

II Curso de Técnicas de Biologia Molecular aplicadas à produção animal



28 de junho a 3 de Julho de 1999

São Carlos, SP

Comissão Organizadora:

Luciana Correia de Almeida Regitano (CPPSE/EMBRAPA)

Luiz Lehmann Coutinho (ESALQ/USP)

Pedro Franklin Barbosa (CPPSE/EMBRAPA)

Colaboração:

Mario Luiz Martinez (CNPGL/Embrapa)

Isabel de Miranda Santos (Cenargen/Embrapa)

Lúcia Elvira Álvares (ESALQ/USP)

Jane Eyre Gabriel (ESALQ/USP)

Luciana Andrea Ribeiro (ESALQ/USP)

Ana Paula Figueira Serrão Semenssato (CPGGEv/UFSCar)

Andréa Pozzi Pereira (CB/UFSCar)

Daniela DeBenedetti Tambasco (CPGGEv/UFSCar)

Érika Cristina Jorge (CB/UFSCar)

Luciana Pimentel Mello Klocker de Vasconcelos (CPGGEv/UFSCar)

Marina DeBenedetti Tambasco (CPGGEv/UFSCar)

Agradecimentos:

Andréa Shibata de Barros (Ilustrações e revisão)

Maria Cristina Campanelli (Editoração)

Apoio:



Amershan Pharmacia Biotech

Índice

	Páginas
Uso de marcadores moleculares na indústria animal.....	04
<i>Luiz Lehmann Coutinho</i>	
Introdução ao uso de marcadores moleculares.....	16
<i>Luciana Correia de Almeida Regitano</i>	
Polimorfismo do sistema imune e sua aplicação em melhoramento para resistência a doenças.....	28
<i>Isabel de Miranda Santos</i>	
Métodos estatísticos para identificação de QTLs em animais.....	45
<i>Mario Luiz Martinez</i>	
Estudo do desenvolvimento muscular em aves.....	98
<i>Luiz Lehmann Coutinho</i>	
RT-PCR no estudo da expressão gênica.....	107
<i>Lúcia Elvira Alvares</i>	
Técnicas empregadas no estudo da expressão gênica.....	111
<i>Jane Eyre Gabriel</i>	
Animais Transgênicos.....	115
<i>Luciana de A. Ribeiro</i>	
 Manual de Laboratório	
Protocolo de extração de DNA de sangue fresco	123
Extração de DNA de sêmem.....	126
Protocolo de amplificação de CSN3.....	127
Digestão do DNA – CSN3.....	128
Protocolo de amplificação de β -Lactoglobulina.....	130
Digestão do DNA - β -Lacto.....	131
Eletroforese em gel de agarose.....	133
Protocolo de RAPD.....	135
Protocolo de amplificação do microssatélite CSFM 50.....	137
Protocolo de amplificação de Texan15.....	139
Preparação das placas para gel de poliacrilamida.....	141
Gel de poliacrilamida desnaturante para microssatélites.....	142
Gel de poliacrilamida não desnaturante para microssatélites.....	143
Eletroforese em gel de poliacrilamida	144
Protocolo de coloração com prata.....	145
Extração de RNA com Trizol LS.....	147
RT-PCR (kit Gibco).....	149
Protocolo experimental da técnica de hibridização “in situ”.....	151

USO DE MARCADORES MOLECULARES NA INDÚSTRIA ANIMAL

Luiz Lehmann Coutinho

Departamento de Zootecnia
Escola Superior de Agricultura Luiz de Queiroz
Universidade de São Paulo

O uso de marcadores moleculares no melhoramento animal vem recebendo cada vez mais atenção, tanto por pesquisadores como por empresas e produtores. Um grande esforço vem sendo concentrado no desenvolvimento de mapas genômicos para suínos, bovinos, ovinos e aves. Alguns exemplos de marcadores já disponíveis ou muito próximos de aplicação comercial são: kappa-caseína, beta-lactoglobulina, associados à composição do leite; gene fec de ovinos, associado à alta taxa de ovulação; receptor de estrogênio em suínos, associado à alta prolificidade; receptor de rianodina, associado a alta capacidade de desenvolvimento muscular; musculatura dupla em bovinos. Existem também alguns exemplos de marcadores moleculares para doenças em animais: BLAD, deficiência de adesão de leucócitos em bovinos, Weaver, doença hereditária observada em bovinos da raça Suiça Parda. O contínuo avanço da pesquisa na área de marcadores moleculares deverá proporcionar um número crescente de marcadores moleculares que poderão ser utilizados em programas de melhoramento animal.

1. Introdução

Avanços bastante significativos vêm sendo obtidos com técnicas tradicionais de melhoramento. Para citar alguns exemplos, em 1968 um frango de corte necessitava de 72 dias para chegar ao peso de abate (2,037 kg) e isto ocorria com uma conversão alimentar de 2,338. Nos últimos anos, tem sido observada redução de um dia por ano para que frangos de corte cheguem ao peso de abate. Em 1994, vários produtores conseguiam que seus animais obtivessem peso de abate de 1,992 kg em apenas 42 dias, com uma conversão alimentar de 1,77. Na suinocultura, um animal de 1958-59 possuía carcaça com 3,05 cm de gordura subcutânea e 29,3 cm² de área de lombo

(Clausen et al., 1970), já um suíno moderno possui 1,2 cm de gordura subcutânea e 35,2 cm² de área de lombo. Na bovinocultura, o avanço na produção de leite também tem sido fantástico nos países desenvolvidos. Na última década, tem-se observado um aumento na produção entre 1 e 2% ao ano (Wiggans et al., 1988). A biotecnologia vem trazer a possibilidade de utilização de novas ferramentas no processo de melhoramento animal. Aliada às metodologias tradicionais, as novas técnicas deverão aumentar ainda mais o progresso genético que vem sendo observado nos animais domésticos. O uso de marcadores moleculares, principalmente de DNA, permite que seja determinado, com maior precisão, o potencial genético de um animal independente do ambiente e antes que o seu fenótipo seja expressado. Em outras palavras, pode-se determinar o potencial genético de um embrião sem que seja necessário avaliar a sua produção ou de sua progênie.

2. Marcadores moleculares para detecção de genótipos superiores

Como descrito por Cheng (1994), a sexagem de frangos através de empenamento diferenciado usa o mesmo princípio de marcadores moleculares. O empenamento lento de frangos está associado ao sexo feminino e desta forma, produtores podem utilizar o fenótipo de empenamento para determinar o sexo do animal logo após o nascimento dos pintinhos. Da mesma maneira, marcadores de DNA associados com características desejáveis podem ser utilizados para seleção. Como será discutido mais adiante, o alelo B da kappa caseína está associado com maior concentração de proteína no leite. Portanto, a seleção de animais que possuam este alelo, resulta em leite de maior valor econômico para a produção de queijo. A vantagem é que o uso de marcadores de DNA permite que se verifique se o animal vai sintetizar a forma B da kappa-caseína antes que este animal inicie a produção de leite. Ainda mais, a análise do DNA permitirá saber se o animal é homozigoto ou heterozigoto para o gene em questão. A possibilidade de se determinar se o gene está presente em homozigose ou heterozigose é fundamental na presença de ação gênica dominante. Nestes casos, um animal heterozigoto apresenta o mesmo fenótipo que um animal homozigoto dominante, e a presença do alelo recessivo só pode ser detectada com um estudo da progênie, ou através de teste de DNA.

A Figura 1 ilustra como marcadores moleculares podem ser selecionados e utilizados em melhoramento animal. Neste exemplo, um gene que influencia o tamanho da leitegada em suínos está presente no cromossomo em uma região desconhecida "C". Vamos assumir que existam duas regiões que apresentem polimorfismo neste cromossomo. Uma próxima, marcador A e outra mais distante, marcador B. Vamos assumir ainda que existam duas linhagens de suínos, uma de maior prolificidade e uma de menor prolificidade. Ao promover o acasalamento destas duas linhagens vamos obter animais heterozigotos para o gene de prolificidade, e ao promovermos o retrocruzamento dos animais heterozigotos com a linhagem menos prolífica produziremos animais com diferentes genótipos. Ao seguirmos a segregação do marcador com a característica prolificidade vamos observar que o marcador B está presente tanto em animais com prolificidade média quanto em animais de prolificidade baixa. Por outro lado, o marcador A, por estar fisicamente próximo ao gene "C", segrega associado a este gene, mesmo no caso de animais em que ocorreu recombinação. Desta forma, podemos selecionar animais com maior leitegada com base no marcador A, mesmo sem saber qual o gene que controla esta característica fenotípica.

Para o desenvolvimento de marcadores moleculares, duas estratégias vêm sendo utilizadas. Uma de certa forma aleatória e outra dirigida. A forma aleatória tem como objetivo inicial o desenvolvimento de um mapa genômico para a espécie em questão. Nos animais domésticos este mapa está baseado principalmente em regiões microssatélites. Nesta estratégia, marcadores estão sendo desenvolvidos para cobrir todo o genoma, para que em uma segunda fase sejam conduzidos estudos de correlação entre estes marcadores e características desejáveis. A segunda estratégia se baseia na detecção de polimorfismo associados a genes sabidamente importantes para a característica que se pretende estudar. Alguns genes que vem sendo investigados são hormônios de crescimento, prolactina, kappa-caseína, beta-lactoglobulina, gene do halotano, e outros potencialmente importantes.

Figura 1. Segregação de marcadores moleculares em retrocruzamento, com recombinação na região II.

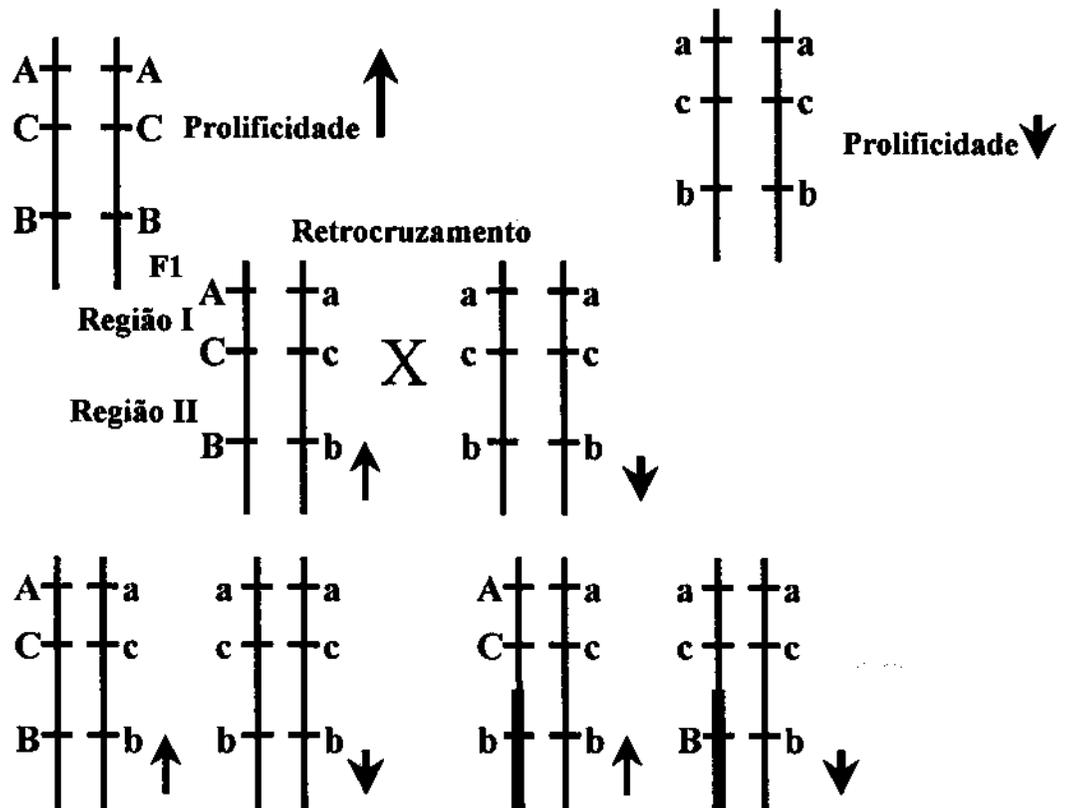


Figura 1. Segregação de marcadores moleculares em retrocruzamentos, Com recombinação na região II.

3. Exemplos de aplicação de marcadores moleculares em melhoramento animal

1) Kappa-caseína e beta-lactoglobulina: O leite de animais da raça holandesa tem em média uma composição de 3,56% de gordura, 2,49% de caseína, 0,53% de proteínas de soro, 4,61% de lactose e 0,73% de cinzas (Jenness, 1985). Entre as proteínas, encontram-se principalmente as caseínas (alpha, beta e kappa), cada uma com pelo menos duas variantes genéticas. Entre as proteínas do soro encontram-se principalmente albumina e beta-lactoglobulina, cada uma também com pelo menos duas variantes genéticas (Eigel et al., 1984).

A importância das proteínas do leite, em relação a sua industrialização (produção de queijo), ligada à presença de diversas variantes genéticas, levou vários

pesquisadores a investigar se existe alguma correlação entre os genótipos existentes e características econômicas de produção (Lin & Lee, 1992). O gene da kappa-caseína está presente em duas formas alélicas (A e B). Em gado holandês, a frequência do alelo A varia entre 0,68 e 0,80 do alelo B entre 0,32 e 0,20 (Lin et al., 1986, McLean, et al., 1984, Ng-Kwai-Hang et al., 1984 e Gonyon et al., 1987). Estudos preliminares conduzidos em rebanhos brasileiros de gado holandês (89 animais) observaram frequências de 0,82 para o alelo A e 0,18 para o alelo B de kappa-caseína (Kemenes & Coutinho, 1995). Por outro lado, estudos também iniciais com gado nelore (45 animais) revelaram uma frequência de 0,93 para o alelo A e 0,07 para o alelo B (Kemenes & Coutinho, 1995). A presença de kappa-caseína B está relacionada com maior produção de queijo (Mariani et al., 1976), talvez porque a presença de kappa-caseína B no leite resulte em coagulação mais rápida e maior firmeza do coágulo (Rampilli et al., 1988 e El-Negoumy, 1972). O gene de kappa-caseína parece afetar também a concentração de proteína no leite. Em um estudo com 6803 novilhas da raça holandesa, Bovenhuis et al., (1992) demonstraram que animais com o genótipo BB para kappa-caseína produziam leite com maior concentração de proteína do que animais AA, embora tenha sido observada uma redução na produção total de leite.

O gene de beta-lactoglobulina também está presente em duas formas alélicas (A e B). A frequência do alelo A é de 0,38 e do alelo B de 0,62 (Lin et al., 1986, McLean, et al., 1984, Ng-Kwai-Hang et al., 1984 e Gonyon et al., 1987). A presença do genótipo BB para beta-lactoglobulina apresentou, em estudo conduzido por Bovenhuis et al. (1992), um efeito positivo significativo na concentração de gordura no leite. No entanto, este genótipo apresentou também menor concentração de proteína menor e redução na produção total de leite.

Com base nos dados discutidos acima, alguns pesquisadores recomendam, quando houver uma remuneração melhor para leite com concentração superior de proteína e gordura, a seleção de animais com os genótipos BB para kappa-caseína e beta-lactoglobulina. DeNise (1993, comunicação pessoal) recomenda, quando da seleção de touros, que os genótipos para kappa-caseína e beta-lactoglobulina também sejam considerados. Finalmente, existe a possibilidade de seleção em gado de corte, de animais que produzam leite com maior concentração de proteína e gordura.

2) Gene Fec dos ovinos: A alta taxa de ovulação encontrada na raça de ovinos australianos Booroola Merino é atribuída a um único gene, o gene Fec. Dois alelos codominantes são conhecidos, o alelo normal, Fec+, e o alelo encontrado na raça Booroola, FecB. Animais homozigotos Fec+ apresentam taxa média de ovulação de 1,5. Nos heterozigotos Fec+/FecB a média é de 2,9 e nos homozigotos FecB essa taxa é de 4,7.

A identificação de fêmeas FecB só pode ser feita através da medida da taxa de ovulação por endoscopia, o que dificulta bastante o trabalho de introgressão desse gene em rebanhos melhorados. Recentes progressos têm sido alcançados na obtenção de marcadores para o loco Fec. Lanneluc et al. (1994) utilizando sondas multilocais, identificaram um grupo de sete fragmentos de DNA minissatélite ligados ao gene Fec. A aplicação desses marcadores é ainda limitada, uma vez que a distância mínima estimada com relação ao gene foi de 15 cM. Entretanto, esses marcadores representam um ponto de partida para a obtenção de marcadores mais próximos ao gene Fec.

3) Receptor de estrogênio: A influência do polimorfismo no gene do receptor de estrogênio (ESR) em características reprodutivas tem sido demonstrada em suínos. Short & McLaren (1995) encontraram associação entre a variação do tamanho da leitegada e polimorfismo do gene ESR em diferentes raças de suínos.

Experimentos com linhagens da raça Chinese Meisham demonstraram que uma forma do gene ESR, encontrada exclusivamente nessa raça, aumenta em 0,8 a 1 leitão por leitegada para cada cópia desse alelo. Assim, indivíduos homozigotos apresentam acréscimo de 1,6 a 2 leitões por leitegada.

Esse polimorfismo está sendo utilizado em programas de melhoramento de suínos da Pig Improvement Company (PIC) com o objetivo de desenvolver linhagens prolíficas através da introgressão do alelo da raça Meishan.

4) Receptor de ryanodina (gene de halotano): Em várias raças de suínos, mas principalmente na raça Landrace, existe uma mutação que causa a síndrome de estresse em suínos (PSS). Esta mutação resulta da substituição de uma única base (C por T) em um gene que controla a concentração de cálcio em células musculares. Animais homozigotos para esta mutação são muito sensíveis a situações de estresse e de difícil manejo. Estes animais apresentam, no entanto, desenvolvimento muscular

superior a animais normais, porém são mais suscetíveis a problemas de qualidade de carne como PSE (carne de coloração pálida, macia e com perda de água). Por outro lado, animais heterozigotos apresentam um fenótipo intermediário, com desenvolvimento muscular superior aos animais normais e susceptibilidade a estresse inferior a animais homozigotos para o gene do halotano.

No passado, esta mutação era detectada pelo uso do gás halotano, visto que animais homozigotos para a mutação apresentavam reação ao anestésico. Infelizmente, este teste não consegue identificar animais heterozigotos, dificultando a erradicação ou o uso controlado desta mutação. Recentemente, Fujii et al. (1991) desenvolveram um teste para a identificação desta mutação. O teste baseia-se na amplificação por PCR (reação de polimerase em cadeia) da região do gene do receptor de ryanodina que pode ou não conter a mutação. Esta região é então digerida com uma enzima de restrição que reconhece a seqüência de bases presente em animais normais e corta o gene amplificado em dois fragmentos. Caso a região amplificado do gene não seja digerida pela enzima, isto indicaria a presença da mutação em homozigose. Por outro lado, a digestão de apenas metade do produto amplificado indica a presença de heterozigose. Este teste possibilita a identificação de animais normais, homozigotos recessivos e heterozigotos, permitindo assim um controle total da presença desta mutação em linhagens de suínos utilizadas em programas de melhoramento ou destinados para o abate.

5) Detecção de doenças:

5.1) BLAD: Recentemente, pesquisadores dos Estados Unidos (Shuster et al., 1992) identificaram uma doença genética em bovinos que causa deficiência na adesão de leucócitos (BLAD). Esta doença é transmitida geneticamente e só ocorre em animais homozigotos recessivos. Animais com esta doença apresentam redução na expressão de beta-2 integrina, resultando em leucócitos anormais que não conseguem penetrar nos tecidos para destruir os agentes patogênicos. Animais homozigotos para esta doença raramente sobrevivem, causando perdas significativas para produtores. A freqüência do gene causador desta doença foi determinada no rebanho de gado holandês norte americano e é de 15% entre os touros e de 6% entre as vacas. Todos

os animais portadores desta doença são relacionados com um único touro, Osborndale Ivanhoe. Como este touro apresenta boas características de produção, o seu sêmen tem sido utilizado em vários países, inclusive o Brasil. Calcula-se que esta doença cause perdas em torno de US\$ 5 milhões por ano nos Estados Unidos. Usando os primers e as condições descritas por Shuster et al. (1992) um fragmento de DNA de 58 pares de bases é amplificado. Este fragmento contém o nucleotídeo 383 do gene que codifica a proteína CD18 em bovinos da raça holandesa. É nesta posição que ocorre a substituição de uma adenina por guanina em animais portadores da doença. Esta substituição causa uma mudança no codon e resulta na substituição de um ácido aspártico por uma glicina no DNA dos animais mutantes. Essa mutação pode ser detectada no DNA amplificado através da digestão com diferentes enzimas de restrição. A digestão do fragmento amplificado com a enzima Taq I resulta em fragmentos de 26 e 32 pares de bases em animais normais. Em animais com a doença (BLAD), o DNA não é digerido por Taq I. Por outro lado, em animais portadores da doença, fragmentos de 58, 32 e 26 pares de bases são observados. Atualmente as empresas de inseminação artificial nos Estados Unidos testam rotineiramente seus reprodutores para esta doença.

5.2 Weaver: Essa doença hereditária é encontrada quase que exclusivamente em bovinos da raça Suíça Parda e caracteriza-se por paresia progressiva dos membros pélvicos e ataxia. A doença foi atribuída a um único gene recessivo de penetrância completa.

O interesse pelo estudo desse gene foi despertado pelo aumento da frequência da mutação em rebanhos melhorados, que poderia ser resultante de uma vantagem seletiva dos indivíduos portadores da mutação. Essa teoria foi reforçada por Hoeschele & Meinert (1990) ao descreverem uma associação entre a doença e produção de leite. A diferença estimada entre indivíduos portadores e não portadores foi de 690 kg de leite e 26,2 kg de gordura. A magnitude do efeito sobre a produção sugere que a mutação deve estar próxima a um gene com grande efeito na produção de leite. Dessa forma, a identificação de indivíduos portadores é essencial para evitar que, através da seleção para produção, o aumento na frequência do gene coloque em risco a sobrevivência da raça.

Recentemente, um marcador de DNA microssatélite foi mapeado muito próximo ao gene responsável pela doença (Georges et al., 1993). Esse marcador deverá permitir a identificação de indivíduos portadores, sem a necessidade de teste de progênie. Além disso, esse marcador deverá contribuir para o mapeamento da região cromossômica responsável pelo acréscimo na produção de leite, que poderá ser realizado mesmo em raças que não apresentam a doença Weaver.

4. Mapeamento genômico de animais domésticos

Vários grupos estão empenhados no desenvolvimento de mapas genômicos para suínos, aves e bovinos. O objetivo é desenvolver marcadores moleculares para todos os cromossomos e a uma distância média de 10-20 cM. Várias estratégias para a geração de marcadores moleculares vêm sendo empregadas. Entre elas isoenzimas, RFLP (polimorfismo de tamanho gerado pelo uso de enzimas de restrição), DNA fingerprinting com sondas multilocais, RAPD e microssatélites. O uso de isoenzimas, bastante comum no passado, tem as desvantagens de não possuir um alto grau de polimorfismo e de ser dependente, em alguns casos, do estágio de desenvolvimento do animal. A técnica de RFLP já apresenta a vantagem de poder ser utilizada mesmo com DNA extraído de células embrionárias; no entanto é trabalhosa e tem um grau de polimorfismo limitado. O uso de sondas multilocais apresenta um alto grau de polimorfismo, no entanto é difícil a determinação precisa dos alelos. O uso de RAPD, também pode apresentar alto grau de polimorfismo, no entanto a detecção de heterozigotos não é possível e o grau de repetibilidade desta técnica é baixo. A amplificação de regiões microssatélites tem sido a maneira preferida para produção de marcadores moleculares. Recentemente, Bishop et al. (1994) publicaram um mapa genético de bovinos com 306 locos marcadores, entre os quais 290 são microssatélites. Devido ao uso da técnica de PCR, a quantidade de amostra necessária é mínima, o grau de polimorfismo é elevado e apenas um loco é amplificado. Para as diferentes espécies citadas acima, existe um número razoável de marcadores disponíveis e espera-se que nos próximos anos todos os cromossomos estejam saturados. Atualmente, mesmo sem um mapa genômico completo, vários pesquisadores já estão estudando a associação de marcadores moleculares com características de

composição de carcaça e crescimento de bovinos de corte (Davis, 1994, comunicação pessoal), produção de leite (Massey & Georges, 1992), e resistência a doenças em aves (Cheng, 1994).

5. Aplicações futuras

Não há dúvida que marcadores moleculares terão um papel fundamental no melhoramento animal. Estes vêm sendo utilizados e terão uma utilização ainda maior para detecção de doenças e de genótipos superiores em animais. Os marcadores poderão também ser utilizados em esquemas de introgressão de genes, através do processo de velogenetics. Como citado por Massey (1992), neste procedimento, fêmeas superiores são superovuladas, inseminadas com sêmen de touros superiores e os embriões transferidos para animais receptores. Estes embriões podem então ser sacrificados entre 90 e 180 dias de desenvolvimento para a retirada dos óvulos que já estejam formados. Estes óvulos são então "amadurecidos" em cultura e fertilizados in vitro. Estes embriões podem então ser classificados de acordo com o seu genótipo (através de marcadores moleculares) e apenas os desejáveis transferidos para animais recipientes. Desta forma, o intervalo entre gerações será reduzido e a velocidade do progresso genético será superior.

6. Considerações finais

O investimento de recursos para o desenvolvimento de marcadores moleculares nos Estados Unidos e Europa, apesar de considerado insatisfatório por pesquisadores destes países, é significativo. Não há dúvida que nos próximos cinco anos deverão ser identificados vários marcadores associados com ETL (locos de interesse econômico). No Brasil, o número de pesquisadores trabalhando nesta área é extremamente pequeno. Tal fato se deve não só à falta de recursos na área, mas também à falta de pesquisadores com treinamento nas duas áreas fundamentais para o desenvolvimento destes estudos, ou seja biologia molecular e genética quantitativa. É importante, portanto, que programas sejam desenvolvidos para capacitação de pesquisadores e desenvolvimento de pesquisa nestas áreas.

