray.cnmcresearch.org/resources.htm>). A única alteração procedida nesse protocolo foi referente à redução do volume da reação para 10 µl. Após a reação de PCR, com auxílio de equipamento robotizado, Genesis RSP 100 (Tecan Systems Inc., San Jose, California, EUA), 2 µl (100 – 200 ng) dessa reação foram transferidos para nova placa de 384 poços. A degradação dos *primers* foi obtida com a adição de *Exonuclease I* (Epicentre, Madison, WI, EUA) seguida de reação de seqüenciamento conforme o protocolo descrito por Sonstegard et al. (2002) empregando seqüenciador automático ABI 3700 *Automated Sequencer* (Applied Biosystems, Foster City, California, EUA).

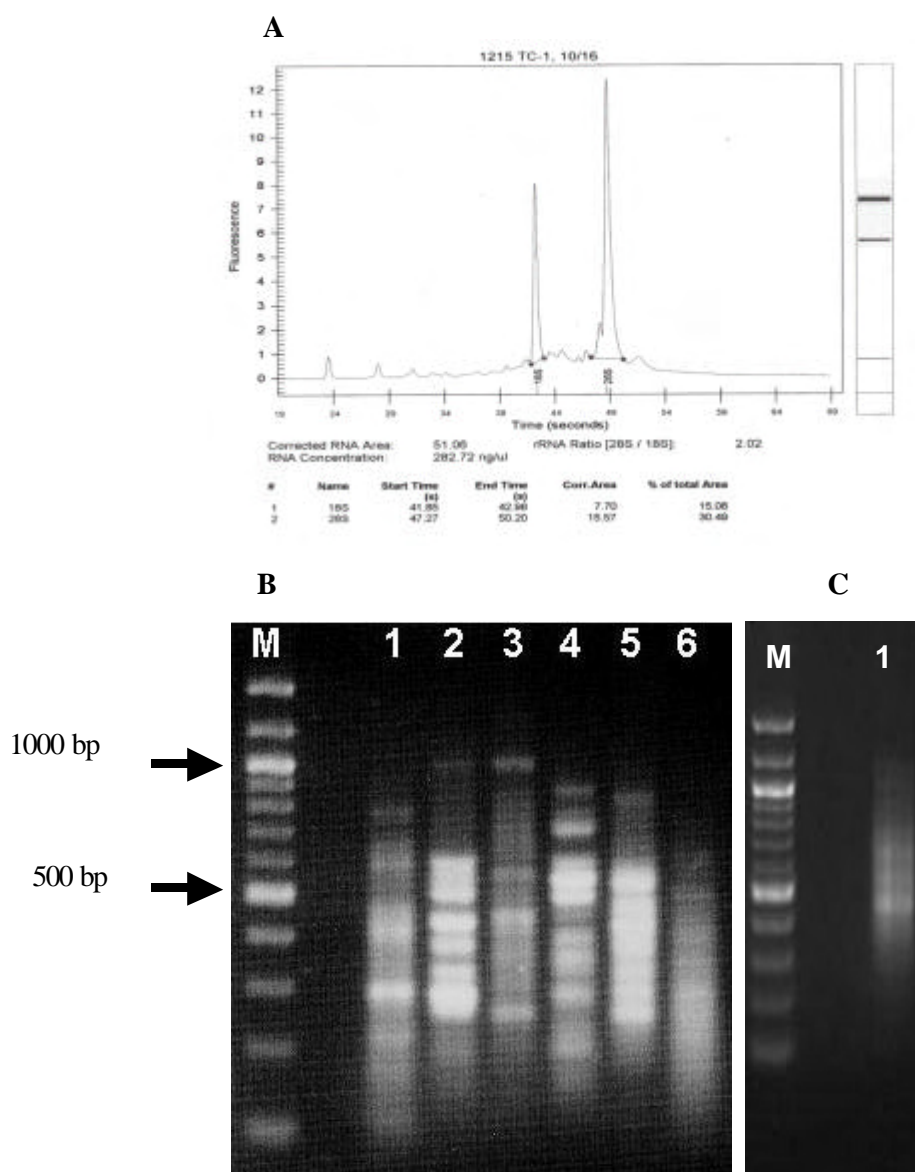


Figura 12 - A. Eletroferograma de amostra de RNA total com visualização da qualidade e concentração. B. Seleção de perfis para a construção de bibliotecas de cDNA. A imagem do gel preparativo é apresentada. A canaleta M contém o marcador de peso molecular (100 bp *ladder*). As amostras de 1 a 6 foram preparadas para seleção. As amostras 1, 3, 5 e 6 foram selecionadas para construção da biblioteca. C. Perfil de cDNA após a recuperação do gel da amplificação por PCR, recomendável para preparação de biblioteca.

7.2.4 Análises computacionais de qualidade e agrupamento

Todas as ESTs foram submetidas a um fluxo automatizado de tratamento de informações por intermédio de uma pipeline ⁷. O processo constou de, inicialmente, eliminação de seqüências com menos de 100 bp e de qualidade inferior a 18 no critério do software Phred, bem como das seqüências dos vetores e seqüências dos *primers*, em três etapas sucessivas. O processo também eliminou seqüências com similaridade ao genoma de *E. coli* superiores ao limite de 1.0×10^{-10} . As seqüências também foram eliminadas caso apresentassem similaridade com elementos repetitivos específicos de *introns* (Sheikh et al., 2002), denotando contaminação com DNA genômico ou presença de transcritos imaturos. As seqüências apresentando baixa complexidade com a presença sucessiva de séries dos mesmos nucleotídeo ou comprimento inferior a 100 bp foram também suprimidas ao final do processo. As demais apresentaram alta-qualidade, suficiente para submissão à base de dados dbEST do GenBank (<http://www.ncbi.nih.gov/dbEST/index.html>). As bibliotecas foram posteriormente submetidas à análise de agrupamento, utilizando o software *Contig Assembling Program* (CAP3) (Huang & Madan, 1999) que produziu seqüências agrupadas em consensos (TC - *Tentative Consensus*) e seqüências isoladas que não puderam ser agrupadas (ST - *Singleton*). Os dados referentes a todas as ESTs e os elementos únicos de seqüência (UNIQUE), que representaram o conjunto formado por TCs e STs, foram a seguir analisados através do software BLAST (BLAST - *Basic Local Alignment Search Tool*) (Altschul et al., 1990). O programa BLASTN foi utilizado para gerar alinhamentos das ESTs com as seqüências do *Bos taurus Gene Index* (BtGI) (ver 7.0) depositadas no *The Institute of Genomic Research* (TIGR) (<http://www.tigr.org>), e com ESTs provenientes de Humanos, camundongo e outras espécies conforme descrito por Quackenbush et al. (2000). BLASTN foi também utilizado para gerar alinhamentos com a base 'nr' do *National Council of Biotechnology Information* do *National Institute of Health* (NCBI/NIH/GenBank) (<http://www.ncbi.nlm.nih.gov/>). BLASTX foi utilizado para produzir alinhamentos com proteínas traduzidas nas bases de dados 'nt' e SWISS-PROT. Os resultados das análises com BLAST foram compatibilizadas em uma base de dados (MYSQL) que também foi alimentada com informações contendo o conjunto de seqüências completas (*full-length*) dos cDNAs Humanos depositadas no *Mammalian Gene Collection* do *National Institute of Health* (<http://mgc.nci.nih.gov/>) para analisar a posição relativa dos alinhamentos das ESTs em relação à seqüência completa dos genes. Informação relativa à ontologia dos genes (GO - "*Gene*

⁷ MATUKUMALLI, L. (BARC-ARS-USDA). Informação pessoal, jun, 2003.

Ontology") também integrou a base de dados para permitir classificar a função molecular, o componente celular e o processo biológico no qual as ESTs estavam envolvidas.

7.3 Resultados e discussão

Estudos de genômica funcional objetivam prover informações sobre a função de seqüências de DNA, identificando padrões de expressão gênica. Um dos recursos mais importantes para a execução desses estudos é o seqüenciamento de clones e a análise de ESTs, que somente é possível mediante a construção e caracterização de bibliotecas de cDNA.

No presente estudo foram construídas 33 bibliotecas de cDNA utilizando exclusivamente amostras de RNA de glândula mamária de fêmeas de origem *Bos taurus* e *Bos indicus*. 9.792 seqüências foram obtidas com a metodologia de ORESTES, empregada para aumentar a freqüência de descoberta de transcritos, especialmente na região central dos genes. As bibliotecas foram construídas com tecidos de glândula mamária para aumentar a representação desses tecidos sobre o transcriptoma bovino. Ao tempo em que este estudo estava em andamento, a representação de seqüências de animais de origem *Bos indicus* estava limitada a 342 seqüências. As bibliotecas construídas adicionaram 6.481 ESTs consideradas de alta qualidade às bases públicas de informações sobre genoma, sendo que 5.950 ESTs são oriundas da raça Gir Leiteira (*Bos indicus*) (Tabela 18).

Tabela 18 . Número de bibliotecas e número de ESTs depositados no GenBank em relação ao status de agrupamento (TC ou ST) e classificação das bibliotecas por raça (Holandesa ou Gir)

	Bibliotecas	ESTs Total	TCs	ESTs/TCs	STs
Holandesa	6	531	86	365	166
Gir	27	5.950	454	4.762	1.188
Total	33	6.481	594	5.124	1.307

Conforme visualizado na Tabela 18, 6.481 seqüências expressas em genes localizados na glândula mamária foram submetidas às bases públicas de informações sobre genomas, sendo a grande maioria oriunda de *Bos indicus* (*accession numbers* CA998517 até CA998586, CB040527 até CB042862 e CB059218 até CB063292). Os dados brutos derivados do equipamento ABI 3700 sofreram análise computacional e as seqüências que não apresentaram qualidade suficiente para submissão foram eliminadas de acordo com as razões apresentadas na Figura 13. O principal motivo de eliminação de seqüências foi devido à qualidade e comprimento das mesmas. Nessa categoria estão incluídas as amostras que apresentaram ausência de insertos, amostras somente com seqüências vetoriais e seqüências curtas, com comprimento inferior a 100 bp. Nessa categoria foram eliminadas 29% das amostras (2.830/9.792 amostras). Foram a seguir também eliminadas 4% das 9.792 seqüências devido à similaridade com o genoma de *E. coli*. Esse percentual de eliminação por similaridade com seqüências de *E. coli* é devido à contaminação por seqüências bacterianas, presumivelmente por causa de contaminação ocorrida em razão do desafio imposto aos animais anteriormente à coleta de amostras. Em duas bibliotecas (321BOV e 322BOV) a concentração de seqüências similares a *E. coli* foi de aproximadamente 80%, indicando ter havido contaminação. As demais bibliotecas apresentaram em média menos de 1% de contaminação com seqüências bacterianas. Outras causas, como a baixa complexidade das seqüências e a presença de elementos repetitivos representaram razões de menor ocorrência (1%) para rejeição das seqüências. Análise da similaridade das ESTs com seqüências de RNA ribossomal e seqüências mitocondriais revelaram mínima contaminação com percentuais de 1,9% e 0,3%, respectivamente, totalizando aproximadamente 2% de contaminação com seqüências ribossomais ou mitocondriais.

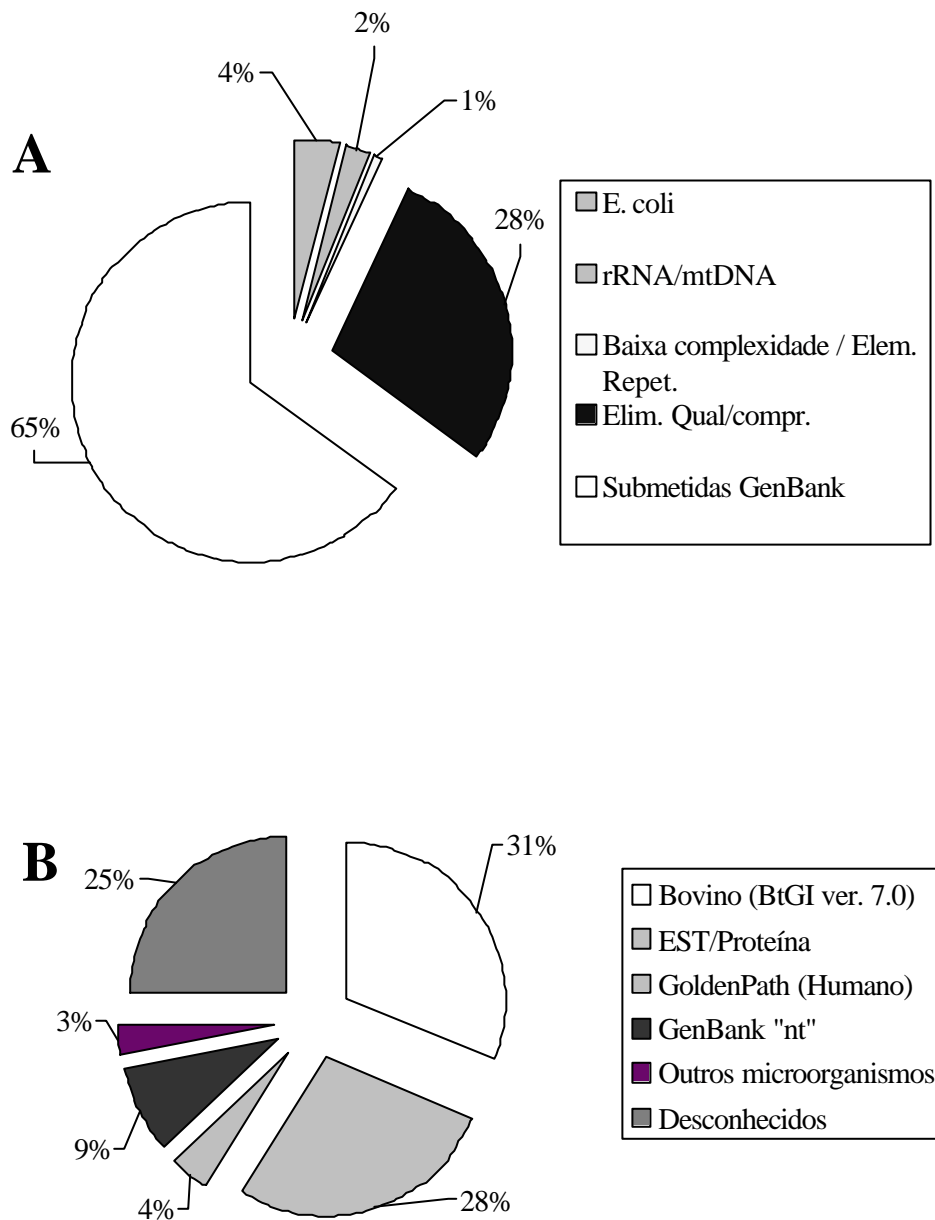


Figura 13 - A. Diagrama esquemático das razões de eliminação das ESTs. Considerando o total de 9.792 ESTs, 6.481 ESTs foram submetidas ao dBEST/GenBank (65%) B. O agrupamento das 6.481 ESTs originou 1901 UNIQUEs (TCs + STs) que foram submetidos à análise de alinhamento (BLAST) A classificação dos alinhamentos obtidos está representada

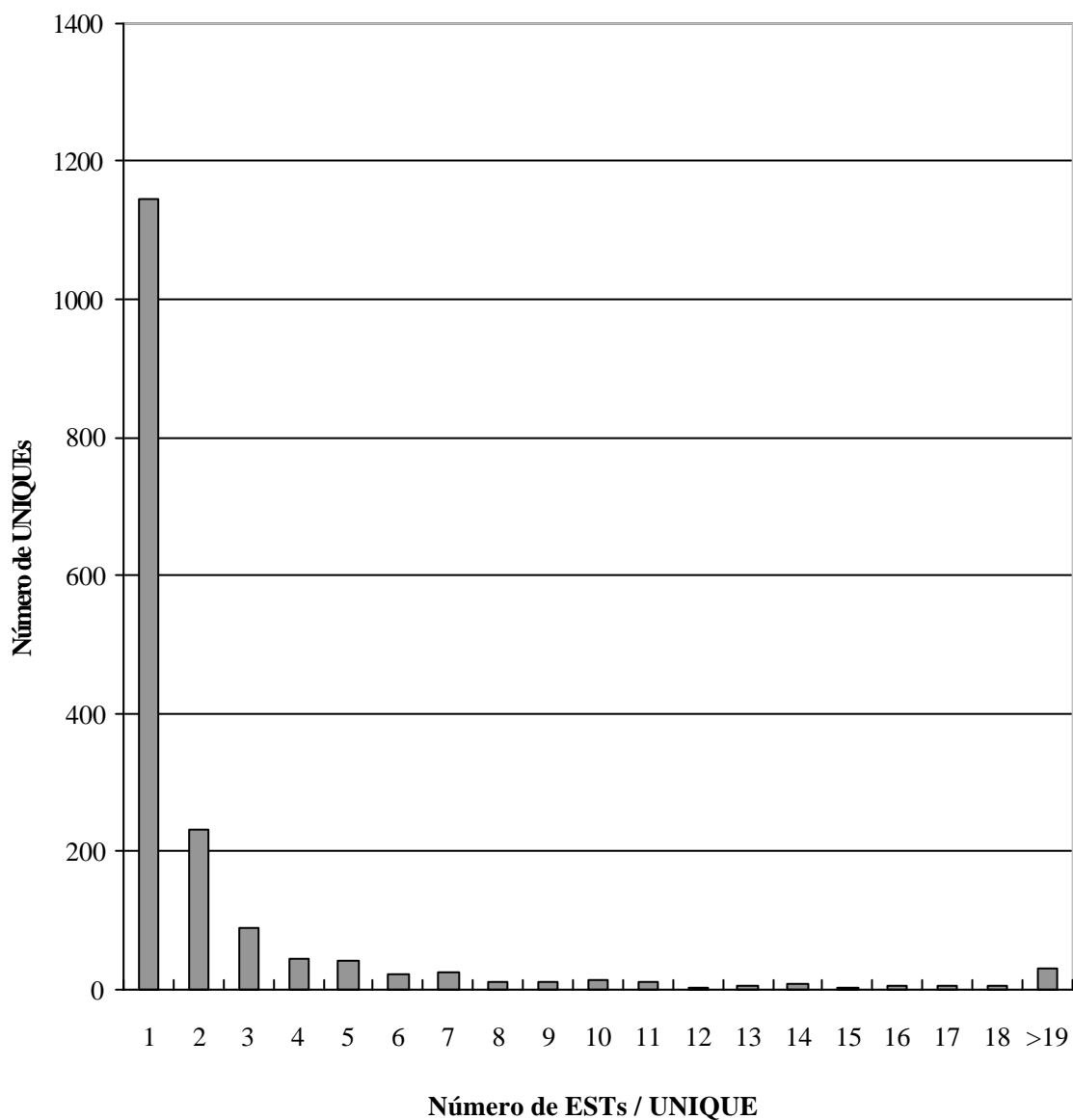


Figura 14 - Gráfico da redundância observada. Número de ESTs nos elementos únicos de seqüência (UNIQUE)

As 6.481 ESTs foram submetidas à análise de agrupamento (*Cluster*) o que resultou em 1.901 (29%) elementos únicos de seqüência (UNIQUE), compreendendo TCs e STs (Figura 14). A análise de alinhamento (BLAST) resultou em 601 seqüências (31%) com alinhamentos

significativos, nos critérios anteriormente mencionados, com o *Bos taurus* Gene Index da TIGR (BtGI) (ver. 7.0). O restante das seqüências (1300 UNIQUES) foi submetido à análise de alinhamento com as bases dBEST de humanos, camundongo e outras espécies, além de alinhamento com proteínas (BLASTX), o que resultou em 534 hits positivos (28% das 1901 UNIQUES). As sequencias restantes, que não apresentaram alinhamentos significantes, foram submetidas, sequencialmente, ao BLASTN com as bases do GoldenPath (humanos) e com a base “nt”do GenBank, o que resultou em 4% e 9% de alinhamentos significantes, respectivamente. No sentido de verificar a natureza das seqüências restantes, que não haviam produzido alinhamentos significativos até então, foi realizada uma análise de alinhamento com informações de genomas de outros microorganismos, o que resultou em 58 alinhamentos significantes (58/1901=3%). O restante das seqüências foi declarado como sendo de natureza desconhecida (509 UNIQUES, 25%), não apresentando alinhamento significativo com nenhuma das bases de dados, representando, potencialmente, nova informação para o transcriptoma bovino. A análise mostrou também alguns aspectos importantes das seqüências em estudo. Considerando os 1901 UNIQUES, 69 dessas apresentaram alinhamento significativo apenas com a base GoldenPath ou com a base “nt”do GenBank (4%). Essas seqüências podem representar contaminação com DNA genômico ocorrido nas bibliotecas. Por outro lado, 268 UNIQUES apresentaram alinhamentos significantes apenas com as bases de ESTs de humanos, camundongos e outras espécies ou com proteínas, sem alinhamentos com o BtGI/TIGR, representando seqüências novas do transcriptoma bovino.

Considerando o total de 1.901 UNIQUES, houve 503 *hits* em relação a ESTs de Humanos (26%), e, 348 (18%), e 706 (37%), produziram alinhamento em relação a ESTs de camundongos (*Mus musculus* e *Mus domesticus*) e ESTs de “outras espécies”, respectivamente. Com relação ao alinhamento com proteínas, a tradução das UNIQUES permitiu alinhamento de 938 dessas (49%) com as bases de dados de proteínas.

Esses percentuais relativamente elevados revelam a descoberta de seqüências novas em relação ao genoma conhecido do bovino, já que trata-se de seqüências conhecidas em genomas de outras espécies e que ainda não estavam disponíveis nas bases de dados públicos.

Após terem sido submetidas ao GenBank, as seqüências passaram a integrar também a base de dados do *The Institute of Genomic Research* (TIGR) e apesar das seqüências serem, na maioria, originadas de animais *Bos indicus*, passaram a integrar a base BtGI daquela instituição. Cada biblioteca recebeu um novo código de identificação da TIGR e as ferramentas de análise disponíveis para as outras espécies, bem como para as demais bibliotecas de bovinos, podem ser

utilizadas para as bibliotecas BARC-EMBRAPA. A Tabela 19 apresenta um sumário contendo a nova identificação das bibliotecas, estatísticas de agrupamento individualizadas realizadas com as bibliotecas e os resultados das análises realizadas pela TIGR. A tabela apresenta ainda cálculos da redundância observada nas bibliotecas. Importante ressaltar que algumas das bibliotecas apresentaram elevada redundância no seqüenciamento, caracterizada por um menor percentual de UNIQUES da biblioteca. Por exemplo, a biblioteca 327BOV (#C63) apresentou apenas cerca de 18% de UNIQUES. Por outro lado, a biblioteca 307BOV (#C5E) apresentou aproximadamente 90% de UNIQUES, representativo de uma reduzida redundância e elevada diversidade das seqüências.

As estatísticas realizadas pela TIGR indicam que foram ainda eliminadas por aquele instituto mais 40 seqüências que apresentavam baixa qualidade ou complexidade. Das 6.441 ESTs restantes depositadas naquele instituto, 5.598 foram alocadas em TCs, enquanto 843 resultaram em STs, não produzindo alinhamento com seqüências depositadas no BtGI. As seqüências participam de 1.818 diferentes TCs.

Cumprindo ainda mencionar que alguns dos TCs agrupados pela TIGR apresentaram apenas seqüências das bibliotecas BARC-EMBRAPA e que destes, alguns apresentam seqüências associadas ao sistema imunológico. Por exemplo, os TCs 169823, 169833, 169825, 169832, 169845, 169831, 169829, 169836, 169833, 154786. Essas são novas seqüências para os bancos de dados de genes bovinos, associadas ao sistema imunológico e que podem ser associadas à resistência à mastite bovina.

A metodologia de ORESTES empregada no presente estudo foi escolhida como forma de maximizar a taxa de descoberta de novas seqüências e para aumentar a freqüência de representação de ESTs na região central dos transcritos, em oposição aos métodos de seqüenciamento das extremidades 5' e 3'. Para verificar se a dispersão da localização das ESTs havia sido mantida na região central dos transcritos apesar da utilização de material diverso do que foi empregado no estudo independente de Camargo et al. (2001), a base de dados foi alimentada com seqüências de clones humanos contendo todo comprimento de genes (*full-length open reading frame* (FL-ORF)) (<http://mgc.nci.nih.gov/>). Calculou-se um percentual relativo para apresentar o resultado do alinhamento de cada UNIQUE considerando a sua posição relativa em relação à posição central do gene *full-length*. Foram calculados percentuais para 990 UNIQUES (52%) e uma curva indicando que a maioria das seqüências estava localizada na região central dos genes (Figura 15). Esses resultados são semelhantes aos resultados de Camargo et al. (2001).

Tabela 19 . Análise de agrupamento (*Cluster*) das bibliotecas BARC-EMBRAPA e resultados das análises realizados pela TIGR

		ANÁLISE DE AGRUPAMENTO			ESTATÍSTICAS BtGI-TIGR			
ID		<i>Contigs</i>	<i>Singletons</i>	UNIQUEs	ESTs			
GenBank	BtGI	(TC)	(ST)	(%)	em TCs	STs	Total	#TCs
300BOV	#C4F	9	6	21.4	64	6	70	13
301BOV	#C58	10	15	47.2	41	12	53	14
302BOV	#C59	28	40	42.2	134	27	161	44
303BOV	#C5A	15	9	27.9	83	3	86	21
304BOV	#C5B	19	13	36.8	79	8	87	25
305BOV	#C5C	15	40	74.3	52	22	74	32
306BOV	#C5D	3	8	78.6	9	5	14	6
307BOV	#C5E	7	45	89.7	37	21	58	31
308BOV	#C5F	13	33	56.1	62	20	82	26
309BOV	#C5G	30	28	31.5	161	23	184	46
311BOV	#C5I	12	17	40.8	58	13	71	15
312BOV	#C5H	7	37	78.6	40	16	56	26
313BOV	#C5J	6	29	74.5	39	8	47	25
314BOV	#C5K	22	62	23.5	334	23	357	62
315BOV	#C5L	50	105	49.7	263	49	312	107
316BOV	#C5M	39	170	60.2	256	91	347	114
317BOV	#C5N	30	94	36.4	289	52	341	77
319BOV	#C5T	17	34	32.3	147	11	158	36
321BOV	#C5U	15	70	59.0	124	20	144	62
322BOV	#C5V	5	29	50.0	54	14	68	20
323BOV	#C60	20	51	22.0	315	7	322	61
324BOV	#C61	21	61	36.6	208	16	224	49
326BOV	#C62	19	39	19.0	295	11	306	54
327BOV	#C63	21	32	17.7	292	8	300	50
328BOV	#C64	21	34	18.3	293	8	301	55
338BOV	#C65	37	151	61.4	240	66	306	117
339BOV	#C68	35	151	62.0	232	68	300	119
351BOV	#C6A	46	95	44.6	277	39	316	102
352BOV	#C69	34	113	64.5	164	64	228	85
353BOV	#C6B	38	108	56.2	207	53	260	92
354BOV	#C6C	16	39	21.1	250	11	261	48
355BOV	#C67	31	73	34.3	286	17	303	90
356BOV	#C66	39	82	49.6	213	31	244	94
Soma/Média		730	1913	46.0	5598	843	6441	1818

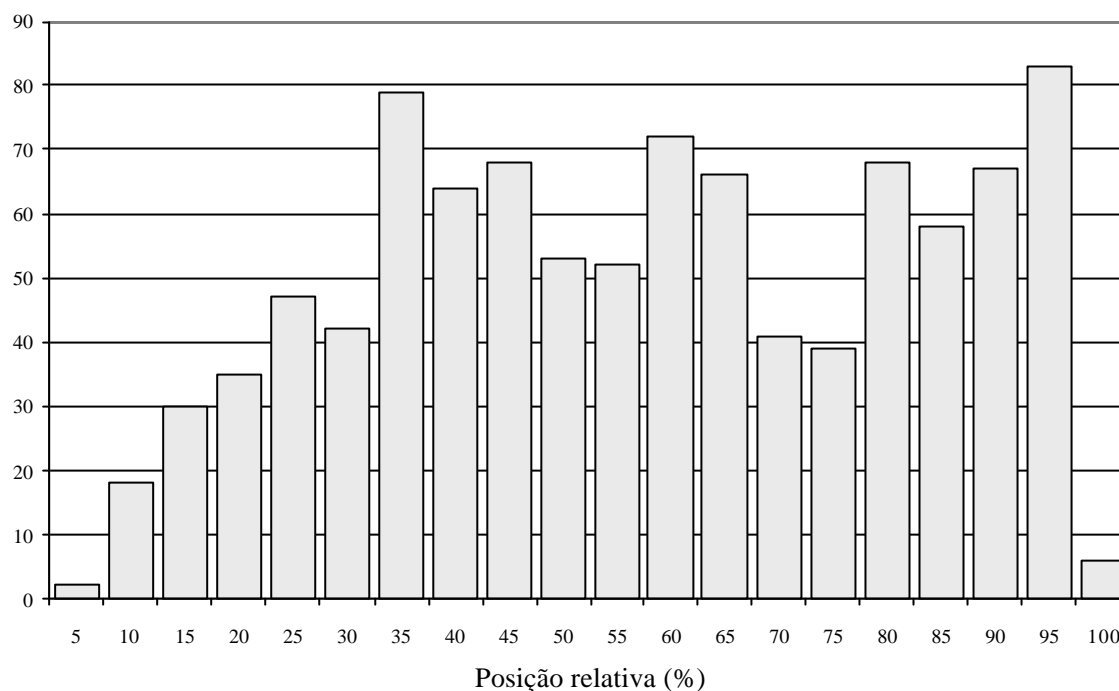


Figura 15 - Distribuição da posição relativa dos elementos de seqüência única (UNIQUE) em relação a seqüências completas (*full-length open reading frames*) de transcritos Humanos. Considerar que o percentual de posição relativa de 100% refere-se à extremidade 3' - do transcrito.

Os resultados apresentados permitem sugerir que a construção de bibliotecas de cDNA por meio de ORESTES pode ser empregada como uma alternativa de extensão do consenso de seqüências compiladas, especialmente em bovinos. ORESTES é uma possibilidade técnica e economicamente razoável de ocasionar o colapso de seqüências do mesmo gene que tenham sido agrupadas separadamente por terem sido obtidas de seqüenciamento das extremidades 5' - ou 3'-. Esses resultados também permitem sugerir que a relativa alta taxa de descoberta observada nas bibliotecas construídas e caracterizadas seja resultado da metodologia de ORESTES empregada na produção de ESTs.

Com relação à natureza das seqüências presentes nas bibliotecas, cabe salientar que as análises permitiram observar que os transcritos mais abundantes estavam associados à resposta do sistema imunológico (*immunoglobulin-lambda light chain variable region, immunoglobulin-gamma heavy chain, and alpha-2-macroglobulin*). Esse fato pode ter ocorrido em razão do desafio imposto aos animais com *Staphylococcus aureus*. As análises com a classificação dos termos *Gene Ontology (GO terms)* permitiu extratificar as UNIQUE e obter ao menos um *GO term* para 542 das UNIQUEs. As informações foram obtidas recuperando *GO terms* das seqüências da TIGR e para o caso em que esta informação não estava disponível, um alinhamento com proteínas depositadas no *SwissProt* recuperou os *GO terms* associados com a seqüência. Os resultados indicaram que para o processo biológico envolvido pelas seqüências com recuperação dos *GO terms*, o processo mais abundante dentre as UNIQUEs foi o de crescimento celular e manutenção (82%). Com relação ao componente celular envolvido, as UNIQUEs apresentaram os maiores percentuais de ocorrência na célula e no ambiente extracelular (90% e 4%, respectivamente). Quanto à função molecular, as maiores ocorrências se deram para a função de enzimas (45%), seguido de ligação (32%), transporte (10%) e seqüências referentes ao sistema imunológico (1%).

A metodologia de ORESTES para construção de bibliotecas de cDNA revelou-se de baixo custo e mais intensiva no uso de mão-de-obra. Um ponto positivo também é representado pela reduzida necessidade de amostras de tecido para extração de RNA total ($\pm 10 \mu\text{g}$). A metodologia também apresenta necessidade de estar associada à utilização de procedimentos especiais de análise computacional, já que é uma metodologia baseada no método de PCR podendo produzir artefatos nas seqüências.

Portanto, baseado nas novas informações gênicas adicionadas nas bases públicas de informações sobre genomas, pode-se concluir que as bibliotecas construídas com o emprego da metodologia de ORESTES apresentaram transcritos raros com contribuição relevante e papel fundamental no futuro próximo.

A identificação de regiões transcritas no genoma bovino é um desafio científico atual e o emprego da metodologia de ORESTES permite o preenchimento de lacunas desse conhecimento, o agrupamento de seqüências até então mantidas separadas, apesar de se referirem ao mesmo gene e a descoberta de genes novos, auxiliando no entendimento sobre a biologia nos bovinos.

7.4 Conclusões

- O estudo das ESTs por meio de bibliotecas de cDNA construídas através da metodologia de ORESTES permitiu descobrir genes novos não descritos anteriormente para bovinos.
- As ESTs obtidas por meio da metodologia de ORESTES se localizam distribuídas ao longo da seqüência dos genes, concentrando-se na região central.
- Foram identificados genes relacionados ao sistema imunológico que ainda não estavam descritos para bovinos. Esses são genes candidatos a conferir resistência à mastite bovina.

Agradecimentos

Os autores agradecem ao USDA-ARS e à Embrapa, que financiaram a realização do projeto internacional colaborativo entre as instituições envolvidas através do LABEX-EUA. Adilson Mota agradece à Diretoria da Embrapa, à chefia da Embrapa Gado de Leite, ao Departamento de Organização e Desenvolvimento e à Secretaria de Cooperação Internacional pelo suporte às ações que resultaram na realização deste projeto. Luiz Lehmann Coutinho e Mario Luiz Martinez agradecem ao CNPq pelo recebimento de bolsa de produtividade em pesquisa.

8 CONCLUSÕES GERAIS

- Uma das importantes alternativas de estudo da variação genética molecular para localizar regiões cromossômicas com genes de interesse é realizada com estudos de populações construídas em desenhos experimentais apropriados e com o emprego de marcadores moleculares.
- Pode-se também chegar aos genes que contribuem para gerar resistência nos bovinos contra a ocorrência de mastite por meio do estudo de genes candidatos. O gene BoLA-DRB3 permanece sendo candidato a conferir resistência à mastite bovina e estudos populacionais podem ser realizados com as técnicas empregadas para genotipagem.
- A principal metodologia de genotipagem do gene BoLA-DRB3 foi TGGE, que permite conhecer a seqüência do *exon* desse gene e – portanto – identificar alelos novos.
- O gene BoLA-DRB3 é altamente polimórfico e os animais da raça Gir leiteira apresentam grande variabilidade e potencial de descoberta de alelos novos desse gene.
- A metodologia de ORESTES é extremamente apropriada para realizar estudos com bibliotecas de cDNA, conduzindo à descoberta de genes novos em bovinos. As seqüências – de modo geral – representam a porção central dos genes.

ANEXOS

ANEXO A

LOCALIZAÇÃO DE LOCOS GÊNICOS DE CARACTERÍSTICAS QUANTITATIVAS (QTLs) EM GADO DE LEITE. PREMISSAS E ESTRATÉGIAS.

Adilson Ferreira da Mota

Empresa Brasileira de Pesquisa Agropecuária – Embrapa - Centro Nacional de Pesquisa de Gado de Leite - Embrapa Gado de Leite - 36038-330 - Juiz de Fora – MG
Email: damota@cnpqgl.embrapa.br

Mário Luiz Martinez

Empresa Brasileira de Pesquisa Agropecuária – Embrapa - Centro Nacional de Pesquisa de Gado de Leite - Embrapa Gado de Leite - 36038-330 - Juiz de Fora – MG
Email: martinez@cnpqgl.embrapa.br

Luiz Lehmann Coutinho

Departamento de Produção Animal - Laboratório de Biotecnologia Animal - ESALQ/USP
CEP 13418-900 - Piracicaba – SP - Brasil
Email: llcouthin@carpa.ciagri.usp.br

RESUMO

Os avanços da biologia molecular têm aberto novas perspectivas de impacto nos sistemas de produção. A identificação dos genes de interesse e a possibilidade de compreensão do seu mecanismo de ação permite empregá-los como “guias gênicos” utilizados na seleção de indivíduos portadores de características de interesse comercial. A premissa básica desta teoria é que a descendência direta de um ancestral diplóide e heterozigoto para um determinado QTL de interesse pode ser dividida em duas classes com níveis de produção diferentes, cada uma delas recebendo um dos alelos parentais. O objetivo deste texto é apresentar uma revisão, na forma de introdução básica compreensível, e uma discussão a respeito de algumas alternativas dentre as diversas metodologias empregadas na localização de QTLs em Gado Leiteiro. Dezenas de milhares de regiões cromossômicas são candidatas a integrar QTLs de interesse econômico. A localização destas regiões cromossômicas e a possibilidade de emprego no melhoramento animal através da seleção assistida por marcadores moleculares (MAS – “Marker Assisted Selection”) representam uma contribuição efetiva que pode ser oferecida aos programas de melhoramento, aos produtores e consumidores em geral.

Palavras-chave: genes principais, mapeamento genético, marcadores moleculares, melhoramento animal, revisão.

MAPPING QTLs (QUANTITATIVE TRAIT LOCI) IN DAIRY CATTLE. PREMISES AND STRATEGY.

ABSTRACT

Recent advances in molecular biology are opening new insights to impact the production systems. The identification of major genes and the understanding of their biological action will allow their use as "gene markers" to select bulls and cows with commercial traits of interest. The basic premiss of this theory is that a diploid heterozygote ancestral has two classes of descendants. Half of the progeny receiving one of the two different production level alleles. The objective of this manuscript is to offer a basic and comprehensive review on the subject and a discussion through the methodologies used in Dairy Cattle for QTL mapping. Tens of thousands of different chromosome regions are putative to be QTLs. Mapping these regions and using them as molecular markers in animal breeding by "marker assisted selection" (MAS) may effectively contribute to breeding programs, farmers and consumers.

Key words: *Animal breeding, major genes, molecular markers, QTL mapping, review.*

INTRODUÇÃO

Os recentes avanços da biologia molecular e o refinamento das técnicas de avaliação do DNA dos animais domésticos têm aberto novas perspectivas de impacto nos sistemas de produção. Sobretudo, essa tecnologia permite compreender a relação biológica existente entre os genes e as características de interesse comercial. Mais especificamente, ocorre que as enzimas e proteínas codificadas no genoma estão relacionadas com os processos bioquímicos e fisiológicos envolvidos na produção, na reprodução e na sanidade dos animais. A identificação dos genes de interesse e a possibilidade de compreensão do seu mecanismo de ação permite empregá-los como "guias gênicos", que podem, dentre outras utilidades, ser empregados na seleção de indivíduos portadores de características de interesse comercial.

Os modelos aplicados em Genética Quantitativa baseiam-se em evidências de que as características métricas de interesse econômico são, em geral, expressas por muitos locos gênicos, denominados Locos de Características Quantitativas (QTL – "Quantitative Trait Loci"). As diferenças nos fenótipos dos animais são resultantes, portanto, da ação individual de vários genes, bem como do efeito das suas interações, na forma de entropia e pleiotropia, e também da ação dos fatores ambientais.

As características quantitativas dos organismos animais apresentam variação contínua nas progênes de ancestrais específicos. O processo de identificação dos genes responsáveis pela variação inclui análises dessas características e o emprego de ferramentas biométricas para localização da região cromossômica (QTL) para a qual os alelos de um ou mais genes de função ainda desconhecida afetam a característica de interesse. A premissa básica desta teoria é que a descendência direta de um

ancestral diplóide e heterozigoto para um determinado QTL de interesse pode ser dividida em duas classes com níveis de produção diferentes, cada uma delas recebendo um dos alelos parentais. Caso os alelos representem efeitos genéticos significativos sobre a característica de interesse, é bastante provável que as duas classes de descendentes também apresentem diferença fenotípica na característica quantitativa.

Apesar da ampla discussão sobre o assunto, a localização de QTLs pressupõe conhecimentos em áreas distintas. O objetivo deste texto é apresentar uma revisão, na forma de introdução básica compreensível, e uma discussão a respeito de algumas alternativas dentre as diversas metodologias empregadas na localização de QTLs em Gado Leiteiro. Espera-se contribuir trazendo o assunto a um número maior de pessoas, já que é indiscutível a importância atual de realizar a localização de QTLs nas espécies domésticas. Dezenas de milhares de regiões cromossômicas são candidatas a integrar QTLs de interesse econômico. A localização destas regiões cromossômicas e a possibilidade de emprego no melhoramento animal através da seleção assistida por marcadores moleculares (MAS – “Marker Assisted Selection”) representam uma contribuição efetiva que pode ser oferecida aos programas de melhoramento, aos produtores e consumidores em geral.

PREMISSAS DAS LEIS DA GENÉTICA

As leis da Genética já estavam consagradas quando a Biotecnologia tornou-se realidade. No entanto, mesmo considerando que estas leis foram formuladas sem a comprovação da existência dos genes, mas somente baseadas em inferências, as mesmas permanecem válidas. A genética mendeliana foi concebida somente através da expressão dos genes na progênie para características visuais, e suas frequências estabelecidas com base nos fenótipos dos ancestrais. Gregor Mendel (1823-1884) formulou duas das três leis que servem de base aos estudos da genética de organismos sexuados.

A 1ª LEI DA GENÉTICA

Em termos atuais, a primeira lei afirma que cada indivíduo carrega duas cópias de cada gene e que somente uma dessas duas cópias é transmitida para cada filho. O filho, portanto, irá receber uma cópia completa de cada gene, recebida de cada progenitor. Assim, recebendo uma cópia do pai e uma da mãe, o filho terá restituído seu genótipo contendo duas cópias de cada gene.

Indivíduos que contêm duas cópias de cada gene são chamados de diplóides. A primeira lei foi formulada por Mendel e é confirmada quando os organismos diplóides produzem células germinativas com apenas uma cópia de cada gene - espermatozóides e óvulos. Durante o processo de divisão celular em que essa transformação ocorre - meiose - as duas cópias de cada gene irão separar-se, ou "segregar", uma da outra. Esse evento produziu o nome da primeira lei de Mendel: "Lei da Segregação".

A segregação pode ser observada na progênie de indivíduos heterozigotos, com dois alelos distintos. Metade da descendência direta receberá um dos alelos, enquanto a outra metade receberá o outro alelo, com a mesma probabilidade.

A primeira lei de Mendel descreve os acontecimentos referentes á herança gênica quando apenas um gene é considerado. A sua Segunda lei foi formulada para descrever o que ocorre quando dois genes diferentes são transmitidos por herança.

A 2ª LEI DA GENÉTICA

A Segunda lei da Genética, também formulada por Mendel, em termos atuais, afirma que a segregação dos alelos em cada loco não exerce influência sobre a segregação em outro loco. Ou seja, as segregações que ocorrem em dois locos são independentes. Assim, a Segunda Lei de Mendel ou a "Lei da Segregação Independente" estabelece que locos independentes produzem segregação também independente. Indivíduos heterozigotos para dois locos, supondo loco A e loco B, produzem gametas com freqüência igual à multiplicação da probabilidade de ocorrência dos gametas para cada loco. No caso em que o indivíduo apresenta genótipo A/a (loco A) e B/b (loco B), a chance de que um gameta seja A (loco A) é de 50% e a chance de que seja B (loco B) também é de 50%. Portanto a chance de que um gameta seja AB (loco A e loco B) será de $0,50 \times 0,50 = 0,25$, ou 25%. A Figura 2.1 apresenta uma ilustração da segregação independente.