Referências Bibliográficas

- BISHOP, M.D., KAPPES, S.M., KEELE, J.W., STONE, R.T., SUNDEN, S.L.F., HAWKINS, G.A., TOLDO, S.S., FRIES, R., GROZS, M.D., YOO, J., BEATTIE, C.W. A genetic linkage map for cattle. *Genetics*, 136:619, 1994.
- BOVENHUIS, H., VAN ARENDONK, J.A.M., KORVER, S. Association between milk protein polymorphisms and milk production traits. *J. Dairy Sci.*, 75:2549, 1992.
- CHENG, H.H. The chicken genetic map: a tool for the future, *Poultry Digest*, June, 24-28, 1994.
- CLAUSEN, A. Annual report, Department of Animal Science, National Agricultural Research Laboratory, Copenhagen, Denmark, 1970.
- EIGEL, W.N., BUTLER, J.E., ERNSTROM, C.A., FARREL JR., H.M., HARKALKAR, V.R., JENNESS, R., WHITNEY, R.M. Nomenclature of proteins of cow's milk: fifth edition. *J. Dairy Sci.*, 67:1599-1631, 1984.
- EL-NEGOUMY, A.M. Effect of polymorphic composition of calcium caseinate sols on their stability to rennin. *J Dairy Research*, v.39, p.373-379, 1972.
- FUJII, J., OTSU, K., ZORZATO, F., DE LEON, S., KHANNA, V.K., WEILER, J.E., O'BRIEN, P.J., MACLENNAN, D.H. *Science*, v.253, p.448-451, 1991.
- GEORGES, M., DIETZ, A.B., MISHRA, A., NIELSEN, D., SARGEANT, L.S., SORENSEN, A., STEELE, M.R., ZHAO, X., LEIPOLD, H., WOMACK, J.E., LATHROP, M. Microsatellite mapping of the gene causing weaver disease in cattle will allow the study of an associated quantitative trait locus. *Proc. Natl. Acad. Sci. USA*, v.90, 1058p., 1993.
- GONYON, D.S., MATHER, R.E., HINES, H.C., HAENLEIN, G.F.W., ARAVE C.W., GAUNT, S.N. Association of bovine blood and milk polymorphisms with lactation traits: Holstein. *J. Dairy Sci.*, v.70, p.2585, 1987.
- HOESCHELE, I., MEINERT, T.R. Association of genetic defects with yield and type traits: The weaver locus effect on yield. *J. Dairy Sci.*, v.73, p.2503, 1990.
- JENNESS, R. Biochemical and nutritional aspects of milk and colostrum in lactation, editado por Larson, B.L., Ames Iowa, USA, Iowa State University Press, 1985.
- LANNELUC, I., DRINKWATER, R.D., ELSER, J.M., HETZEL, D.J.S., NGUYEN, T.C., PIPER, L.R., THIMONIER, J., HARRISON, B., GELLIN, J. Genetic markers for the Booroola fecundity (Fec) gene in sheep. *Mammalian Genome*, v.5, 26p., 1994.

- LIN C.Y., MCALLISTER, A.J., NG-KWAI-HANG, K.F., HAYES, J.F. Effects of milk protein loci on first lactation production in dairy cattle. *J. Dairy Sci.*, 74(Suppl. 1):284, 1986.
- LIN, C.Y, LEE, A.J. Direct typing fo milk proteins as an aid for genetic improvement of dairy bulls and cows: A review. *Animal Breeding Abstracts*, v.60, n.1, p.1-10, 1992.
- MARIANI, P., LOSI, G., RUSSO, V., CASTAGNETTI, G.B., GRAZIA, L., MORINI, D., FOSSA, A. Caseification tests made with milk characterized by variants A and B of kappa-casein in the production of Parmigiano-Reggiano cheese. *Scienza e Tecnica Lattiero-Casearia*, v.27, p.208-227, 1976.
- MASSEY, J.M., GEORGE, M. Genmark's approach to marker assisted selection. *Animal Biotechnology*, v.3, n.1, p.95-109, 1992.
- MCLEAN, D.M., GRAHAM, E.R.B., PONZONI, R.W., MCKENZI, H.A. Effects of milk protein genetic variants on milk yield and composition. *J. Dairy Research*, v.51, p.531-546, 1984.
- NG-KWAI-HANG, K.F., HAYES, J.F., MOXLEY, J.E., MONARDES, H.G. Association of genetic variants of casein and milk serum proteins with milk fat and protein production in dairy cattle. *J. Dairy Sci.*, v.67, p.835, 1984.
- RAMPILLI, M., CAROLI, A., BOLLA, P., PIRLO, G. Relazioni tra genotip lattoproteici, composizione caseinica e attitudine alla coagulazione del latte nel corso della lattazione. *Scienza e Tecnica Lattiero-Casearia*, v.39, p.262-279, 1988.
- SHORT, T.H., McLAREN, D.G. Hal-1843, the ESR gene test and beyond: the future of marker assisted selection in pig breeding. *American Association of Swine Practitioners*: 133, 1995.
- SHUSTER, D.E., MARCUS, E., KEHRLI, JR., ACKERMANN, M.R., GILBERT, O.R. Identification and prevalence of a genetic defect that causes leukocyte adhesion deciciency in holstein cattle, *Proc. Natl. Acad. Sci., USA*, v.89, p.9225-9229, 1992.
- WIGGANS, G.R., MISTZAL, I., VAN VLEK, L.D. Implementation of an animal model for genetic evaluation of dairy cattle in the United States, *J. Dairy Sci.*, 71:(Suppl. 2) 54-69, 1988.

INTRODUÇÃO AO USO DE MARCADORES MOLECULARES

Luciana Correia de Almeida Regitano

Embrapa-CPPSE

Introdução

Marcadores genéticos são características de herança mendeliana simples que permitem a inferência do genótipo a partir do fenótipo do indivíduo, fornecendo informações importantes para a análise genética de uma espécie. Para a comparação entre populações, os marcadores devem ser neutros com relação a adaptação e a sobrevivência do animal, preferencialmente codominantes e estáveis ao longo do desenvolvimento.

Os primeiros marcadores disponíveis foram as mutações que produziam alterações morfológicas como, por exemplo, o nanismo, a ausência de asas em *Drosophila* e a ausência de pelos em camundongos. Entretanto, tais mutações são pouco freqüentes nas populações naturais, nas quais a maior parte da variação genética é de caráter contínuo, e freqüentemente comprometem a adaptação do indivíduo (Tanksley, 1993).

A partir da década de 30, o polimorfismo de antígenos eritrocitários passou a ser investigado, fornecendo uma nova classe de marcadores em animais. A análise de marcadores foi ampliada com o desenvolvimento de técnicas de eletroforese de proteínas associadas a métodos de coloração histoquímica, que permitiram que a variação genética das isoenzimas passasse a ser estudada. Esses marcadores têm sido intensamente utilizados para investigar a estrutura das populações, elucidando questões como fluxo gênico em populações naturais, dispersão e filogenia. Apesar disso, apenas uma parte da variação genética pode ser observada, aquela resultante de mutações que ocorrem em regiões codificadoras do genoma e resultam em alteração detectável do produto gênico. Além disso, as isoenzimas podem sofrer variações de acordo com a idade e o tecido utilizado.

Essa limitação foi superada com o desenvolvimento das técnicas de análise de DNA e o conseqüente desenvolvimento de diversos tipos de marcadores moleculares, que serão discutidos a seguir.

Polimorfismo de comprimento de fragmentos de restrição

A caracterização do polimorfismo de DNA tornou-se possível com a descoberta das endonucleases de restrição da Classe II, capazes de cortar a molécula de DNA em sítios específicos, denominados sítios de restrição. O número de cortes efetuados por uma determinada enzima de restrição é função do número de sítios de restrição presentes ao longo da molécula de DNA. A ocorrência de variação individual no número e tamanho dos fragmentos formados pela digestão do DNA com uma enzima de restrição foi demonstrada por Grodzicker *et al.* (1974), em adenovírus. Essa variação é o resultado de mutações que eliminam ou criam sítios de restrição para uma determinada enzima e foi denominada Polimorfismo de Comprimento de Fragmentos de Restrição ou RFLP.

O método desenvolvido para adenovírus não produziu resultados satisfatórios em eucariotos superiores, em função do tamanho do genoma. Nessas espécies, o número de fragmentos resultantes da digestão com enzimas de restrição é muito grande. Consequentemente, tentativas de separá-los por tamanho em gel de eletroforese produziram um rastro ao invés de bandas discretas. A solução para esse problema surgiu com o desenvolvimento da técnica de Southern blot (Southern, 1975), na qual os fragmentos separados são imobilizados em um suporte sólido e desnaturados. Explorando a tendência da molécula de DNA de formar duplas-fitas complementares, Southern utilizou uma seqüência de DNA fita-simples, marcada com ^{32}P , como sonda. Como resultado, apenas os fragmentos complementares à seqüência utilizada como sonda foram revelados (Figura 1).

Estas sondas de DNA podem ser de regiões codificadoras ou não. O emprego de sonda de DNA complementar ao RNA mensageiro (cDNA) é o mais comum e apresenta a vantagem de não conter seqüências repetitivas. As sondas anônimas, obtidas a partir de fragmentos de DNA ao acaso, possibilitam a rápida obtenção de marcadores para construção de mapas de ligação e permitem a análise de regiões não codificadoras do

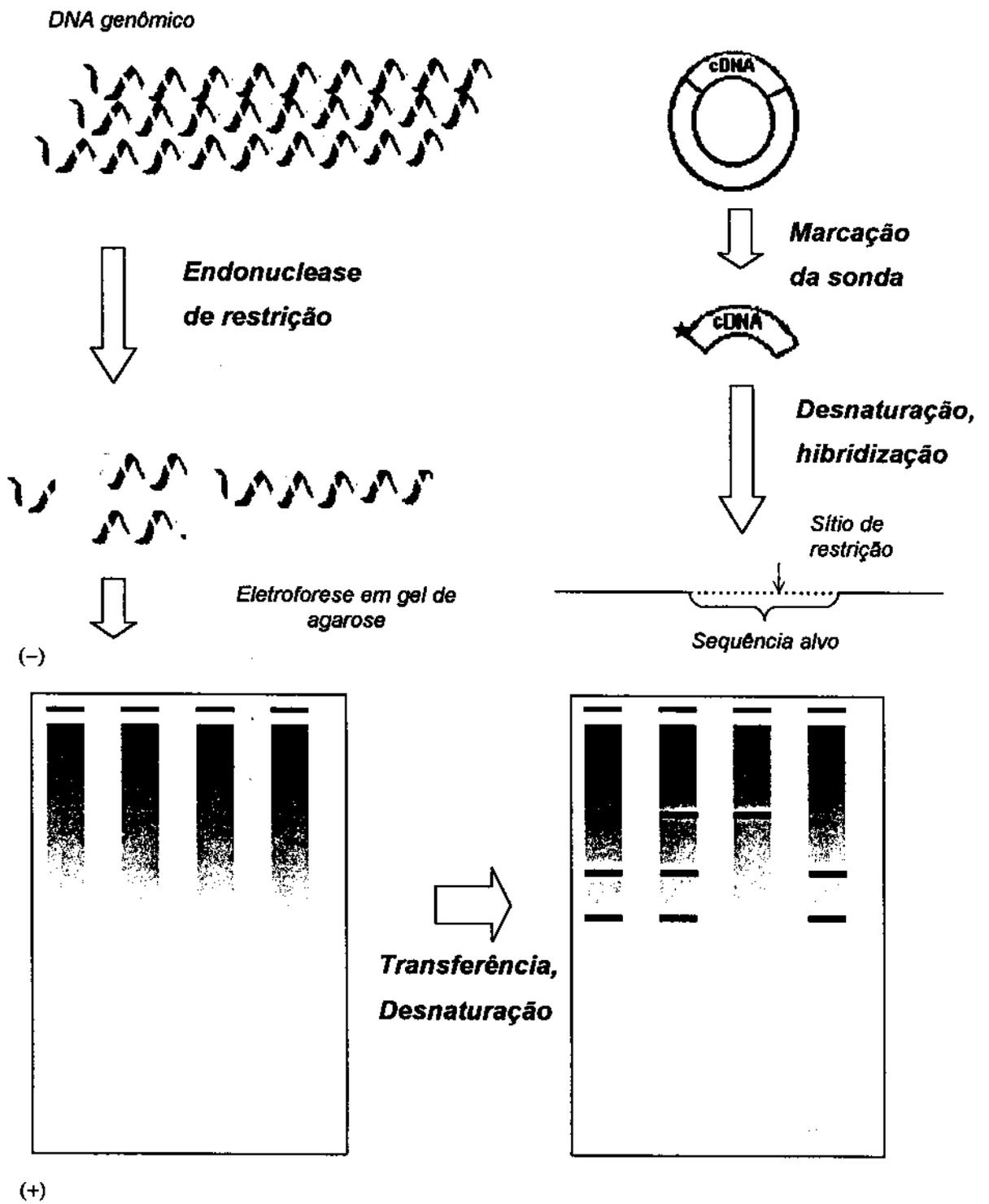


Figura 1. Análise de RFLP pela técnica de Southern blot.

DNA. Porém, em virtude da abundância de DNA repetitivo no genoma dos eucariotos, quando o objetivo é a obtenção de marcadores unilocais, essas sondas devem ser submetidas a um processo de seleção, de forma a garantir que apenas aquelas que representam seqüências de cópia única sejam empregadas.

Um aspecto importante dos marcadores RFLP é a prevalência de dialelismo. Muitos dos locos marcadores descritos são dialélicos, quer originados de mutações de ponto ou de mutações estruturais. Marcadores dialélicos são bastante limitados para o mapeamento genético em populações segregantes, em função do número de cruzamentos informativos. Outra desvantagem é a presença de alelos raros, fazendo com que muitas populações apresentem um padrão monomórfico. Apesar dessa limitação, em 1990 mais de 120 locos haviam sido mapeados em bovinos (Kennedy et al., 1990).

A técnica de PCR

Uma das mais marcantes contribuições ao estudo de marcadores moleculares foi o desenvolvimento da técnica de Reação em Cadeia da Polimerase (PCR). Essa técnica foi desenvolvida por Mullis em 1983 (Mullis, 1990), mas sua importância só ficou demonstrada com a publicação dos primeiros trabalhos de aplicação da PCR em diagnóstico de doenças (Saiki *et al.*, 1985). A técnica consiste na replicação do DNA "in vitro", catalisada por uma DNA polimerase termoestável (Figura 2). A reação requer a presença dos quatro tipos de nucleotídeos e de oligonucleotídeos complementares às extremidades da região do DNA que se deseja amplificar. Os oligonucleotídeos ("iniciadores" ou "primers") funcionam como ponto de início para a síntese de uma fita de DNA complementar à fita molde. Cada ciclo de PCR envolve a desnaturação da molécula de DNA alvo, obtida pela elevação da temperatura para 92 a 95°C, o anelamento dos "primers" pela redução da temperatura até o ponto ideal para cada par de "primer", e a extensão da síntese da nova fita de DNA. Como as fitas recém sintetizadas funcionam como molde para o próximo ciclo, ao final de vários ciclos obtém-se um acúmulo exponencial de cópias da região delimitada pelos "primers". A temperatura ideal de anelamento depende basicamente do tamanho e da seqüência dos oligonucleotídeos utilizados. Como regra geral, "primers"

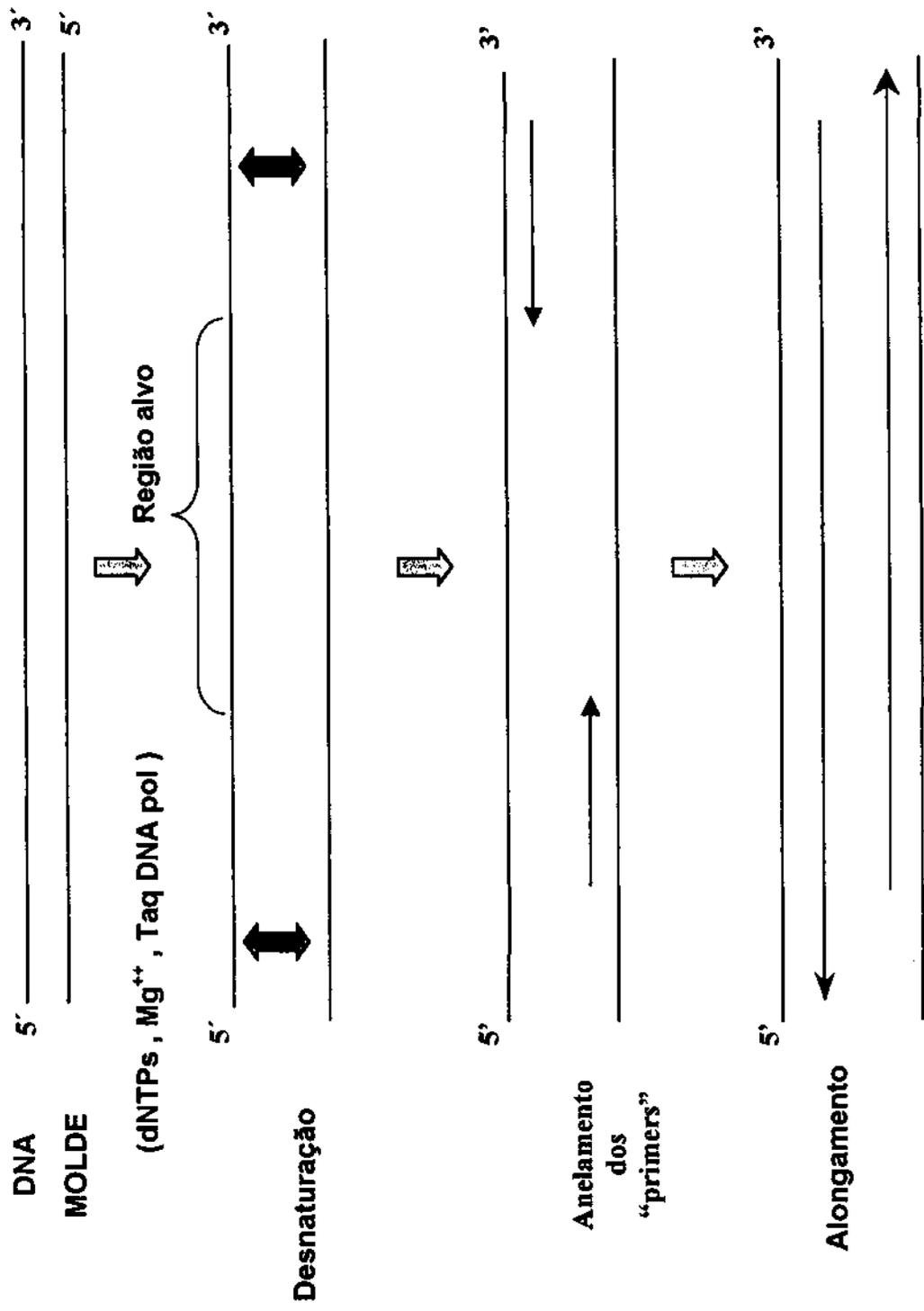


Figura 2. Representação esquemática da reação em cadeia da polimerase (PCR)

2º CICLO

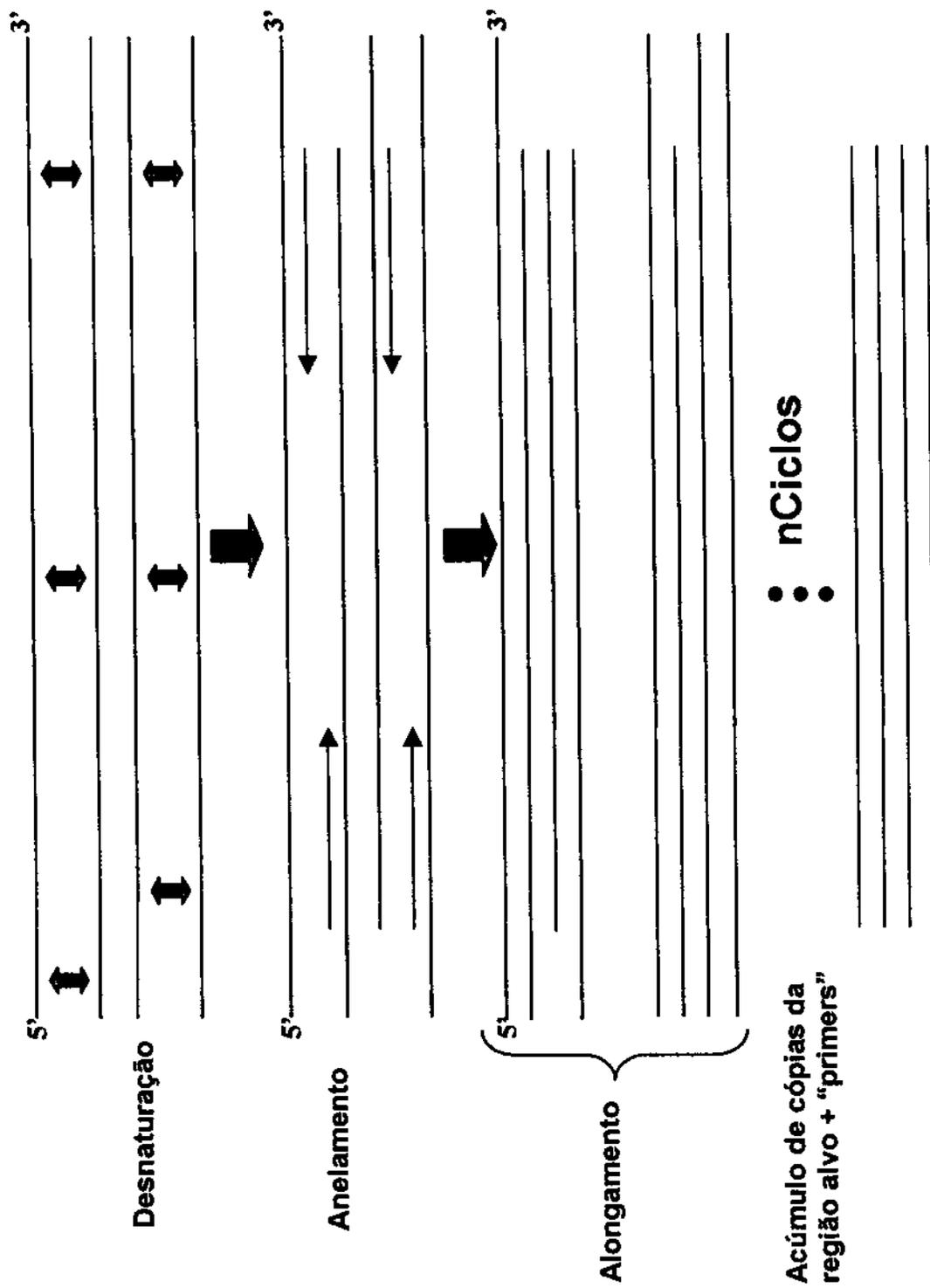


Fig.2. Continuação

mais longos e com maior conteúdo de guanina (G) e citosina (C) apresentam anelamento mais estável.

A extensão da síntese da nova fita de DNA era originalmente catalisada pelo fragmento Klenow da DNA-polimerase I que, por ser termolábil, exigia adição de enzima a cada novo ciclo de PCR.

A utilização de uma DNA polimerase (Taq polimerase) isolada da bactéria *Thermus aquaticus*, capaz de catalisar a síntese de DNA a 72°C e se manter estável a 92°C, permitiu a automação da técnica de PCR (Saiki *et al.*, 1988).

O impacto da técnica de PCR na análise de marcadores de DNA deve-se principalmente à simplicidade e sensibilidade. A quantidade de DNA molde necessária para a detecção do produto de amplificação torna possível a análise do DNA obtido de uma única célula. O princípio de amplificação de DNA *in vitro* tem sido utilizado para o desenvolvimento de técnicas de sequenciamento, clonagem e para o desenvolvimento de novos marcadores, como os RAPDs (Polimorfismos de DNA amplificados ao acaso), AFLPs (Polimorfismos de Comprimento de Fragmentos Amplificados) (Ferreira e Grattapaglia, 1995) e microssatélites.

Além disso, muitos dos polimorfismos descritos pelo método de Southern têm sido transformados em marcadores baseados em PCR específica.

Para a análise de RFLP utilizando PCR, as extremidades da sonda que revelou o polimorfismo são sequenciadas e as informações são utilizadas para a síntese de "primers" complementares. Esses "primers" são utilizados para amplificar a região que contém os sítios polimórficos de restrição. Os fragmentos produzidos pela digestão dos produtos de amplificação podem ser então analisados em géis de agarose (Ferreira e Grattapaglia, 1995). O PCR-RFLP tem sido aplicado no diagnóstico de doenças hereditárias em humanos e em animais.

Com o crescente desenvolvimento dos projetos de sequenciamento do genoma das principais espécies animais, os "primers" podem também ser desenvolvidos a partir de seqüências publicadas.

Microssatélites

Os genomas dos eucariotos abrigam grande quantidade de DNA repetitivo, classificado de acordo com o número de nucleotídeos e a complexidade da seqüência que compõe as repetições. As diferentes classes são também caracterizadas por comportamento típico quanto à distribuição no genoma e quanto ao mecanismo envolvido em suas origens (Jobse *et al.*, 1995).