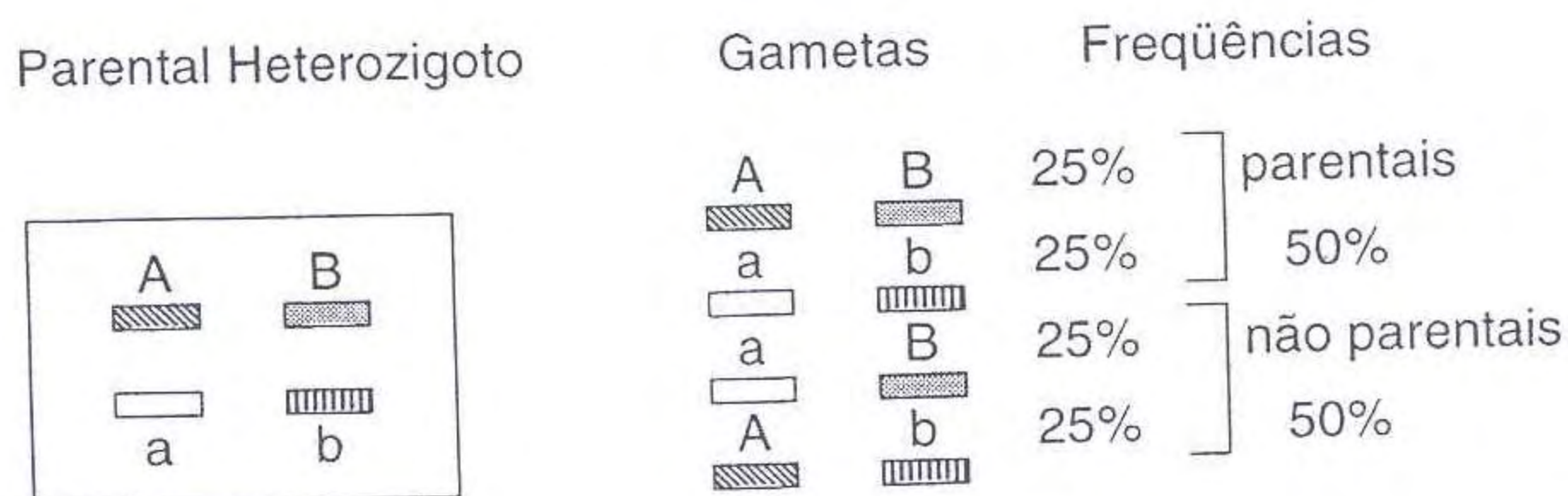


Figura 2.1 Diagrama da segregação independente entre locos sem ligação

Sabe-se nos dias atuais que a Segunda Lei de Mendel somente é válida no caso de genes que não estão ligados no mesmo cromossomo. Quando dois genes com dois alelos têm ligação cromossômica, as freqüências para cada uma das possíveis combinações dos alelos será desviada de 25%. Duas combinações de alelos irão representar os arranjos de ligação dos cromossomos parentais e estas serão, cada uma, transmitidas a uma freqüência superior a 25%. As duas classes restantes representam os arranjos recombinantes e irão ocorrer a uma freqüência inferior a 25%. No caso extremo de ligação absoluta, somente as duas classes parentais serão transmitidas, cada uma à freqüência de 50%.

A 3ª LEI DA GENÉTICA

Para os casos de genes em ligação, as fases de associação ou repulsão seriam consequência da localização dos genes no mesmo cromossomo. Alelos em fase de associação estariam na metade homóloga do cromossomo parental, ao contrário daqueles em fase de repulsão.

A figura 2.2 apresenta um diagrama das freqüências dos gametas produzidos por um indivíduo heterozigoto para os locos A e B, com uma freqüência de recombinação de 10%.

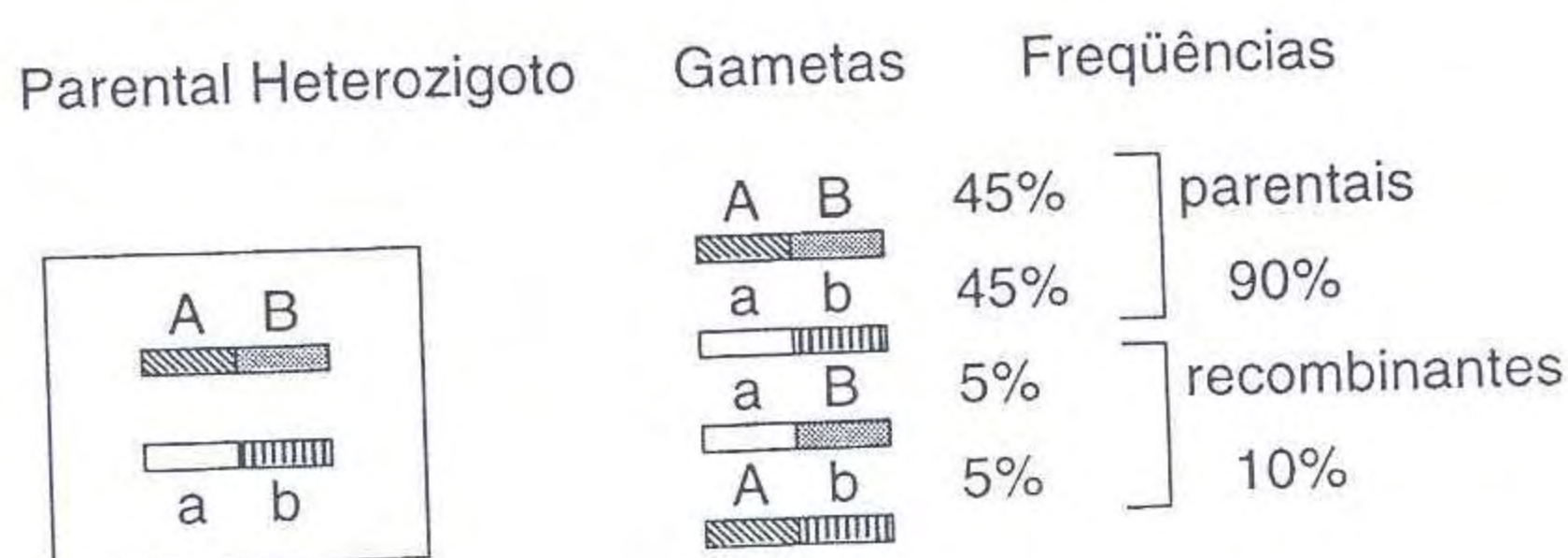


Figura 2.2 Diagrama da segregação entre locos com ligação e freqüência de recombinação de 10%.

Na ilustração apresentada na Figura 2.2 um indivíduo heterozigoto cujo genótipo pode ser descrito como sendo AB/ab produz 90% de seus gametas da mesma forma apresentada no seu genoma e 10% dos gametas produzidos apresentam uma forma não-parental, ou recombinante. Isso é consequência da ocorrência de crossing-over durante a divisão celular, meiose, que deu origem aos gametas. Quanto menor a distância entre os genes ligados no mesmo cromossomo (Figura 2.2), menores são as chances de que os eventos de "crossing-over" possam ocorrer entre os dois genes. Portanto, a freqüência de recombinação permite estimar a distância genética entre os genes.

As distâncias genéticas são medidas em centimorgans (cM), sendo que 1 cM é definido como a distância entre dois locos cuja recombinação ocorre em 1% dos casos. Importante salientar que mesmo sendo a recombinação proporcional à quantidade de DNA que separa dois locos no mesmo cromossomo, quando esta distância for grande, a freqüência de recombinação irá aproximar-se de 50%, tornando os resultados similares aos observados entre locos sem ligação.

Há ainda que considerar-se a situação em que mais de um evento de "crossing-over" ocorre entre dois genes durante a meiose. No entanto, à medida que a distância genética é reduzida, a freqüência de ocorrência de mais de um evento de "crossing-over" é reduzida exponencialmente. Para dois genes distantes 10 cM entre si, as chances de ocorrência de dois eventos de "crossing-over" entre eles é de apenas 1%. Assim, neste texto, com a finalidade de simplificação, não serão consideradas as ocorrências de mais de um evento de "crossing-over" entre o marcador genético e o QTL.

As três Leis da genética apresentadas resumem a compreensão necessária de como emprega-se o conhecimento do desequilíbrio de ligação e da frequência de recombinação dos alelos para utilizar marcadores genéticos na localização de QTLs de interesse. O raciocínio é equivalente para o emprego de marcadores moleculares polimórficos e exemplos serão apresentados a seguir para os desenhos experimentais especialmente construídos para localizar novos QTLs em Gado Leiteiro.

Outra questão importante dentro do contexto de localização de QTL, e que deverá ser alvo de escolha durante o processo, diz respeito aos marcadores empregados.

MARCADORES MOLECULARES

Como a localização e identificação de um QTL não são conhecidas *a priori*, não há como discriminar os genótipos do QTL dos indivíduos parentais ou de sua progênie. Assim, são empregadas marcas com localização genômica conhecida, marcadores moleculares, estrategicamente dispostos ao longo do genoma. A hipótese empregada para a localização de um QTL é de que estas marcas, ou marcadores moleculares polimórficos, estarão próximas e portanto suficientemente ligadas ao QTL, de forma que, na maioria da progênie de um indivíduo heterozigoto para o marcador, será possível associar as variações na característica quantitativa de interesse com o genótipo do marcador, que será indicativo do genótipo do QTL. A associação entre os genótipos do QTL e do marcador não ocorrerá quando tiver ocorrido "crossing-over" na meiose da formação de gametas da geração anterior, indicativo do desequilíbrio da ligação. A frequência de ocorrência de recombinação causada por "crossing-over" dos cromossomos dos indivíduos parentais permitirá estabelecer a estimativa da distância genética entre o marcador e o QTL. Como a localização do marcador é conhecida, desta forma será possível conhecer a estimativa da localização do QTL.

MARCADORES MICROSATÉLITES

Sem a existência da reação de PCR, a maioria das regiões microsátélites não teria utilidade como marcadores genéticos (SILVER, 1995). As regiões microsátélites são marcadores moleculares especiais que consistem de pequenas sequências com 1 a 4 nucleotídeos, repetidas em tandem ao longo do genoma. São mais frequentes nos genomas de organismos eucariotos e são bem distribuídas ao acaso. Os marcadores microsátélites constituem a classe de marcadores que apresenta o maior grau de polimorfismo entre os marcadores disponíveis (FERREIRA & GRATTAPAGLIA, 1996). Regiões do DNA podem ser amplificadas por PCR com "primers" específicos, complementares a regiões flanqueadoras às sequências repetidas da região microsátélite. Esses fragmentos amplificados apresentam polimorfismo extenso, resultante de diferentes números de sequências de nucleotídeos repetidas. Cada fragmento amplificado de tamanho diferente representa um alelo diferente do mesmo loco. A Figura 3.2 apresenta um diagrama

representando os possíveis alelos de um microsatélite e como os genótipos podem ser detectados. A diferença existente entre o comprimento das seqüências na região microsatélite produz fragmentos amplificados com diferentes mobilidades, o que permite identificar padrões ou bandas através da eletroforese em gel de poliacrilamida.

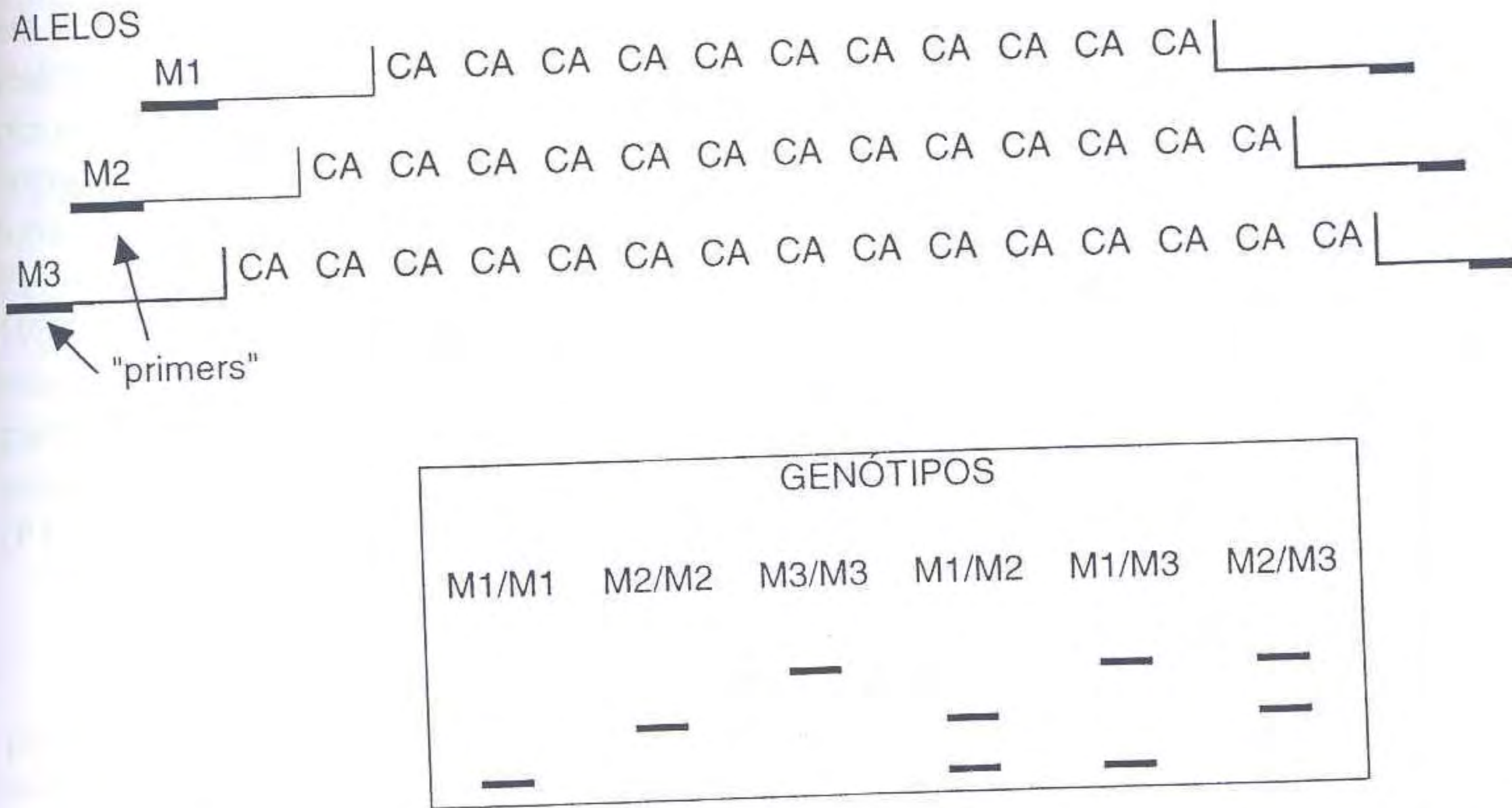


Figura 3.2 Alelos do microsatélite e sua detecção por PCR e eletroforese. Três diferentes alelos compostos da repetição dos nucleotídeos "CA". Os "primers" específicos empregados na amplificação estão representados junto à identificação dos alelos. Os padrões do gel representam possíveis genótipos com combinações dos alelos indicados.

Considerando a estimativa de que nos genomas das espécies de animais domésticos podem estar disponíveis de 50.000 a 150.000 regiões microsatélites (HILLEL *et al.*, 1992), a facilidade em obter a genotipagem de animais através da reação de PCR e também o elevado conteúdo de informação de polimorfismo, qualquer população onde ocorre segregação pode ser alvo de estudo de ligação ou de mapeamento genético empregando microsatélites.

Segundo FERREIRA & GRATTAPAGLIA (1996) a limitação básica que existe para aplicação mais ampla destes estudos empregando microsatélites recai na necessidade de trabalho técnico especializado e equipamentos laboratoriais sofisticados. Segundo os mesmos autores, no entanto, esta técnica será cada vez mais utilizada em um número cada vez maior de espécies.

Baseado no exposto, considerando que é possível empregar populações construídas nos estudos para detectar e mapear QTL de interesse por meio da análise de ligação e através do emprego de marcadores moleculares microsatélites, resta estabelecer o desenho experimental a ser utilizado na construção da população e a metodologia necessária para estimar a localização precisa do QTL.

POPULAÇÕES CONSTRUÍDAS

Há diversos métodos de detecção de QTL, como: Distribuição Multimodal em seleção divergente; Distribuição Bimodal em retrocruzamento com seleção; Distribuição não-normal; Heterogeneidade da variância dentro de famílias; e Semelhança entre a progênie e os pais. No entanto, a abordagem com maior poder em detectar um gene de efeito principal que afeta uma característica quantitativa é realizada através da Análise Complexa da Segregação. Uma das soluções para permitir a análise da segregação é a construção de populações formadas a partir do cruzamento entre linhagens consanguíneas ou entre raças puras (P1 x P2), gerando uma produção de indivíduos teoricamente heterozigotos (F1). A partir da geração F1, há duas opções de novos cruzamentos e progênies (SOLLER & BECKMANN, 1983; KNAPP *et al.*, 1990). A primeira delas é o desenho experimental de retrocruzamento (BC), quando a geração F1 será cruzada novamente com os parentais (P1 e P2, sendo BC1 e BC2, respectivamente). A segunda delas é a produção de uma geração F2, obtida pelo cruzamento entre indivíduos da geração F1 (F1 X F1).

APENAS UM LOCO MARCADOR

Além da escolha do desenho experimental, diferentes modelos genéticos podem ser utilizados (ARÚS & MORENO-GONZÁLES, 1993). Supondo a situação em que se emprega apenas um marcador molecular microsatélite (M) e que este possui apenas dois alelos ("M" e "m") e considerando o QTL de interesse também possuindo apenas dois alelos ("Q" e "q"), supondo também uma frequência de recombinação (r), os indivíduos de F1 obtidos do cruzamento entre P1 e P2 teriam o diagrama de seus cromossomos conforme apresentado na Figura 4.1.

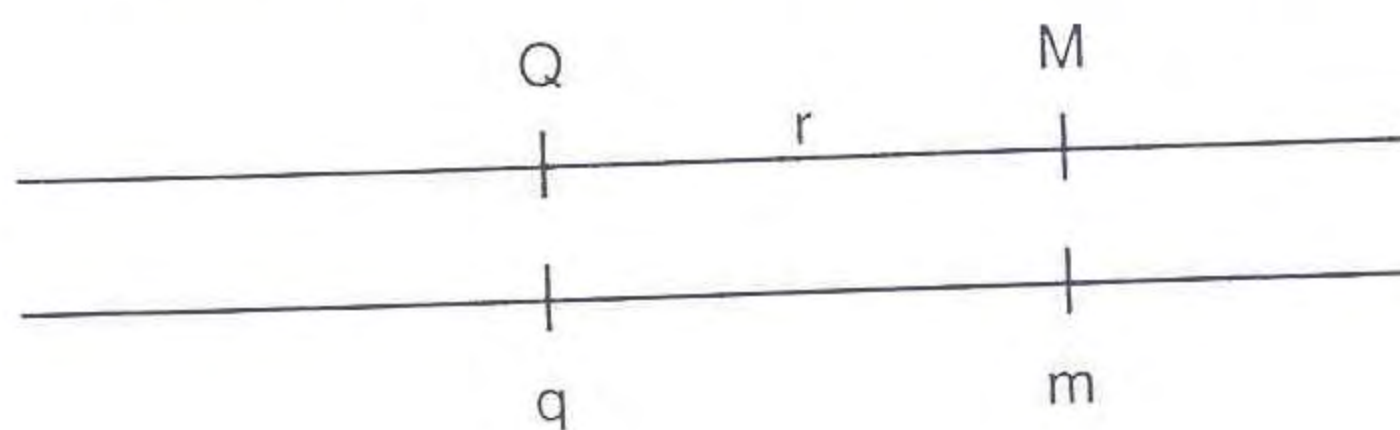


Figura 4.1 Diagrama do segmento cromossômico de indivíduos da geração F1 com genótipo Qq/Mm para o QTL (Q) e o marcador (M), respectivamente, quando r é a frequência de recombinação

Os tipos de gametas em F1 e as frequências de ocorrência são apresentados a seguir na Tabela 4.1.

Tabela 4.1: Tipos de gametas e suas frequências derivados da meiose em uma geração F1.

Gametas	freqüências
QM	$(1 - r) / 2$
qM	$r / 2$
Qm	$r / 2$
qm	$(1 - r) / 2$

DESENHO EXPERIMENTAL F2 PARA MODELO COM UM MARCADOR

Supondo um desenho experimental de F2, haveria o cruzamento entre si dos gametas descritos na Tabela 4.1. Seguindo SOLLER & BECKMANN (1983) e assumindo que +a, d e -a sejam os efeitos genotípicos para os genótipos QQ, Qq e qq, respectivamente, os genótipos da geração F2 produzida, bem como suas frequências e valor genotípico são apresentados na tabela apresentada a seguir (Tabela 4.2).

Tabela 4.2: Genótipos, frequências e valor genotípico esperados na geração F2 obtida do cruzamento entre F1 x F1.

Genótipo	freqüências	valor genotípico
QQMM	$(1 - r)^2 / 4$	+ a
QQMm	$r(1 - r) / 2$	+ a
QQmm	$r^2 / 4$	+ a
QqMM	$r(1 - r) / 2$	d
QqMm	$(1 - 2r + 2r^2) / 2$	d
Qqmm	$r(1 - r) / 2$	d
qqMM	$r^2 / 4$	- a
qqMm	$r(1 - r) / 2$	- a
qqmm	$(1 - r)^2 / 4$	- a

O valor da estimativa da taxa de recombinação, indicativa da distância genética existente entre o marcador e o QTL, é obtido por cálculo algébrico, agrupando-se os indivíduos em três classes de marcadores: MM, Mm e mm. No entanto, esta estimativa só tem sentido quando houver diferença significativa entre as médias observadas para a característica de interesse para as classes de marcadores mencionadas. As médias esperadas para as três classes de marcadores, obtidas através da combinação entre as frequências e os valores genotípicos apresentados na Tabela 4.2, são apresentadas na Tabela 4.3, apresentada a seguir.

Tabela 4.3: Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular.

Classe do marcador	média esperada da característica
MM	$\mu + a(1 - 2r) + 2rd(1 - r)$
Mm	$\mu + d(1 - 2r + 2r^2)$
mm	$\mu + a(2r - 1) + 2rd(1 - r)$

Contrastes ortogonais realizados entre as médias observadas nos indivíduos classificados para o marcador permitem estimativas dos efeitos do QTL, referentes aos efeitos aditivos e de dominância. Considerando [MM] como sendo a média da característica de interesse observada nos indivíduos classificados como apresentando o genótipo MM para o marcador. Considerando, de forma análoga, a média [Mm] para os indivíduos Mm e a média [mm] para os da classe mm, os contrastes estão apresentados a seguir (Tabela 4.4).