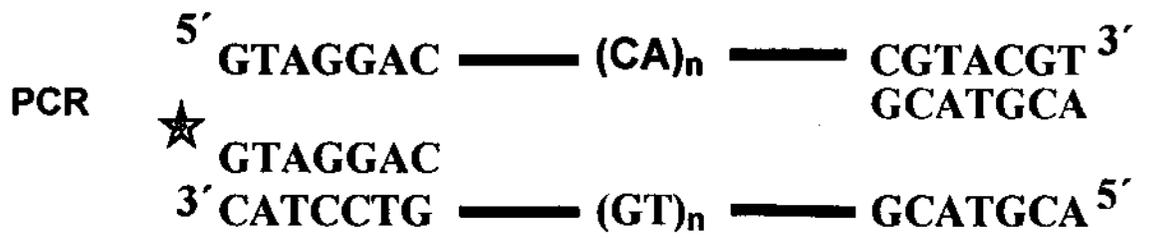
A variabilidade de seqüências repetitivas foi demonstrada por Bell *et al.* (1982), e posteriormente utilizada para a produção de perfis característicos de cada indivíduo pela técnica de "DNA fingerprint" (Jeffreys *et al.*, 1985).

Um avanço importante na obtenção de marcadores unilocais veio da utilização de uma classe de DNA repetitivo, as seqüências microssatélites. Os microssatélites são caracterizados por repetições em tandem de um mono, di ou trinucleotídeo mais freqüente, localizado dentro de regiões de seqüência única. Cada bloco de repetições é geralmente menor do que 100 pares de nucleotídeos (Tautz, 1989). Do mesmo modo que outras regiões repetitivas do genoma, a variação do número de repetições em cada loco é resultante de erros no deslocamento da DNA polimerase durante a replicação do DNA.

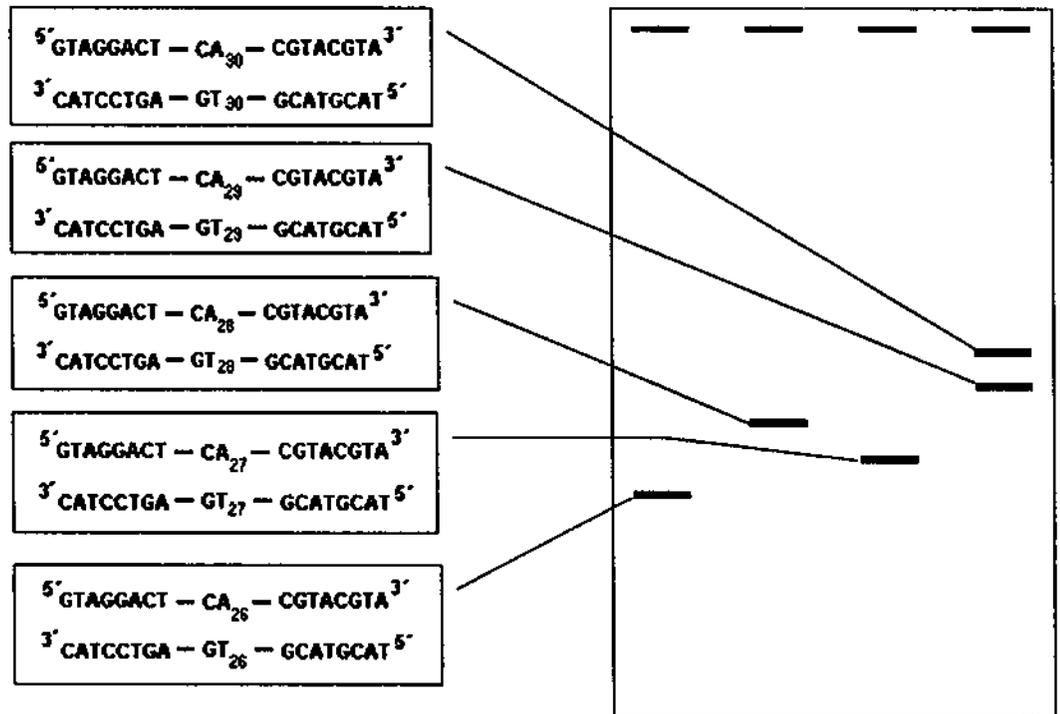
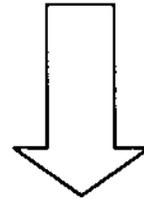
Os microssatélites poli(G) e poli(A) são os mais simples enquanto que poli(GT) são os mais freqüentes, aparecendo em aproximadamente $5 \text{ a } 10 \times 10^4$ locos individuais no genoma dos mamíferos (Soller, 1990). Muitos outros microssatélites foram identificados até o momento, e é possível que qualquer seqüência de poucos nucleotídeos possa representar um microssatélite no genoma de eucariotos.

Os microssatélites podem ser amplificados de maneira específica pela técnica de PCR utilizando "primers" que contém parte da seqüência flanqueadora de cópia única (Figura 3). A amplificação resulta em produtos de diferentes tamanhos, em função do número de cópias da seqüência microssatélite delimitada pelos "primers". Uma vez que os oligonucleotídeos utilizados são complementares às seqüências de cópia única que flanqueiam os microssatélites, obtém-se marcadores unilocais, altamente polimórficos e de herança codominante.

Distorções da segregação podem surgir em decorrência da "expansão" ou "contração" do microssatélite durante a meiose. A freqüência desses eventos é da ordem



Eletroforese em gel de poliacrilamida



Autorradiografia, coloração com prata ou sequenciador automático

Figura 3. Análise de marcadores microssatélites utilizando a técnica de PCR

de 10^{-4} a 10^{-5} por loco por gameta (Holmes, 1994), suficientemente baixa para permitir seu uso como marcador genético. Entretanto, em presença de alterações do sistema de reparo, a taxa de mutação pode aumentar de 100 a 1000 vezes. A instabilidade resultante tem sido associada com o desenvolvimento de alguns tipos de câncer (Simpson, 1996) e de doenças hereditárias, como a síndrome do cromossomo X-frágil.

Considerando a frequência dessas seqüências e um tamanho médio do genoma de mamífero de 3×10^9 pares de nucleotídeos, seria possível construir mapas genéticos com aproximadamente 10.000 locos microssatélites (Soller, 1990). Em bovinos, os mapas genéticos compostos por uma combinação de marcadores RFLP e locos microssatélite, fornecem uma cobertura de aproximadamente 90% do genoma (Bishop et al., 1994; Barendse et al., 1994; Barendse et al., 1997).

Recentes progressos têm sido alcançados em decorrência do grande estímulo ao programa Genoma Humano, cujo objetivo é mapear e sequenciar todos os genes da espécie. O mapeamento comparativo deverá resultar em avanços significativos também em outras espécies de mamíferos, uma vez que a conservação de padrões cariotípicos, seqüências codificadoras e de ordem de genes nos cromossomos tem sido descrita em muitas espécies (Womack e Moll, 1986).

A utilização de PCR para a análise desses marcadores representa uma importante redução no tempo despendido para a identificação dos genótipos, além de permitir a identificação do genótipo de embriões. Esse fato, aliado ao intenso polimorfismo, torna os microssatélites marcadores bastante adequados para a construção de mapas genéticos e identificação de locos que controlam caracteres quantitativos (QTLs), principalmente em espécies onde o número de progênies é limitado (Massey e Georges, 1992).

Referências bibliográficas

BARENDSE, W.; ARMITAGE, S. M.; KOSSAREK, L.M.; SHALOM, A.; KIRKPATRICK, B. W.; RYAN, A. M.; CLAYTON, D.; LI, L.; NEIBERG, H. L.; ZHANG, N.; GROSSE, W. M.; WEISS, J.; CREIGHTON, P.; McCARTHY, F.; RON, M.; TEALE, A. J.; FRIES, R.; McGRAW, R. A.; MOORE, S. S.; GEORGES, M.; SOLLER, J. E.; WOMACK, J. E.; HETZEL, D. J. S. A genetic linkage map of the bovine genome. **Nature Genetics**, 6: 227-244, 1994.

- BARENDSE, W., VAIMAN, D., KEMP, S.J., SUGIMOTO, Y., ARMITAGE, J.L., WILLIAMS, J.L., SUN, H.S., EGGEN, A., AGABA, M., ALEYASIN, S.A., BAND, M., BISHOP, M.D., BUITKAMP, J., BYRNE, K., COLLINS, F., COOPER, L., COPETTIERS, W., DENYS, B., DRINKWATER, R.D., EASTERDAY, K., ELDUQUE, C., ENNIS, S., ERHARDT, G., FERRETTI, L., FLAVIN, N., GAO, Q., GEORGES, M., GURUNG, R., HARZILIUS, B., HAWKINS, G., HETZEL, J., HIRANO, T., HULME, D., JORGENSEN, C., KESSLER, M., KIRKPATRICK, B.W., KONFORTOV, B., KOSTIA, S., KUHN, C., LENSTRA, J.A., LEVEZIEL, H., LEWIN, H.A., LEYHE, B., LIL, L., BURRIEL, I.M., MCGRAW, R.A., MILLER, J.R., MOODY, D.E., MOORE, S.S., NAKANE, S., NIJMAN, I.J., OLSAKER, I., POMP, D., RANDO, A., RON, M., SHALOM, A., TEALE, A.J., THIEVEN, U., URQUHART, B.G.D., VAGE, D.I., VAN DE WEGHE, A., VARVIO, S., VELMALA, R., VILKKI, J., WEIKARD, R., WOODSIDE, C., WOMACK, J.E., ZANOTTI, M., ZARAGOZA, P. A medium-density genetic linkage map of the bovine genome. *Mammalian Genome*, **8**: 21-28, 1997.
- BELL, G.I.; SELBY, M.J. & RUTTER, W.J. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating units. *Nature*, **295**: 31, 1982.
- BISHOP, M. D.; KAPPES, S. M.; KEELE, J. W.; STONE, R. T.; SUNDEN, S.L.F.; HAWKINS, G. A.; TOLDO, S. S.; FRIES, R.; GROSZ, M. D.; YOO, J.; BEATTIE, C. W. A genetic linkage map for cattle. *Genetics*, **136**: 619-639, 1994.
- BRUFORD, M.W.; HANOTTE, O.; BROOKFIELD, J.F.Y.; BURKE, T. Single-locus and multilocus DNA fingerprinting. In: HOEZEL, A.R. (ed.). *Molecular Genetic Analysis of Populations. A Practical Approach*. IRL Press, Oxford, 1992. cap. 8, p. 225-269.
- FERREIRA, M. E. & GRATTAPAGLIA, D. *Introdução ao uso de marcadores RAPD e RFLP em análise genética*. EMBRAPA-CENARGEN, Brasília, 1995. 220p.
- GRODZICKER, T.; WILLIAMS, J.; SHARP, P.; SAMBROOK, J. Physical mapping of temperature - sensitive mutations of adenoviruses. *Cold Spring Harbor Symposium on Quantitative Biology*, **39**: 439-446, 1974.
- HOLMES, N.G. Microsatellite markers and the analysis of genetic disease. *British Veterinary Journal*, **150**: 411-421, 1994.
- JEFFREYS, A.J.; WILSON, V. & THEIN, S.L. Individual-specific "fingerprints" of human DNA. *Nature*, **316**: 76, 1985.
- JOBSE, C.; BUNTJER, J.B.; HAAGSMA, N.; BREUKELMAN, H.J.; BEINTEMA, J.J.; LENSTRA, J.A. Evolution and recombination of bovine DNA repeats. *Journal of Molecular Evolutions*, **41**: 277-283, 1995.
- KENNEDY, B.W.; GIBBINS, A.M.V.; GIBSON, J.P.; SMITH, C. Coalescence of molecular and quantitative genetics for livestock improvement. *Journal of Dairy Science*, **73**: 2619-2627, 1990.

- MASSEY, J.M. & GEORGES, M. Genmark's approach to marker-assisted selection. **Animal Bio-technology**, **3**: 95-109, 1992.
- MULLIS, K. B.. The unusual origin of the polymerase chain reaction. **Scientific American**, **262**: 36 - 42, 1990.
- SAIKI, R. K.; SCHARF, S.; FALOONA, F.; MULLIS, K. B.; HORN, G. T.; ERLICH, H. A.; ARNHEIN, N. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of Sickle Cell Anemia. **Science**, **230**: 1350-1354, 1985.
- SAIKI, R.K.; GELFAND, D.H.; STOFFE, S.J.; HIGUCHI, R.; HORN, G.T.; MULLIS, K.B.; ERLICH, H.A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. **Science**, **239**: 487, 1988.
- SIMPSON, A.J.G. Mini Review. Microsatellite instability in human cancer. **Brazilian Journal of Genetics**, **19**: 171-174, 1996.
- SOLLER, M. Genetic mapping of the bovine genome using deoxyribonucleic acid-level markers to identify loci affecting quantitative traits of economic importance. **Journal of Dairy Science**, **73**: 2628-2646, 1990.
- SOUTHERN, E.M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. **Journal of Molecular Biology**, London, **98**: 503-517, 1975.
- TANKSLEY, S. D. Mapping polygenes. **Annual Review of Genetics**, **27**: 205-233, 1993.
- TAUTZ, D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. **Nucleic Acids Research**, **17**: 6463-6471, 1989.
- WOMACK, J. E. & MOLL, Y. D. Gene map of the cow: conservation of linkage with mouse and man. **The Journal of Heredity**, **77**: 2-7, 1986.

POLIMORFISMOS DO SISTEMA IMUNE E SUA APLICAÇÃO NO MELHORAMENTO ANIMAL PARA RESISTÊNCIA A DOENÇAS INFECCIOSAS E PARASITÁRIAS.

Isabel K. F. de Miranda Santos
Laboratório de Imunologia
Embrapa Recursos Genéticos e Biotecnologia
Brasília DF

Demandas para marcadores para resistência a doenças na produção pecuária

As doenças infecciosas e parasitárias impõem perdas importantes em vários níveis da produção pecuária do País. A nova ordem no comércio mundial substituiu as barreiras tarifárias pelas barreiras sanitárias afetando o comércio de produtos contaminados seja com microorganismos seja com os produtos farmaco-químicos empregados no controle de doenças e infestações. Há o problema igualmente significativo das perdas de animais e produtos pelas doenças e infestações em si. Embora os dados do Ministério da Agricultura referentes a infecções e parasitoses sejam incompletos assim como são escassos os estudos avaliando o impacto econômico dessas doenças, apenas alguns números ilustram a magnitude das perdas. Os prejuízos causados pelo carrapato são estimados em mais de US\$ 1 bilhão anuais somente em perdas na qualidade de couros. Juntamente a essas perdas somam-se as causadas por parasitas transmitidos por esse vetor: a babesiose causa prejuízos de mais de US\$1.365 milhões por ano para a América Latina¹.

O Brasil é o terceiro maior mercado de defensivos e insumos veterinários². Na estimativa de membros do SINDAN uma parte significativa dessa soma destina-se a produtos que poderiam ser substituídos por tecnologias mais eficazes e baratas como, por exemplo, animais geneticamente resistentes a parasitas. Atualmente poucos grupos exploram os polimorfismos do sistema imune para obter animais resistentes a doenças sendo que apenas um marcador imunológico é empregado no melhoramento de

¹ Montenegro-James S. Prevalence and control of babesiosis in the Americas. Mem. Inst. Oswaldo Cruz 87 Suppl. 3: 27-36, 1992.

² Fonte: SINDAN

animais de pecuária, aquele do complexo de histocompatibilidade de aves associado com resistência à Doença de Marek³.

O significado biológico dos polimorfismos do sistema imune no contexto do melhoramento animal

O sistema imune presta-se particularmente bem para a exploração de marcadores para melhoramento pois é o sistema fisiológico mais polimórfico de todo o reino animal. Tal variabilidade advém de sua função que é a de reconhecer um número enorme de patógenos de composição antigênica variada. Por sua vez os patógenos, durante a co-evolução com hospedeiros superiores, desenvolveram mecanismos de escape à resposta imune. Esse embate constante é a força seletiva que mantém o polimorfismo dos elementos do sistema imune⁴.

O significado biológico desse polimorfismo é ilustrado por meio do estudo do Complexo Maior de Histocompatibilidade (CMH) que tem o maior nível de polimorfismo e de heterozigosidade descrito para qualquer organismo conhecido⁵. A função das moléculas codificadas nesse complexo, glicoproteínas expressas em superfícies celulares, é a de apresentar peptídeos antigênicos a linfócitos T, um passo essencial para a indução da resposta imune. A vantagem seletiva do grande polimorfismo é confirmada através das seguintes observações: **a)** a distribuição uniforme da frequência de alelos o que não seria esperado caso a seleção fosse neutra; **b)** a heterozigosidade observada, também maior do que o esperado; **c)** a taxa de substituições não-sinônimas de bases excede em muito a taxa de substituições sinônimas apenas na região codificante para o sítio de ligação de peptídeos antigênicos; **d)** os polimorfismos são mantidos através das barreiras de espécies e antecedem a especiação; **e)** os introns são homogêneos em relação aos exons o que sugere que a seleção balanceadora mantém a diversidade nos exons.

³ Tizzard

⁴ Haldane JBS. 1949. Disease and Evolution La Ricerca Sci. 19, suppl.68-76. Citado em Infectious Diseases of Humans. Dynamics and Control. Anderson RM e May RM. Oxford University Press. 1992. p. 641.

Em relação à heterozigosidade elevada, propõe-se que há uma vantagem seletiva (seleção sobredominante) para os portadores porque podem apresentar uma gama maior de peptídeos antigênicos⁵. Essa vantagem foi confirmada em relação à infecção pelo HIV onde indivíduos apresentando heterozigosidade máxima nos locos HLA-A, B e C do CMH humano evoluíram muito mais lentamente para a síndrome da imunodeficiência do que os indivíduos que não apresentavam esse grau de heterozigosidade⁷. Um estudo desenvolvido em populações naturais de ovinos sem manejo indicou que de fato ocorre uma seleção balanceada dos polimorfismos do CMH e que um parasita pode ser uma das forças seletivas: a população apresentou uma associação significativa entre alelos do OLA, o CMH de ovinos e resistência a nematóides intestinais e sobrevivência de cordeiros⁸.

Os patógenos podem, de modo geral, ser divididos em dois grandes grupos conforme os mecanismos imunes efetores do hospedeiro empregados para combatê-los: os chamados patógenos intracelulares, melhor controlados pela imunidade humoral (i.e. Por anticorpos, complemento e demais componentes séricos;), e os patógenos extracelulares, melhor controlados por mecanismos celulares (i.e., Células citotóxicas, macrófagos ativados, etc) de imunidade. Estes mecanismos são geralmente, porém nem sempre, mutuamente excludentes⁹. Cada vez mais evidências acumulam-se mostrando que os polimorfismos do sistema imune afetam de maneira positiva ou negativa a capacidade de controlar patógenos¹⁰. Assim, conforme o repertório genético de um hospedeiro (entre outros fatores) predomina a imunidade humoral ou celular.

⁵ Hughes AL e Yeager M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. *Ann. Ver. Genetics* 32: 415-435.

⁶ Doherty PC e Zinkemagel RM. 1975. Enhanced immunologic surveillance in mice heterozygous at the H-2 gene complex. *Nature* 256: 50-52.

⁷ Carrington M, Nelson GW, Martin MP et al. 1999. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* 283: 1748-1752.

⁸ Paterson S, Wilson K e Pemberton JM. 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aires L.*). *Proc. Natl. Acad. Sci. USA* 95: 3714-3719.

⁹ Janeway CA e Travers P. *Immunobiology. The immune system in health and disease.* 3rd Edition. Current Biology. 1997. New York.

Isso posto, vale ressaltar que, devido à própria fisiologia do sistema imune, não obteremos o animal universalmente resistente: num mesmo indivíduo a resistência a determinados grupos de patógenos pode significar sua susceptibilidade a outros. Alelos de HLA-D (CMH humano) associados com resistência a formas clínicas graves de infestação por uma filária, a oncocercose, são os mesmos alelos associados com susceptibilidade a formas clínicas graves de malária¹¹. O polimorfismo do sistema imune visa, na verdade, a proteção do rebanho e não do indivíduo. Ainda, a seleção dirigida para um ou poucos alelos de MHC carrega o risco associado de selecionar concomitantemente patógenos mais virulentos. Tal possibilidade é sugerida pela observação de que a frequência do alelo HLA-A11 quando é alta em populações humanas, acarreta a perda de epitopos do vírus de Epstein-Barr que são apresentadas no contexto desse alelo; essa variante antigênica do vírus não ocorre em populações que apresentam frequências baixas desse mesmo alelo¹². Os atuais processos de produção de aves contam com animais com pouca variabilidade genética, inclusive a nível de MHC. Não há, porém, estudos que abordem a pressão seletiva exercida por frequências altas de alelos do MHC sobre patógenos. Qualquer programa de melhoramento animal para resistência a doenças deve considerar o impacto fisiológico dos alelos/polimorfismos eleitos para assistir as seleções e também o contexto epidemiológico e produtivo no qual os animais melhorados serão empregados.

A grande e necessária variabilidade dos elementos imunes tem sofrido com os processos de melhoramento animal para características produtivas os quais estreitam a base genética das populações alvo. Com efeito, o número de alelos observado para um loco muito polimórfico do bola (cmh bovino) é maior em populações de bovinos africanas do que européias pois estas últimas contam com uma base populacional menor devido ao pequeno número de touros empregados na reprodução¹³. As

¹⁰ Hill AVS. 1998. The immunogenetics of human infectious diseases. *Ann. Rev. Immunol.* 16: 593-617.

¹¹ Meyer CG, Gallin M, Ertmann KD et al. 1994. HLA-D alleles associated with generalized disease, localized disease and putative immunity in *Onchocerca volvulus* infection. *Proc. Natl. Acad. Sci. USA* 91: 7515-7519.

¹² CAMPOS DE LIMA PO, GAVIOLI R, ZHANG QJ, WALLACE LE, DOLCETTI R, ROWE M, RICKINSON AB, MASUCCI MG. HLA-A11 epitope loss isolates of Epstein-Barr virus from a highly A11⁺ population. *Science* 260: 98-100, 1993.

¹³ Mikko S e Andersson L. 1995. Extensive MHC class II diversity in African and European cattle. *Immunogenetics* 42: 408-413.

consequências fisiológicas desse estreitamento são ilustradas por meio da reconhecida rusticidade de populações naturais ou naturalizadas, não comerciais. Um exemplo documentado é a resistência de gado n'dama aos tripanosomas africanos que impedem a bovinocultura em vastas regiões da África. Perus melhorados para peso de carcaça tornam-se mais suscetíveis a infecções por *Pasteurella multocida* em relação a animais não melhorados¹⁴. Outro exemplo é a extrema sensibilidade de guepardos de populações da África do Sul ao vírus da peritonite felina¹⁵. Esses animais são endogâmicos e monomórficos para diversos locos do CMH (cmh felino) o que dá uma idéia da extensão do estreitamento genético e, também, do papel do CMH na resistência a doenças. O primata *Saguinus oedipus* apresenta polimorfismo muito limitado no CMH e alta susceptibilidade com mortalidade a várias infecções virais¹⁶.

Essas observações ressaltam o valor de bancos de germoplasma animal os quais são fonte de polimorfismos úteis para o melhoramento para resistência. Por outro lado, um estudo documentou uma baixa diversidade no CMH de alces europeus e norte americanos os quais, apesar de um pequeno número de alelos, mantêm populações viáveis¹⁷. Há, porém, que se considerar se a flora de patógenos prevalente na área de distribuição geográfica circumpolar dessa espécie não é adequadamente coberta pelos alelos presentes nessa espécie. Ultimamente de altas taxas de mortalidade causadas por um nematóide e um retrovírus tem sido registradas nessas populações.

Abordagens para obter marcadores do sistema imune para resistência a doenças.

Nos últimos 5 anos a imunogenética tem descrito um número cada vez maior de polimorfismos para componentes do sistema imune que não o CMH. As consequências imunobiológicas dos polimorfismos podem ser distintas porém a força seletiva é a

¹⁴ NESTOR KE, SAIF YM, ZHU J, NOBLE DO. Influence of growth selection in turkeys on resistance to *Pasteurella multocida*. *Poultry Sci* 75: 1161-1163, 1996

¹⁵ O'BRIEN SJ, ROELKE ME, MARKER L, ET AL. Genetic basis for species vulnerability in the cheetah. *Science* 227(4693):1428-1434, 1985

¹⁶ Evans DT, Piekarczyk MS, Allen TM et al. 1997. Immunodominance of a single CTL epitope in a primate species with limited MHC class I polymorphism. *J. Immunol.* 159: 1374-1382.

¹⁷ Mikko S e Andersson L. 1995. Low major histocompatibility complex class II diversity in European and North American moose. *Proc. Natil. Acad. Sci. USA* 92: 4259-4263.

mesma. Qualquer que seja a metodologia empregada para detectar os polimorfismos, a pergunta básica é: qual é a extensão do componente genético de susceptibilidade e quais alelos afetam a susceptibilidade a doenças infecciosas? Embora o objetivo desse artigo seja a descrição de polimorfismos do sistema imune com potencial para melhoramento animal para resistência a patógenos, as abordagens para obter marcadores comentadas aqui pois, conforme a metodologia empregada, afetam diretamente a qualidade da informação obtida e sua aplicabilidade imediata em programas com seleção assistida por marcadores. Referimos os leitores a várias revisões sobre o assunto indicados ao longo do texto e também aos artigos de Weeks e Lathrop e Risch e Merikangas¹⁸.

A descoberta de alelos e marcadores envolvidos com resistência e potencialmente úteis ao melhoramento tem como ponto de partida a seguinte observação: a grande heterogeneidade da resposta imune observada em populações infectadas com o mesmo patógeno ou imunizadas com o mesmo imunógeno. Durante a interação de um patógeno com uma população há grande variação individual nos desfechos clínicos observados, alguns chegam ao óbito enquanto outros controlam o mesmo patógeno em níveis toleráveis ou até subclínicos. Esses diferentes fenótipos clínicos dependem tanto de fatores externos aos hospedeiros como grau de virulência do patógeno, estado nutricional, tamanho do inóculo/nível de exposição e outros quanto da composição da genética do sistema imune de cada hospedeiro. Os estudos clínicos-epidemiológicos conduzidos adequadamente distinguem estes diversos fatores e fornecem o material para documentar o impacto fisiológico de alelos/marcadores do sistema imune sobre a resistência do hospedeiro.

Para buscar marcadores para resistência pode-se optar entre duas abordagens básicas, nenhuma excludente da outra e cada um apresentando vantagens e desvantagens: a do **gene candidato** associado com o fenótipo desejável de resistência

¹⁸ Weeks DE e Lathrop GM. 1995. Polygenic diseases: methods for mapping complex disease traits. TIG 11: 513-519.

Risch N e Merikangas K. 1996. The future of genetic studies of complex human diseases. Science 273: 1516-1517.

Comentários sobre o artigo de Risch e Merikangas: 1997. Science 275: 1327-1263.

ou a dos **estudos de ligação** entre marcadores de função desconhecida e o fenótipo de interesse.

Gene candidato: A opção do gene candidato, é válida desde que se tenha um candidato forte, ou seja, um gene de cujo produto conhece-se a fisiologia e cujos alelos têm um impacto igualmente conhecido ou fácil de prever¹⁹. Segundo Rothschild e Soller a probabilidade de um gene candidato estar associado a uma característica de importância econômica é teoricamente muito pequena mas, apesar disso, na prática essa abordagem tem sido bem sucedida para identificar genes de interesse. Na pecuária muitos genes candidatos são escolhidos baseados em conhecimentos obtidos através do estudo do efeito de uma mutação no homem ou em animais experimentais. Polimorfismos do sistema imune de animais de pecuária que estão correlacionados com fenótipos de resistência e susceptibilidade a doenças infecciosas e parasitárias são descritos no tópico final. Sempre que houver uma relevância potencial para animais, os polimorfismos humanos ainda não descritos em animais serão discutidos.

Uma vez escolhido o gene candidato, existem várias etapas até que sua associação com o fenótipo desejável seja desmonstrado¹⁶ as quais são resumidas a seguir. As sequências de primers devem ser identificadas para proceder à amplificação do gene o que é, obviamente, facilitado se houver informação sobre a sequência do gene na espécie de interesse ou em espécies relacionadas. Parte-se para a identificação dos polimorfismos no gene candidato, de preferência por meio de um processo que permita genotipagem em grande escala. A próxima etapa é a identificação de uma população para desenvolver estudos de associação e a obtenção de informação sobre os fenótipos de interesse. Os estudo de associação em si podem envolver um único sítio polimórfico ou então haplotipos polimórficos que porventura tenham sido descritos na etapa inicial e os quais conferem maior especificidade ao estudo. Vários modelos podem ser contruídos para comprovar a associação do gene candidato com o fenótipo de interesse como, por exemplo, verificar as frequências dos alelos entre raças que diferem entre si quanto ao fenótipo. As associações observadas

¹⁹ Rothschild MF e Soller M. 1997. Candidate gene analysis to detect genes controlling traits of economic importance in domestica livestock. Probe 8: 13-20.

precisam ser confirmadas estudando-se outras populações e outras gerações da população do estudo inicial.

As vantagens da abordagem do gene candidato são: o maior poder estatístico obtido com menor número de famílias e indivíduos do que o necessário para estudos de ligação, a fácil aplicabilidade porque não se necessita de populações de animais com duas ou três gerações, custo baixo, a simplicidade de se trabalhar com um único gene e, finalmente, a aplicação imediata em programas de seleção e estabilidade do efeito do gene candidato sobre a característica desejada. Genes candidatos de efeito mais modesto no fenótipo de interesse, que dificilmente seriam detectados em estudos de ligação, podem ser avaliados com sucesso. Entre as desvantagens estão o efeito da pleiotropia de outros genes sobre o gene candidato e a não participação do gene candidato no fenótipo desejado ou seja, a escolha inicial de um gene candidato fraco, sem fundamento fisiológico. Uma última desvantagem é a necessidade de identificação prévia do gene e do mecanismo de ação de seu produto. Em relação a animais de pecuária esse é um fator menos limitante do que na área vegetal devido aos maiores conhecimentos disponíveis sobre a fisiologia de fenótipos de interesse.¹⁰⁰

Estudos de ligação e análise de genoma inteiro: Muitos genes causadores de doenças já foram identificados por meio de estudos que verificam a co-segregação em famílias do fenótipo (i.e., doença) com marcadores aleatórios de um loco genético, ou seja, estudos de ligação. Em relação a essa abordagem, a varredura do genoma inteiro por meio de marcadores microssatélites constitui a tecnologia mais moderna. No caso do homem, mais de 300 desses marcadores informativos permitem a dissecção do genoma em segmentos de 10 centimorgans ou 10 megabases. Em animais de pecuária o número de microssatélites aumenta constantemente e ainda há a possibilidade de se compartilhar muitos microssatélites entre diferentes espécies, por exemplo entre bovinos, ovinos e caprinos, o que permite a varredura desses genomas com resolução muito boa.

A vantagem dos estudos de ligação é a de que não exigem conhecimento prévio do(s) gene(s) envolvido(s) na expressão do fenótipo de interesse. Permitem, assim, a identificação de genes novos que não seriam apontados por estudos envolvendo genes candidatos. Também permitem a identificação de genes que não codificam produtos

envolvidos no fenótipo de interesse mas cujo efeito sobre esse fenótipo é pleiotrópico. Entre suas desvantagens estão o alto custo e quantidade de indivíduos e famílias necessários para uma análise fiel. Marcadores ou regiões cromossômicas identificadas por meio dos estudos de ligação dificilmente podem ser aplicados imediatamente a programas de melhoramento por seleção assistida por marcadores (SAM). A exceção ocorre nos casos onde mapas de ligação muito detalhados são obtidos como nos casos de famílias com grande número de irmãos ou meio-irmãos e onde a SAM pode ser diretamente implementada.

Dois estudos independentes^{20,21} que empregaram a abordagem dos estudos de ligação para identificar regiões cromossômicas envolvidos com resistência à leishmaniose murina, uma doença parasitária, encontraram resultados totalmente discrepantes o que indica que essa tecnologia ainda tem limitações. Um desses estudos identificou regiões nos cromossomos 6, 7, 10, 11, 15 e 16 do camundongo enquanto que o outro, empregando o mesmo patógeno, identificou os cromossomos 9 e 17. Uma das razões para essa discrepância é que alguns loci envolvidos com resistência teriam baixa representatividade na população inicial. Outra explicação para as diferenças entre os dois estudos é que a expressão do fenótipo, no caso tamanho de lesão induzida pela *Leishmania major*, depende fortemente do tamanho do inóculo e estes foram diferentes nos dois estudos. Esses resultados também ilustram que nos estudos envolvendo fenótipos de resistência, a penetrância do fenótipo de resistência depende do grau de virulência e do tamanho do inóculo do patógeno sendo preferível usar inóculos de baixa virulência mesmo que alguns animais suscetíveis manifestem resistência.

Atualmente há em curso um estudo para identificar genes de bovinos associados com tolerância a tripanosomas africanos, parasitas que causam perdas econômicas importantes à bovinocultura no continente africano. Nesse empreendimento empregou-se a ovulação múltipla com transferência de embriões para produzir um grande número de famílias de irmãos e meio-irmãos de F2 de N'Dama (resistente) X Boran (suscetível).

²⁰ ROBERTS LJ, BALDWIN TM, CURTIS JM, HANDMAN E, FOOTE SJ Resistance to *Leishmania major* is linked to the H2 region on chromosome 17 and to chromosome 9. *J Exp Med* 185:1705-1710, 1997.

Estes animais serão desafiados com *T. congolense* sob condições controladas e concomitantemente três gerações de animais terão seus genomas varridos com microssatélites na esperança de identificar regiões ligadas ao fenótipo de tripanotolerância²². Uma abordagem semelhante num modelo murino de tripanosomíase revelou três regiões ligadas a este fenótipo nos cromossomos 5, 17 e 1, sendo o primeiro estudo a mapear QTLs de uma doença infecciosa de importância econômica²³. Nas regiões dos cromossomos 17 e 1 ligados a resistência existem genes candidatos como o CMH e apolipoproteína A2 envolvida na atividade lítica do soro humano contra *T. brucei brucei*.

O valor para o melhoramento de animais selecionados como altos ou baixos respondedores.

Independentemente da presença de fenótipos de resistência específica a uma infecção é possível, segundo alguns autores, selecionar animais resistentes de acordo com o padrão de resposta imune a antígenos não relacionados a patógenos tais como proteínas da lisozima da clara do ovo, hemácias de carneiro e outros. Embora haja efetivamente uma correlação entre títulos de anticorpos e índices de proliferação e ativação celular com a função de macrófagos para produção de radicais tóxicos de oxigênio²⁴ não há indício de que esse suposto melhor desempenho do macrófago para um elemento imune efetor (i.e, O₂⁻) esteja correlacionado com real proteção contra um patógeno. O valor da seleção para títulos de anticorpos ou índice de proliferação celular ainda não provou ser um marcador ou indicador de resistência contra uma ou mais doenças infecciosas.

²¹ BEEBE AM, MAUZE S, SCHORK NJ, COFFMAN RL. Serial backcross mapping of multiple loci associated with resistance to *Leishmania major* in mice. *Immunity* 6:551-557, 1997

²² Kemp SJ e Teale A. 1998. Genetic basis of trypanotolerance in cattle and in mice. *Parasitol. today* 14: 450-454

²³ Kemp SJ, Iraqi F, Darvazi A, Soller M e Teale A. 1997. Localization of genes controlling resistance to trypanosomiasis in mice. *Nat. Genetics* 16: 194-196

Por outro lado, animais selecionados para resposta inflamatória mínima ou máxima apresentam níveis bastante diferentes de resistência para patógenos intracelulares não relacionados tais como *Salmonella typhimurium* e *Listeria monocytogenes*²⁵. Nesse caso alguns dos mecanismos envolvidos na resposta inflamatória podem ser os mesmos envolvidos na resposta inata a patógenos intracelulares. Mas no caso de anticorpos específicos, a associação de altos títulos com função macrofágica elevada pode apenas revelar a ação de complexos imunes, presentes em maior quantidade, sobre esta célula.

Genética funcional e análise do proteoma de células linfóides.

O termo Proteoma foi empregado pela primeira apenas recentemente e refere-se ao complemento protéico total de um genoma, ou seja, o genoma expresso. Envolve a caracterização de proteínas e a ligação destas aos seus genes codificantes (para uma revisão veja Humphrey-Smith et al.²⁶). A tecnologia básica envolvida é a eletroforese bidimensional devido ao seu grande poder de resolução de misturas complexas de proteínas. Uma análise comparativa dos géis 2-D de misturas de proteínas expressas por genomas idênticos porém estimulados diferencialmente pode revelar moléculas novas ainda não caracterizadas e envolvidas em processos celulares como, por exemplo, mecanismos imunes efetores envolvidos com resistência a parasitas. Após a eletroforese 2-D as proteínas de interesse são sujeitas a digestão e a massa dos peptídeos resultantes é submetido a "fingerprinting" por meio de espectrometria de massa em MALDI-TOF. As informações sobre massa molecular, ponto isoelétrico e posição de ligações peptídicas e sequência de amino ácidos são cruzadas em bancos de dados e a proteína é então caracterizada. A análise de proteomas, ou seja, a genética funcional promete ser uma nova fonte de prospecção de marcadores para resistência a doenças.

²⁴ Groves TC, Wilkie BN, Nennedy BW e Mallard BA. 1993. Effect of selection of swine for high and low immune responsiveness on monocyte superoxide anion production and class II MHC antigen expression. *Vet. Immunol. Immunopathol.* 36: 347-358.

²⁵ Araujo, LMM, Ribeiro OG, Siqueira M et al. 1998. Innate resistance to infection by intracellular bacterial pathogens differs in mice selected for maximal or minimal acute inflammatory response. *Eur. J. Immunol.* 28: 2913-2920.

Alguns genes candidatos do sistema imune e sua associação com resistência a doenças infecciosas e parasitárias e resposta a vacinas.

Alguns polimorfismos do sistema imune são descritos aqui que apresentam utilidade potencial em programas de SAM porque podem afetar de maneira positiva a resistência a patógenos. Muitos ainda não foram detectados em animais de pecuária mas é razoável supor que venham a ter um papel em mecanismos de resistência devido aos mecanismos gerais, trans-específicas que regem a resposta imune. Alguns polimorfismos são comentados porque podem afetar a resposta a antígenos vacinais. Os polimorfismos descritos aqui são apenas representativos do que ocorre no sistema imune e não compreendem todos aqueles que são candidatos potenciais.

Polimorfismos do CMH úteis para o melhoramento. O estudo da associação de alelos do CMH humano com doenças é uma área muito produtiva e que revela muitos aspectos sobre a interação patógeno/hospedeiro a nível molecular. Não é, porém, o escopo deste artigo enumerar todas essas associações mas apenas descrever aquelas com potencial para o melhoramento.

O mapeamento de epitopos de células T de um antígeno derivado de *Babesia bovis*, o 12D3, revelou a existência de determinantes que induzem proliferação de linfócitos T em apenas alguns bovinos, proliferação esta associada a certos alelos do BoLA-DRB3²⁷. Isso indica que a eficácia de algumas vacinas peptídicas pode ser comprometida pela restrição genética exercida pelo MHC. Por outro lado, no homem alguns epitopos derivados de *P. falciparum* e caracterizados como degenerados são capazes de ser apresentados por alelos múltiplos de HLA-A e HLA-B, alelos estes que compõem uma família de supertipos. Este dado sugere que, mesmo em populações com variabilidade grande no MHC, é factível obter vacina feita de peptídeos que seja universalmente eficaz²⁸. De fato, estudos recentes sobre a seleção de epitopos derivados do vírus da febre aftosa em bovinos revelaram que alelos múltiplos do BoLA

²⁶ Humphrey-Smith I, Cordwell SJ e Blackstock WP. 1997. Proteome research: complementarity and limitations with respect to the RNA and DNA worlds. *Electrophoresis* 18: 1217-1242.

²⁷ COURT RA, SITTE K, OPDEBEECK JP, EAST IJ. Mapping the T cell epitopes of the *Babesia bovis* antigen 12D3: implications for vaccine design. *Parasite Immunol.* 20: 1-8, 1998.

²⁸ DOOLAN DL, HOFFMAN SL, SOUTHWOOD S, et al. Degenerate cytotoxic T cell epitopes from *P. falciparum* restricted by multiple HLA-A and HLA-B supertype alleles. *Immunity* 7: 97-112, 1997

eram capazes de induzir respostas a alguns peptídeos virais conservados²⁹.

AARESTRUP e colegas³⁰ verificaram uma associação entre haplotipos de classe I do BoLA com susceptibilidade a mastite subclínica em vacas leiteiras. XU e colegas³¹ correlacionaram um polimorfismo no exon 2 do BoLA DRB3 com resistência à linfocitose persistente, condição muito prevalente em gado leiteiro. Sugerem seu emprego como marcador para melhoramento para resistência a essa condição.

Polimorfismos de citocinas, receptores de quimiocinas e resistência a infecções.

As citocinas, quimiocinas e seus receptores são moléculas envolvidas em aspectos fundamentais da regulação da resposta imune. Regulam o crescimento, diferenciação e ativação de diversas populações de células do sistema imune ou, no caso das quimiocinas, afetam a migração e ativação de células fagocíticas e linfócitos. O tipo de resposta imune induzido, se predominantemente celular ou humoral, depende tanto de aspectos qualitativos quanto quantitativos do perfil de citocinas e quimiocinas induzido. Em outras palavras, isso implica que o tipo e a quantidade de cada citocina e/ou quimiocina presente durante a indução de uma resposta imune afetará a resistência de hospedeiro a um patógeno. Qualquer polimorfismo genético que afete este perfil também afetará o grau de resistência do hospedeiro.

O TNF- α é uma citocina inflamatória com importante papel na regulação da função de células fagocíticas e assessorias da resposta imune. O TNF- α humano foi a primeira citocina para a qual um polimorfismo foi descrito que afetava seus níveis constitutivos e induzidos e que era associado com efeitos clínicos. As duas variantes genéticas conhecidas localizam-se na região do promotor e são causadas por uma única substituição de base na posição -238 ou -308. Afetam o desfecho clínico das infecções causadas pelo *Plasmodium falciparum*, *Leishmania brasiliensis*,

²⁹ VAN LIEROP MJ, NILSSON PR, WAGENAAR JP, VAN NOORT JM, CAMPELL JD, GLASS EJ, JOOSTEN I, HENSEN EJ. The influence of MHC polymorphism on the selection of T-cell determinants of FMDV in cattle. *Immunology* 84: 79-85, 1995.

³⁰ AARESTRUP FM, JENSEN NE, OSTERGARD H. Analysis of associations between major histocompatibility complex (BoLA) class I haplotypes and subclinical mastitis in dairy cows. *J. Dairy Sci.* 78: 1684-1692.

³¹ XU A, van EIJK MJT, PARK C & LEWIN HA. Polymorphism in BoLA DRB3 exon 2 correlates with resistance to persistent lymphocytosis caused by bovine leukemia virus. *J. Immunol* 151: 6977-6985, 1993.

*Mycobacterium leprae*³², entre outras infecções. Há relato de que bovinos apresentam polimorfismo semelhante com impacto sobre a parasitemia causada por infecções com o *Trypanosoma vivax*³³.

Outra citocina para a qual polimorfismos começam a ser descritos é a IL-10. Essa citocina é importante na modulação de subpopulações de linfócitos T e seus níveis determinam o perfil dominante do mecanismo efetor se humoral ou celular. No homem, variantes genéticas da IL-10 estão associadas a resistência a infecções meningocócicas fulminantes³⁴. Polimorfismos para esta citocina também foram identificados recentemente em bovinos³⁵ mas ainda não existe nenhum estudo que tenha observado correlações clínicas com os diferentes alelos. Esta citocina tem um importante papel na imunorregulação de infecções com *Babesia bovis* e *Fasciola hepática* em bovinos³⁶ o que torna seus polimorfismos candidatos a marcadores para melhoramento.

Há um relato descrevendo uma variação nos níveis de IL-1 produzidos por indivíduos normais que está relacionada com um polimorfismo em um receptor do antagonista de IL-1 (IL-1Ra)³⁷. Embora esse polimorfismo não tenha sido descrito em bovinos, tal observação tem aplicação potencial no melhoramento para resistência a carrapatos pois bovinos resistentes a esse ectoparasita produzem níveis maiores dessa

³² CABRERA M, SHAW AM, SHARPLES C, WILLIAMS H, CASTES M, CONVIT J, BLACKWELL JM. Polymorphism in tumor necrosis factor genes associated with mucocutaneous leishmaniasis. *J Exp Med* 182: 1259-1264, 1995.

McGUIRE W, HILL ASV, ALLSOP CEM, GREENWOOD BM, KWIATKOWSKI D. Variation in the TNF- α promoter region associated with susceptibility to cerebral malaria. *Nature* 371: 508-511, 1994.

ROY S, MCGUIRE W, MASCIE-TAYLOR CG, SAHA B, HAZRA SK, HILL AV, KWIATKOWSKI D. Tumor necrosis factor promoter polymorphism and susceptibility to lepromatous leprosy. *J Infect Dis* 176:530-532, 1997

³³ ILRI Annual Report. 1997.

³⁴ WESTENDORP RG, LANGERMANS JA, HUIZINGA TW, VERWEIJ CL, STURK A. Genetic influence on cytokine production in meningococcal disease. *Lancet* 349:1912-1913, 1997

³⁵ BEEVER JE, FISHER SR, LEWIN HA. Polymorphism identification in the ACADM, AT3, IL10, MYOG and TSHB genes of cattle. *Anim Genet* 28:373-374, 1997

³⁶ CHITKO-McKOWN CG, RUEF BJ, RICE-FICHT AC, BROWN WC. Interleukin-10 downregulates proliferation and expression of interleukin-2 receptor p55 chain and interferon-gamma, but not interleukin-2 or interleukin-4, by parasite-specific helper T cell clones obtained from cattle chronically infected with *Babesia bovis* or *Fasciola hepatica*. *J. Interferon Cytokine Res.* 15: 915-922, 1995.

³⁷ DANIS VA, MILLINGTON M, HYLAND VJ, GRENNAN D. Cytokine production by normal human monocytes: intersubject variation and relationship to an IL-1 receptor antagonist (IL-1Ra) gene polymorphism. *Clin Exp Immunol* 99: 303-310, 1995

citocina³⁸.

A análise de segregação em indivíduos infectados com *Schistosoma mansoni* indicou a existência de um gene que afeta a produção de IL-5, uma citocina com papel importante na regulação da resposta imune a este parasita³⁹ e outros helmintos inclusive os de importância para a produção animal. Um estudo envolvendo a varredura dos genomas de famílias apresentando intensidade variável de infecção por esse parasita revelou uma região no cromossomo 5 ligado a esse fenótipo, região esta que codifica a IL-5 e outras citocinas de importância na imunorregulação de parasitas intestinais⁴⁰.

Um polimorfismo que afeta o receptor celular do interferon-gama no homem está associado a susceptibilidade genética a infecções generalizadas por micobactérias que normalmente são pouco patogênicas. Esse defeito impede a ativação de macrófagos para a produção de TNF- α ⁴¹. Há indícios indiretos de que polimorfismos que afetam os níveis de IFN- γ possam ocorrer em bovinos⁴² pois há uma variação intraespecífica muito grande para a produção dessa citocina após estímulo padronizado.

As quimiocinas C-C participam de um importante mecanismo protetor contra a infecção pelo HIV pois indivíduos hemofílicos que são hiperprodutores de beta quimiocinas não desenvolveram imunodeficiência apesar de terem sido inoculados continuamente com este vírus⁴³. Outra quimiocina, a SDF-1, também apresenta polimorfismo que está associado com um adiamento significativo do desenvolvimento

³⁸ RAMACHANDRA RN, WIKEL SK. Effects of *Dermacentor andersoni* (Acari: Ixodidae) salivary gland extracts on *Bos indicus* and *B. taurus* lymphocytes and macrophages: in vitro cytokine elaboration and lymphocyte blastogenesis. *J. Med. Entomol.* 32:338-345, 1995.

³⁹ RODRIGUES V JR, ABEL L, PIPER K, DESSEIN AJ. Segregation analysis indicates a major gene in the control of interleukine-5 production in humans infected with *Schistosoma mansoni* *Am J Hum Genet* 59:453-461, 1996.

⁴⁰ MARQUET S, ABEL L, HILLAIRE D, DESSEIN H, KALIL J, FEINGOLD J, WEISSENBACH J, DESSEIN AJ Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Nat Genet* 14:181-184, 1996.

⁴¹ NEWPORT MJ, HUXLEY CM, HUSTON S, HAWRYLOWICZ CM, OOSTRA BA, WILLIAMSON R, LEVIN M. A mutation in the interferon-gamma-receptor gene and susceptibility to mycobacterial infection. *N Engl J Med* 335:1941-1949, 1996.

⁴² EDFORS-LILJA I, WATTRANG E, MAGNUSSON U, FOSSUM C. Genetic variation in parameters reflecting immune competence in swine. *Vet Immunol Immunopathol* 40: 1-16, 1994.

⁴³ ZAGURY D, LACHGAR A, CHAMS V, et al. C-C chemokines, pivotal in protection against HIV type 1 infection. *Proc Natl Acad Sci USA* 95: 3857-3861, 1998.

de imunodeficiência em indivíduos infectados pelo HIV⁴⁴. Por outro lado, mutações nos receptores celulares de quimiocinas, CCR5 e CCR2, também têm um papel crítico no desfecho clínico da infecção pelo HIV⁴⁵. Tais associações tomam os polimorfismos das quimiocinas e de seus receptores candidatos a marcadores para o melhoramento para resistência ao vírus da leucose bovina, um retrovírus que possivelmente emprega mecanismos de invasão e virulência semelhantes ao HIV e que causa perdas importantes ao setor de gado leiteiro.

Polimorfismos de alotipos de imunoglobulinas e resistência a infecções. Como o homem, bovinos apresentam polimorfismos na região constante de suas imunoglobulinas (alotipos) e existem alguns estudos que associam alotipos distintos com susceptibilidade a infecções. O alelo IgG2b, por exemplo, está associado com maior susceptibilidade a mastite clínica enquanto que o alelo IgG2a está associado a resistência a mastite⁴⁶. No homem certos alotipos estão relacionados com resposta diminuída ou ausente para antígenos de carboidratos.

Polimorfismos do NRAMP1 e resistência a infecções intracelulares. Em várias espécies já foi documentada a resistência, chamada natural, para infecções com patógenos intracelulares. SKAMENE, GROS e colaboradores demonstraram que essa resistência é mediada por uma proteína de membrana expressa exclusivamente em macrófagos e que pertence à família das proteínas de transporte⁴⁷. Portadores de variantes genéticas dessa proteína são suscetíveis a essa classe de patógenos porque, provavelmente, a proteína defeituosa não consegue manter no macrófago um microambiente hostil ao patógeno intracelular. Essa proteína já foi descrita em

⁴⁴ WINKLER C, MODI W, SMITH MW, ET AL. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort. *Science* 279:389-393, 1998.

⁴⁵ SMITH MW, DEAN M, CARRINGTON M, ET AL. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* 277:959-965 1997.

⁴⁶ KELM SC, DETILLEUX JC, FREEMAN AE, KEHRLI ME JR, DIETZ AB, FOX LK, BUTLER JE, KASCKOVICS I, KELLEY DH Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. *J Dairy Sci* 80:1767-1775, 1997.

⁴⁷ VIDAL SM, MALO D, VOGAN K, SKAMENE E, GROS P Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell* 73:469-485, 1993.

bovinos⁴⁸ e está associada com mecanismos efetores para o controle de *Brucella abortus*, *Mycobacterium bovis* BCG e *Salmonella* sp. Bovinos apresentam polimorfismos para este marcador cujas associações com resistência a infecções intracelulares começam a ser estudadas (SKAMENE, comunicação pessoal). No homem recentemente foi descrito pela primeira vez uma relação de variantes do gene NRAMP 1 com susceptibilidade a tuberculose⁴⁹. Em aves a resistência à salmonelose também está ligada ao NRAMP 1⁵⁰.