Tabela 4.4: Contrastes ortogonais empregados no cálculo dos efeitos genéticos aditivo (a) e devido à dominância (d) do QTL.

Contrastes Ortogonais

$$([MM] - [mm]) / 2 = a(1 - 2r)$$

$$[Mm] - (([MM] - [mm]) / 2) = d(1 - 2r)^2$$

$$(2[Mm] - [MM] - [mm]) / ([MM] - [mm]) = (d / a)(1 - 2r)$$

Para a situação em que emprega-se um desenho experimental de retrocruzamento (BC), através do retrocruzamento com P1 e P2, respectivamente BC1 e BC2, com um marcador molecular, há alterações nas frequências dos

genótipos, nos valores genotípicos esperados e nas médias esperadas para o marcador, bem como para os contrastes.

DESENHO EXPERIMENTAL DE RETROCRUZAMENTO (BC) PARA MODELO COM UM MARCADOR

Considerando um retrocruzamento realizado com o cruzamento entre uma geração F1 (QqMm) e as linhas ou raças parentais P1 (QQMM) e P2 (qqmm), as freqüências genotípicas que seriam observadas estão dispostas na Tabela 4.5.

Tabela 4.5: Genótipos e freqüências genotípicas no desenho BC, sendo BC1= (QqMm x QQMM) e BC2 = (QqMm x qqmm).

Genótipo	Freqüências Genotípicas	
	BC1	BC2
QQMM	$(1 - r) / 2$	
QQMm	$r / 2$	
QqMM	$r / 2$	
QqMm	$(1 - r) / 2$	
		$(1 - r) / 2$
QqMm		$r / 2$
Qqmm		$r / 2$
qqMm		$(1 - r) / 2$
qqmm		

Tabela 4.6: Valores esperados das médias da característica de interesse de acordo com a classe do marcador molecular para desenho BC com um marcador.

Classe do marcador	média esperada da característica
MM (BC1)	$\mu_{BC1} + a(1 - r) + rd$
Mm(BC1)	$\mu_{BC1} + d(1 - r) + ra$
Mm(BC2)	$\mu_{BC2} + d(1 - r) - ra$
mm(BC2)	$\mu_{BC2} + a(r - 1) + rd$

A separação dos indivíduos em classes de genótipo para os marcadores, de acordo com o retrocruzamento, BC1 ou BC2, estão apresentadas na Tabela 4.6, onde μ_{BC1} e μ_{BC2} representam as médias dos valores genotípicos para os demais genes que

contribuem para a característica de interesse, respectivamente para BC1 e BC2. Os contrastes ortogonais estão apresentados na Tabela 4.7.

Tabela 4.7: Contrastes ortogonais empregados no cálculo dos efeitos genéticos aditivo (a) e devido à dominância (d) do QTL, para desenho BC com um marcador.

Contrastes Ortogonais

$$([MM(BC1)] - [Mm(BC1)] + [Mm(BC2)] - [mm(BC2)]) / 2 = a (1 - 2r)$$

$$([Mm(BC1)] + [Mm(BC2)] - [MM(BC1)] - [mm(BC2)]) / 2 = d (1 - 2r)$$

Segundo ARÚS & MORENO-GONZÁLES (1993), o método empregando apenas um marcador molecular apresenta algumas desvantagens, dentre elas destaca-se que pode ocorrer que os valores de "a" e "d" sejam subestimados e quando o valor de "r" for maior, esse viés também será maior. Outra desvantagem é que com apenas um marcador, há indicações da distância entre o marcador e o QTL, no entanto sua localização não permanece estabelecida, podendo este estar a jusante ou a montante do marcador. Baseado nas desvantagens e limitações do método utilizando apenas um marcador, LANDER & BOTSTEIN (1989) e KNAPP *et al.* (1990) sugeriram o emprego de marcadores flanqueando a região candidata ao QTL.

LOCALIZAÇÃO DO QTL EMPREGANDO MAIS DE UM MARCADOR FLANQUEANDO A REGIÃO

Com a utilização de marcadores microsátélites estrategicamente dispostos ao longo do genoma, diversos estudos têm sido desenvolvidos com maior densidade na localização de QTLs (GEORGES, 1995).

O método consiste em cruzar duas linhagens ou raças (P1 x P2), de genótipos MMNN (P1) e mmnn (P2) para os marcadores M e N localizados no mesmo cromossomo e com relativa proximidade. Existindo QTL (Q) localizado entre os marcadores, o genótipo da geração F1 está apresentado na Tabela 4.2, disposta a seguir, onde r_1 é a frequência de recombinação entre o marcador M e o QTL e r_2 a frequência de recombinação entre o marcador N e o QTL.

DESENHO EXPERIMENTAL BC PARA MODELO COM MAIS DE UM MARCADOR FLANQUEANDO A REGIÃO

Considerando a alternativa de empregar retrocruzamento com os parentais P1 e P2, as frequências genotípicas e os genótipos para BC1 (F1 x P1) e BC2 (F1 x P2) são apresentados na Tabela 4.8. Conforme mencionado anteriormente, as tabelas consideram desprezíveis as chances de ocorrência de mais de um evento de crossing-over entre os marcadores ($r_1 \times r_2 = 0$).

Tabela 4.8: Genótipos e frequências genótípicas no desenho BC, sendo BC1= (MmQqNn x MMQQNN) e BC2 = (MmQqNn x mmqqnn). r_1 e r_2 são as frequências de recombinação entre o loco M e o QTL (Q) e a frequência de recombinação entre os locos N e Q, respectivamente.

Genótipo	Frequências Genótípicas	
	BC1	BC2
MMQQNN	$(1 - r_1 - r_2) / 2$	
MMQQNn	$r_2 / 2$	
MMQqNn	$r_1 / 2$	
MmQQNN	$r_1 / 2$	
MmQqNN	$r_2 / 2$	
MmQqNn	$(1 - r_1 - r_2) / 2$	
MmQqNn		$(1 - r_1 - r_2) / 2$
MmQqnn		$r_2 / 2$
Mmqqnn		$r_1 / 2$
mmQqNn		$r_1 / 2$
mmqqNn		$r_2 / 2$
mmqqnn		$(1 - r_1 - r_2) / 2$

A Tabela 4.9 apresentada a seguir permite visualizar as classes dos marcadores detectáveis nos indivíduos e as médias esperadas para a característica de interesse, sendo "a" e "d" efeitos genéticos aditivo e devido à dominância do QTL, respectivamente.

Considerando a distância entre os marcadores M e N como sendo r , de forma que na ausência de duplas recombinações $r = r_1 + r_2$, segundo Arús & Moreno-González (1993), teremos que a proporção r_1/r será dada por: $r_1/r = (\theta_1 - \theta_2 + \theta_3 - \theta_4) / 4 (\theta_1 - \theta_4) + (\theta_7 - \theta_8 + \theta_5 - \theta_6) / 4 (\theta_5 - \theta_8)$; e os valores de "a" e "d", calculados de acordo com as expressões: $a = (\theta_1 - \theta_4 + \theta_5 - \theta_8) / 2$; $d = (\theta_5 - \theta_8 - \theta_1 + \theta_4) / 2$. Com

base nas expressões apresentadas pode-se realizar a estimativa de localização do QTL, bem como seu valor genético aditivo e de dominância.

Tabela 4.9: Valores esperados das médias da característica de interesse de acordo com as classes dos marcadores moleculares para desenho BC com marcadores flanqueando o QTL.

Retrocruzamento	Classe do marcador	média esperada da característica
BC1	MMNN	$\theta_1 = \mu_{BC1} + a$
	MMNn	$\theta_2 = \mu_{BC1} + d r_1 + a r_2$
	MmNN	$\theta_3 = \mu_{BC1} + d r_2 + a r_1$
	MmNn	$\theta_4 = \mu_{BC1} + d$
BC2	MmNn	$\theta_5 = \mu_{BC2} + d$
	Mmnn	$\theta_6 = \mu_{BC2} + d r_2 - a r_1$
	mmNn	$\theta_7 = \mu_{BC2} + d r_1 - a r_2$
	mmnn	$\theta_8 = \mu_{BC2} - a$

O modelo de F2 encontra-se descrito em CARBONELL *et al.* (1992), com marcadores flanqueando a região do QTL, desenvolvido atendendo o mesmo raciocínio empregado para o desenho BC. Vários outros autores publicaram trabalhos considerando os desenhos mencionados e indicando testes apropriados (LANDER & BOTSTEIN, 1989; SOLLER & GENIZI, 1978; LEBOWITZ *et al.*, 1987; HALEY & KNOTT, 1992; WELLER, 1986; VISSCHER & HALEY, 1996).

Há também de ser considerada a possibilidade de ser realizada a genotipagem seletiva dos animais produzidos a partir do cruzamento de linhagens ou raças. A genotipagem seletiva foi proposta por LEBOWITZ *et al.* (1987) e constitui uma alternativa interessante na redução de custos mediante a genotipagem dos indivíduos distribuídos nos extremos da curva de distribuição dos valores da característica quantitativa avaliada. DARVASI & SOLLER (1992) mostraram que a genotipagem seletiva pode reduzir o número de genotipagens para um mesmo poder do teste. BOVENHUIS & SPELMAN (1996) encontraram resultados semelhantes e segundo HILLEL *et al.* (1992) empregando-se 5% da população total com a genotipagem de indivíduos distantes ao menos dois desvios-padrão da média, essas informações contribuem com 28% do total da informação de ligação disponível.

DESENHO DE FILHAS (DAUGHTER DESIGN)

Especialmente em gado de leite, os desenhos experimentais são fundamentais. Os touros não expressam a característica de interesse dos animais, que é a lactação. Há também uso freqüente de inseminação artificial e a conseqüente multiplicação em muitos rebanhos de parte do genoma dos touros em níveis elevados.

A teoria do Desenho de Filhas consiste na pressuposição de existência de um touro heterozigoto para um determinado marcador e que possui um grande número de descendentes. É possível, portanto, testar a presença de ligação entre marcador e QTL na sua progênie através da diferença entre as médias dos valores da característica de interesse nas filhas, sendo que estas podem ter recebido um ou outro dos alelos do marcador do touro. Na maioria dos casos é interessante incluir no estudo um maior número de touros, para os casos em que o touro seja heterozigoto para o marcador e homozigoto para o QTL. Uma análise da variância entre os subgrupos de filhas que receberam distintos alelos permite avaliar a significância da diferença entre as médias dos fenótipos observados. O diagrama da Figura 5.1 ilustra o Desenho de Filhas.

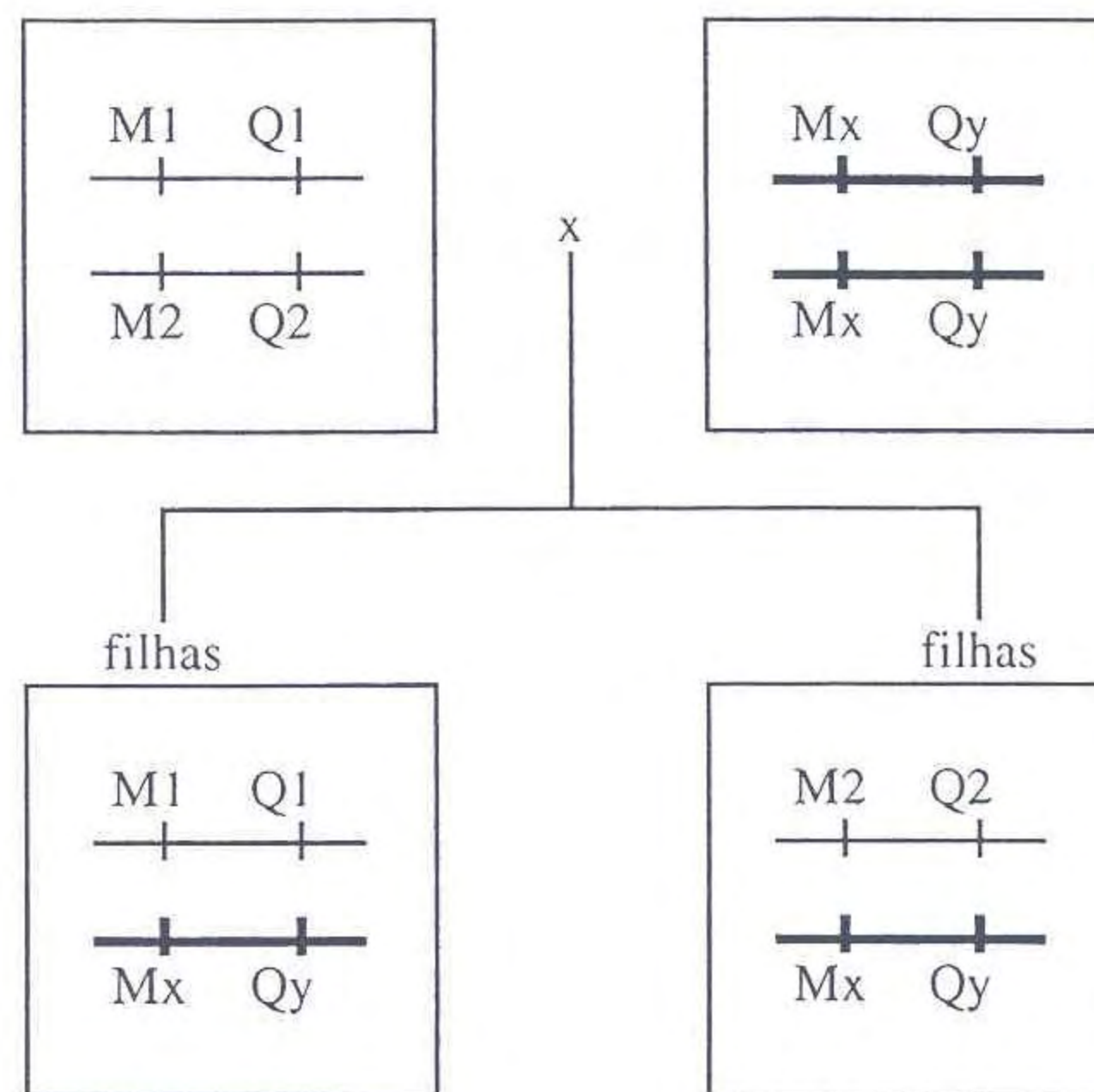


Figura 5.1: Diagrama do Desenho de Filhas. M1 e M2 alelos do marcador molecular do touro. Mx alelos dos marcadores das matrizes. Q1 e Q2 alelos do QTL do touro. Qy alelos do QTL das matrizes. cromossomo do touro; — cromossomo das matrizes. —

No Desenho de Filhas, todos os animais são genotipados para o marcador e são coletadas informações da produção da característica de interesse das filhas.

Segundo SOLLER & GENIZI (1978) e WELLER *et al.* (1990) para se detectar um QTL de efeito médio de substituição de alelos inferior a $\frac{1}{2}$ desvio-padrão será necessário genotipar e acompanhar a produção de milhares de filhas de diversos touros.

DESENHO DE NETAS (GRANDDAUGHTER DESIGN)

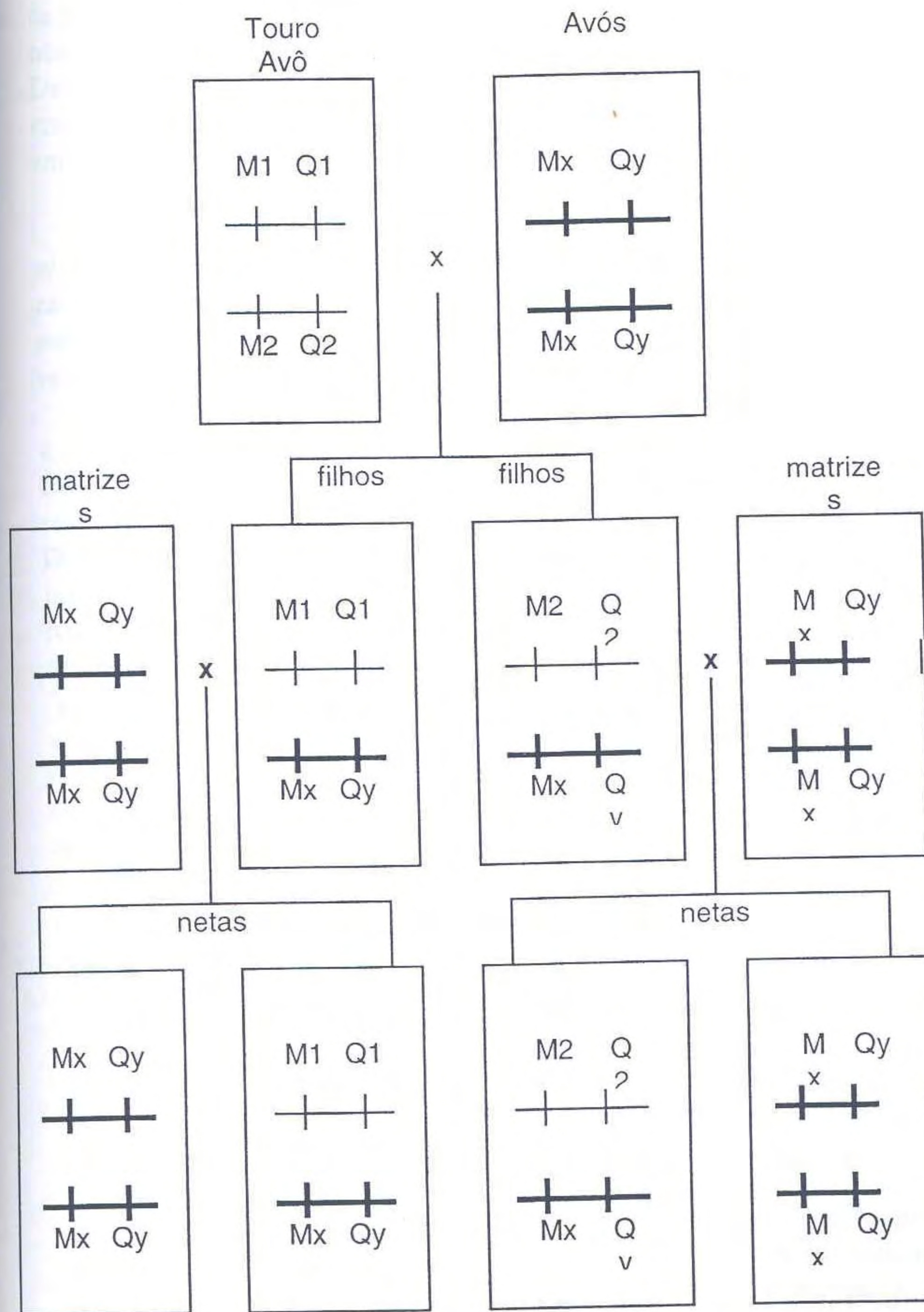


Figura 5.2 Diagrama do Desenho de Netas. M1 e M2 alelos do marcador molecular do touro. Mx alelos dos marcadores das matrizes. Q1 e Q2 alelos do QTL do touro e dos filhos. Qy alelos do QTL das matrizes. cromossomo do touro; — cromossomo das matrizes - - - .

Neste desenho, os touros possuidores de filhos que também são utilizados nos programas de inseminação artificial, são genotipados em busca de marcadores

heterozigotos. Os seus filhos também são genotipados e são coletadas informações da característica de interesse somente nas netas. Segundo WELLER *et al.* (1990) o número de genotipagens necessárias no Desenho de Netas é inferior ao necessário no Desenho de Filhas, mantendo-se o mesmo poder do teste. Os autores também consideram vantajoso utilizar apenas sêmen dos touros para a genotipagem dos animais, sem a necessidade de coleta de sangue das matrizes para essa finalidade.

Na Figura 5.2 está ilustrado o diagrama do Desenho de Netas.

O Desenho de Netas utiliza registros genealógicos e segundo GEORGES *et al.* (1995) com a redução da necessidade de genotipagens dos animais, para características de baixa herdabilidade, esse desenho experimental proporciona a redução de aproximadamente 3,5 vezes no tempo destinado para genotipagem em relação ao Desenho de Filhas.

Segundo WELLER *et al.* (1990), uma questão importante reside em conhecer a estrutura da população para julgar o desenho mais apropriado para análise. O Desenho de Filhas requer entre 5 e 20 touros com 200 a 2000 filhas cada um para sua realização com razoável poder de detecção. Segundo este autor, a aplicação do Desenho de Filhas seria interessante em rebanhos numerosos e com uso intensivo da inseminação artificial, com distâncias reduzidas entre os locais de criação e com reduzido número de touros provados usados na inseminação a cada ano. O Desenho de Netas seria mais apropriado em rebanhos nos quais os touros classificados como "elite" possuíssem filhos também empregados nos programas de inseminação artificial e preferencialmente pudesse ser obtido sêmen desses animais, em uma única coleta, em centrais de inseminação artificial. Esse desenho também seria mais vantajoso nos casos em que há um grande número de matrizes cujas informações zootécnicas estejam reunidas e organizadas.

Com relação aos métodos de análise do QTL, podem ser realizadas análises pelo método da análise de variância (WELLER *et al.*, 1990), pelo método da máxima verossimilhança (WELLER, 1986), ou pelo método do Qui-quadrado (GELDERMANN, 1975).

O PODER PARA SE DETECTAR QTLs

Uma das preocupações recentes entre os autores com relação aos distintos desenhos e modelos experimentais empregados para a localização de QTLs refere-se ao poder do teste para detectar a ligação deste com o marcador. Nos estudos com vários marcadores moleculares distribuídos ao longo do genoma, têm sido empregado com frequência o método da máxima verossimilhança. A principal diferença em relação ao método dos quadrados mínimos é que a máxima verossimilhança independe da distribuição dos dados. As estimativas de máxima verossimilhança (MLE - "Maximum Likelihood Estimation") são obtidas em análises realizadas por computadores e alguns programas foram desenvolvidos por grupos de pesquisa e estão disponíveis mediante solicitação. Usualmente as análises destas estimativas geram gráficos de "LOD scores" (LOD, para "log odds") que apresentam a verossimilhança da ligação entre dois locos geneticamente próximos um do outro.

O "LOD score" é o logarítmo na base dez de uma razão, na qual o numerador corresponde às chances de que os dados provenham de um QTL, enquanto o denominador corresponde às chances de que não exista nenhum QTL ligado aos marcadores. Valores de "LOD scores" que evitam a detecção de um QTL falso-positivo com 95% de confiabilidade quando são testados 60 marcadores distribuídos em uma região do genoma com 1200 cM foram estimados como sendo de 2,4 (LANDER & BOTSTEIN, 1989). O gráfico é construído calculando-se os valores de "LOD scores" para posições definidas dentro dos intervalos entre os marcadores.

A atenção devida à possibilidade de que sejam detectados QTLs falsos-positivos têm levado vários autores a realizarem estudos para estabelecer o poder dos testes estatísticos relativos a diferentes números de animais utilizados na análise (MARTINEZ, 1998; WELLER *et al.*, 1990; SOLLER & GENEZI, 1978; SIMPSON, 1989; MOODY *et al.*, 1997).

CONSIDERAÇÕES FINAIS

Uma abordagem tradicional em genética quantitativa é avaliar os fenótipos dos animais e dados a respeito de sua genealogia e inferir sobre a base genética que causa a variação fenotípica. A abordagem alternativa trazida pela genética molecular inclui a genotipagem direta dos animais nesse contexto, mas também a avaliação de como a localização dos genes e a elaboração de mapas genéticos podem auxiliar na explicação da variação fenotípica.

A identificação de genes de efeito principal responsáveis por características métricas comercialmente importantes em animais domésticos irá requerer a definição de locos candidatos. A localização pode ser realizada em populações abertas ou em populações construídas para este fim, no entanto uma das dificuldades nas metodologias para se detectar QTLs envolve o tamanho dos experimentos, mesmo para detectar efeitos genéticos principais. Outra complicação é que ainda que animais apresentem o mesmo genótipo para os principais genes relacionados a determinada característica quantitativa, verifica-se variação nos níveis de produção devida a fatores não-genéticos. Há também que considerar-se que nos animais de grande porte os intervalos entre cada geração são grandes e assim, as populações construídas com grandes animais trazem a necessidade de experimentos com múltiplas gerações, o que pode ser atingido somente após muitos anos.

Dentre as principais possibilidades de utilização do conhecimento sobre o genoma dos animais domésticos pode-se destacar o uso da seleção de animais assistida por marcadores moleculares, especialmente quando as características de interesse apresentarem baixa herdabilidade ou correlações genéticas indesejáveis. Também pode-se destacar a seleção precoce de animais destinados à reprodução. Outro importante desafio a mencionar refere-se ao isolamento, clonagem e manipulação direta de genes que controlem características quantitativas desejáveis. A engenharia genética de animais como forma de alterar suas características produtivas pode multiplicar muitas vezes os níveis de produção. Tudo isso depende inicialmente da localização correta e eficiente dos QTLs.

Não obstante às dificuldades, experimentos dessa natureza são conduzidos, pois não há como declinar de tarefa científica tão fundamental quanto a de contribuir na construção dos mapas genéticos das espécies mais importantes. As perspectivas existentes na vasta literatura sobre o assunto são de que as próximas décadas irão se beneficiar muito das informações básicas a respeito da estrutura e organização dos genomas. Inúmeras possibilidades advindas desse conhecimento e da exploração eficiente da variabilidade genética estão previstas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ARÚS, P.; MORENO-GONZÁLES, J. 1994. Molecular Markers. In: *Plant Breeding: Principles and Prospects*. Hayward M.D., Bosenmark N.O. and Romagosa I.(Ed.). Chapman & Hall, London. p. 314.
- BOVENHUIS, H.; SPELMAN, R. J. 1996. *Selective genotyping to detect QTL for multiple traits in outbred populations*. (compact disc). In: VI Congress on Genetics Applied to Livestock Production, Trabalhos, Armidale, Australia.
- CARBONELL, E. A.; GERIG, T. M.; BALANSARD, E.; ASINS, M. J. 1992. Interval mapping in the analysis of nonadditive quantitative trait loci. *Biometrics*, v. 48, p. 305-315.
- DARVASI, A.; SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theoretical Applied Genetics*, v. 85, p. 353-359, 1992.
- FUJII, J., OTSU, K., ZORZATO, F., DE LEON, S., KHANNA, V. K., WEILER, J. E., O'BRIEN, P. J., MACLENNAN, D. H. 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hypothermia. *Science*, v. 253, p. 448-451.
- FALCONER, D. S.; MACKAY T. F. C. 1996. *Introduction to Quantitative Genetics*. 4th ed. Essex, England. Longman. 464 p.
- GELDERMANN, H. 1975. Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. *Theoretical Applied Genetics*, v.46, p. 319-330.
- GEORGES, M.; NIELSEN, D.; MACKINNON, M.; MISHRA, A. OKIMOTO, R.; PASQUINO, A.T.; SARGEANT, L.S.; SORENSEN, S. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics*, v. 139, p. 907-920.
- FERREIRA, M. E.; GRATAPAGLIA, D. 1996. *Introdução ao uso de marcadores moleculares em análise genética*. Embrapa. 2.ed. Brasília: EMBRAPA-CENARGEN.
- HALEY, C. S.; KNOTT, S. A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers, *Heredity*, v. 69, p. 315-324.
- HILLEL, J.; DUNNINGTON, E. A.; SIEGEL, P. B. 1992. DNA markers in poultry breeding and genetic analyses. *Poultry Science Review*, v. 4, p. 169-186.

- KNAPP, S. J.; BRIDGES, W. C. JR.; BIKERS, D. 1992. Mapping quantitative trait loci using molecular marker linkage maps. *Theoretical Applied Genetics*, v. 79, p. 583-592.
- LANDER, E. S.; BOTSTEIN D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, v.121, p. 185-199.
- LEBOWITZ, R. J.; SOLLER, M.; BECKMANN, J. S. 1987. Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. *Theoretical Applied Genetics*, v. 73, p. 556-562.
- MARTINEZ M. L. 1998. Efeitos da substituição de alelos, da frequência gênica e da taxa de recombinação sobre o poder de detectar ligação entre um marcador genético e loci de característica quantitativa de baixa herdabilidade. *Revista Brasileira de Zootecnia*, v.27, n.4, p.715-721.
- MCPHERRON, A.; LEE, S. J. 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proceedings of the National Academy of Sciences*, v. 94, p. 12457-12461.
- MOODY, D. E.; POMP, D.; BUCHANAN, D. S. 1997. Feasibility of the grandprogeny design for quantitative trait loci (QTL) detection in purebred beef cattle. *Journal of Animal Science*, v. 75, p. 941-949.
- SAX, K. 1923. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, v. 8, p. 552-560.
- SHOOK, G. E. 1989. Selection for disease resistance. *Journal of Dairy Science*, v. 72, p.1349-1362.
- SILVER, L. M. 1995. *Mouse Genetics: Concepts and Applications*. Oxford University Press. 1995. New York.
- SIMPSON, S. P. 1989. Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. *Theoretical Applied Genetics*, v. 77, p. 815-819.
- SOLLER, M.; BECKMANN, J. S. 1983. Genetic polymorphism in varietal identification and genetic improvement. *Theoretical Applied Genetics*, v. 67, p. 25-33.
- SOLLER, M.; GENIZI, A. 1978. The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. *Biometrics*, v. 34, p. 47-55.
- VISSCHER, P. M.; HALEY, C. S. 1996. Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. *Theoretical Applied Genetics*, v. 93, p. 691-702.
- *WEIR, B. S. 1996. *Genetic Data Analysis II*. 2. ed. Sinauer Associates. Sunderland, MA, USA.
- WELLER, J. I. 1986. Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. *Biometrics*, v.42, p. 627-640.
- WELLER, J. I.; KASHI, Y.; SOLLER, M. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *Journal of Dairy Science*, v.73, p. 2525-2537.

ANEXO B

Distribution of bovine lymphocyte antigen (BoLA-DRB3) alleles in Brazilian dairy Gir cattle (*Bos indicus*)

A. F. da Mota*†, J. E. Gabriel*, M. L. Martinez† & L. L. Coutinho*

Summary

Brazilian dairy Gir (*Bos indicus*) cattle are a tropical, well-adapted breed, for which no information on the major histocompatibility complex (MHC) is presently available. The second exon of the bovine lymphocyte antigen (BoLA-DRB3) was amplified by polymerase chain reaction (PCR) of DNA samples from 28 Brazilian dairy Gir cows. Two experimental genotyping approaches were used: direct sequencing of PCR gene products (PCR-DS) and sequencing of cloned PCR fragments (S-CLO). Results demonstrate the viability of both typing approaches. PCR-DS allowed typing of 39% of the animals while the remainder were genotyped by S-CLO. Seventeen BoLA-DRB3 alleles were assigned, including some that were only recently described for zebu cattle. Allelic frequencies ranged from 0.02 to 0.18. The most frequent alleles were *3601 (frequency = 0.18), *2201 (0.14) and *2101 (0.11).

Introduction

Brazilian Gir cattle were imported from India at the beginning of the 20th Century. In Brazil, these animals were bred and selected for compact body, pendulous jowl and navel, convex head shape and red fur were selected. According to Santiago (1987), in 1938 selection of Gir dairy cattle took place at the Experimental Station of Umbuzeiro, Paraíba state, Brazil. Nowadays, the animals are well adapted to tropical conditions and current milking performance with purebred animals is near 2700 kg in a lactation of 305 days (<http://www.cnpqgl.embrapa.br/gir>).

Mastitis is the most common disease affecting Brazilian dairy herds (Veiga *et al.*, 1994). The inflammation of the

mammary gland reduces milk production, market price, and is a risk to human consumption. Resistance to mastitis is associated with the bovine lymphocyte antigen gene (BoLA), the major histocompatibility complex (MHC) found in cattle (Dietz *et al.*, 1997b). There are several studies in the literature in which the genetic association of the BoLA-DRB3 gene with resistance to mastitis has been investigated (Giovambattista *et al.*, 1996; Dietz *et al.*, 1997a,b; Kelm *et al.*, 1997; Starkenburg *et al.*, 1997; Sharif *et al.*, 1998; Gilliespie *et al.*, 1999). The BoLA-DRB3 gene product is a binding protein involved in the formation of an antigen-antibody complex, and is thus an important molecule associated with the specific immune response. To date, more than 70 different alleles have been reported for exon 2 of the BoLA-DRB3 gene (<http://www.ri.bbsrc.ac.uk/bola>) and correctly genotyping the animals is not a trivial task. The first studies used restriction fragment length polymorphism of amplified DNA fragments (PCR-RFLP) for allele assignments (Van Eijk *et al.*, 1992; Gelhaus *et al.*, 1995; Sitte *et al.*, 1995). This procedure is well characterized; however, as a result of the high polymorphism observed in this gene, several alleles are not detected by this methodology. As there are no studies of BoLA-DRB3 genotyping in the Gir breed, the purpose of this study was to identify alleles of the BoLA-DRB3 gene from a tropical zebu dairy breed (Gir dairy cattle) using two experimental genotyping approaches: direct sequencing of PCR gene products (PCR-DS) and sequencing of cloned amplified fragments (S-CLO).

Materials and methods

Animals and blood collection

Twenty-eight Gir dairy cattle (*Bos indicus*) from the Brazilian Organization for Agricultural Research, National Dairy Cattle Research Center (Embrapa Gado de Leite), Juiz de Fora, MG, Brazil, were used. Blood samples (5 ml) were collected in 0.5% EDTA and frozen at -20 °C for further analysis.

DNA extraction

The DNA extraction protocol consisted of adding 7 ml of lysis buffer (0.32 M sucrose, 12 mM Tris-HCl, pH 7.5, 5 mM MgCl₂ and 1% Triton × 100) to 5 ml of blood.

* Laboratory of Animal Biotechnology, Department of Animal Production, ESALQ, University of São Paulo, Avenue Pádua Dias 11, 13418-900, Piracicaba, SP, Brazil, and † Embrapa Gado de Leite, Brazilian Organization for Agricultural Research, National Dairy Cattle Research Center, Rua Eugênio do Nascimento 610, 36038-330, Juiz de Fora, MG, Brazil.

Received 15 January 2001; revised 23 July 2001; accepted 24 July 2001

Correspondence: Luiz Lehmann Coutinho, Laboratório de Biotecnologia Animal, Departamento de Produção Animal, ESALQ/USP, Av. Pádua Dias 11, 13418-900, Piracicaba, SP, Brazil. E-mail: llcouthin@carpa.ciagri.usp.br

Samples were then centrifuged at 2000 g for 20 min. The supernatant was discarded and another 3 ml of lysis buffer was added to the samples and centrifuged at 2000 g for 3 min. This step was repeated until a clean pellet was obtained. DNA was extracted by resuspending the pellet in 900 µl of a solution containing 450 µl of proteinase K buffer (0.375 M NaCl and 0.12 M EDTA, pH 8.0), 45 µl of proteinase K (20 mg ml⁻¹) and 20 µl of 20% sodium dodecyl sulphate. Samples were incubated overnight at 55 °C. On the following morning, protein was precipitated with 500 µl of 5 M NaCl and centrifuged at 2000 g for 25 min. The supernatant (1050 µl) was transferred to three 1500-µl Eppendorf tubes (350 µl each) and 1000 µl of absolute ethanol was added to each tube. The DNA was precipitated at 10 000 g for 15 min. The supernatant was discarded, and the samples washed in 70% ethanol. Then samples were dried at 37 °C for 1 h, and each DNA pellet was resuspended in 100 µl of milli-Q water (Millipore Corporation, Bedford, MA).

Amplification of BoLA-DRB3 exon 2

Oligonucleotide primers used for amplification of the second exon of BoLA-DRB3 were previously published in Van Eijk *et al.* (1992). Primers HL030 (5'-ATCCTCTCTCTGCAGCACATTTCC-3') and HL031 (5'-TTTAATTCGCGCTCACCTCGCCGCT-3') were used in the first amplification round. Amplification reactions were carried out with 100 ng of DNA (5 µl) in a 25-µl total volume containing 1 × PCR buffer, 10 mM dNTP mix, 50 mM MgCl₂, 0.5 µM of each primer and 1 unit of *Taq* DNA polymerase (Gibco BRL, New York, NY). The thermal cycling profile for the first round of amplification was an initial denaturation step of 3 min at 94 °C followed by 10 cycles of 25 s at 94 °C, 30 s at 60 °C, 30 s at 72 °C and a final extension step of 5 min at 72 °C. After the first round, a hemi-nested second PCR reaction was carried out with 10 µl of first-round product distributed (2 µl/tube) into five new tubes containing the same volume and concentration as described above except with primers HL030 and HL032 (5'-TCGCCGCTGCACAGTGAACTCTC-3'). Primer HL032 is internal to the sequence of the amplified product of the first-round PCR and has eight bases that overlap with primer HL031 (underlined in the text above).

The thermal cycling profile for the second round was 25 cycles of 40 s at 94 °C and 30 s at 65 °C, followed by a final extension step of 5 min at 72 °C. Reactions were performed in a 96-well thermal cycler (Perkin-Elmer Applied Biosystems, Foster City, CA).

Elution of the PCR products and direct sequencing

Briefly, electrophoresis was carried out on 2% agarose (Gibco BRL, New York, NY) gels with 100 µl of PCR reaction total volume combined from four tubes. Gel slices containing the specific BoLA-DRB3 PCR products were eluted and purified using the Sephaglas BandPreP Kit (Amersham-Pharmacia Biotech Inc., Piscataway, NJ).

Two µl of purified products containing 200–500 ng of DNA template was cycle-sequenced in a 15-µl reaction containing 2 µl of ABI Prism Dye Terminator Cycle Sequencing Core Kit with AmpliTaq DNA Polymerase (Perkin-Elmer Applied Biosystems), 1 µl of 5 µM primer HL030 or HL031 (consistent with forward or reverse sequencing reaction) and 2 µl of sequencing buffer (200 mM Tris-HCl, pH 9.0, and 5 mM MgCl₂). Reactions were performed according to the following cycling profile: 96 °C for 3 min, 25 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min. The sequenced samples were precipitated with 90% isopropanol, washed with 75% ethanol, and dried at room temperature. Samples were resuspended in 4 µl of deionized formamide: 50 mM EDTA, pH 8.0 (5 : 1), and 2 µl was loaded in an ABI 377 Automated Sequencer (Perkin-Elmer Applied Biosystems) and submitted to electrophoresis for 7 h.

Cloning procedure of eluted PCR products

Samples that revealed the possibility of identifying novel alleles, or whose genotyping from the sequenced PCR product was not reliable, were cloned in the *Sma*I site of pUC18 vector using SureClone Ligation Kit (Amersham-Pharmacia Biotech Inc.) and transformed into competent JM109 *Escherichia coli* cells (Hanahan, 1983). Clones containing BoLA-DRB3 alleles were sequenced under the same conditions as described above, except for the use of forward or reverse M13 universal primers (Perkin-Elmer Applied Biosystems).

Allele assignments

Alleles were initially assigned by examining the sequence data derived from PCR-DS reads with forward and reverse primers. The consensus sequence of both reads and chromatograms were compared with all previously identified alleles using SEQUENCHER assignment software, version 3.0 (Gene Code Corporation, Ann Arbor, MI). In order to resolve ambiguities in heterozygous samples, only alleles that allowed possible matches were considered. Allele designation was made only when the sequences of two published alleles matched precisely — at all positions — with the reads. Samples were cloned and sequenced in cases where direct typing of PCR product was ambiguous.

Results and Discussion

Comparison of different typing methods

We tested two approaches for sequencing products of BoLA-DRB3 exon 2 alleles. In the first, PCR products were directly sequenced (PCR-DS) and in the second, the amplified gene products were cloned before sequencing (S-CLO). When direct sequencing of the PCR products was carried out without elution of the fragments from agarose gels, correct assignment of the alleles was impossible. Results showed that, although PCR is a

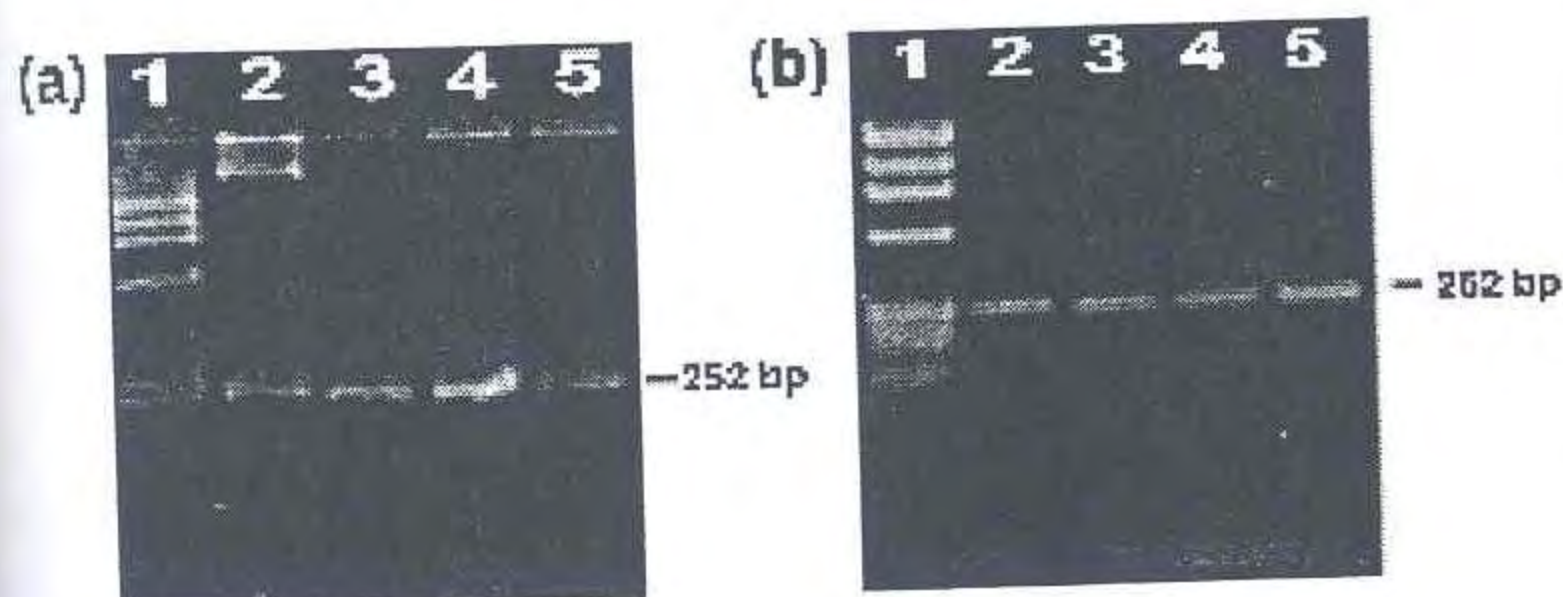


Figure 1. (a) Electrophoresis on 2% agarose gel of 8 µl of PCR from the BoLA-DRB3 gene (252 bp) after two rounds of hemi-nested reactions with specific primers. Background bands and samples that showed excess genomic DNA (lane 2), smear formation (lane 4) and a variable quantity of amplification of the BoLA-DRB3 gene (lanes 3, 4 and 5) are shown. (b) The same samples and volumes of PCR after band isolation, DNA elution and concentration (lane 1 = φ X174 *Hae*III, 1.5 µg).