Conclusões

Os genes envolvidos em defesa do hospedeiro são os que mais rapidamente evoluem e por isso o sistema imune é fonte de grande número de polimorfismos que podem ser empregados em programas de melhoramento assistida por marcadores. Uma vez identificado um marcador ligado a resistência, o impacto fisiológico do produto do gene escolhido sobre outros patógenos que não o envolvido no fenótipo sob avaliação deve ser levado em consideração. O animal universalmente resistente dificilmente será obtido. Ao mesmo tempo, novos marcadores identificados por meio de estudos de ligação ou análise de proteomas podem revelar componentes imunes ainda desconhecidos nos mecanismos efetores de proteção e na interação parasito/hospedeiro. Com o advento dos mapas de genomas baseados em microssatélites os estudos de ligação estão bastante facilitados e devem render novos marcadores assim como confirmar a participação de genes candidatos nos fenótipos de interesse.

⁴⁸ FENG J, LI Y, HASHAD M, SCHURR E, GROS P, ADAMS LG, TEMPLETON JW. Bovine natural resistance associated macrophage protein 1 (Nramp 1) gene. *Genome Res* 6: 956-964, 1996.

QURESHI T, TEMPLETON JW, ADAMS LG. Intracellular survival of *Brucella abortus*, *Mycobacterium bovis* BCG, *Salmonella dublin* and *Salmonella typhimurium* in macrophages from cattle genetically resistant to *Brucella abortus*. *Vet Immunol Immunopathol* 50: 55-65, 1997.

⁴⁹ BELLAMY R, RUWENDE C, CORRAH T, MCADAM KP, WHITTLE HC, HILL AV Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. *N Engl J Med* 338:640-644, 1998.

⁵⁰ Hu J, Bumstead N, Barrow P. et al. 1997. Resistance to salmonellosis in the chicken is linked to NRAMP1 and TNC. *Genome Research* 7: 693-704.

MÉTODOS ESTATÍSTICOS PARA IDENTIFICAÇÃO DE QTLS EM ANIMAIS

Mario Luiz Martinez
Embrapa Gado de Leite

Delineamentos Experimentais

A.1. Introdução

Um grande número de delineamentos experimentais e de metodologias estatísticas tem sido proposto para se detectar genes que afetam características quantitativas com o auxílio de marcadores genéticos. Todos os delineamentos sugeridos têm vários elementos em comum. Considera-se que um *locus* de característica quantitativa (QTL) esteja ligado ao *locus* de um marcador com uma frequência de recombinação r . A priori, assume-se que apenas dois alelos estão segregando na população para ambos o *locus* do marcador, M e o *locus* do QTL, Q. Os genótipos do marcador serão denominados de M1M1, M1M2 e M2M2. Os genótipos do QTL serão denominados por Q1Q1, Q1Q2 e Q2Q2, com efeitos esperados de a , d e $-a$, respectivamente, sobre a característica quantitativa. Assim, se um indivíduo for heterozigoto para ambos os *loci*, metade da progênie receberá o alelo M1 e a outra metade o alelo M2. Desde que M e Q estejam ligados, aquelas progênies que receberem M1 também receberá Q1, enquanto aquelas que receberem M2 também receberá Q2, exceto pelos indivíduos recombinantes que receberão M1 com Q2 ou M2 com Q1. Dessa forma, o efeito do QTL pode ser detectado através da comparação das médias dos grupos de progênies que receberem alelos marcadores alternativos de pais heterozigotos. Os diversos delineamentos propostos diferem nos métodos utilizados para se "criar" pais heterozigotos para ambos os *loci* e os cruzamentos realizados.

A.2. Delineamento experimental para detecção de QTL em cruzamentos de linhagens consangüíneas.

A maioria das análises realizadas para se detectar QTL tem sido baseada em cruzamentos planejados, embora alguns estudos tenham sido realizados com populações de animais existentes, especialmente gado de leite. O caso mais simples é do cruzamento entre linhagens consangüíneas. Como mencionado anteriormente, o primeiro passo é cruzar duas linhagens que diferem em marcadores genéticos para produzir progênes heterozigotas F-1. Após este passo, as seguintes progênes têm sido consideradas para análises:

- i) Retrocruzamento (BC) de indivíduos F-1 com uma das linhagens parentais;
- ii) Indivíduos F-2 produzidos pelo cruzamento entre indivíduos F-1 ou autofecundação de F-1;
- iii) Linhagens consangüíneas recombinates (RILF) produzidas pela autofecundação de indivíduos F-2, ou pelo acasalamento de irmão com irmãs, nos casos em que autofecundação não é possível (RILF);
- iv) Linhagens consangüíneas recombinantes (RILB) produzidas com indivíduos retrocruzados, ou acasalamentos de irmãos com irmãs (RILB);
- v) Linhagens duplas de haplóides (DH) produzidas pela autofecundação de duplos haplóides derivados de F-1;
- vi) Progênes testes (TC) produzidas pelo acasalamento de indivíduos F-1 com uma terceira linhagem consangüínea.

Existem outros delineamentos mais complicados, mas todos são variações dos delineamentos acima listados.

A escolha do delineamento a ser usado deve levar em consideração os aspectos biológicos, econômicos, genéticos e estatísticos. Considerações biológicas estão relacionadas com o fato de que não são todos os delineamentos que podem ser usados com todas as espécies, por exemplo, o DH. Para certas espécies com quase total autofecundação é mais fácil produzir grande número de indivíduos F-2 do que BC ou TC que requerem fertilização cruzada; para espécies em que os acasalamentos só

ocorrem dentro da população, a consangüinidade pode resultar na redução do valor adaptativo devido à presença de genes deletérios recessivos. Semelhantemente, o valor econômico dos delineamentos que são possíveis pode ser bem diferente. Por exemplo; se o objetivo é a introgressão de genes para uma característica específica de uma variedade selvagem em uma cultivar previamente selecionada, o BC com a cultivar selecionada terá um valor econômico muito maior do que o F-2. As considerações genéticas estão associadas com os parâmetros genéticos que desejamos estimar. Assim, os delineamentos BC ou TC não permitem estimar relações de dominância, mas já o F-2 sim. As considerações estatísticas estão relacionadas com os delineamentos que maximizem o poder de detectar os efeitos do QTL dentro de uma determinada situação. Por exemplo, nos delineamentos RI e DH todos os indivíduos dentro de uma linhagem terão o mesmo genótipo. Assim, será necessário genotipar apenas um único indivíduo de cada linhagem, enquanto o desempenho fenotípico de todos os indivíduos pode ser utilizado para determinar o efeito do QTL. Dessa forma, o poder estatístico será maior por indivíduo genotipado.

A.3. Delineamento de retrocruzamento (BC)

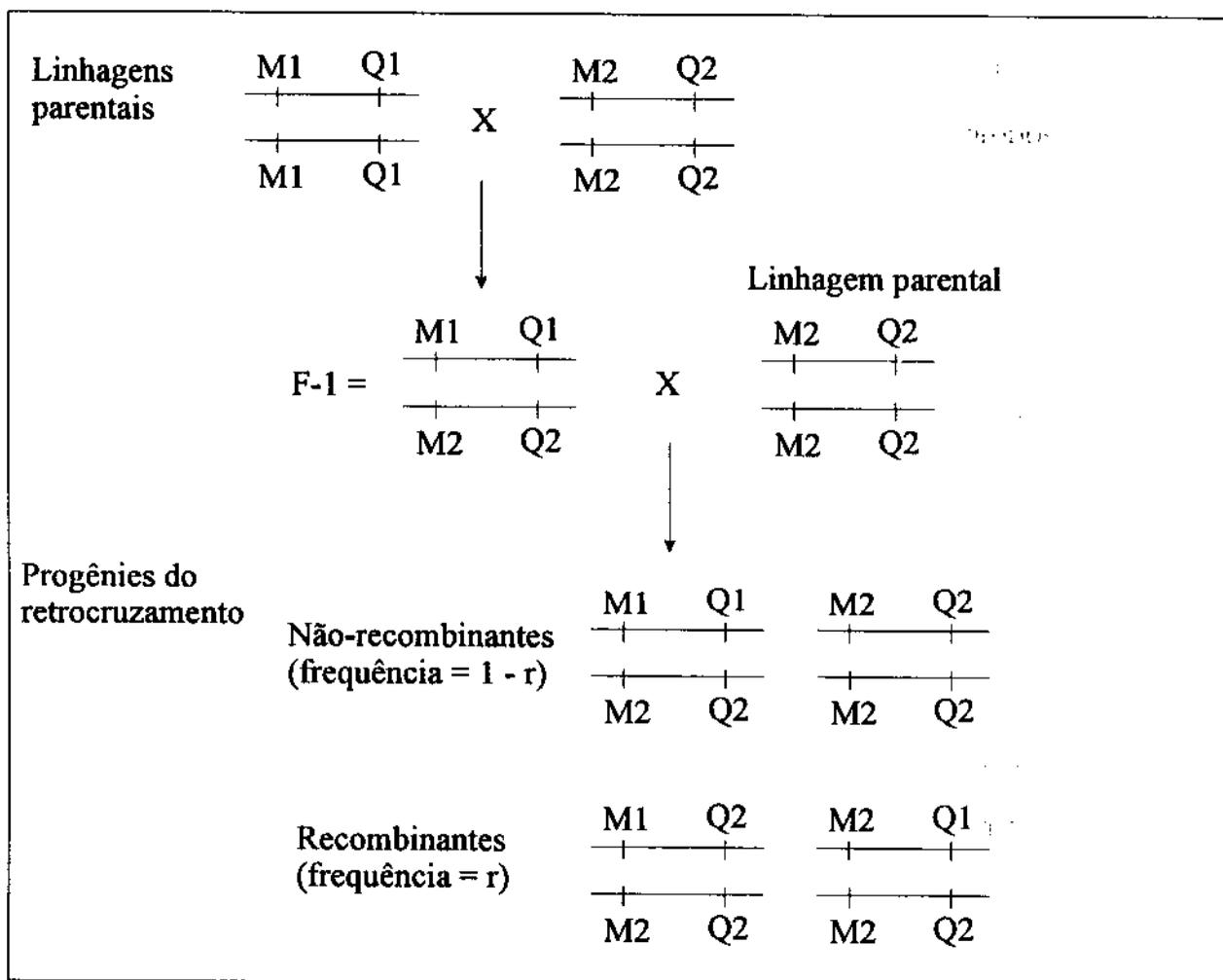
Na Figura 1 está ilustrado o delineamento de retrocruzamento com uma das raças parentais. A produção de F-1 é obtida através do acasalamento entre duas linhagens que diferem em genótipos do QTL e do marcador. Normalmente, assume-se que as duas linhagens parentais são homocigotas para os alelos alternativos de ambos os *loci*. Dessa forma, todos os indivíduos F-1 terão o mesmo genótipo heterocigoto. Estes indivíduos F-1 são então acasalados com uma das linhagens parentais. A base genética destes indivíduos é composta de $\frac{3}{4}$ do pai recorrente e $\frac{1}{4}$ do outro pai. As progênes BC são então divididas em dois grupos baseados nos genótipos dos marcadores. À semelhança do que será considerado nos outros delineamentos experimentais, todos os demais *loci* que não estiverem ligados ao marcador genético em estudo, estarão distribuídos aleatoriamente entre os grupos dos genótipos marcadores. Quando apenas um marcador é considerado, haverá apenas dois grupos genéticos marcadores para este tipo de delineamento e apenas um contraste poderá

ser testado. Este contraste é a diferença entre as médias dos dois grupos de progênie que têm genótipos marcadores M1M2 e M2M2. Para este delineamento, a maioria dos estudos tem usado variações simples do seguinte modelo:

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

aonde Y_{ijk} é a medida da característica para o indivíduo k, do "bloco" j e de genótipo i; A é o efeito do genótipo marcador i (M1M2 ou M2M2); B é o efeito do "bloco" j e e_{ijk} é o efeito residual ao acaso associado a cada indivíduo.

Figura 1. Delineamento de retrocruzamento (BC).



O efeito do "bloco" representa todos os efeitos de meio ambiente, que os grupos de indivíduos possam ter em comum, tais como linhagens, área, rebanho, estação de crescimento, etc. Apenas por simplicidade não se incluiu no modelo o efeito da média geral. Este efeito pode ser considerado incluído no efeito de bloco. Observa-se no modelo acima que se assume que apenas marcador e bloco estejam afetando a média da característica.

O teste de significância para o efeito de genótipo pode ser realizado por dois métodos: análise de variância (ANOVA), ou o teste de t de um contraste estimável. No caso da ANOVA, a razão entre o quadrado médio do marcador e o quadrado médio do resíduo é calculada e sob a hipótese de nulidade (não há QTL segregando), esta razão terá uma distribuição central de F. Um desvio significativo desta forma estatística da distribuição central de F é indicativo de QTL segregando. No caso do teste t, a diferença entre as médias das duas classes de genótipos é dividida pelo erro padrão deste contraste. Sob a hipótese de nulidade, essa estatística terá uma distribuição central de t com graus de liberdade igual ao número total de indivíduos incluídos na comparação menos dois.

A vantagem principal deste modelo linear de análise é que ele pode ser executado pela maioria dos pacotes estatísticos existentes. As desvantagens são:

- 1) efeito estimado está confundido com o efeito da recombinação entre o QTL e o marcador, e conseqüentemente a estimativa do efeito do QTL é viesada;
- 2) Assume-se que os resíduos seguem uma distribuição normal;
- 3) Assume-se que as variáveis independentes não são correlacionadas. Portanto, o método não é apropriado para múltiplos marcadores que estão ligados, que são correlacionados;
- 4) O método não diferencia entre o efeito de um QTL ligado e o efeito pleotrópico do marcador genético.

No Quadro 1 são apresentadas as possibilidades dos genótipos do QTL dados o genótipo do marcador e o valor da característica para cada genótipo do marcador. Assumindo-se ligação incompleta entre QTL e marcador, cada classe do genótipo do marcador consiste de dois genótipos do QTL com freqüências r e $1 - r$. O valor

esperado da característica para cada classe de genótipo do marcador é calculado como a soma das probabilidades condicionais de cada genótipo do QTL vezes o seu valor. A estimativa para o contraste entre as médias do genótipo heterozigoto e do homozigoto do marcador é calculada como a diferença entre os seus valores esperados. Este valor é: $(1-2r)(d+a)$. O termo $1 - 2r$ aparece na maioria dos contrastes dos genótipos dos marcadores. Quando $r = 0$, o contraste será $d + a$, e significa um efeito total do QTL, e quando $r = 0,5$, o contraste é zero, e significa que não há ligação entre marcador e QTL. Como já mencionado anteriormente, o delineamento BC tem apenas um contraste que é estimável. Desde que este contraste seja função de a , d e r , estes parâmetros estão confundidos no modelo linear de análises e portanto não podem ser estimados separadamente.

Quadro 1. Probabilidades dos genótipos no delineamento BC e valores esperados da característica quantitativa.

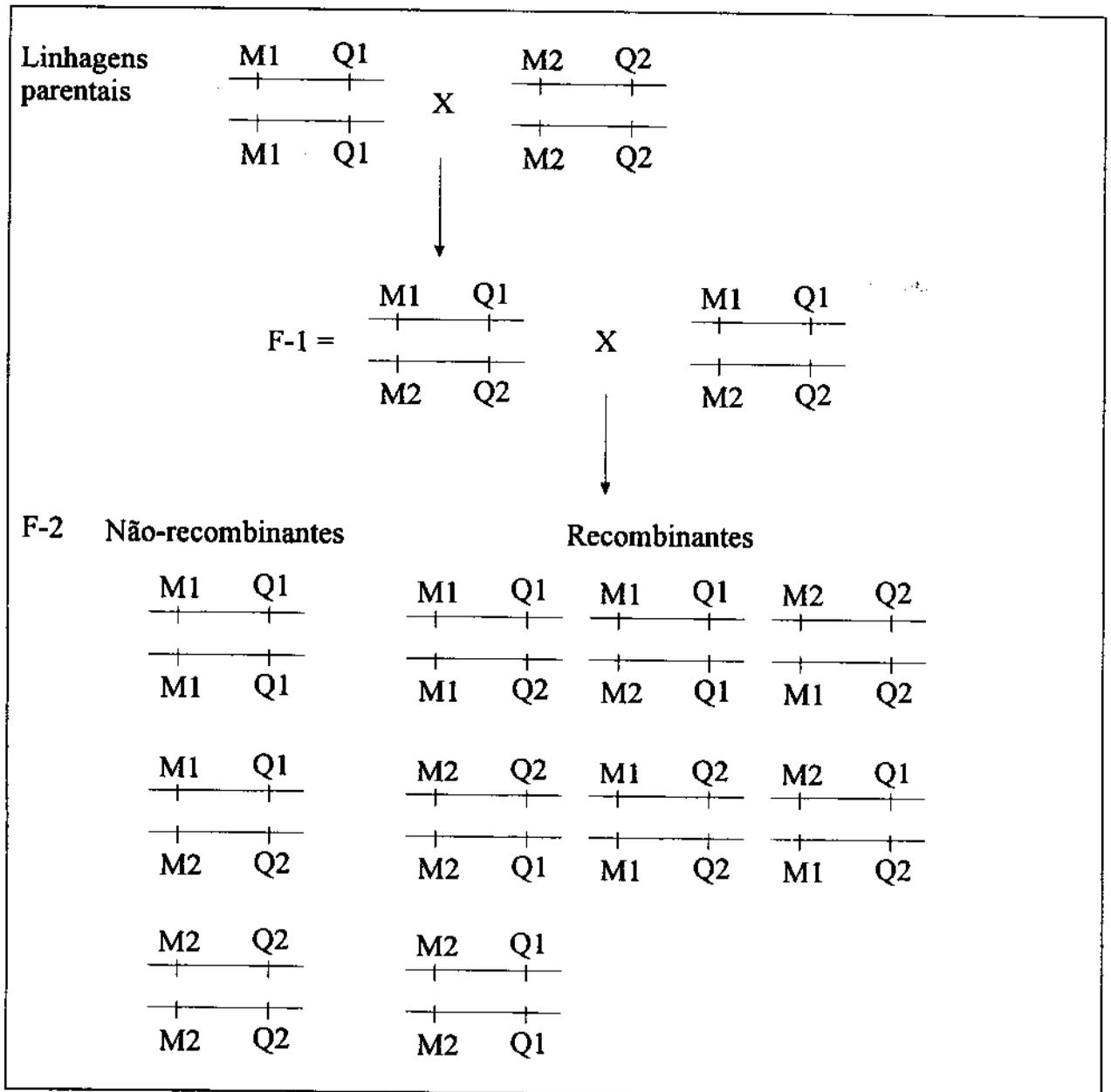
Genótipo do Marcador	Genótipo do QTL	Probabilidade do QTL, dado o genótipo do marcador	Valor da característica	Valor esperado da característica para o genótipo do marcador
M1M2	Q1Q2	$1 - r$	d	$d - r(d + a)$
	Q2Q2	r	$-a$	
M2M2	Q2Q2	$1 - r$	$-a$	$-a + r(d + a)$
	Q1Q2	r	d	

A.4. Delineamento de F-2

Na Figura 2 está ilustrado o delineamento de F-2. À semelhança do delineamento BC, dois pais homozigotos são cruzados para se obter o indivíduo heterozigoto F-1. As progênies F-2 são produzidas através da autofecundação dos indivíduos F-1 ou pelo cruzamento entre eles. Como resultado, as progênies F-2 serão de um dos três genótipos do marcador. Neste delineamento, a recombinação em

qualquer um dos cromossomos afetará o valor esperado da característica. No Quadro 2 são apresentados os genótipos possíveis, suas probabilidades e seus valores esperados para a característica quantitativa.

Figura 2. Delineamento de F-2.



Quadro 2. Probabilidades dos genótipos no delineamento F-2 e valores esperados da característica quantitativa.

Genótipo do Marcador	Probabilidade do genótipo marcador	Genótipo do QTL	Probabilidade do QTL, dado o genótipo do marcador	Valor da característica	Valor esperado da característica para o genótipo do marcador
M1M1	1/4	Q1Q1	$(1 - r)^2$	a	$a(1-2r) + 2dr(1-r)$
		Q1Q2	$2r(1 - r)$	d	
		Q2Q2	r^2	-a	
M1M2	1/2	Q1Q1	$r(1 - r)$	a	$d(1 - 2r + r^2)$
		Q1Q2	$1 - 2r + r^2$	d	
		Q2Q2	$r(1 - r)$	-a	
M2M2	1/4	Q1Q1	r^2	a	$-a(1-2r) + 2dr(1-r)$
		Q1Q2	$2r(1 - r)$	d	
		Q2Q2	$(1 - r)^2$	-a	

Neste delineamento a probabilidade de indivíduos heterozigotos para o *locus* marcador é 0,5, mas a maior parte da informação relativa à detecção do QTL está contida nos indivíduos homozigotos. Considerando-se ligação incompleta, são possíveis três genótipos do QTL para cada um dos três genótipos do marcador.

O teste para se detectar a presença de QTL pode ser realizado pela ANOVA, colocando-se no modelo todos os três genótipos. Além disso, vários contrastes podem ser testados através do teste de t. O contraste de maior interesse é o realizado entre as médias dos dois genótipos homozigotos do marcador. O valor deste contraste é: $2a(1 - 2r)$. É importante notar que à semelhança do delineamento BC, este contraste também inclui o termo $1 - 2r$, e portanto os efeitos de a e r estão confundidos. No F-2, este contraste não é função de d. O teste de t pode também ser utilizado para se testar a significância dos contrastes entre cada um dos genótipos homozigotos com o genótipo heterozigoto. A estimativa da diferença entre o homozigoto M1M1 e o heterozigoto M1M2 será $a(1 - 2r) - d(1-2r)^2$. Neste contraste os efeitos de a, d e r estão

confundidos e portanto é de pouco interesse prático. Todavia, este delineamento permite se testar o efeito de dominância através do contraste:

$$\frac{1}{2}(M1M1 + M2M2) - (M1M2) = -d(1 - 2r)^2.$$

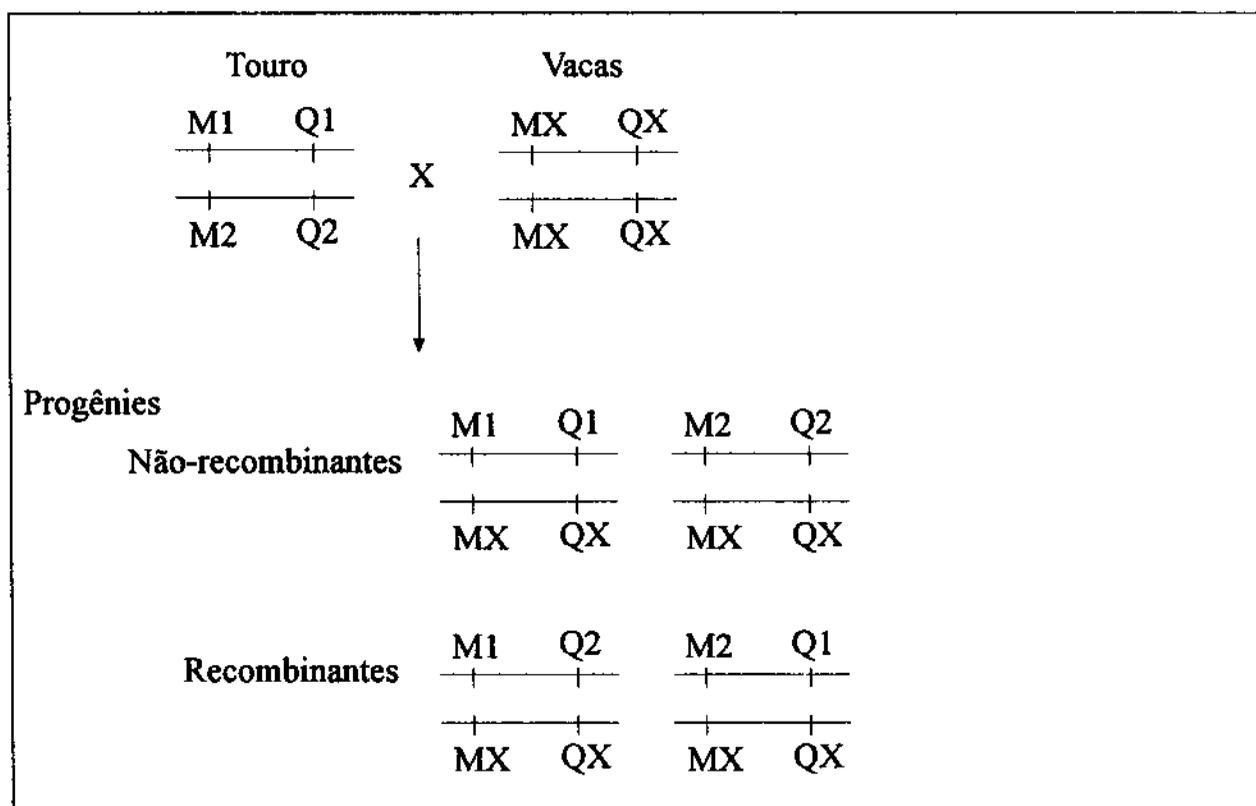
Observa-se que este contraste é também uma função de $1 - 2r$, mas não é dependente de a . Embora se possa testar para os efeitos aditivos e de dominância do QTL, ambas as estimativas não podem ser estimadas sem viés através das análises de modelo linear.

A.5. Delineamento para detecção de QTL em grandes populações segregantes

Em humanos, na maioria das espécies domésticas de animais e em árvores frutíferas perenes, é impraticável se produzir linhagens consanguíneas, que são a base dos delineamentos descritos em 3 e 4. Nessas situações, as análises têm sido executadas em famílias existentes dentro de uma população. Três tipos de análises têm sido propostas para estes casos: a análise de par de irmãos para o caso de muitas famílias pequenas; a análise do delineamento de irmãos completos para o caso de grandes famílias de irmãos completos e a análise de meio-irmãs ou delineamento de filhas para o caso de grandes famílias de meio-irmãs.