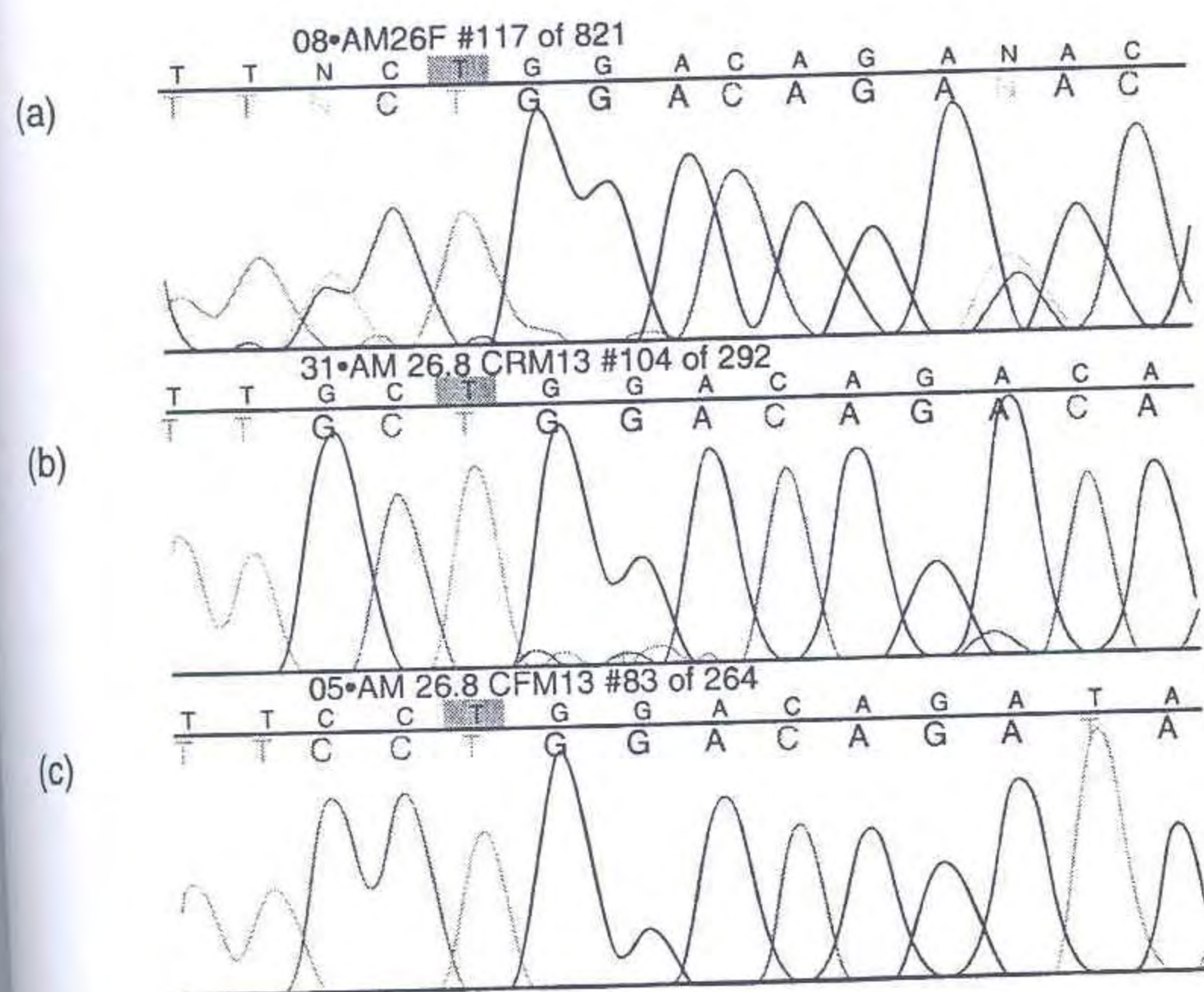


Figure 2. (a) Chromatogram of sequence reaction with nucleotide sequence from PCR-DS showing multiple mismatches positions (N). (b) Same samples from S-CLO showing the sequence from one of the alleles. (c) Chromatogram from S-CLO showing the sequence from the other allele. Nucleotide sequence of BoLA-DRB3*1703 allele deposited in GenBank (accession number AF209710): 5'-tggagtattc tacggcgag tgcatttct tcaacgggac cgagcgggtg cggctactgg acagatactt ctataatgga gaagagtacg cgcgcttcca cagcactgg ggcgagtcc gggcgggtgac cgagctgggg cggccggacg ccaagtactg gaacagccag aaggagatcc tggagcggga gcgggcctat gtggacacgt actgcagaca caactacggg gtcggtgaga gtttactgt gc-3'.

rapid and effective technique for amplifying specific DNA fragments, sequencing of PCR products requires the isolation and purification of templates (Fig. 1). When PCR products were eluted and purified from agarose gels, the primers were taken off and the sequencing reaction was unique for the primer added to the reaction.

Genotyping was feasible by direct sequencing of gel-eluted fragments. However, the multiple mismatching of sequences attributable to excess heterozygous differences resulted in time-consuming and sometimes ambiguous allele identification. Sequencing of clones produced better quality sequences and immediate genotyping (Fig. 2). All animals had at least two sequences analysed by PCR-DS, one with the forward primer and the other with the reverse primer. For the 28 cows studied, 11 (39.3%) were correctly genotyped by PCR-DS. Two cows out of these 11 animals needed four reads from direct sequencing for reliable typing. Material from the other 17 animals had to be cloned in order to obtain reliable typing. S-CLO was also conducted on the 11 animals genotyped by PCR-DS to confirm the results.

PCR-DS has both advantages and limitations. First, the sequence is readily obtained because this approach does not depend on the cloning of PCR fragments. Also,

only 64 reads were necessary to genotype 11 animals. However, when gel-purified BoLA-DRB3 products were directly sequenced, the high level of polymorphism at this locus demanded extremely time-consuming DNA sequence analysis, and in some cases cloning was still necessary.

S-CLO was more labour intensive. Two hundred and ninety-seven reads were necessary to genotype the remaining 17 animals. When the PCR products were cloned, several high-quality sequence reads were needed to form a consensus and distinguish possible mutations from misincorporated nucleotides introduced by *Taq* DNA polymerase errors during PCR (data not shown). Misincorporation of nucleotides and possible errors by *Taq* DNA polymerase have been discussed in the literature (Gyllestein & Allen, 1993).

To circumvent the limitations associated with both the PCR-DS and S-CLO approaches, a possible solution discussed in the literature is the separation of alleles prior to DNA sequencing. Non-denaturing single-strand conformation polymorphism (SSCP) (Mok *et al.*, 1993) and denaturing gradient gel electrophoresis (DGGE) (Aldridge *et al.*, 1998) have provided separation of alleles. BoLA-DRB3 exon 2 has extensive polymorphisms, and the

Table 1. Allelic frequencies of bovine lymphocyte antigen BoLA-DRB3 from 28 Gir Dairy cows (*Bos indicus*) typed by direct sequencing of PCR products or by cloning-sequencing protocol

BoLA-DRB3 allele	Number of copies	%
*0201	1	
*1001	1	1.79
*14011	1	1.79
*1602	1	1.79
*25011	1	1.79
*4101	1	1.79
*1703	1	1.79
*2704	2	3.57
*3001	2	3.57
*2802	2	3.57
*3101	3	5.36
*1801	3	5.36
*2801	4	7.14
*3301	4	7.14
*2101	4	7.14
*2201	6	10.71
*3601	8	14.29
	10	17.86

methodologies mentioned above could be used to separate the alleles prior to sequencing.

Allelic frequencies

Allelic frequencies of Brazilian dairy Gir cows, a tropical *Bos indicus* cattle breed, ranged from 0.02 to 0.18 (Table 1). The least frequent alleles were *0201, *1001, *14011 and *1602 (frequencies = 0.0179). Conversely, the most frequent alleles were *3601 (0.1786), *2201 (0.1429) and *2101 (0.1071), corresponding to almost 43% of allelic frequencies.

The nucleotide sequence of the BoLA-DRB3*1703 allele deposited in GenBank (<http://www.ncbi.nlm.nih.gov>) has 252 bp (accession number AF209710), which is 3 bp larger at the 3-prime end than the latest allele submission recently deposited in GenBank from an independent study on Brahman cattle (Maillard *et al.*, 1999).

Other studies have investigated allelic frequencies of the BoLA-DRB3 gene. Kelm *et al.* (1997) genotyped 131 Holstein cows and, according to PCR-RFLP patterns and nomenclature described by Van Eijk *et al.* (1992) and Gelhaus *et al.* (1995), observed that alleles DRB3.2*8 and *11 were the most frequent (21.31 and 17.62%, respectively). Sharif *et al.* (1998) found allelic frequencies of 20.1% for allele DRB3.2*8 and 14.9% for *11 in Holsteins, while in Jersey cattle allele *8 was present at a frequency of 1.7% and allele *11 was absent. Gillespie *et al.* (1999), in a study of Jersey cows, found an allelic frequency of 11.3% for allele DRB3.2*8 and an absence of allele *11. The pattern for allele DRB3.2*8, obtained by PCR-RFLP, corresponds to those for alleles DRB3.2*1201 and DRB3.2*1202 (<http://www2.ri.bbsrc.ac.uk/bola/dr3tab.htm>). In our study with Brazilian dairy Gir cattle, alleles *1201 and *1202 were not observed. Allele DRB3.2*11, which corresponds to alleles DRB3.2*0901 and *0902, was also not observed in our study. In

contrast, allele DRB3.2*2201 (DRB3.2*06 for PCR-RFLP pattern), which had a frequency of 0.003 in the study of Giovambattista *et al.* (1996) on Argentinian Creole cattle, was the second most frequent allele in our study (0.1429). In addition, alleles *3601, *2201, *2101, *1801 and *2801 (PCR-RFLP alleles *20, *06, *35, *18 and *31, respectively), which were responsible for almost 60% of the allelic frequencies in our study, were not found in American Holsteins (Kelm *et al.*, 1997) and represented only 2% of the allelic frequencies of Canadian Holsteins (Sharif *et al.*, 1998).

A source of uncertainty concerning the zebu allelic frequencies at the BoLA-DRB3.2 locus is the lack of studies on zebu cattle in tropical regions. This is the first study of the BoLA-DRB3 gene in Brazilian dairy Gir cattle. Results in the literature indicate that allelic frequencies of BoLA-DRB3 may, at least to some extent, depend on the breed, as a result of the founder population and selection pressure. At present, almost all previously identified BoLA-DRB3 alleles were recognized in European breeds. Even for a relatively small population, this study allowed typing of 17 alleles, some of which were only recently described for zebu cattle (Maillard *et al.*, 1999). Other new alleles probably remain to be discovered for the BoLA-DRB3 gene in future studies on tropical cattle.

We demonstrated the viability of both typing approaches. PCR-DS allowed the typing of BoLA-DRB3 alleles. The S-CLO procedure was used to determine reliable nucleotide sequences of the BoLA-DRB3 gene. Other genotyping studies with genes with lower polymorphism should be conducted by PCR-DS. This technique is substantially less labour intensive and can give correct assignments of the alleles. In situations of ambiguous typing, cloning and sequencing of isolated alleles will lead to the correct genotype.

Acknowledgements

A. F. da Mota acknowledges EMBRAPA for a PhD fellowship, L. L. Coutinho and M. L. Martinez acknowledge the research productivity scholarship received from the National Research Council of Brazil (CNPq), and J. E. Gabriel acknowledges FAPESP for a PhD fellowship.

References

- Aldridge, B.M., McGuirk, S.M., Clark, R.J., Knapp, L.A., Watkins, D.I. & Lunn, D.P. (1998) Denaturing gradient gel electrophoresis: a rapid method for differentiating BoLA-DRB3 alleles. *Animal Genetics*, **29**, 389.
- Dietz, A.B., Denilleux, J.C., Freemann, A.E., Kelley, D.H., Stabel, J.R. & Kehrl, M.E., Jr (1997a) Genetic association of bovine lymphocyte antigen DRB3 alleles with immunological traits of Holstein cattle. *Journal of Dairy Science*, **80**, 400.
- Dietz, A.B., Cohen, N.D., Timms, L. & Kehrl, M.E., Jr (1997b) Bovine lymphocyte antigen class II alleles as risk factors for high somatic cell counts in milk of lactating dairy cows. *Journal of Dairy Science*, **80**, 406.
- Gelhaus, L., Schnittger, L., Mehlitz, D., Horstmann, R.D. & Meyer, C.G. (1995) Sequence and PCR-RFLP analysis of 14 novel BoLA-DRB3 alleles. *Animal Genetics*, **26**, 147.

- Gilliespie, B.E., Jayarao, B.M., Dowlen, H.H. & Oliver, S.P. (1999) Analysis and frequency of bovine lymphocyte antigen DRB3.2 alleles in Jersey cows. *Journal of Dairy Science*, **82**, 2049.
- Giovambattista, G., Golijow, C.D., Dulout, F.N. & Lojo, M.M. (1996) Gene frequencies of DRB3.2 locus of Argentine Creole cattle. *Animal Genetics*, **27**, 55.
- Gyllestein, U.B. & Allen, M. (1993) Sequencing of *in vitro* amplified DNA. *Methods in Enzymology*, **18**, 3.
- Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, **166**, 557.
- Kelm, S.C., Dettileux, J.C., Freeman, A.E., Kehrl, M.E., Jr, Dietz, A.B., Fox, L.K., Buttler, J.E., Kascovics, I. & Kelley, D.H. (1997) Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. *Journal of Dairy Science*, **80**, 1767.
- Maillard, J.C., Renard, C., Chardon, P., Chantal, I. & Bensaid, A. (1999) Characterization of 18 new BoLA-DRB3 alleles. *Animal Genetics*, **30**, 200.
- Mok, S.C., Lo, K.W. & Tsao, S.W. (1993) Direct cycle sequencing of mutated alleles detected by PCR single-strand conformation polymorphism analysis. *Biotechniques*, **14**, 790.
- Santiago, A.A. (1987) *Cruzamentos na Pecuária Bovina: Raças Taurinas, Raças Zebuínas, Cruzamentos*, p. 549. Institute Zootecnia, São Paulo.
- Sharif, S., Mallard, B.A., Wilkie, B.N., Sargeant, J.M., Scott, H.M., Dekkers, J.C.M. & Leslie, K.E. (1998) Associations of the bovine major histocompatibility complex DRB3 (BoLA-DRB3) alleles with occurrence of disease and milk somatic cell score in Canadian dairy cattle. *Animal Genetics*, **29**, 185.
- Sitte, K., East, I.J., Lavin, M.F. & Jazwinska, E.C. (1995) Identification and characterization of new BoLA-DRB3 alleles by heteroduplex analysis and direct sequencing. *Animal Genetics*, **26**, 413.
- Starkenburg, R.J., Hansen, L.B., Kehrl, M.E., Jr & Chester-Jones, H. (1997) Frequencies and effects of alternative DRB3.2 alleles of bovine lymphocyte antigen for Holsteins in milk selection and control lines. *Journal of Dairy Science*, **80**, 3411.
- Van Eijk, M.J.T., Stewart-Haynes, J.A. & Lewin, H.A. (1992) Extensive polymorphism of the BoLA-DRB3 gene distinguished by PCR-RFLP. *Animal Genetics*, **23**, 483.
- Veiga, V.M.O., Teixeira, M., Brito, M.A.V.P. & Brito, J.R.F. (1994) *Controle da Mamite dos Bovinos*. Empresa Brasileira de Pesquisa Agropecuária-Embrapa, Centro Nacional de Pesquisa de Gado de Leite, Juiz de Fora, Embrapa, Minas Gerais, Brazil.

ANEXO C

Adilson F. da Mota • Mário L. Martinez • Luiz L. Coutinho

Genotyping BoLA-DRB3 alleles in Brazilian Dairy Gir cattle (*Bos indicus*) by temperature-gradient gel electrophoresis (TGGE) and direct sequencing

Received:

A. F. da Mota • M. L. Martinez

Embrapa Gado de Leite, Brazilian Organization for Agricultural Research, National Dairy Cattle Research Center, Juiz de Fora - MG, Brazil

L. L. Coutinho (✉)

Laboratory of Animal Biotechnology - Dept. of Animal Production - ESALQ - University of São Paulo - Av. Pádua Dias 11, 13418-900, Piracicaba - SP, Brazil

e-mail: llcoutin@esalq.usp.br

tel: +55 19 32494434, fax: +55 19 34294284

Genotyping BoLA-DRB3 alleles in Brazilian Dairy Gir cattle (*Bos indicus*) by temperature-gradient gel electrophoresis (TGGE) and direct sequencing

Abstract

BoLA-DRB3 is a gene of the Major Histocompatibility Complex (MHC) in cattle. The product of the BoLA-DRB3 is a protein related to formation of an antigen-antibody complex associated with immune response. Genotyping BoLA-DRB3 gene is relatively complex due to the extensive polymorphism associated with this locus. Current techniques for assignment of genotypes are restriction fragment length polymorphism of DNA amplified fragments (PCR-RFLP), direct sequencing of PCR products, cloning-sequencing, and denaturant-gradient gel electrophoresis. These techniques are either time-consuming, do not discriminate all possible alleles, or are not readily reproducible. The objective of this study was genotyping BoLA-DRB3 by use of temperature-gradient gel electrophoresis (TGGE) to separate alleles before sequencing. PCR reactions made out of 28 DNA samples from Gir Dairy cows (*Bos indicus* Brazilian breed) were submitted to TGGE. New PCR products were generated from separated alleles, purified, and sequenced. Allele separation was possible in 21 out of 26 heterozygote samples (81%). Results indicate that two sequence reads (forward and reverse) were sufficient to accurately genotype BoLA-DRB3 alleles. Separation of alleles by TGGE provides a high-throughput reliable typing of BoLA-DRB3, critical to performing disease association studies in cattle.

Keywords: BoLA-DRB3 gene • Bovine Lymphocyte Antigen • Immune system • Temperature-Gradient Gel Electrophoresis

Introduction

The Bovine lymphocyte Antigen gene BoLA-DRB3 product is a binding protein involved in formation of DR-beta chain highly expressed antigen-antibody complexes, playing a critical role in the specificity of immune response. To date, more than 100 different alleles from exon 2 of gene BoLA-DRB3 have been deposited in GenBank (<http://www.ncbi.nlm.nih.gov/>). Genotyping BoLA is relatively complex because the genes within this family are extremely polymorphic. This parallels reports for HLA-DRB1 where more than 290 alleles (<http://www.anthonynolan.org.uk/HIG/lists/class2list.html>) have been identified. The diversity of bovine DRB3 alleles increases the difficulty of generating accurate genotypes for this locus, especially for association studies lacking pedigree information.

Many studies have attempted to investigate the association between BoLA-DRB3 with resistance to mastitis (Van Eijk et al., 1992, Gelhaus et al., 1995, Dietz et al., 1997, Kelm et al., 1997, Sharif et al., 1998, Gilliespie et al., 1999) and achieved different conclusions. In the earliest studies, genotyping was determined by restriction fragment length polymorphism of DNA amplified fragments (PCR-RFLP) for assignment of alleles. This methodology cannot accurately determine differences between all current alleles and this may have led to the different conclusions in disease association studies. Techniques that are more current indicate that direct sequencing of PCR products and cloning-sequencing allow for the correct determination of BoLA-DRB3 typing (da Mota et al., 2002). However, cloning-sequencing is a time-consuming technique and direct sequencing of PCR products cannot always resolve alleles distinguished by insert and deletion events. A more robust method for genotyping used denaturing gradient gel electrophoresis (DGGE) to separate alleles prior to sequencing (Aldridge, 1998) based on shift caused by different secondary structure of amplimers.

The drawback of this technique is reproducibility of the gels.

In this study the principal aim was to genotype BoLA-DRB3 alleles by temperature-gradient gel electrophoresis (TGGE) to separate the alleles prior to sequencing, using DNA samples from Brazilian Dairy Gir cattle (*Bos indicus*).

Material and Methods

DNA extraction

Five-ml of blood samples were collected from twenty-eight Dairy Gir cows (*Bos indicus*) raised at the Brazilian Organization for Agricultural Research, National Dairy Cattle Research Center (Embrapa Gado de Leite), Juiz de Fora, Minas Gerais state, Brazil.

For DNA extractions, 7 ml of lysis buffer (0.32 M sucrose, 12 mM Tris-HCl pH 7.5, 5 mM MgCl₂, 1% Triton 100X) were mixed to 5 ml of blood samples and centrifuged at 2,000 x g for 20 minutes. The supernatant was discarded. Another 3 ml of lyses buffer were mixed to the samples and centrifuged at 2,000 x g for 3 minutes. Washes with the lyses buffer were repeated until a clean pellet of lymphocytes was obtained. Then, cells were resuspended in 900 μ l of a solution containing 450 μ l of proteinase K buffer (0.375 M NaCl, 0.12 M EDTA pH 8.0), 45 μ l proteinase K (20 mg/ml) and 20 μ l of 20% sodium dodecyl sulfate (SDS). Samples were incubated overnight at 55°C. Protein was precipitated by addition of 500 μ l of 5M NaCl and centrifugation at 2,000 x g for 25 minutes. The supernatant was transferred to clean eppendorf tubes and DNA was precipitated with addition of 1,000 μ l of absolute ethanol to each tube and centrifugation at 10,000 x g for 15 minutes. The supernatant was discarded and the DNA pellet was washed in 70% ethanol. After drying at 37°C for one hour, DNA was

resuspended in 100 μ l of milli-Q water.

Amplification of BoLA-DRB3 exon 2

Oligonucleotide primers used for amplification of the second exon of BoLA-DRB3 were previously published primer sequences (Van Eijk et al., 1992). Primers HL030 [5'-ATCCTCTCTCTGCAGCACATTTCC-3'] and HL031 [5'-TTTAATTCGCGCTCACCTCGCCGCT-3'] were used in a first amplification round. Reactions were carried out on 100 ng of DNA (5 μ l) in a 25 μ l total volume containing 1 x PCR buffer (50 mM KCl, 10mM TRIS-HCl, 0.11% gelatin), 0.8 mM dNTP mix, 1.5 mM MgCl₂, 0.5 μ M of each primer and 1 unit of *Taq* DNA polymerase (GIBCO BRL, New York, NI). The thermal cycling profile for the first round of amplification was an initial denaturation step of 3 minutes at 94°C followed by 10 cycles of 25 seconds at 94°C, 30 seconds at 60°C, 30 seconds at 72°C and a final extension step of 5 minutes at 72°C. After the first round, a hemi-nested second PCR reaction was carried out with 10 μ l of first-round product distributed (2 μ l/tube) on five new tubes containing the same volume and concentration as described above, except for using primers HL030 and HL032 [5'-TCGCCGCTGCACAGTGAACTCTC-3']. Primer HL032 is internal to the sequence of the amplified product of the first round PCR and has an overlapping palindrome eight base box that is underlined in the text above. The thermal cycling profile for the second round was 25 cycles of 40 seconds at 94°C and 30 seconds at 65°C, followed by a final extension step of 5 minutes at 72°C. The reactions were performed in a 96-well thermal cycler (Perkin-Elmer-Applied Biosystems, Foster City, CA).

TGGE

Electrophoresis was carried out on a horizontal TGGE system with 9 cm x 9 cm gel size (Biometra, Göttingen, Germany). Optimal temperature gradient was established by perpendicular TGGE and by considering the thermal denaturation profiles and gel mobility shift of double stranded DNA (Steger, 1994). Electrophoresis was carried out with an 8% acrylamide gel (29:1 acrylamide:bis-acrylamide) run in 0.2 x Na-TAE buffer pH 8.4 (1 M Sodium Acetate, 10 mM EDTA, 400 mM TRIS). A thermal gradient from 35°C to 75°C was tested using amplification products mixed 2:1 with loading buffer (0.2 x Na-TAE pH 8.4, 0.1% Triton-X 100, 0.01% Bromophenol Blue dye, 0.01% Xylene Cyanol dye). Samples were pre-run for 15 minutes at 100 V and run for 70 minutes at 250 V. Gels were stained in ethidium bromide solution (0.5 µg/ml) and visualized using an ultra-violet transilluminator.

PCR samples containing approximately 3 µg of the amplified DNA fragment of BoLA-DRB3 gene were run on parallel TGGE under the electrophoresis conditions listed above. Gel plugs containing the separated alleles were excised from the acrylamide gel using a wide bore pipette tip and were transferred to a microcentrifuge tube containing 50 µl of the elution solution for polyacrylamide gels (10 mM TRIS-HCl, 50 mM KCl, 1.5 mM MgCl₂, 0.1% Triton X-100) and incubated overnight at room temperature to elute the DNA.

Sequence analysis

Two-µl (50 ng) of the solution containing the eluted DNA were used as template in another PCR using the same conditions listed for the second round of amplification of BoLA-DRB3 gene. PCR amplified product was purified using a GFX™ Purification Kit (Amersham Pharmacia Biotech, Piscataway, NJ). Two-µl of purified

products containing approximately 200-500 ng of template DNA were cycle-sequenced in a reaction total volume of 15 μ l containing 2 μ l of ABI Prism Dye Terminator Cycle Sequencing Core Kit with AmpliTaq DNA Polymerase (Perkin-Elmer-Applied Biosystems, Foster City, CA), 1 μ l of 3.2 μ M primer HL030 or HL032 (consistent with forward or reverse sequencing reaction) and 2 μ l of sequencing buffer. Reactions were performed according to the manufacturer cycling profile. Sequenced samples were precipitated in 90% isopropanol, washed with 75% ethanol, and dried at room temperature. Samples of the different alleles were resuspended in 4 μ l of deionized formamide : 50 mM EDTA pH 8.0 (5 : 1) and 2 μ l were loaded on an ABI 377 Automated Sequencer (Perkin-Elmer-Applied Biosystems, Foster City, CA) and submitted to electrophoresis for 7 hours.

Allele assignments

Alleles separated by TGGE were assigned by examining the sequence data derived from direct sequencing with forward and/or reverse primers. The sequences of both reads and chromatograms were compared to all the identified alleles using Sequencher™ assignment software (Gene Code Corporation, Ann Arbor, MI). New amplimers were generated for all the genotyped samples and were cloned and sequenced to confirm the previous TGGE typing.

Results and Discussion

We have developed a new approach to improve accuracy and throughput of genotyping exon 2 of BoLA-DRB3 alleles. All the samples were amplified by PCR and optimal temperature-gradient gel electrophoresis conditions were established. Best

results were visualized when the electrophoresis time was 70 minutes at 250 V and the temperature-gradient was from 35°C to 75°C (Figure 1).

Animals were genotyped by TGGE methodology herein described. There were twenty-six heterozygote animals. The first time the samples were submitted to optimized electrophoresis, separation of the alleles by TGGE and consequent typing was possible in twenty-one out of the twenty-six heterozygote samples (81%). Some heterozygote could not be genotyped using TGGE because the alleles were not completely separated (Figure 1A, lanes 1 and 6). TGGE presented consistent results with highly reproducible gel images. In Figure 1A and 1B, sample 4 presents the same gel electrophoresis pattern in two independent runs.

Theoretically, all different alleles should migrate with different bands pattern shift in a TGGE gel because even double strand DNA (dsDNA) fragments of the same length show different secondary conformation caused by point mutations. The increase in temperature denatures dsDNA by opening defined melting domains and generating a melting profile with different melting domains. This results on the reduction in electrophoretic mobility when dsDNA migrates from the lower melting domain to the higher melting domain. Due to the presence of heteroduplexes formed by the annealing of different alleles, hetero/homoduplexes have different melting patterns according to base composition, and private electrophoretic behaviors. Consequently, it is possible to visualize four bands (Figure 1A, lane 3) in a heterozygote sample. The two lower bands are homoduplexes while the other two bands correspond to heteroduplexes.

For some samples, it was possible to obtain complete separation of the homoduplexes (Figure 1A, samples 2 and 5, and Figure 1B, samples 1, 2, 3, 4, and 6) that were further eluted and sequenced.

Table 1 allows the comparison of some details on the different alleles. Some of

the alleles that could not be separated had the same length, similar melting temperature, and G-C content (ID=7 and ID=27). On the other hand, alleles presenting larger differences in G-C content, sequence of nucleotides, and molecular weight were not separated (ID= 6 and ID=28). It is possible to observe also that alleles with events of insertion/deletion were separated in some cases (ID= 4, 21, 22, and 24), but samples ID=10 and ID=12, with insertion/deletion of 3 nucleotides, were not separated. Table 1 also shows that even samples containing alleles most similar in nucleotide sequence were separated by TGGE (ID= 5 and ID=13). Our results suggest that it is not possible to infer if one amplicon from BoLA-DRB3 will be separated, based solely on G-C content, length of the alleles, melting temperature, and molecular weight.

Simulations can be used to visualize DNA secondary structures that could result from increasing temperature (<http://www.biophys.uni-duesseldorf.de/local/POLAND/poland.html>) when there are differences in the sequences. We tested some of the alleles considered as having similar sequences, using the website with POLAND methodology (data not shown). Some of the alleles presented simulated secondary structures that would justify their separation for the electrophoretic conditions set up in this study.

When base pairs vary along the sequences, different DNA secondary structure happens, and forms differences within a double strand of DNA. The secondary structure forms DNA loops or bulges. Also, a possible explanation for the relation between secondary structure and the sequence itself would be the angle of interaction between base pairs, resulting in major or minor grooves and helix diameter. Simulation using Poland methodology for secondary structures of the alleles did not allow a definite conclusion.

The formation obtained from a specific combination of alleles is not predictive

before genotyping the alleles. Alleles that seem to be very similar can present differences in secondary structure and melting domains that are not explained only by length, G-C content, molecular weight, or software simulation of secondary structures. As the alleles are not known *a priori*, simulations are not useful in this case for specific alleles combination. However, electrophoretic conditions used in this study were optimized to produce the better separation of BoLA-DRB3 alleles.

In the cases where the alleles were not separated by TGGE, it would be possible to add a GC clamp during PCR, which would generate a second melting domain and a different band migration pattern. The addition of G-C clamps for separation of alleles is presented in the literature (Kappes et al., 1995). Cloning-sequencing procedure also generates reliable genotyping in these cases (da Mota et al., 2002).

Conclusion

TGGE was considered a fast and reliable technique for BoLA-DRB3 typing, as the alleles of 81% of the samples could be separated for sequence analysis. Our results indicate that two sequence reads (forward and reverse) were sufficient to determine the correct nucleotide sequence and typing of exon 2 of BoLA-DRB3 alleles. Separation of BoLA-DRB3 alleles by TGGE saves time and allows for better detection rate of haplotypes. The reproducibility of the temperature-gradient is consistent. This technique is amenable to high-throughput, allowing disease association studies and discovery of new alleles, especially useful for *Bos indicus* breeds as Brazilian Dairy Gir cattle.

Acknowledgements

Adilson F. da Mota acknowledges EMBRAPA for a PhD fellowship; Luiz Lehmann Coutinho and Mário Luiz Martinez acknowledge the research productivity

scholarship received from the National Research Council of Brazil (CNPq). The authors acknowledge Dr. Tad S. Sonstegard for critical review on the manuscript. All the experiments with animals comply with current Brazilian laws.

References

- Aldridge, B. M., McGuirk, S. M., Clark R. J. et al. (1998) Denaturing gradient gel electrophoresis: a rapid method for differentiating BoLA-DRB3 alleles. *Animal Genetics* 29: 389-394.
- da Mota, A. M., Gabriel, J. E., Martinez, M. L., Coutinho, L. L. (2002) Distribution of the bovine lymphocyte antigen (BoLA-DRB3) alleles from Brazilian Dairy Gir cattle (*Bos indicus*). *European Journal of Immunogenetics*. 29(3): 213-7.
- Dietz, A. B., Cohen, N. D., Timms, L., Kehrl Jr., M. E. (1997) Bovine lymphocyte antigen Class II alleles as risk factors for high somatic cell counts in milk of lactating dairy cows. *Journal of Dairy Science* 80: 406-412.
- Gelhaus, L.; Schnittger, L.; Mehlitz, D.; Horstmann, R.D.; Meyer, C.G. (1995) Sequence and PCR-RFLP analysis of 14 novel BoLA-DRB3 alleles. *Animal Genetics* 26: 147-153.
- Gilliespie, B. E., Jayarao, B. M., Dowlen, H. H., Oliver, S. P. (1999) Analysis and frequency of bovine lymphocyte antigen DRB3.2 alleles in Jersey cows. *Journal of Dairy Science* 82: 2049-2053.
- Kappes, S., Milde-Langosch, K., Kressin, P., Passlack, B., Dockhorn-Dworniczak, B., et al. (1995) p53 mutations in ovarian tumors, detected by temperature-gradient gel electrophoresis, direct sequencing and immunohistochemistry. *International Journal of Cancer* 64: 52-59.
- Kelm, S. C., Dettileux, J. C., Freeman, A. E., et al. (1997) Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. *Journal of Dairy Science* 80: 1767-1775.
- Sharif, S., Mallard, B. A., Wilkie, B. N., Sargeant, J. M., Scott, H. M., Dekkers, J. C. M., Leslie, K. E. (1998) Associations of the bovine major histocompatibility complex DRB3 (BoLA-DRB3) alleles with occurrence of disease and milk somatic cell score in Canadian dairy cattle. *Animal Genetics* 29: 185-193.
- Steger, S. (1994) Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic Acids Research* 22: 2760-2768.

Van Eijk, M. J. T., Stewart-Haynes, J.A., Lewin, H. A. (1992) Extensive polymorphism of the BoLA-DRB3 gene distinguished by PCR-RFLP. *Animal Genetics* 23: 483-496.

Table 1: Animal identification, assignment of alleles, differences on the sequences, insertions/deletions, G-C content (%), melting temperature (°C), molecular weight (Daltons) and molarities equivalent to optical density equal to 1 for the 26 heterozygote samples.

Animal ID	BoLA-DRB3 Alleles	Diff. in Seq. (bp)	Ins./ Del. (bp)	G-C (%)	Melting Temp. (°C)	Mol. Weight (Daltons)	picoMolar equivalent to OD=1
1	*14011	26	-	61	87	78252	362
	*25011			61	87	77774	367
2	*2101	20	-	60	87	77924	363
	*3601			60	87	77497	365
3	*2802	25	-	61	87	73494	385
	*3601			60	87	77497	365
4	*4101	17	3	60	87	76551	369
	*2101			60	87	77924	363
5	*2801	16	-	61	87	77861	365
	*2201			62	87	77948	363
6**	*1801	21	-	62	88	77759	368
	*2802			61	87	73494	385
7**	*1703	23	-	60	87	77494	366
	*2801			61	87	77861	365
8	*1801	22	-	62	88	77759	368
	*4201			61	87	77629	363
9	*1001	18	-	61	87	77774	367
	*2704			60	87	77969	363
10/12**	*3301	21	3	60	87	77157	369
	*3601			60	87	77497	365
11	*1602	26	-	58	86	77482	363
	*3001			60	87	77879	362
13	*2704	13	-	60	87	77969	363
	*3101			61	87	78234	363
14	*2201	30	-	62	87	77948	363
	*3601			60	87	77497	365
15	*2101	20	-	60	87	77924	363
	*3601			60	87	77497	365
16	*4201	23	-	61	87	77629	363
	*3601			60	87	77497	365
17	*1801	26	-	62	88	77759	368
	*2201			62	87	77948	363
18	*3001	28	-	60	87	77879	362
	*2201			62	87	77948	363
19	*2801	16	-	61	87	77861	365
	*2201			62	87	77948	363
21	*0201	23	3	62	88	76867	372
	*2201			62	87	77948	363
22	*1703	18	3	60	87	77494	366
	*3301			60	87	77157	369
23	*2801	16	-	61	87	77861	365
	*2101			60	87	77924	363
24	*3301	25	3	60	87	77157	369
	*3101			61	87	78234	363
26	*2101	13	1	60	87	77924	363
	*2802			61	87	73494	385
27**	*3101	23	-	61	87	78234	363
	*3601			60	87	77497	365
28**	*1801	19	-	62	88	77759	368
	*2101			60	87	77924	363

**samples that could not be separated by TGGE.

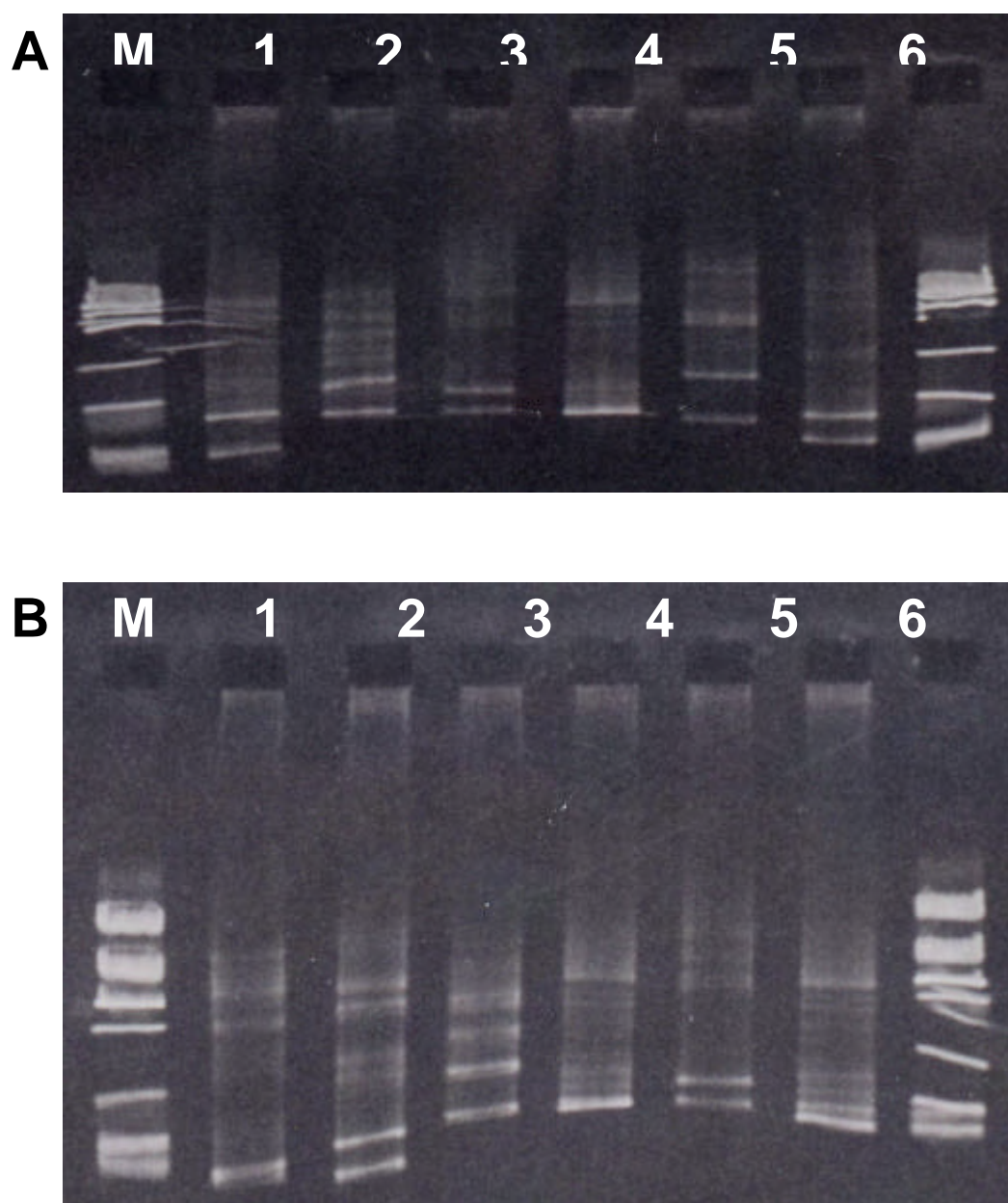


Figure 1: BoLA-DRB3 alleles separated by TGGE using 35°C - 75°C for the temperature gradient and submitted to electrophoresis during 70 minutes at 250 V. A) Gel showing separation of BoLA-DRB3 alleles in five heterozygote samples (lanes 1, 2, 3, 5, 6). B) Gel showing the same homozygote sample (lane 4) and some heterozygote samples that could not be genotyped by TGGE separation of the alleles (lanes 1, 6). Lanes designated as "M" are ϕ X174 *Hae*III, 1.5 μ g.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALDRIDGE, B.M.; MCGUIRK, S.M.; CLARK, R.J.; KNAPP, L.A.; WATKINS, D.I. & LUNN, D.P. Denaturing gradient gel electrophoresis: a rapid method for differentiating BoLA-DRB3 alleles. **Animal Genetics**, v.29, p.389, 1998.
- ANUALPEC. FNP Consultoria & Comércio **ANUALPEC 2003**. Anuário da Pecuária Brasileira. São Paulo, 2003. 400 p.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic Local Alignment Search Tool. **Journal of Molecular Biology**, v.215, p.403-410, 1990.
- ARÚS, P.; MORENO-GONZÁLES, J. Molecular markers. In: **Plant Breeding: Principles and Prospects**. Edited by M.D. Hayward, M. D.; Bosenmark N.O, Romagosa, I. (Ed.). London: Chapman & Hall. 1994. 550p.
- BAKAY, M.; CHEN, Y. W.; REHANNAH, B.; ZHAO, P.; NARAGAJU, K.; HOFFMAN, E. Sources of variability and effect of experimental approach on expression profiling data interpretation. **BMC Bioinformatics**, v.3, n.1, p.4, 2002.
- BERNOCO, D.; LEWIN, H. A. The bovine lymphocyte antigen (BoLA) system: Importance and relationship to disease in cattle. **Revista Brasileira de Genética**, v.12, p.107-122, 1989. Supplement, 3.
- BOVENHUIS, H. ; SPELMAN, R. J. **Selective genotyping to detect QTL for multiple traits in outbred populations**. (compact disk) In: VI Congress on Genetics Applied to Livestock Production, Trabalhos, Armidale, Australia. 1996.

- BRACKETT, B. G.; BOUSQUET, D.; NICE, M. L.; DONAWICK, W. J.; EVANS, J. F.; DRESSEL, M. A. Normal development following in vitro fertilization in the cow. **Biology of Reproduction**, v.27, p.147-158, 1982.
- BURKE, M.G.; STONE, R.T.; MUGGLI-COCKETT, N.E. Nucleotide-sequence and northern analysis of a bovine major histocompatibility class II DR- beta-like cDNA. **Animal Genetics**, v.22, p.343-352, 1991.
- CAMARGO A.A.; SAMAIA H.P.; DIAS-NETO E.; SIMAO D.F.; MIGOTTO I.A.; BRIONES M.R.; COSTA F.F.; NAGAI M.A.; VERJOVSKI-ALMEIDA S.; ZAGO M.A.; ANDRADE L.E. ET AL. The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome. **Proceedings of the National Academy of Sciences of the United States of America**, v.98, n.21, p.12103-8, 2001.
- CARBONELL, E. A.; GERIG, T. M.; BALANSARD, E.; ASINS, M. J. Interval mapping in the analysis of nonadditive quantitative trait loci **Biometrics**, v.48, p.305-315, 1992.
- CHUDIN, E.; WALKER, R.; KOSAKA, A.; WU, S. X.; RABERT, D.; CHANG, T. K.; KREDER, D. E. Assessment of the relationship between signal intensities and transcript concentration for Affymetrix GeneChip arrays. **Genome Biology**, v.3, n.1, e0005.1, 2001.
- DA MOTA, A. F.; MARTINEZ, M. L.; COUTINHO, L. L. Localização de locos gênicos de características quantitativas (QTLs) em gado de leite. Premissas e estratégias. **Multiciência**, v.4, n.1, pp. 18-37, 2000.
- DA MOTA, A. F.; GABRIEL, J. E.; MARTINEZ, M. L.; COUTINHO, L. L. Distribution of the bovine lymphocyte antigen (BoLA-DRB3) alleles from Brazilian Dairy Gir cattle (*Bos indicus*). **European Journal of Immunogenetics**, v.29, n.3, p.213-7, 2002.

- DARVASI, A. ; SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theoretical Applied Genetics**, v.85, p.353-359, 1992.
- DAVIES J.L.; KAWAGUCHI Y.; BENNETT S.T.; COPEMAN J.B.; CORDELL H.J.; PRITCHARD L.E.; REED P.W.; GOUGH S.C.; JENKINS S.C.; PALMER S.M. A genome-wide search for human type 1 diabetes susceptibility genes. **Nature**, v.371, n.6493, p.104-5, 1994.
- DENNIS J.A.; HEALY P.J.; BEAUDET A.L.; O'BRIEN WE. Molecular definition of bovine argininosuccinato synthetase deficiency. **Proc. Natl. Acad. Sci. USA**, v.86, p.7947-7951, 1989.
- DE SOUZA, S. J.; CAMARGO, A. A.; BRIONES, M. R.; COSTA, F. F.; NAGAI, M. A.; VERJOVSKI-ALMEIDA, S.; ZAGO, M. A.; ANDRADE, L. E.; CARRER, H.; EL-DORRY, H. F.; et al. **Proceedings of the National Academy of Sciences of the United States of America**, v.97, p.12690–12693, 2000.
- DIAS NETO E.; GARCIA CORREA R.; VERJOVSKI-ALMEIDA S.; BRIONES M.R.; NAGAI M.A.; DA SILVA W. JR.; ZAGO M.A.; BORDIN S.; COSTA F.F.; GOLDMAN G.H. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. **Proceedings of the National Academy of Sciences of the United States of America**, v. 97, n.7, p.3491-6, 2000.
- DIETZ, A. B.; DETILLEUX, J. C.; FREEMANN A. E.; KELLEY, D. H.; STABEL, J. R.; KEHRLI JR.; M. E Genetic association of bovine lymphocyte antigen DRB3 alleles with immunological traits of Holstein cattle. **Journal of Dairy Science**, v.80, p.400-405, 1997a.
- DIETZ, A. B.; COHEN, N.D.; TIMMS, L.; KEHRLI JR.; M. E. Bovine lymphocyte antigen Class II alleles as risk factors for high somatic cell counts in milk of lactating dairy cows. **Journal of Dairy Science**, v.80, p.406-412, 1997b.