O delineamento de filhas foi inicialmente proposto por Neimann-Soressen e Roberson (1961) e tem sido usado principalmente em gado de leite em que um touro pode ter centenas de progênies com medidas em características quantitativas de interesse econômico. Na Figura 3 é ilustrada o delineamento de filhas para o caso de uma única família.

Figura 3. Delineamento de filhas.



Neste delineamento as filhas de um touro heterocigoto para o marcador serão genotipadas e mensuradas para as características quantitativas de interesse. Considerando-se que o genótipo da mãe é geralmente desconhecido e difere de animal para animal, os alelos do marcador e do QTL são representados por MX e QX, respectivamente.

Na situação em que apenas as progênies de um touro é considerada, a detecção de um QTL ligado a um marcador pode ser realizada utilizando o modelo dado na equação (1). A diferença é que A_i representa agora o efeito da substituição de alelos recebidos do touro. Se assumirmos que apenas os mesmos dois alelos do QTL estão presentes na população de vacas, com frequência p e $1 - p$, então é possível se calcular o valor esperado da característica para cada grupo de progênie que recebeu os alelos alternativos do marcador. O valor esperado para cada grupo é mostrado no Quadro 3. O contraste entre os dois grupos de progênies é: $a(1 - 2r) + d(1 - 2r)(1 - 2p)$.

Se a frequência dos dois alelos do QTL for igual a 0,5, então o contraste se torna igual a: $a(1 - 2r)$. Para o caso de ligação total ($r = 0$), o contraste se torna: $a + d(1 - 2p)$. Se definirmos a frequência do segundo alelo como $q = 1 - p$, o contraste se torna igual a: $a + d(q - p)$, que é a fórmula geral para o efeito da substituição de alelos (Falconer, 1981).

Embora um QTL possa estar segregando na população, um touro em específico pode ser homozigoto para o QTL. Devido a isto, a maioria dos estudos têm sido baseada na análise de vários touros heterozigotos ao mesmo tempo. Neste caso a análise utilizando-se a equação (1) pode conduzir a conclusões erradas, pois mesmo que vários indivíduos possam ser heterozigotos para um marcador ligado ao QTL, a relação de ligação pode ser diferente para cada indivíduo. Dessa forma, quando somados os efeitos de todas as progênie nos dois grupos de alelos do marcador, pode não se observar nenhum efeito associado aos alelos do marcador. O modelo apropriado nessa situação de múltiplas famílias é:

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

onde S_i é o efeito do touro i , A_{ij} é o efeito do alelo j , aninhado dentro do touro i , e os demais termos são os mesmos da equação (1). Neste caso, o grupo de progênie que recebeu o alelo M1 do touro será comparado com o grupo de progênie que recebeu o alelo M2. Se as pressuposições feitas anteriormente se mantiverem e se a distribuição das mães entre os dois grupos forem ao acaso, então a diferença entre os dois grupos de progênie para a característica quantitativa será devido ao QTL ligado ao marcador M do touro. Isto assume que a origem do alelo do marcador pode ser determinada para as filhas. O teste de detecção do QTL pode ser realizado pela ANOVA. Considerando-se a razão entre os quadrados médios do efeito do marcador e do resíduo.

Quadro 3. Probabilidades dos genótipos no delineamento de filhas e valores esperados da característica quantitativa.

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

Se assumirmos que há apenas dois alelos do QTL presentes com a mesma frequência na população, e a distribuição dos genótipos em equilíbrio Hardy-Weinberg, apenas a metade dos touros serão heterozigotos para o QTL. Dessa forma, a variância do termo A_{ij} será $a^2/4$. Além do teste pela ANOVA, o teste do qui-quadrado pode ser utilizado para se testar a significância do efeito do QTL. A diferença entre as médias dos dois grupos de progênies (M1 e M2) é calculada e dividida pelos seus erros padrões. Sob a hipótese de nulidade, a soma de quadrado desta estimativa tem uma distribuição de qui-quadrado com o número de graus de liberdade igual ao número de touros. O poder estatístico para este tipo de delineamento foi calculado por Soller e Genizi (1978) pela ANOVA e por Weller et al. (1990) pelo qui-quadrado.

Neste tipo de delineamento não é possível determinar inequivocadamente a origem do alelo marcador para todas as progênies, ao contrário do que ocorre com o cruzamento entre linhagens consanguíneas. Se uma filha tem o mesmo genótipo do pai, ela pode ter recebido qualquer um dos alelos de qualquer um dos pais. Quando apenas dois alelos estão presentes na população, a metade das filhas terá o mesmo

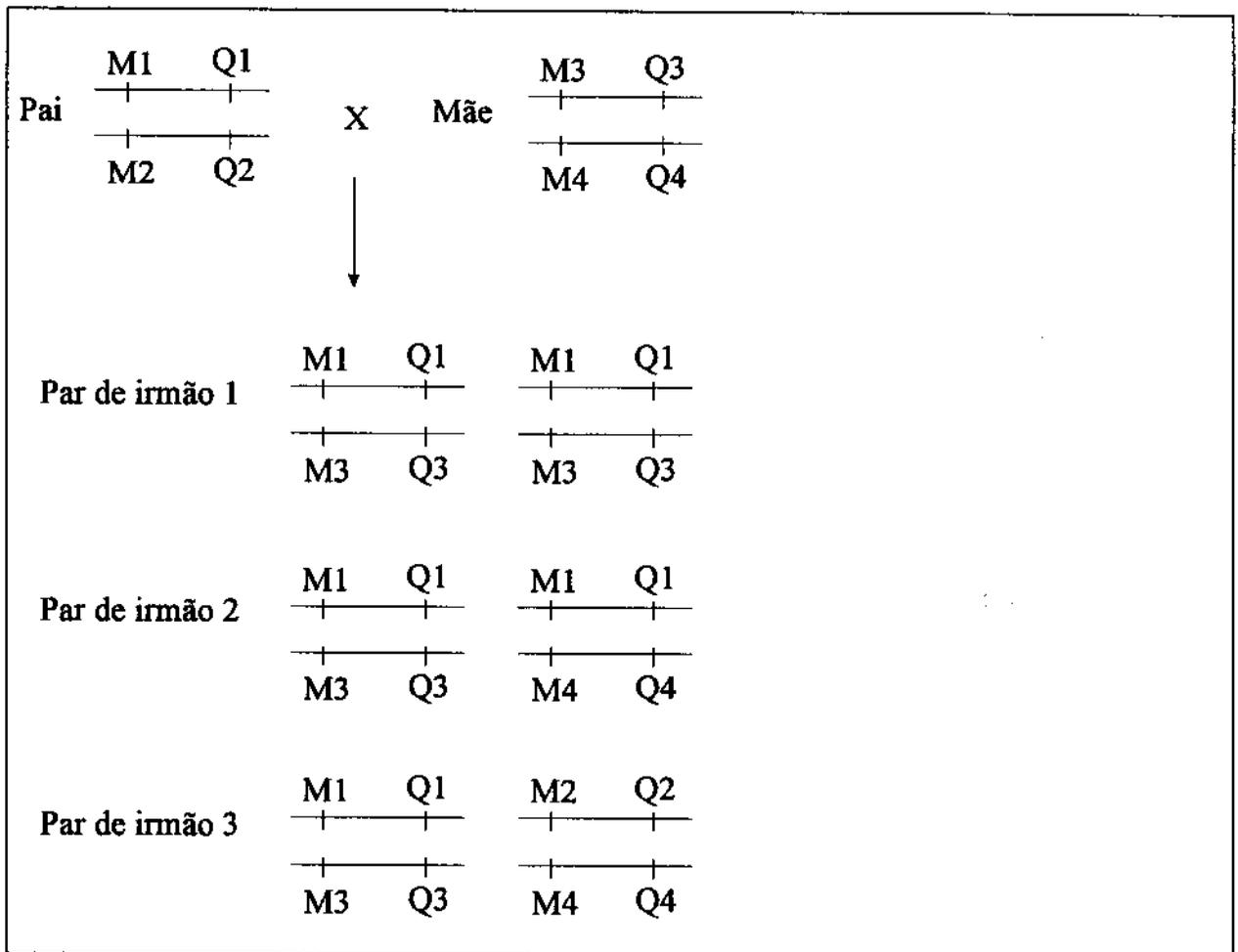
genótipo do pai, independente da frequência dos alelos entre as mães. Em geral, a frequência esperada de indivíduos para os quais a origem do alelo paterno pode ser determinada é $1 - (p + q)/2$, onde p e q são as frequências dos dois alelos paternos do marcador. Para *loci* multialélicos, $p + q$ pode ser bem menor do que 1. Em geral, as informações dessas progênes não são utilizadas nas análises. Embora, alguma informação adicional possa ser extraída desses indivíduos, ela será consideravelmente menor do que as obtidas dos indivíduos cuja origem dos alelos é conhecida.

A análise de uma “grande” família de irmãos-completos pode ser realizada no melhor dos casos, quando os dois pais tem três ou quatro alelos marcadores diferentes e as progênes podem ser divididas em quatro grupos diferentes de genótipos do marcador. O teste para se detectar QTL pode ser realizado pela ANOVA considerando-se a razão entre o efeito do alelo paterno e o do resíduo. É possível também se estimar o efeito da substituição de alelo para cada pai, mas não um efeito geral para todos os pais. Uma estimativa da variância genética total associada com o marcador também pode ser obtida. A ANOVA com efeito do marcador dentro de família não é apropriada para famílias pequenas de irmãos completos, como ocorre em humanos, devido ao pequeno número de graus de liberdade. Uma análise alternativa é dada pelo delineamento seguinte.

A.6. Delineamento para detecção de QTL em famílias pequenas de populações segregantes

Haseman e Elston propuseram o método de análise de par de irmãos em 1972. Este método é mais apropriado à população de humanos no qual famílias pequenas de irmãos-completos são mais comuns. O delineamento básico é mostrado na Figura 4. Na situação mais simples, os dois pais têm quatro alelos marcadores diferentes: M1, M2, M3 e M4. Assume-se também que este *locus* está ligado a um QTL heterozigoto para ambos os pais. Na figura são apresentados apenas dois tipos de par de irmãos. Para efeito de simplicidade, apenas os não-recombinantes são apresentados.

Figura 4. Delineamento de par de irmãos-completos.



No par de irmão 1, os dois alelos marcadores de ambos os irmãos são idênticos por descendência (IBD). Ambos os irmãos receberam os mesmos dois alelos marcadores de ambos os pais. No par de irmão 2, ambos os irmãos receberam o alelo M1 do pai, mas alelos diferentes da mãe. Assim, um par de alelos é IBD, enquanto o outro não é. No par de irmão 3, nenhum dos alelos marcadores é IBD. Assumindo-se que o *locus* marcador está ligado ao QTL como mostrado na Figura 4, os indivíduos com mais *loci* marcador IBD deverá também ser mais similar para a característica quantitativa. À semelhança do caso de grandes famílias de meio-irmãos, a análise deve levar em consideração a estrutura da família, ou seja, um QTL ligado ao marcador pode

não resultar em um efeito geral do genótipo marcador. Haseman e Elston (1972) propuseram o seguinte modelo estatístico:

$$\begin{aligned} X_{1j} &= \mu + g_{1j} + e_{1j} \\ X_{2j} &= \mu + g_{2j} + e_{2j} \end{aligned} \quad (3)$$

Inicialmente, assume-se que os genótipos do QTL podem ser observados e posteriormente será considerada a análise baseada no marcador ligado ao QTL. Na equação acima, X_{1j} e X_{2j} são os valores da característica para os irmãos 1 e 2 da família j , μ é a média geral, g_{1j} e e_{1j} são os efeitos direto do QTL e do resíduo para o irmão i . Nos modelos anteriores, consideraram-se as diferenças entre as médias dos genótipos. Neste caso, quando há muitas famílias pequenas isso não é viável. Como já mencionado anteriormente, algumas das diferenças serão positivas e outras negativas. Devido a isto, Haseman e Elston (1972) basearam suas análises no quadrado das diferenças, que são sempre positivos.

Considere $Y_j = (X_{1j} - X_{2j})^2 = (g_{1j} + e_{1j} - g_{2j} - e_{2j})^2$. Se os dois irmãos têm ambos os alelos QTL IBD, então $g_{1j} = g_{2j}$ e $Y_j = (e_{1j} - e_{2j})^2$, que será menor que Y_j para irmãos que recebam diferentes alelos do QTL de seus pais. Quando há co-dominância ou dominância parcial, a estimativa de Y_j para indivíduos que recebam um alelo IBD será intermediária entre indivíduos com zero e dois alelos IBD. A presença de um QTL ligado pode ser testada através da seguinte regressão:

$$Y_j = \alpha + \beta \pi_j \quad (4)$$

onde π_j é a fração de alelos marcadores (0, 1/2 ou 1) IBD para o par de irmãos j ; α é o ponto de interseção no eixo y ; e β é o coeficiente de regressão, que será sempre negativo. A detecção do QTL ligado ao marcador pode ser testada através do valor de β . Mesmo no caso de ligação incompleta entre QTL e marcador, β ainda será negativo. No caso de não existir ligação (hipótese de nulidade), β será igual a zero. A estimativa de β dependerá de vários fatores, incluindo a taxa de recombinação entre QTL e marcador, o número de alelos do QTL que estão segregando na população e a relação de dominância entre os alelos.

Inicialmente, assume-se ligação completa entre QTL e marcador e que há apenas dois alelos do QTL segregando na população com frequências p e q . Considere ainda que os genótipos estão distribuídos segundo Hardy-Weinberg e que a população está em equilíbrio de ligação entre QTL e marcador. Assim como nos exemplos anteriores, os três genótipos do QTL tem valores de a , d e $-a$. Para estes três genótipos é possível se terem nove combinações de pares de irmãos. A frequência esperada dessas combinações e suas probabilidades são apresentadas no Quadro 4. A estimativa para cada valor de π_j é calculada multiplicando-se a probabilidade pela estimativa de Y_j . Estes valores estão também no Quadro 4. Na última linhagem do quadro, essas estimativas são dadas em termos das variâncias genéticas aditivas (σ_a^2) e de dominância (σ_d^2) devido ao QTL e da variância do resíduo (σ_e^2). As variâncias aditivas e de dominância são calculadas como (Falconer, 1981):

$$\begin{aligned} \sigma_a^2 &= 2pq [a - (p - q)d]^2 \\ \sigma_d^2 &= 4p^2q^2d^2 \end{aligned} \quad (5)$$

Se há um QTL segregando que esteja ligado ao marcador, $E(Y_j | \pi_j)$ decrescerá monotonicamente com π_j , ao menos que $\sigma_d^2 = 0$. Se $\sigma_d^2 = 0$, então $E(Y_j | \pi_j)$ pode ser representada pela seguinte função linear de π_j :

$$E(Y_j | \pi_j) = (\sigma_e^2 + 2\sigma_a^2) - 2\sigma_a^2\pi_j \quad (6)$$

Dessa forma $\alpha = \sigma_e^2 + 2\sigma_a^2$, e $\beta = -2\sigma_a^2\pi_j$. Haseman e Elston (1972) demonstraram que, com dominância e para o caso de grandes amostras, β tende a $-2\sigma_a^2$ sendo $\sigma_e^2 = \sigma_a^2 + \sigma_d^2$.

Quadro 4. Frequência esperada para pares de irmãos das combinações dos genótipos do QTL, e esperança das diferenças ao quadrado (Y_j).

Par de irmãos	Probabilidade condicional (π_j)			Y_j
	0	1/2	1	
Q1Q1 - Q1Q1	p^4	p^3	p^2	e_j^2
Q2Q2 - Q2Q2	q^4	q^3	q^2	e_j^2
Q1Q2 - Q1Q2	$4p^2q^2$	pq	$2pq$	e_j^2
Q1Q1 - Q1Q2	$2p^3q$	p^2q	0	$(a - d + e_j)^2$
Q1Q2 - Q1Q1	$2p^3q$	p^2q	0	$(-a + d + e_j)^2$
Q1Q2 - Q2Q2	$2pq^3$	pq^2	0	$(a + d + e_j)^2$
Q2Q2 - Q1Q2	$2pq^3$	pq^2	0	$(-a - d + e_j)^2$
Q1Q1 - Q2Q2	p^2q^2	0	0	$(2a + e_j)^2$
Q2Q2 - Q1Q1	p^2q^2	0	0	$(-2a + e_j)^2$
$E(Y_j \pi_j)$	$e_j^2 + 4pq [a-(p-q)d]^2 + 2pqd^2 [1-(p-q)^2] = \sigma_e^2 + 2\sigma_a^2 + 2\sigma_d^2$	$e_j^2 + 2pq [a-(p-q)d]^2 + 2pqd^2 [1-(p-q)^2] = \sigma_e^2 + \sigma_a^2 + 2\sigma_d^2$	$e_j^2 = \sigma_e^2$	

Até este ponto assumiu-se ligação completa ($r = 0$) entre QTL e marcador. Para o caso de ligação parcial, sem dominância e informação completa dos pais, Haseman e Elston (1972) demonstraram que a esperança de $\beta = -2(1 - 2r)^2 \sigma_e^2$. Os interessados em conhecerem como se chega a este valor devem consultar o trabalho de Haseman e Elston. É importante esclarecer que o termo "informação completa dos pais" significa dizer que o número de alelos marcadores IBD pode ser determinado para cada par de irmãos. Este é o caso ilustrado na Figura 4, em que os dois pais têm quatro alelos diferentes. Isto é verdade também para o caso em que ambos os pais são heterozigotos mas tem um mesmo alelo em comum. Todavia para outras combinações de alelos marcadores não será possível se determinar inequivocadamente o número de alelos IBD para todas as combinações possíveis de pares de irmãos. Para estes casos, Haseman e Elston (1972) demonstraram que π_j pode ser estimado como:

$\hat{\Pi}_j = f_{j2} + \frac{1}{2} f_{j1}$, onde f_{j1} e f_{j2} são as probabilidades que o par de irmãos j tenha um ou dois alelos IBD, respectivamente.

É importante observar, ao contrário dos modelos anteriores, que a esperança das diferenças entre as classes genotípicas é função da variância genética e não mais

de *a* e *d*. Na maioria das situações práticas *a* será menor do que o desvio padrão do resíduo e conseqüentemente o delineamento de par de irmãos dará menor poder estatístico do que os outros delineamentos anteriormente mencionados.

Além do modelo de regressão, Y_j pode também ser analisado com o número de alelos IBD como um efeito de classe. No caso de dominância incompleta, o modelo de regressão terá maior poder.

A.7. Delineamento baseado em gerações adicionais

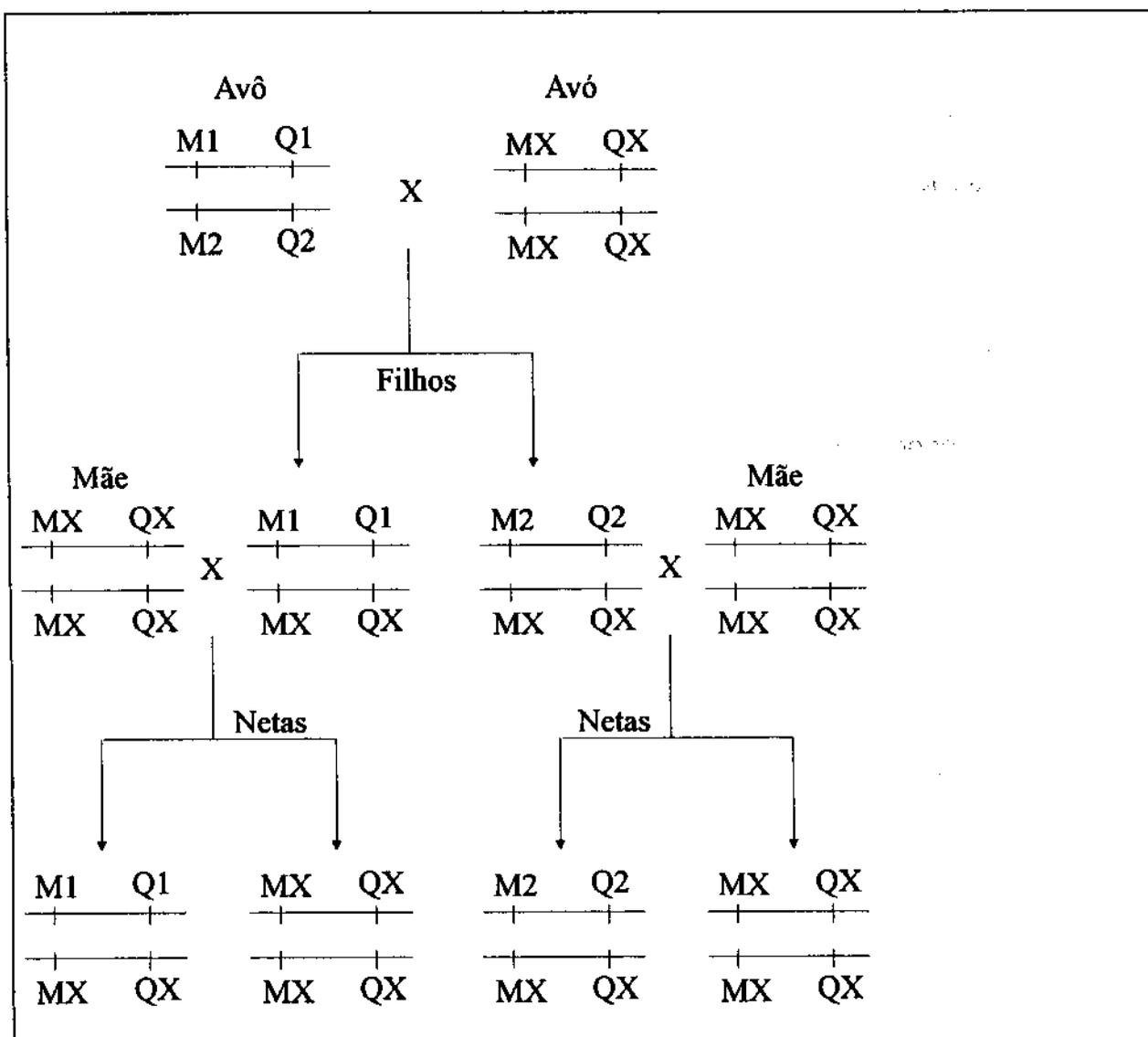
Todos os delineamentos considerados até o momento são baseados na análise de uma única geração após a produção de indivíduos heterozigotos para ambos os *loci*. Todavia, informações adicionais podem ser obtidas com a análise de mais gerações. Os delineamentos de multigerações podem ser divididos em dois casos; um em que os indivíduos das gerações futuras são medidos para as características quantitativas e também genotipados para os marcadores, e outro em que os indivíduos das gerações futuras são medidos para as características quantitativas mas não são genotipados para os marcadores. Inicialmente consideraremos o último caso e posteriormente o primeiro.

Em muitas espécies de plantas, um grupo de progênie F-3 pode ser facilmente produzido de cada indivíduo F-2. Na ausência de seleção ou viabilidade diferenciada, o grupo de progênie F-3 terá, na média, a mesma freqüência de alelos dos pais F-2. Dessa forma o contraste será igual ao do delineamento F-2, mas a variância do resíduo pode ser significativamente reduzida, pois vários indivíduos F-3 podem ser medidos para cada genótipo de F-2. Todas as progênies de um específico indivíduo F-2 terão um mesmo componente genético igual à metade da variância genética, que não estará incluída na variância do resíduo. Dessa forma, este delineamento e o delineamento de netas são muito úteis para características com baixa herdabilidade.

Para espécies autocruzantes com limitada fertilidade nas fêmeas, como é o caso de gado de leite, os genótipos podem ser determinados em uma amostra da progênie de um pai heterozigoto, e as características quantitativas podem ser medidas nas netas (isto é, as progênies descendentes de um pai heterozigoto genotipado). Este delineamento de netas (Weller et al. 1990) está representado na Figura 5. Os filhos de

touros heterozigotos para os marcadores são genotipados e as filhas destes filhos (isto é, as netas dos touros) são medidas para as características quantitativas. Assume-se que os acasalamentos com os touros (avós) e os filhos são ao acaso. Dessa forma, apenas a metade das netas receberá o alelo paterno, e a esperança do contraste entre os grupos de progênie (netas) é apenas a metade do que é no delineamento de filhas. Todavia, um número muito maior de registro de desempenho pode mais do que compensar a redução do valor do contraste.

Figura 5. Delineamento de netas.



Este delineamento tem a vantagem, que para certas espécies, especialmente gado de leite, a população comercial de gado já tem a estrutura apropriada, e as características quantitativas de interesse já são medidas rotineiramente pela indústria.

Um QTL ligado ao marcador pode ser detectado neste tipo de delineamento através da análise pelo seguinte modelo linear:

$$Y_{ijklm} = GS_i + A_j + SO_{ijk} + B_l + e_{ijklm} \quad (7)$$

onde GS_i é o efeito do avô i , SO_{ijk} é o efeito do indivíduo k com o alelo marcador j , filho do avô i , e os demais termos são os mesmos já definidos para a equação (2). À semelhança do delineamento de filhas, um efeito significativo do alelo marcador é indicativo de um QTL ligado ao marcador. A significância do efeito pode ser testada através da ANOVA, utilizando-se no numerador o quadrado médio do marcador. Todavia, este quadrado médio inclui também um componente devido à diferença entre filhos, e assim, o denominador para um teste de F correto será uma função dos quadrados médios dos filhos e dos resíduos (Ron et al. 1994). Uma forma alternativa de se testar a significância deste efeito é realizar o teste do qui-quadrado, à semelhança do que é realizado para o delineamento de filhas.