- DOMINGUES, O. **Gado leiteiro para o Brasil: gado europeu, gado indiano, gado bubalino.** 4.ed. Sao Paulo: Nobel, 1974. 112p.(Biblioteca rural, 3).
- EVANS, H. M.; LONG, J. A. The effect of the anterior lobe of the hypophysis administered intraperitoneally on growth, maturity, and estrous cycles of the rat. **Anat. Res.;** v.21, p.61-63, 1921.
- FALCONER, D. S. & MACKAY **Introduction to quantitative genetics** 4.ed. Essex, England: Longman, 1996. 464 p.
- FERREIRA, M. E. ; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética.** Embrapa. 2.ed. Brasília: Embrapa-CENARGEN, 1996.
- FOX, L. K.; GERSHMAN, M.; HANCOCK, D. D.; HUTTON, C. T. Fomites and reservoirs of staphylococcus aureus causing intramammary infections as determined by phage typing: the effect of milking time hygiene practices. **Cornell Veterinary**, v.81, p.183-193, 1991.
- FUJII, J.; OTSU, K.; ZORZATO, F.; DE LEON, S.; KHANNA, V.K.; WEILER, J. E.; O'BRIEN, P.J.; MACLENNAN, D. H. Identification of a mutation in porcine ryanodine receptor associated with malignant hypothermia. **Science**, v.253, p.448-451. 1991.
- FRASER, DC.; CRAIGMILE, S.; CAMPBELL, JDM.; OLIVER, RA.; BROWN, DJ.; RUSSELL, GC.; SPOONER, RL.; GLASS, EJ. Functional expression of a cattle MHC class II DR- like antigen on mouse L-cells. **Immunogenetics**, v.43, p.296-303, 1996.
- GELDERMANN, H. Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. **Theoretical Applied Genetics**, v.46, p.319-330, 1975.

- GELHAUS, L.; SCHNITTGER, L.; MEHLITZ, D.; HORSTMANN, R.D. ; MEYER, C.G. Sequence and PCR-RFLP analysis of 14 novel BoLA-DRB3 alleles. **Animal Genetics**, v.26, p.147, 1995.
- GEORGES M.; DIETZ A.B.; MISHRA A.; NIELSEN D.; SARGEANT L.S.; SORENSEN A.; STEELE M.R.; ZHAO X.; LEIPOLD H.; WOMACK J.E.; et al. Microsatellite mapping of the gene causing weaver disease in cattle will allow the study of an associated quantitative trait Loci. **Proceedings of the National Academy of Sciences of the United States of America**, v.90, p.1058-1062, 1993.
- GEORGES M, NIELSEN D, MACKINNON M, MISHRA A, OKIMOTO R, PASQUINO AT, SARGEANT LS, SORENSEN A, STEELE MR, ZHAO X. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. **Genetics**, v.139(2), p.907-20. 1995.
- GILLIESPIE, B.E.; JAYARAO, B.M.; DOWLEN, H.H. ; OLIVER, S.P. Analysis and frequency of bovine lymphocyte antigen DRB3.2 alleles in Jersey cows. **Journal of Dairy Science**, v.82, p.2049, 1999.
- GIOVAMBATTISTA, G.; GOLIJOW, C.D.; DULOUT, F.N.; LOJO, M.M. Gene frequencies of DRB3.2 locus of Argentine Creole cattle. **Animal Genetics**, v.27, p.55, 1996.
- GROENEN, M. A. M.; VAN DER POEL, J. J.; DIJKOF, R. J. M.; GIPHART, M. J. The nucleotide sequence of bovine MHC Class II DQB and DRB genes. **Immunogenetics**, v.31, p.37-44, 1990.
- GRISART B.; COPPIETERS W.; FARNIR F.; KARIM L.; FORD C.; BERZI P.; CAMBISANO N.; MNI M.; REID S.; SIMON P.; SPELMAN R.; GEORGES M.; SNELL R.. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. **Genome Research**, v.12, n.2, p.222-31, 2002.

- GYLLESTEIN, U.B. & ALLEN, M. Sequencing of *in vitro* amplified DNA. **Methods in Enzymology**, v.18, p.3, 1993.
- HALEY, C. S. ; KNOTT, S. A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers, **Heredity**, v.69, p.315-324, 1992.
- HANAHAN, D. Studies on transformation of *Escherichia coli* with plasmids. **Journal of Molecular Biology**, v.166, p.557, 1983.
- HARMON, R. J. Physiology of mastitis and factors affecting somatic cell counts. **Journal of Dairy Science** , v.77, p.2103-2112, 1994.
- HERWIG, R.; AANSTAD, P.; CLARK, M.; LEHRACH, H. Statistical evaluation of differential expression on cDNA nylon arrays with replicated experiments. **Nucleic Acids Research**, v.29, n.23, e117, 2001.
- HILLEL, J.; DUNNINGTON, E. A.; SIEGEL, P.B. DNA markers in poultry breeding and genetic analyses. **Poultry Science Review**, v.4, p.169-186, 1992.
- HUANG X.; MADAN A. CAP3: A DNA sequence assembly program. **Genome Research**, v.9, p.868-877, 1999.
- IRITANI, A.; NIWA, K Capacitation of bull spermatozoa and fertilization in vitro of cattle follicular oocytes matured in culture. **Journal of Reproduction and Fertility**, v.50, p.119-121, 1977.
- JOHNSON, L. A.; FLOOK, J. P.; LOOK, M. V.; PIKEL, D. Flow sorting of X- and Y-bearing spermatozoa into populations. **Gamete Research**, v.16, p.1-9, 1987.

- KAPPES, S.; MILDE-LANGOSCH, K.; KRESSIN, P.; PASSLACK, B.; DOCKHORN-DWORNICZAK, B.; et al. p53 mutations in ovarian tumors, detected by temperature-gradient gel electrophoresis, direct sequencing and immunohistochemistry. **International Journal of Cancer**, v.64, p.52-59, 1995.
- KELM, S. C.; DETTILEUX, J. C.; FREEMAN, A. E.; KEHRLI JR.; M. E.; DIETZ, ^a B.; FOX, L. K.; BUTTLER, J. E.; KASCOVICS, I.; KELLEY, D. H. Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. **Journal of Dairy Science**, v.80, p.1767-1775, 1997.
- KNAPP, S. J.; BRIDGES, W. C. JR.; BIKERS, D. Mapping quantitative trait loci using molecular marker linkage maps. **Theoretical Applied Genetics**; v.79, p.583-592, 1990.
- LANDER, E. S. ; BOTSTEIN D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v.121, p.185-199, 1989.
- LEBOWITZ, R. J.; SOLLER, M.; BECKMANN, J. S. Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. **Theoretical Applied Genetics**, v.73, p.556-562, 1987.
- LEWIN, H. A. Disease resistance and immune response genes in cattle: Strategies for their detection and evidence of their existence. **Journal of Dairy Science**, v.72, p.1334-1348, 1989.
- LOOFT C.; REINSCH N.; KARALL-ALBRECHT C.; PAUL S.; BRINK M.; THOMSEN H.; BROCKMANN G.; KUHN C; SCHWERIN M.; KALM E. A mammary gland EST showing linkage disequilibrium to a milk production QTL on bovine Chromosome 14. **Mammalian Genome**, v.12, n.8, p.646-50, 2001.
- MCCARTHY, F. D.; LINDSEY, J. B.; GORE, M. T.; NOTTER, D. R. Incidence and control of sub clinical mastitis in intensively managed ewes. **Journal of Animal Science**, v.66, p.2715-2721, 1988.

- MACLENNAN, D. H. The genetic basis of malignant hyperthermia. **Trends in Pharmacological Science**, v.13, p.330-334, 1992.
- MAILLARD, J.C.; RENARD, C.; CHARDON, P.; CHANTAL, I. & BENSAID, A. Characterization of 18 new BoLA-DRB3 alleles. **Animal Genetics**, v.30, p.200, 1999.
- MARCUS E.; KEHRLI, JR. Factors affecting milk somatic cells and their role in health of the bovine mammary gland. **Journal of Dairy Science**, v.77, p.619-627, 1994.
- MARTINEZ M. L. Efeitos da substituição de alelos, da frequência gênica e da taxa de recombinação sobre o poder de detectar ligação entre um marcador genético e *loci* de característica quantitativa de baixa herdabilidade. **Revista Brasileira de Zootecnia**, v.27, n.4, p.715-721, 1998.
- MCPHERRON, A. & LEE, S. J. Double muscling in cattle due to mutations in the myostatin gene. **Proceedings of the National Academy of Sciences of the United States of America** v.94, p.12457-12461, 1997.
- MCMANUS, M. T.; SHARP, P.A. Gene silencing in mammals by small interfering RNAs. **Nature Reviews. Genetics**, v.3, p.737-747, 2002.
- MCPHERRON, A. ; LEE, S. J. Double muscling in cattle due to mutations in the myostatin gene. **Proceedings of the National Academy of Sciences of the United States of America**, v.94, p.12457-12461, 1997.
- MEISSNER N, RADKE J, HEDGES JF, WHITE M, BEHNKE M, BERTOLINO S, ABRAHAMSEN M, JUTILA MA. Serial analysis of gene expression in circulating gamma delta T cell subsets defines distinct immunoregulatory phenotypes and unexpected gene expression profiles. **Journal of Immunology**, v.170, n.1, p.356-64, 2003.
- MOODY, D. E.; POMP, D. BUCHANAN, D. S. Feasibility of the grandprogeny design for quantitative trait loci (QTL) detection in purebred beef cattle. **Journal of Animal Science**, v.75, p.941-949, 1997.

- MOK, S.C.; LO, K.W. ; TSAO, S.W. Direct cycle sequencing of mutated alleles detected by PCR single-strand conformation polymorphism analysis. **Biotechniques**, v.14, p.790, 1993.
- NG-KWAI-HANG, K.F.; H.G. MONARDES; J.F. HAYES. Association between genetic polymorphism of milk proteins and production traits during three lactations. **Journal of Dairy Science** , v.73, p.3414-3420, 1990.
- NONNECKE, B. J. , HARP, J. A. Function and regulation of Lymphocyte-Mediated responses: Relevance to bovine mastitis. **Journal of Dairy Science**, v.72, n.5, p.1313-1327, 1988.
- NOVAK, J.; SLADEK, R.; HUDSON, T. J. Characterization of variability in large-scale gene expression data: **Genomics**, v.79, n.1, p.104-113, 2001.
- OTSU, K.; PHILLIPS, M. S.; KHANNA, V.K.; DE LEON, S.; MACLENNAN, D. H. Refinement of diagnostic assays for a probable causal mutation for porcine and human malignant hyperthermia. **Genomics**, v.13, p.835-837, 1992.
- ØSTERGÅRD, H.; KRISTENSEN, B. ANDERSEN, S. Investigations in farm animals of associations between the MHC system and disease resistance and fertility. **Livestock Production Science**, v.22, p.49-67, 1989.
- PHILLIPS, P.H. Preservation of bull semen. **Journal of Biology and Chemistry**, v.130, p.415-, 1939.
- POLGE, C.; SMITH, A.; PARK, A. S. Revival of spermatozoa after vitrification and dehydration at low temperature. **Nature**, v.164, p.666, 1949.
- RUSSELL, G.C.; MARELLO, K.L.; GALLAGHER, A.; MCKEEVER, D.J.; SPOONER, R.L. Amplification and sequencing of expressed DRB second exons from *Bos indicus*. **Immunogenetics**, v.39, p.432-436, 1994.

- SAIKI R.K.; SCHARF S.; FALOONA F.; MULLIS K.B.; HORN G.T.; ERLICH H.A.; ARNHEIM N. Enzymatic amplification of b-globulin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. **Science**, v.230, p.1350-1354, 1985.
- QUACKENBUSH J.; CHO J.; LEE D.; LIANG F.; HOLT I.; et al. The TIGR gene indices, analysis of gene transcript sequences in highly sampled eukaryotic species. **Nucleic Acids Research**, v.29, p.159-164, 2001.
- RAMDAS, L.; COOMBES, K. R.; BAGGERLY, K.; ABRUZZO, L.; HIGHSMITH, W. E.; KROGMANN, T.; HAMILTON, S. R.; ZHANG, W. Sources of nonlinearity in cDNA microarray expression measurements. **Genome Biology**, v. 2, n.11, e0047.1, 2001.
- ROBL, J. M.; PRATHER, R. S.; BRANES, F. EYESTONE, W.; NORTHEY, D.; GILLIGAN, B.; FIRST, N.L. Nuclear transplantation in bovine embryos. **Journal of Animal Science**, v.64, p.642-647, 1987.
- RON M.; BAND M.; YANAI A.; WELLER J.I. Mapping quantitative trait loci with DNA microsatellites in a commercial dairy cattle population. **Animal Genetics**, v.25, p.259-264, 1994.
- SANGER, F.; THOMPSON, E. O. P.; KATAI, R. The amide groups of Insulin. **Biochemistry Journal** v.59, p.509-514, 1955.
- SANGER, F. Chemistry of Insulin. **Science**, v.129, p.1340-1344, 1959.
- SANTIAGO, A.A. **Os cruzamentos na pecuaria bovina: racas taurinas, raças zebuinas, cruzamentos**. São Paulo: Instituto de Zootecnia, 1975. 549 p.
- SAX, K. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. **Genetics**, v.8, p.552-560, 1923.
- SCHUTZ, M. M. Genetic evaluation of somatic cell scores for United States dairy cattle. **Journal of Dairy Science**, v.77, n.7, p.2113-2129, 1993.

- SCHUKKEN, Y. H.; LAM, T. J. G. M.; BARKEMA, H. W. Biological basis for selection on udder health traits. In: Proc. International workshop on genetic improvement of functional traits in cattle health. Uppsala, Sweden, June 8-10, Interbull Bulletin n° 15, p.27-32, 1997.
- SCHUKKEN, Y.H.; LESLIE, K.E.; BARNUM, D.A.; MALLARD, B.A.; LUMSDEN, J.H.; DICK, P.C.; VESSIE, G.H.; KEHRLI, M.E. Experimental *Staphylococcus aureus* intramammary challenge in late lactation dairy cows: quarter and cow effects determining the probability of infection. **Journal of Dairy Science**, v.82, p.2393-2401, 1999.
- SEYFERT, H. M.; HENKE, M.; INTERTHAL, H, *et al.* Defining candidate genes for mastitis resistance in cattle: The role of lactoferrin and lysosyme. **Journal of Animal Breeding and Genetics**, v.113, p.269-276, 1996.
- SHARIF, S.; MALLARD, B.A.; WILKIE, B.N.; SARGEANT, J.M.; SCOTT, H.M.; DEKKERS, J.C.M. ; LESLIE, K.E. Associations of the bovine major histocompatibility complex DRB3 (BoLA-DRB3) alleles with occurrence of disease and milk somatic cell score in Canadian dairy cattle. **Animal Genetics**, v.29, p.185, 1998.
- SHEIKH, F. G.; MUKHOPADHYAY, S. S.; GUPTA, P. *PstI* repeat: a family of short interspersed nucleotide element (SINE)-like sequences in the genomes of cattle, goat, and buffalo. **Genome**, p.45, p.45-50, 2002.
- SHOOK, G. E. Selection for disease resistance. **Journal of Dairy Science**, v.72, p.1349-1362, 1989.
- SHUSTER D.E.; KEHRLI M.E. JR, ACKERMANN M.R.; GILBERT R.O. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. **Proceedings of the National Academy of Sciences of the United States of America**, v.89, p.9225-9229, 1992.
- SILVER, L. M. **Mouse Genetics: Concepts and Applications**. New York: Oxford University Press. 1995. p.362.

- SIMPSON, S.P.; WEBB, A. J. Growth and carcass performance of British Landrace pigs heterozygous at the Halothane locus. **Animal Production**, v.49, p.503-509. 1989.
- SIMPSON, S. P. Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. **Theoretical Applied Genetics**, v.77, p.815-819, 1989.
- SITTE, K.; EAST, I.J.; LAVIN, M.F. & JAZWINSKA, E.C. Identification and characterization of new BoLA-DRB3 alleles by heteroduplex analysis and direct sequencing. **Animal Genetics**, v.26, p.413, 1995.
- SMITH, K. L.; TODHUNTER, D. A.; SCHOENBERGER, P.S. Environmental pathogens and intramammary infection during the dry period. **Journal of Dairy Science**, v.68, p.402-417, 1985.
- SMITH T.P.; CASAS E.; REXROAD C.E. 3RD, KAPPES S.M.; KEELE J.W. Bovine CAPN1 maps to a region of BTA29 containing a quantitative trait locus for meat tenderness. **Journal of Animal Science**, v.78, p.2589-2594, 2000.
- SMITH T.P.L, GODTEL R.A, LEE R.T PCR-based setup for high-throughput cDNA library sequencing on the ABI 3700 automated DNA sequencer. **Biotechniques**, v.29, p.698-700, 2000.
- SMITH T.P.; GROSSE W.M.; FREKING B.A.; ROBERTS A.J.; STONE R.T.; CASAS E.; WRAY J.E.; WHITE J.; CHO J.; FAHRENKRUG S.C.; BENNETT G.L.; HEATON M.P.; LAEGREID W.W.; ROHRER G.A.; CHITKO-MCKOWN C.G.; PERTEA G.; HOLT I.; KARAMYCHEVA S.; LIANG F.; QUACKENBUSH J.; KEELE J.W. Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle. **Genome Research**, v.11, n.4, p.626-30, 2001.

- SMITH T.P.L, GROSSE W.M.; FREKING B.A.; ROBERTS A.J.; STONE R.T.; et al. Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle. **Genome Research**, v.11, p.626-630, 2001.
- SOLLER, M.; GENIZI, A. The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. **Biometrics**, v.34, p.47-55, 1978.
- SOLLER, M. & BECKMANN, J. S. Genetic polymorphism in varietal identification and genetic improvement. **Theoretical Applied Genetics**, v.67, p.25-33, 1983.
- SONSTEGARD T.S.; GARRETT W.M.; ASHWELL MS.; BENNETT G.L.; KAPPES S.M.; VAN TASSELL C.P. Comparative map alignment of BTA27 and HSA4 and 8 to identify conserved segments of genome containing fat deposition QTL. **Mammalian Genome**, v.11, n.8, p.682-688, 2000.
- SONSTEGARD, T.S.; CAPUCO, A. V.; WHITE, J.; VAN TASSELL, C. P.; CONNOR, E. E.; CHO, J.; SULTANA, R.; SHADE, L.; WRAY, J. E.; WELLS, K. D.; QUACKENBUSH, J. Analysis of bovine mammary gland EST and functional annotation of the *Bos taurus* gene index. **Mammalian Genome**, v.13, n.7, p.373-9, 2002.
- SORDILLO, L. M.; SHAFER-WEAVER, K.; DEROSA, D. Immunobiology of the mammary gland. **Journal of Dairy Science**, v.80, p.1851-1865, 1997.
- STARKENBURG, R. J.; HANSEN, L. B.; KEHRLI JR.; M. E.; CHESTER-JONES, H. Frequencies and effects of alternative DRB3.2 alleles of bovine lymphocyte antigen for Holsteins in milk selection and control lines. **Journal of Dairy Science**, v.80, p.3411-3419, 1997.
- STEGER, S. Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. **Nucleic Acids Research**, v.22, p.2760-2768, 1994.

- TAKASUGA, A.; HIROTSUNE, S.; ITOH, R.; ITOHZONO, A.; SUZUKI, H.; ASO, H.; SIGIMOTO, Y. Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs. **Nucleic Acids Research**, v. 29, n.22, e108, 2001.
- VAN EIJK, M.J.T.; STEWART-HAYNES, J.A. & LEWIN, H.A. Extensive polymorphism of the BoLA-DRB3 gene distinguished by PCR-RFLP. **Animal Genetics**, v.23, p.483, 1992.
- VEIGA, V.M. O.; TEIXEIRA, M.; BRITO, M. A. V.P.; BRITO, J. R. F. **Controle da mamite dos bovinos**. Empresa Brasileira de Pesquisa Agropecuária-Embrapa, Centro Nacional de Pesquisa de Gado de Leite. Juiz de Fora. MG. 1994. 70 p.
- VENTER J.C.; ADAMS M.D.; MYERS E.W.; LIP.W.; MURAL R.J.; SUTTON G.G.; SMITH H.O.; YANDELL M.; EVANS C.A.; HOLT R.A.; et al. The Sequence of the Human Genome. **Science**, v.291, p.1304-1351, 2001.
- VISSCHER, P.M.; ; HALEY, C. S. Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. **Theoretical Applied Genetics**, v.93, p.691-702.; 1996.
- WEIR, B. S. **Genetic Data Analysis II**. 2. Sunderland, MA: Sinauer Associates. 1996. 445 p.
- WELLER, J. I. Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. **Biometrics**, v.42, p.627-640, 1986.
- WELLER, J. I.; KASHI, Y.; SOLLER, M. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. **Journal of Dairy Science**, v.73, p.2525-2537, 1990.
- WILKINS, R. J. & KUYS, Y. M. Rapid beta-lactoglobulin genotyping of cattle using the polymerase chain reaction. **Animal Genetics**, v.23, p.175-178, 1992.

WINTER A.; KRAMER W.; WERNER F.A.; KOLLERS S.; KATA S.; DURSTEWITZ G.; BUITKAMP J.; WOMACK J.E.; THALLER G.; FRIES R. Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. **Proceedings of the National Academy of Sciences of the United States of America**, v.99, n.14, 9300-5, 2002.

ZADWORNYY, D. & KUHNLEIN, U. The identification of the kappa-casein genotype in the Holstein dairy cattle using the polymerase chain reaction. **Theoretical and Applied Genetics**, v.80, p.631-640, 1990.