Neste delineamento, aumentando-se o número de netas, reduz-se a variação residual, sem no entanto reduzir a variação genética entre grupos de progênies. Isto faz com que a vantagem deste delineamento seja maior para as características de baixa herdabilidade.

Embora o poder estatístico para todos os delineamentos aqui considerados seja discutido posteriormente com mais detalhes, apresenta-se no Quadro 5 a variância esperada devido ao QTL para os vários delineamentos. Em todos os casos assumiu-se $r = 0$. Para os delineamentos em populações segregantes, assumiram-se também apenas dois alelos com igual frequência para o QTL. Os resultados para o F-2 foram calculados considerando-se $d = a$ e $d = 0$. As variâncias para os demais delineamentos foram obtidas considerando-se $d = 0$.

Quadro 5. Variância esperada devido ao QTL.

Variância do QTL	Delineamentos					
	F-2 (d = a)	F-2 (d = 0)	BC	Irmãos completos	Meio-irmãos	Netas
	$3a^2/2$	$a^2/2$	$a^2/4$	$a^2/8$	$a^2/8$	$a^2/32$

Exceto pelo delineamento de irmãos completos, a variância devido ao QTL é uma função de a^2 . Considerando que o efeito do QTL geralmente é menor do que o desvio padrão do resíduo, o delineamento de irmãos completos terá uma variância devido ao QTL muito menor do que os outros delineamentos. À exceção deste delineamento, os demais estão listados no Quadro 5, em ordem descendente da variância do QTL. A variância do QTL é aproximadamente proporcional ao inverso do tamanho da amostra necessária para se ter um determinado poder de detecção de QTL. Dessa forma, o delineamento de netas exigirá aproximadamente 16 vezes mais registros para se obter o mesmo poder que o delineamento F-2 com $d = 0$.

Quando gerações adicionais são genotipadas e medidas para a característica quantitativa, o poder de detecção não aumenta por indivíduo genotipado, mas a precisão no mapeamento do QTL aumenta, devido a uma geração adicional de recombinação. A vantagem da geração adicional com relação ao mapeamento do QTL será considerado em maiores detalhes posteriormente.

A.8. Modelos de efeito genético para análise completa da população

Fernando e Grossman (1989) desenvolveram um método para estimar os valores genéticos para todos os indivíduos de uma população, incluindo QTL via ligação com marcadores genéticos, desde que a herdabilidade e a taxa de recombinação entre QTL e marcador fosse conhecida. Este método, que é uma extensão do Best Linear Unbiased Prediction (BLUP) via modelo animal, é apropriado para qualquer estrutura de população, e também permite incorporar efeitos poligênicos não ligados ao marcador e outros efeitos tais como rebanhos ou blocos.

Em todos os modelos considerados anteriormente, o QTL foi considerado como efeito fixo. Neste modelo, Fernando e Grossman (1989) assumem que o QTL é um

efeito ao acaso com variância conhecida. Dessa forma cada indivíduo com ancestrais desconhecidos tem dois únicos alelos para o QTL, que são uma amostra de uma população infinita de alelos. Para cada indivíduo, o modelo gamético para o QTL, com efeitos separados dos alelos dos touros e vacas, é:

$$Y_i = B_i + V_i^p + V_i^m + u_i + e_i \quad (8)$$

onde V_i^p e V_i^m são os efeitos aditivos do QTL recebidos do touro e da vaca, u_i é o efeito poligênico ao acaso, não explicado pelo marcador, e os demais termos são os mesmos já definidos anteriormente. Este modelo pode ser escrito em notação matricial como:

$$Y = XB + ZV + u + e \quad (9)$$

onde X e Z são as matrizes de incidência relacionando indivíduos aos efeitos de um bloco específico e aos de um gameta. Devido ao fato de que cada indivíduo tem um efeito poligênico diferente, não se requer uma matriz de incidência para u . Assim, um conjunto de equações mistas pode ser estabelecido como:

$$\begin{bmatrix} X'X & X'Z & X' \\ Z'X & Z'Z + G_v^{-1}\sigma_e^2 & Z' \\ X & Z & I + G_u^{-1}\sigma_e^2 \end{bmatrix} \begin{bmatrix} B \\ V \\ u \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ y \end{bmatrix}$$

Para se resolver essas equações, é necessário primeiramente inverter G_v e G_u . Para os animais que não são genotipados, a probabilidade de receber um dos alelos de um dos pais é a mesma, mas se ambos os pais e a progênie forem genotipados para o marcador, então a probabilidade de receber um alelo paterno específico do QTL ligado ao marcador será uma função do genótipo do marcador na progênie e da taxa de recombinação r . Baseados nessas probabilidades, Fernando e Grossman (1989) demonstraram como a matriz de variância - co-variância para os efeitos gaméticos do QTL pode ser construída. Eles também descrevem um algoritmo simples para se obter a inversa dessas matrizes, à semelhança do método de Henderson de calcular a inversa da matriz de parentesco. Assim, BLUP de a , o vetor de valores genéticos aditivos, é estimado como: $Z\hat{v} + \hat{u}$.

Uma vantagem adicional deste método é que, ao assumir que o efeito do QTL é ao acaso, ao contrário de ser fixo, os efeitos do QTL com as maiores estimativas não são viesadas para cima (Georges et al. 1995). Teoricamente este método pode ser ampliado para analisar múltiplos marcadores e características, embora a complexidade se torne muito maior. Cantet e Smith (1991) demonstraram que o número de equações pode ser significativamente reduzidos através da absorção das equações dos indivíduos não-parentes.

Embora este método tenha propriedades atrativas, ele não tem sido usado ainda em dados reais. Além do grande requerimento computacional, este método assume que ambos, r e a variância, devido ao QTL, sejam conhecidos a priori. Estudos de simulação têm mostrado que, embora a Máxima Verossimilhança Restrita (REML) possa ser usada para estimar estes parâmetros, eles estão completamente confundidos um com o outro no caso de um único *locus* marcador. Além disso, como cada indivíduo com pais desconhecidos assume-se ter dois únicos alelos, a predição da variância dos erros dos efeitos para qualquer indivíduo será bem grande, e, conseqüentemente, pouco informativa. Um outro ponto é que a pressuposição de uma distribuição normal dos efeitos dos possíveis QTL pode não ser realista. Todavia, não está claro qual seria o efeito de assumir que existam apenas dois alelos do QTL segregando na população, se o número real for maior.

B) Métodos Estatísticos para a detecção de QTL e estimativa de parâmetros

B.1. Introdução

Os parâmetros principais que serão considerados são as médias e variâncias dos genótipos do QTL e a taxa de recombinação entre os marcadores e o QTL. É possível perceber pela seção anterior que a estimação de parâmetros do QTL não é algo trivial. O primeiro problema com que se depara é que, em quase todos os casos de interesse, a variância devido ao QTL que está segregando é apenas uma fração muito pequena da variância total. Além disso, desde que se esteja geralmente tratando com dados de campo, haverá sempre outras variáveis que podem causar confundimento,

tais como rebanho ou bloco. As estimativas dos parâmetros podem ser viesadas se estes fatores não forem incluídos no modelo de análise. Como já visto anteriormente, em geral a ligação entre marcador e QTL é incompleta. Dessa forma, a taxa de recombinação deve ser incluída nas análises e conseqüentemente o modelo de análise não será uma função linear dos parâmetros. Nas análises em populações segregantes, será necessário também se considerar a variância poligênica devido a outros *loci* que não estão ligados aos marcadores. Além disso, muitas análises incluem múltiplos marcadores e múltiplas características, o que gera maiores dificuldades na estimativa dos parâmetros.

B.2. Propriedades desejáveis das estimativas dos parâmetros do QTL

Existem quatro propriedades desejáveis principais: sem viés, variância mínima do erro, estimativas dentro do espaço dos parâmetros e consistência. Para situações simples, é possível derivar estimadores com todas essas propriedades, mas nos casos mais complicados não será possível se obter estimativas com todas essas propriedades desejáveis e aí tem-se que decidir o que é mais importante em determinada situação.

Viés: Considere que $\hat{\theta}$ seja um estimador de θ . Então, $\hat{\theta}$ é sem viés se $E(\hat{\theta}) = \theta$, ou seja, a esperança do estimador é igual ao valor do parâmetro. Por exemplo: quando se estima a variância com base na média da amostra, divide-se por $n-1$, sendo n igual ao tamanho da amostra. Dividindo-se por n , o estimador da variância será uma estimativa viesada da variância.

Variância mínima do erro: É definida como o valor de $\hat{\theta}$ para o qual $E[(\hat{\theta} - \theta)^2]$ é mínimo. Esta propriedade é a base da estimação pelos mínimos quadrados. O estimador com estimativa mínima da variância do erro é também chamado da "melhor" estimativa.

Estimativa dentro do espaço do parâmetro: Embora este requerimento possa parecer trivial, freqüentemente isto não ocorre. Em muitas situações não é possível se obter uma estimativa que seja ao mesmo tempo sem viés e dentro do espaço do parâmetro. Alguns exemplos simples de estimadores fora do espaço do parâmetro são

as estimativas negativas dos componentes de variância, ou estimativas de taxa de recombinação menores do que zero ou maiores do que 0,5. As estimativas por ML estão sempre, por definição, dentro do espaço do parâmetro.

Consistência: Um estimador $\hat{\theta}$ é dito consistente se $\hat{\theta}$ tende para θ quando o tamanho da amostra tende ao infinito. Um estimador pode ser consistente, mesmo se ele for viesado. Considere o exemplo anterior da variância de uma amostra. Se dividimos por n em vez de por $n-1$, o estimador é viesado, mas é consistente, pois quando n tende a infinito, $\hat{\theta}$ tende para θ . Embora essa propriedade também pareça trivial, ela é especialmente importante para a detecção de QTL. Considere o caso da seção anterior dos delineamentos F-2 e BC. Os contrastes entre as classes dos genótipos do marcador não é um estimador consistente do contraste entre as médias dos genótipos do QTL. Esta estimativa é viesada pelo fator $1-2r$, independente do tamanho da amostra.

B.3. Estimação pelos Quadrados Mínimos

Assuma que há uma série de observações para uma variável y , a qual se deseja modelar em função de outras variáveis para as quais existam dados. Considere que y é a variável dependente e que as outras sejam as variáveis independentes. O objetivo é explicar a variável dependente em termos de uma série de estimativas de parâmetros que façam a ligação entre a variável dependente e as independentes, ou seja, estabelecer uma função das variáveis independentes que se aproximem das observações de y . Em geral não é possível explicar completamente y em termos das variáveis independentes. A diferença entre as estimativas de y , baseadas nas variáveis independentes, e o valor observado de y , é denominada de erro ou resíduo.

Estimação pelos quadrados mínimos é baseada em se gerar estimativas de parâmetros que minimizem a esperança da soma de quadrado dos erros. Assim, por definição, este método tem estimativa mínima da variância do erro. Definindo-se $\theta =$ vetor de parâmetros, $y =$ vetor de observações e $e =$ vetor de resíduos, sendo que y é também a variável dependente, o modelo geral em notação matricial pode ser escrito como:

$$y = f(\theta') + e \quad (\text{B.1})$$

onde $f(\theta')$ é alguma função de θ . A solução ($\hat{\theta}$) pelos mínimos quadrados é o vetor que minimiza $[y - f(\hat{\theta})]^2 = e^2$. Para um modelo linear, a equação (B.1) pode ser escrita como:

$$y = X\theta + e \quad (\text{B.2})$$

onde X é a matriz de coeficientes de θ . Os efeitos em modelos lineares podem ser contínuos ou de classes. Efeitos tais como rebanho, bloco, ou sexo são denominados de efeito de classe. Embora os níveis destes efeitos possam ser numerados, não há relação entre o número de um rebanho específico e seu efeito. Para o caso de efeitos contínuos, uma relação linear é assumida entre o valor para a variável independente e a variável dependente. Cada linhagem de X corresponde ao coeficiente de θ para um registro específico de y . Para os efeitos de classe, os elementos de X será zero ou um. Para os efeitos contínuos, cada elemento de X corresponde ao valor observado da variável independente.

Para um modelo linear, é fácil demonstrar que a solução pelos mínimos quadrados para θ pode ser obtida pelas equações normais: $\hat{\theta} = (X'X)^{-1} X'y$. Neste caso as estimativas dos parâmetros será sem viés, consistente e dentro do espaço dos parâmetros.

Como já apresentado na seção anterior, as médias e variâncias dos genótipos e a taxa de recombinação não podem ser descritas por um modelo linear. Além disso, como já mencionado anteriormente, dentro do contexto de estimação pelos mínimos quadrados, as médias dos genótipos e a taxa de recombinação entre um QTL e um único marcador estão completamente confundidas e não podem ser estimadas separadamente. Todavia, usando-se dois marcadores flanqueando o QTL, é possível se derivar estimativas separadas das médias dos genótipos e das taxas de recombinação, mas não é possível se construir um modelo linear que descreva com precisão a relação entre as observações e os parâmetros do QTL.

O método não-linear pelos quadrados mínimos da estimação dos parâmetros do QTL usando-se dois marcadores foi desenvolvido independentemente em 1992 por Haley e Knott, e Martinez e Curnow. Este método será demonstrado utilizando-se o delineamento BC, embora ele já esteja adaptado para a maioria dos delineamentos apresentados na seção anterior. O delineamento BC com dois marcadores é ilustrado na Figura B.1. Para as progênies do BC apenas o cromossomo do pai F-1 é apresentado na figura. Existem oito haplo-tipos (incluindo o QTL): dois não recombinantes, quatro recombinantes simples, e dois recombinantes duplos. A taxa de recombinação entre os dois marcadores M e N será considerada igual a R, e as taxas de recombinação entre M e Q e entre Q e N são, respectivamente r_1 e r_2 . Assumindo-se interferência zero, $R = r_1 + r_2 - 2 r_1 r_2$. Assim, o seguinte modelo pode ser escrito:

$$Y_{ij} = \mu_1 (1 - p_i) + \mu_2 p_i + e_{ij} \quad (B.3)$$

onde Y_{ij} é o registro de produção do indivíduo j com genótipo marcador i , μ_1 é a média dos indivíduos com genótipo Q1Q2, μ_2 é a média dos indivíduos com genótipo Q2Q2, p_i é a probabilidade que um indivíduo com genótipo marcador i tenha genótipo Q2Q2 e e_{ij} é o resíduo. Este modelo pode ser reescrito como:

$$Y_{ij} = \mu_1 + (\mu_2 - \mu_1) p_i + e_{ij} \quad (B.3a)$$

p_i é uma função dos parâmetros das taxas de recombinação e pode ser estimado para cada um dos quatro haplotipos do marcador como:

$$p_{M1N1} = r_1 r_2 / (1 - R) \quad (B.4)$$

$$p_{M1N2} = r_1 (1 - r_2) / R \quad (B.5)$$

$$p_{M2N1} = r_2 (1 - r_1) / R \quad (B.6)$$

$$p_{M2N2} = 1 - r_1 r_2 / (1 - R) \quad (B.7)$$

Se r_1 for conhecido, é possível substituir estes valores na equação (B.3a) e resolver como uma simples regressão linear, com μ_1 como sendo a intersecção em y e $\mu_2 - \mu_1$, como a declividade da reta. Como r_1 não é conhecida, a equação (B.3a) pode

ser considerada como sendo quatro equações distintas, uma para cada haplotipo do marcador. Mesmo considerando interferência igual a zero, e resolvendo r_2 em função de R , haverá quatro equações que são funções complicadas da média do QTL e de r_1 . Dessa forma, este modelo é não-linear para r_1 . A solução pelos quadrados mínimos para os três parâmetros (μ_1 , μ_2 e r_1) serão os valores que minimizem a soma de quadro dos resíduos. Este modelo pode ser facilmente resolvido por um algoritmo não-linear, tal como o PROC NLIN do SAS. As declarações do SAS para isto são dadas na Figura B.2.

Figura B.1. Delineamento BC com dois marcadores flanqueando o QTL.

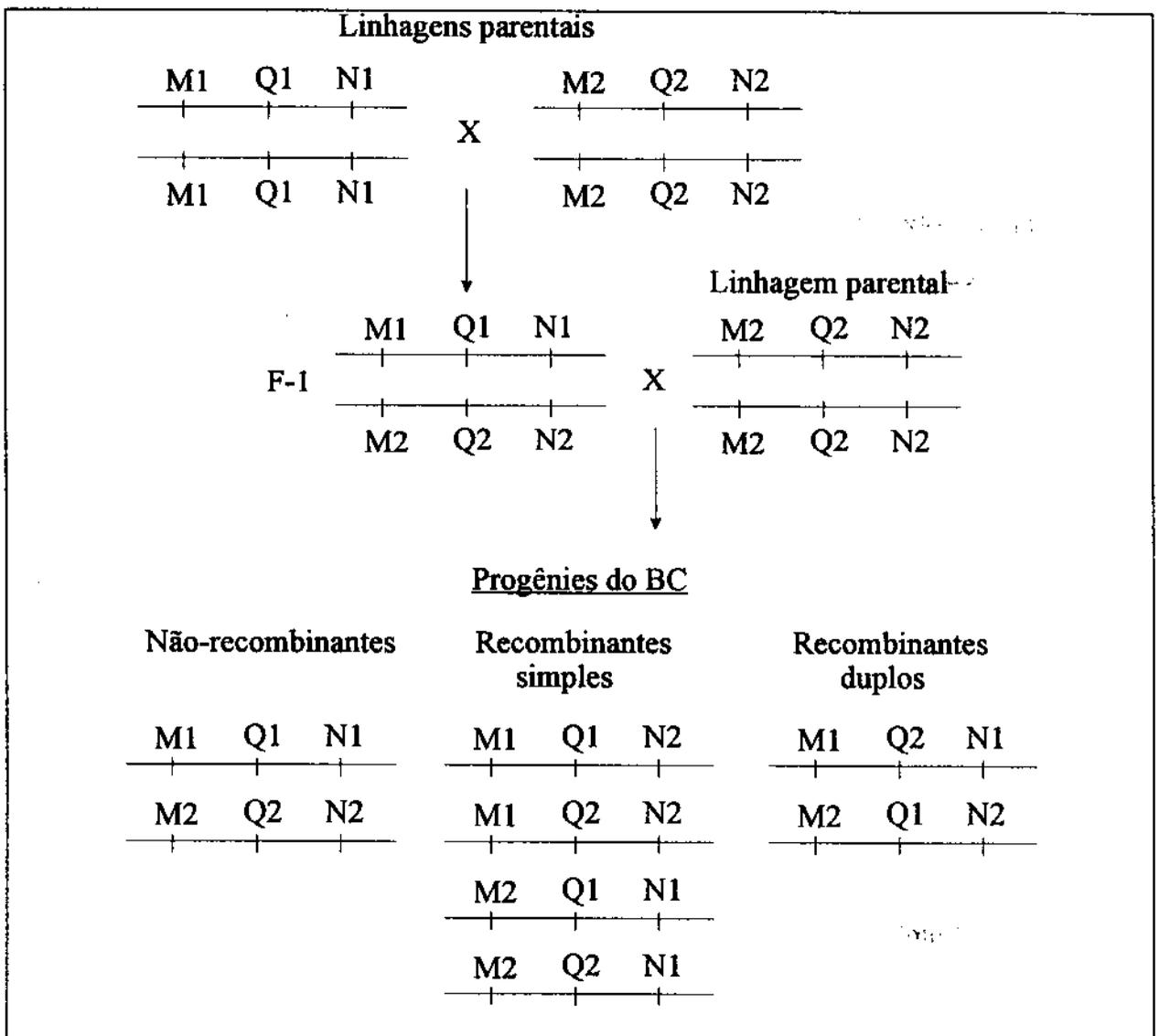


Figura B.2. Declarações do SAS para mapeamento do QTL pela regressão não-linear em um delineamento BC.

```

proc nlin;
R = 0,3; (por exemplo)
if m = 1 and n = 1 then p = r1 * r2/(1 - R);
if m = 1 and n = 2 then p = r1 * (1 - r2)/R;
if m = 2 and n = 1 then p = (1 - r1 * r2)/(1 - R);
if m = 2 and n = 2 then p = r2 * (1 - r1)/R;
r2 = (R - r1)/(1 - 2 * r1);
parameters mμ1 = -.1 mμ2 = -.1
           r1 = 0 to 0.3 by 0.05;
model trait = mμ1 + (mμ2 - mμ1) * p;
           bounds -0.3 < mμ1 < 0.3;
           bounds -0.3 < mμ2 < 0.3;
           bounds 0 < r1 < 0.3;

run;

```

7/25/2000

A principal vantagem deste método é que ele pode ser executado por vários pacotes estatísticos e a significância do efeito do QTL pode ser testada pelo teste de F do quadrado médio do modelo versus o quadrado médio do resíduo, o que é mais conhecido da maioria dos pesquisadores, do que o teste da razão entre funções de verossimilhança. As desvantagens do método é que ele é aplicável em apenas certas situações. Ele não pode ser aplicado para estimar a taxa de recombinação entre QTL e um único marcador, para estimar a variância do efeito do QTL ou em populações segregantes.

B.4. Estimação de um único parâmetro por Máxima Verossimilhança

A estimação por Máxima Verossimilhança (ML) é muito mais flexível do que pelos quadrados mínimos (LS), mas requer programação mais complexa.

Como a estimação por ML é menos familiar do que pelos LS, a metodologia básica para a estimação por ML será ilustrada usando um exemplo de uma distribuição binomial. Assuma que de uma amostra de dez observações três são "sucessos" e as outras sete são "falhas". Querendo obter a estimativa ML (MLE) da probabilidade de

sucessos p , escreve-se a probabilidade binomial, em função de p , para se obter este resultado:

$$L = \frac{10!}{3!7!} (p)^3 (1 - p)^7 \quad (\text{B.8})$$

onde L é a probabilidade de se obter este resultado condicional em p . L é denominado de função de verossimilhança. A MLE para p é o valor de p que maximiza L . A MLE é calculada através da derivada de L em relação a p , igualando a função derivada a zero e resolvendo para p . Na prática é mais fácil calcular e diferenciar a função $\log L$. Isto é equivalente a diferenciar L , pois a função e seu logaritmo terão valores máximos para os mesmos valores da variável. Assim, é possível derivar a MLE de p como:

$$\text{Log } L = \log (10!) - \log (3! 7!) + 3 (\log p) + 7 [\log (1 - p)] \quad (\text{B.9})$$

$$d \text{Log}L/dp = 3/p - 7/(1 - p) = 0 \quad (\text{B.10})$$

$$p = 3/10 \quad (\text{B.11})$$

Este resultado é obviamente a proporção de sucessos obtidos na amostra. Assim, para este caso simples, a MLE é o valor estimado intuitivamente. Deve estar claro, das discussões anteriores, porque a MLE deve estar dentro do espaço do parâmetro. Uma estimativa do parâmetro fora do espaço terá por definição $L = 0$ e não pode portanto ser a MLE.

Para o caso de distribuição contínua, a função de verossimilhança é calculada como a função de densidade da distribuição, condicionada na amostra. A função de densidade, $f(y)$, para uma variável contínua y , é definida como o valor da ordenada da função de distribuição, para um dado valor de y . Por exemplo, assumamos que a amostra foi obtida de uma distribuição normal. Para se obter a MLE da média, é necessário que se calcule a função de densidade conjunta da amostra. Para uma única observação a verossimilhança será:

$$L = \frac{e^{-(y-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma}} \quad (\text{B.12})$$

onde σ é o desvio padrão, e é a base do logaritmo natural e aproximadamente igual a 2,72, μ é a média e y é o valor da variável. Para uma amostra de N observações, a função de verossimilhança será o produto de L para cada observação individual. Como anteriormente, a MLE para μ pode ser obtida calculando-se a derivada de $\log L$ em relação a μ , igualando a zero e resolvendo para μ . A derivada de $\log L$ para uma amostra de uma distribuição normal é calculada como:

$$d(\log L)/d\mu = \sum (y_i - \mu) \quad (\text{B.13})$$

Igualando esta função a zero, verifica-se que a MLE da média da população é a média da amostra, o que é, novamente, o resultado intuitivamente correto.

A MLE da variância pode ser obtida da mesma maneira e novamente gerar um resultado intuitivo da variância da amostra. Embora nos dois exemplos dados a ML tenha sido usada para se obter estimativas que poderiam ser obtidas por outros métodos, será mostrado adiante, que para problemas mais complicados que a ML, é o único método de estimação que pode utilizar de todos os dados disponíveis.

B.5. Estimação de multiparâmetros por Máxima Verossimilhança

A ML pode também ser usada para estimar vários parâmetros simultaneamente, como, por exemplo, estimar a média e a variância em uma distribuição normal. Neste caso será necessário maximizar L com relação a ambos os parâmetros. Isto é feito calculando-se a derivada parcial do $\log L$ com relação a cada parâmetro e igualando a função derivada igual a zero. O próximo passo é resolver o sistema de equações que será igual ao número de parâmetros que se pretende estimar. Em geral a função de verossimilhança para a estimação de m parâmetros $(\theta_1, \theta_2, \dots, \theta_m)$, de uma amostra de n observação (y_1, y_2, \dots, y_n) , pode ser escrita como:

$$\begin{aligned} L &= p(y_1, y_2, \dots, y_n \mid \theta_1, \theta_2, \dots, \theta_m) = \\ &= p(y_1 \mid \theta_1, \theta_2, \dots, \theta_m) p(y_2 \mid \theta_1, \theta_2, \dots, \theta_m) \dots p(y_n \mid \theta_1, \theta_2, \dots, \theta_m) \\ &= \prod p(y_i \mid \theta_1, \theta_2, \dots, \theta_m) \end{aligned} \quad (\text{B.14})$$

onde $p(y|\theta)$ representa a probabilidade de se obter y , condicional em θ , e Π significa o produto $p(y_i|\theta_1, \theta_2, \dots, \theta_m)$ de y_1 até y_n . Se a distribuição é contínua, então $p(y|\theta)$ é substituído por $f(y|\theta)$, isto é, a função de densidade de y , condicional em θ . Desta forma, qualquer problema que puder ser representado pela equação (B.14) pode ser resolvido pela ML.

B.6. Predição da Variância do erro para a MLE

Além de se estimar os parâmetros, é também importante se conhecer qual é a precisão das estimativas. Em geral se usa o erro padrão da estima para esta finalidade. O erro padrão ao quadrado é conhecido como a predição da variância do erro. A equação abaixo pode, em geral, ser usada para se obter a predição da variância do erro para a MLE de um único parâmetro:

$$\text{Var}(\hat{\theta}) = \frac{-1}{E[d^2(\log L)/d\theta^2]} \quad (\text{B.15})$$

onde $\hat{\theta}$ é a MLE de θ e $E[d^2(\log L)/d\theta^2]$ é a esperança da segunda derivada de L com relação a θ . A equação (B.15) será correta se a primeira derivada de θ é um múltiplo da diferença entre o valor real do parâmetro e sua estimativa. Caso contrário, a predição da variância do erro será ligeiramente maior do que o lado direito da equação (B.15). Sob uma grande variedade de condições, a equação (B.15) será assintoticamente correta, ou seja, à medida que o tamanho da amostra aumenta, a diferença entre os lados direito e esquerdo da equação tende a zero. A raiz quadrada da predição da variância do erro, o erro padrão da estimativa, pode ser usada para se calcular o intervalo de confiança da estimativa.

A predição das variáveis dos erros para o caso de estimação de vários parâmetros pode ser obtida de maneira semelhante à descrita na equação (B.14). As estimativas dos parâmetros e as derivadas de primeira ordem serão cada uma constituída por um vetor em que o número de elementos será igual a m . A derivada de segunda ordem e a predição das variâncias dos erros serão ambas matrizes quadradas

de mxm. Usando-se colchetes para representar matrizes e vetores, a matriz da predição das variâncias dos erros pode ser calculada pela seguinte equação:

$$\text{Var}(\hat{\theta}) = - \left[\frac{d^2 \log L}{d[\theta]^2} \right]^{-1} \quad (\text{B.16})$$

onde o lado direito é a matriz inversa das derivadas parciais de segunda ordem com relação a $[\theta]$. Os elementos da diagonal serão as predições das variâncias dos erros das estimativas, e os elementos fora da diagonal as predições das co-variâncias dos erros entre as estimativas. Estes elementos são necessários para se testar hipóteses baseadas em funções lineares dos parâmetros.

Mesmo quando a predição da variância do erro não é calculada, a ML pode ainda ser usada para se testar hipóteses através do teste da razão entre funções de verossimilhança. No teste da razão, as ML obtidas sob as hipóteses alternativas são comparadas. No caso da hipótese de nulidade, um ou mais parâmetros que são maximizados na hipótese alternativa é mantido fixo. Por exemplo, fixando $r = 0,5$. Considerando-se que a hipótese de nulidade é aninhada dentro da hipótese alternativa, sob a pressuposição de não haver diferença, o log natural da razão entre funções de verossimilhança será assintoticamente distribuída como $(-1/2) \chi^2$, onde χ^2 é a estatística de qui-quadrado. O termo "hipótese aninhada" significa dizer que, embora alguns parâmetros estejam fixos na hipótese de nulidade e são maximizados sob a hipótese alternativa, todos os parâmetros que forem fixos na hipótese alternativa também o serão na hipótese de nulidade. O número de graus de liberdade será igual ao número de parâmetros que está sendo maximizado na hipótese alternativa, mas fixo na hipótese de nulidade.

Embora geralmente seja possível se escrever a função de verossimilhança e diferenciar $\log L$ com relação aos diferentes parâmetros, para os modelos de detecção de QTL não é possível derivar soluções analíticas para o sistema de equações que se produza. Nestes casos, os métodos iterativos devem ser utilizados para se obterem as soluções.

B.7. Estimação de parâmetros do QTL via ML para cruzamentos entre linhagens consangüíneas e um único marcador

O delineamento de BC (Figura B.1) será utilizado para se ilustrar o uso da ML quando apenas um QTL está ligado a apenas um marcador. Assuma que a característica quantitativa tenha distribuição normal e que os dois genótipos do QTL tenham a mesma variância. Será mostrado posteriormente que a ML é uma metodologia muito mais flexível do que a dos quadrados mínimos e que pode ser utilizada em muitas situações nas quais as pressuposições acima não sejam verdadeiras. A função de densidade para um indivíduo com genótipo M1M2 será:

$$f(y) = \frac{(1-r)e^{-(y-\mu_1)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} + \frac{(r)e^{-(y-\mu_2)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \quad (B.17)$$

onde y é o valor da característica, σ é o desvio padrão, μ_1 é a média dos indivíduos com genótipo Q1Q2, e μ_2 é a média dos indivíduos com genótipo Q2Q2. Os indivíduos com genótipo M2M2 terão a mesma função de verossimilhança, apenas invertendo-se as médias μ_1 e μ_2 na equação (B.17). A função completa para uma amostra de indivíduos pode ser escrita como:

$$L = \prod_{i=1}^{n_1} [f(y_i, M1M2)] \prod_{j=1}^{n_2} [f(y_j, M2M2)] \quad (B.18)$$

onde Π representa o produto de uma série, $f(y_i, (M1M2))$ e $f(y_j, M2M2)$ são as funções de densidades para as observações i e j com genótipos M1M2 e M2M2, respectivamente; e n_1 e n_2 são os números de indivíduos com os dois genótipos, respectivamente.

Para se obter as estimativas dos parâmetros, o log desta função deve ser derivada com relação aos quatro parâmetros (μ_1 , μ_2 , σ e r). As derivadas parciais devem ser igualadas a zero e o sistema de quatro equações solucionadas. Este sistema não pode ser resolvido analiticamente e métodos iterativos devem ser utilizados.

A ligação do marcador com o QTL pode ser testada comparando-se a função expressa em (B.18) no ponto em que ela tem o valor máximo com a ML obtida com $r = 0,5$ (que representa a hipótese de nulidade e significa que o QTL e o marcador não estão ligados). Como já mencionado anteriormente sob a hipótese de nulidade de não haver ligação, com um parâmetro fixo (neste caso, r), a razão entre os log das funções será assintoticamente distribuída como $(-1/2) \chi^2$, onde χ^2 é o valor do qui-quadrado com um grau de liberdade.

A equação (B.17) pode facilmente ser modificada para o caso em que as variâncias sejam diferentes para cada genótipo do QTL. Neste caso será necessário estimar cinco parâmetros em vez de quatro. A hipótese de variâncias heterogêneas pode ser testada contra a hipótese de variâncias homogêneas através da razão entre os log das ML com variâncias heterogêneas e homogêneas.

Para o caso do delineamento F-2 apresentado na Figura 2, cada genótipo do marcador será constituído de uma mistura de três distribuições normais, duas para os QTL homocigotos e uma para o QTL heterocigoto. As probabilidades de cada um dos três genótipos do QTL dentro de cada genótipo do marcador estão no Quadro 2. Desta forma, serão necessários se estimar cinco parâmetros, as médias dos três genótipos do QTL, r e a variância do resíduo. Este modelo pode também ser modificado para se considerar variâncias diferentes para cada genótipo do QTL. Neste caso já serão necessários se estimar sete parâmetros; três médias, três variâncias e r (Weller, 1986).

B.8. Estimação de parâmetros do QTL via ML para cruzamentos entre linhagens consanguíneas e dois marcadores flanqueando o QTL

A função de verossimilhança para o delineamento BC com um QTL entre dois marcadores foi descrita por Lander e Botstein (1989) como:

$$L = \prod_{f_{M1N1}}^{n_{M1N1}} \prod_{f_{M1N2}}^{n_{M1N2}} \prod_{f_{M2N1}}^{n_{M2N1}} \prod_{f_{M2N2}}^{n_{M2N2}} \quad (B.19)$$

onde n_{M1N1} , n_{M1N2} , n_{M2N1} , n_{M2N2} são os números de indivíduos com os genótipos M1N1/M2N2, M1N2/M2N2, M2N1/M2N2 e M2N2/M2N2, respectivamente, e f_{M1N1} , f_{M1N2} , f_{M2N1} e f_{M2N2} , são as funções de densidade dos quatro genótipos possíveis dos

marcadores. Considerando-se que todos os indivíduos recebem um segmento idêntico M2N2 da linhagem parental recorrente, apenas os segmentos do cromossomo recebido do F-1 será indicado. Assim, as funções de densidades para os genótipos possíveis são calculados como:

$$f_{M1N1} = (1 - \alpha) f(Q1) + \alpha f(Q2) \quad (B.20)$$

$$f_{M1N2} = (1 - \beta) f(Q1) + \beta f(Q2) \quad (B.21)$$

$$f_{M2N1} = (1 - \beta) f(Q2) + \beta f(Q1) \quad (B.22)$$

$$f_{M2N2} = (1 - \alpha) f(Q2) + \alpha f(Q1) \quad (B.23)$$

onde $\alpha = r_1 r_2 / (1 - R)$, $\beta = r_1 (1 - r_2) / R$, e $f(Q1)$ e $f(Q2)$ são as funções de densidade para cada observação com desvio padrão σ , e médias μ_1 para o genótipo Q1Q2 e μ_2 para o genótipo Q2Q2. Assim, a função de verossimilhança pode ser obtida, multiplicando-se as funções de densidade para cada indivíduo, dependendo de seu genótipo. À semelhança do modelo de regressão não linear da seção A.3, assume-se interferência zero e assim $R = r_1 + r_2 - 2 r_1 r_2$. Dessa forma será necessário derivar estimativas ML para apenas quatro parâmetros, r_1 , σ , μ_1 e μ_2 .

B.9. Análise de modelos mais complexos usando-se a ML

Alguns outros modelos mais complexos cujas soluções são possíveis pela ML serão brevemente discutidos nesta seção. Mackinnon e Weller (1995) usaram a ML para estimar parâmetros do QTL no delineamento de filhas, sob a pressuposição que apenas dois alelos do QTL estariam segregando na população. Neste caso, além dos parâmetros que são estimados no delineamento BC, é necessário se estimar a frequência relativa dos dois alelos do QTL. A função de verossimilhança para o modelo de Mackinnon e Weller (1995) é:

$$L = \prod_{k=1}^K \sum_{v=1}^4 P_v \prod_{i=1}^3 \prod_{l=1}^{L_j} \sum_{j=1}^3 c_{j|i,v} f(y_{ikl} - \mu_j) \quad (B.24)$$

onde K é o número de touros, P_v é a probabilidade do genótipo do QTL do touro v , $c_{j|i,v}$ é a probabilidade da progênie ter o genótipo j para o QTL, condicionada na combinação do genótipo do QTL do touro v e genótipo da progênie i , L_i é o número de filhas com genótipo marcador i , y_{ikl} é o valor da característica da progênie l , do touro k , com genótipo marcador i , μ_j é a média para as progênies com genótipo do QTL j , e $f(y_{ikl} - \mu_j)$ é a função de densidade normal para a progênie l do touro k , condicionada no genótipo j para o QTL. Mackinnon e Weller (1995) assumiram também que todos os touros eram heterozigotos para o marcador e que apenas estes touros seriam incluídos nas análises. Dessa forma, quatro genótipos para os touros são estabelecidos com relação ao QTL e ao marcador; os homozigotos Q1Q1 e Q2Q2, o heterozigoto Q1 ligado ao M1, e o heterozigoto Q1 ligado ao M2. P_v pode ser calculada assumindo-se a distribuição Hardy-Weinberg para os genótipos. Nesta situação é necessário se definir apenas três classes de genótipos do marcador nas progênies; aquelas que recebem M1 mas não M2, aquelas que recebem M2 mas não M1, e aquelas que recebem ambos os alelos paternos. A fórmula para $c_{j|i,v}$ é dada no trabalho de Mackinnon e Weller (1995). Para as progênies com o mesmo genótipo marcador paterno, não se sabe qual alelo ela recebeu do pai e qual ela recebeu da mãe. Desta forma, $c_{j|i,v}$ é uma função da frequência do alelo marcador entre as mães. Se considerarmos que o genótipo do QTL não afeta a variância, será necessário se estimar apenas seis parâmetros; as três médias dos genótipos do QTL, a variância do resíduo, a taxa de recombinação entre QTL e marcador, e a frequência do alelo Q1. Segundo Mackinnon e Weller (1995), quando as amostras são grandes, é possível se obter razoáveis estimativas para todos os parâmetros.

Mackinnon e Weller (1995) modificaram este modelo para incluir o efeito poligênico do touro. Neste caso a função de densidade se torna $f(y_{iklm} - u_j - g_k)$, onde g_k é o efeito poligênico do touro k em suas progênies. Esta modificação aumenta o número de parâmetros que deve ser estimado pelo número de touros. Bovenhuis e Weller (1994) modificaram este mesmo modelo para incluírem um efeito direto do genótipo do marcador, além do ligado ao QTL.

Vários autores têm derivadas fórmulas para o delineamento de netas. Por ex.: Hoeschele e VanRaden (1993a) e Hackett e Weller (1995).

B.10. Estimativas Baysiana dos efeitos do QTL e amostragem de Gibbs

A estimação Baysiana difere da ML porque em vez de se maximizar a função de verossimilhança, se maximiza a probabilidade posterior de θ , $p(y|\theta)$ como função da função de verossimilhança e da distribuição anterior de θ . Em termos gerais, o teorema de Bayes, para múltiplos parâmetros e observações, pode ser calculado como:

$$p(\theta_1, \theta_2, \dots, \theta_m | y_1, y_2, \dots, y_n) = p(\theta_1, \theta_2, \dots, \theta_m) p(y_1, y_2, \dots, y_n | \theta_1, \theta_2, \dots, \theta_m) \quad (B.25)$$

onde $p(\theta_1, \theta_2, \dots, \theta_m)$ é a probabilidade posterior dos parâmetros e $p(y_1, y_2, \dots, y_n | \theta_1, \theta_2, \dots, \theta_m)$ é a função de verossimilhança. Se assumirmos que se tenha informações anteriores dos parâmetros, a estimação Baysiana deve ser preferida em relação a ML, que ignora qualquer informação anterior dos parâmetros. Todavia, existem dois problemas com a estimação Baysiana. O primeiro é que em geral não se tem informações anteriores sobre os parâmetros e não é possível matematicamente representar essas informações sem ter que se fazer pressuposições adicionais que pode não serem verdadeiras. O segundo ponto é que se muitos registros são incluídos nas análises, então a função de verossimilhança tende a ser melhor que a distribuição anterior de θ e as estimativas Baysianas tendem às estimativas da ML. Considerando-se que na maioria dos casos de estimação dos efeitos do QTL as amostras são grandes, pouco se ganha em se utilizar a metodologia Baysiana. Hoeschele e VanRaden (1993) derivaram a distribuição anterior dos parâmetros e a função de verossimilhança para as análises do delineamento de netas. Eles consideraram que distribuição anterior do efeito do QTL era exponencial e impuseram um limite inferior para a magnitude do efeito do QTL. Assumiram também que havia apenas dois alelos do QTL segregando na população e uma frequência de alelos distribuída uniformemente variando de zero a unidade. Dadas a distribuição do efeito do QTL e suas frequências alélicas, e a variância genética aditiva total, eles estimaram o número esperado de QTL que poderiam ser detectados. A distribuição destes QTL no genoma

foi assumida ser uniforme. As análises Baysiana e ML deram resultados semelhantes e ambas foram muito próximos dos valores simulados.

Na amostragem de Gibbs, um valor é gerado para cada parâmetro e para cada dado perdido através da distribuição condicionada nos dados observados e nos valores de outras amostras. Após muitas amostras repetidas, distribuições empíricas posteriores dos parâmetros são obtidas. Estas são usadas para se estimar os valores dos parâmetros e os intervalos de confiança. Em geral, amostras muito grandes (10000 observações ou mais) são necessárias para se obter resultados que não sejam dependentes dos valores iniciais (Hoeschele, 1994).

C. Poder estatístico para detectar QTL

C.1. Introdução

O poder estatístico para detectar QTL depende do número de indivíduos genotipados para os marcadores e medidos para as características quantitativas, a magnitude do erro tipo I permitido, o efeito do QTL em relação às variâncias genéticas e do meio ambiente, a taxa de recombinação entre QTL e marcador, o delineamento experimental usado e o método de análise. Considerando-se que o número de combinações possíveis é muito grande, serão apresentados alguns exemplos da literatura e discutidos de forma geral os efeitos de alguns parâmetros sobre o poder estatístico do experimento.

Embora intuitivamente possa parecer que as metodologias estatísticas que proporcionem estimativas mais precisas dos parâmetros sejam as mesmas que aumente o poder de detecção, os resultados com simulação não têm geralmente mostrado isso. A ML que utiliza todas as informações dos dados, aparentemente deveria ter maior poder de detecção do que a ANOVA que utiliza apenas a média e a variância da distribuição. Resultados de simulação (Simpson, 1992) não têm indicado esta superioridade. Na maioria dos casos o poder baseado no teste de t ou ANOVA para análise de uma característica tem sido semelhante ao obtido com a ML.

C.2. Estimativa do poder estatístico em cruzamentos entre linhagens consangüíneas

Diferentes estudos sobre estimativa do poder de detecção do QTL tem considerado o efeito de substituição dos alelos do QTL em termos do desvio padrão fenotípico, residual ou genético. Contudo, considerando que as variâncias genéticas e do resíduo só são conhecidas a posteriori, os efeitos do QTL serão expressos em unidades do desvio padrão fenotípico (SDU). Soller et al. (1976) calcularam o número de progênies necessárias para se obter determinado poder de detecção para os delineamentos BC e F-2, baseando-se no teste de t. No caso do F-2, apenas os indivíduos homozigotos foram considerados. O número de progênies (n) necessário é calculado por:

$$n = \frac{2(Z_{\alpha} + Z_{\beta})^2}{(\delta/\sigma)^2} \quad (C.1)$$

onde Z_{α} e Z_{β} são os valores para os erros tipo I (α) e II (β) obtidos na distribuição normal padrão, δ é a esperança do contraste entre as médias dos grupos de marcadores, e σ é o desvio padrão do resíduo. A esperança do contraste e o número de progênies necessário para se ter um poder de detecção ($1 - \beta$) de 90%, quando $2a = 0,282\sigma$, $r = 0$, $\alpha = 0,05$ e $\beta = 0,1$, são apresentados no Quadro C1. Um locus com essa magnitude de efeito é responsável por 1% da variância fenotípica em um população de F-2.

Quadro C1. Esperança dos contrastes e número de progênies necessário para se obter um poder de detecção de 90% nos delineamentos BC e F-2.

Cruzamentos	Contraste	Tamanho da amostra	Dominância		
			d = -a	d = 0	d = a
BC	$(a-d)(1-2r)$	2n	525	2100	-
F-2	$2a(1-2r)$	4n	1050	1050	1050

O efeito da magnitude do QTL e a taxa de recombinação entre QTL e marcador sobre o tamanho da amostra para se obter um determinado poder é quadrático. Assim, para se obter o mesmo poder quando o efeito do QTL é a metade de um dado valor, torna-se necessário medir-se quatro vezes mais indivíduos. Em ambos os delineamentos (BC e F-2), a magnitude do efeito do QTL medido decrescerá proporcionalmente a $1-2r$ quando comparado a uma ligação completa ($r = 0$). Neste caso, para se obter o mesmo poder de detecção do que quando $r = 0$, será necessário se aumentar o tamanho do experimento equivalente ao fator $1/(1-2r)^2$. Por ex., para $r = 0,1$, o tamanho da amostra deve aumentar de 1,5625, ou seja, passar de 1.050 para 1.641 indivíduos.

No caso de F-2, o poder também pode ser estimado através da ANOVA, incluindo-se todos os três genótipos. A probabilidade da hipótese alternativa é calculada baseada na distribuição não-central de F. Segundo Soller et al. (1976), o poder de detecção ao incluir os heterozigotos só será maior se $|d| > a/2$.

Quando o QTL está entre dois marcadores, o efeito medido do QTL não é reduzido pela taxa de recombinação, exceto no caso de dupla recombinação, mas em uma análise com um modelo linear simples, os indivíduos recombinantes serão eliminados. A proporção de recombinantes para os delineamentos F-2 e BC será de $(1-R)^2$ e $(1-R)$, respectivamente, onde R é a taxa de recombinação entre os dois marcadores. Assim, o poder, quando o QTL está entre dois marcadores, será reduzido por este fator quando comparado à situação de completa ligação. Considerando $R = 2r_1$, que é a situação ótima para o caso de QTL entre dois marcadores, o poder com marcadores flanqueando o QTL será aumentado em $(1-R)$ no caso do delineamento BC e será igual ao da análise com um único marcador para o delineamento F-2 (Weller, 1992). Darvasi et al. (1993) compararam o poder obtido pelo teste de t com um único marcador, com o poder obtido pela razão entre verossimilhança com marcadores flanqueando o QTL. A diferença máxima em poder foi obtida quando os dois marcadores estavam bem distantes um do outro e o QTL localizado bem no meio dos marcadores. Mesmo com uma distância de 50 cM entre os dois marcadores, a diferença

máxima entre poder foi de 8%. Resultados semelhantes também foram obtidos por Haley e Knott (1992).

C.3. Replicação de progênes em cruzamento entre linhagens consangüíneas

Soller e Beckmann (1990) consideraram gerações F-3 e F-4, linhagens consangüíneas recombinantes (RIL), clones vegetais e linhagens haplóides duplas (DHL). No delineamento F-3 cada indivíduo F-2 é autofecundado. Assume-se que apenas os indivíduos F-2 são genotipados, mas as características quantitativas são medidas nos indivíduos F-3. Similarmente no delineamento F-4, os indivíduos F-3 são autofecundados e os indivíduos F-4 são medidos mas não genotipados. Os dados dos genótipos da geração F-2 são analisados. Em ambos os delineamentos, a taxa de recombinação nas gerações futuras (F-3 ou F-4) não afetam as análises porque apenas os indivíduos F-2 são genotipados. No caso da RIL os indivíduos são produzidos após várias gerações de autofecundação, começando-se com os indivíduos F-2. Na geração final, um grupo de indivíduos de cada linhagem será medido para a característica quantitativa, mas apenas um único indivíduo será genotipado para os marcadores. Desde que cada linhagem seja praticamente homozigota e isogênica, é necessário genotipar apenas um indivíduo por marcador, e a variância genética dentro de cada RIL tenderá a zero. Todavia, porque os indivíduos RIL são genotipados após várias gerações de consangüinidade, a taxa de recombinação entre QTL e marcadores, relativa aos F-2, irá aumentar. Os clones vegetais produzidos de indivíduos F-2 são semelhantes às RIL, com relação à variância genética que será zero dentro de cada clone, mas a taxa de recombinação não aumentará, como ocorre com as RIL. As DHL têm as mesmas características estatísticas das RIL quando a taxa de recombinação entre QTL e marcador tende a zero.

O efeito da replicação de progênes sobre o poder estatístico para estes delineamentos pode ser derivado pela análise do seguinte modelo:

$$Y_{ijkl} = A_i + B_j + L_{ik} + e_{ijkl} \quad (C.2)$$

onde Y_{ijkl} é a mensuração do indivíduo l da linhagem k com genótipo i e no bloco j , A_i é o efeito do genótipo i , B_j é o efeito do bloco j , L_{ik} é o efeito da linhagem k aninhada no

genótipo i e e_{ijk} é o efeito do resíduo ao acaso. Este modelo difere do modelo da equação (1) apenas devido à inclusão do efeito da linhagem. A significância do efeito do QTL pode ser testada através do teste de F da razão entre os quadrados médios de A e L vezes o número de linhagens consangüíneas. Isto pode ser calculado como: a esperança do quadrado médio de L será $\sigma_G^2 + \sigma_e^2/n$, onde σ_G^2 é a variância genética entre linhagens, σ_e^2 a variância do resíduo e n o número de indivíduos por linhagem; a esperança do quadrado médio para A para o delineamento F-2 com alelos co-dominantes do QTL e completa ligação será $a^2 + \sigma_G^2/m + \sigma_e^2/(mn)$, onde m é igual ao número de linhagens consangüíneas. Assim, a razão entre o quadrado médio de A e L vezes m terá uma distribuição central de F sob a hipótese de nulidade de $a^2 = 0$ (Um teste de F da razão entre os quadrados médios de A e do resíduo como é feito no modelo da equação (1) dará resultados errados).

A variância genética entre linhagens (σ_G^2) é função da herdabilidade, dominância e da estratégia específica de acasalamento utilizada. A maior vantagem das linhagens consangüíneas é quando a σ_e^2 é grande quando comparada com a σ_G^2 . Baseado em Soller e Beckmann (1990), os componentes de variância entre e dentro de grupos de progênies e o número de linhagens necessárias, relativo ao delineamento F-2, para se obter o mesmo poder estatístico serão apresentados no quadro C2.

Quadro C2. Componentes de variâncias entre e dentro de grupos de progênie e o número de linhagens necessárias, relativo ao delineamento F-2, para se obter o mesmo poder, como função da herdabilidade (h^2) e do número de indivíduos por linhagem (n).

Tipo de Progênie	Componente de Variância		Número de linhagens necessárias relativo ao F-2
	Entre linhagens	Dentro de linhagens	
F-2	h^2	$(1 - h^2)$	1
F-3	h^2	$(1 - h^2/2)/n$	$h^2 + (1 - h^2/2)/n$
F-4	h^2	$(1 - h^2/4)/n$	$h^2 + (1 - h^2/4)/n$
Clones Vegetais	h^2	$(1 - h^2)/n$	$h^2 + (1 - h^2)/n$
RIL	$2h^2$	$(1 - h^2)/n$	$[h^2 + (1 - h^2)/2n][1 - 2r_L]^2 / [1 - 4r/(1 + 2r_L)]^2$
DHL	$2h^2$	$(1 - h^2)/n$	$h^2 + (1 - h^2/2)/2n$

A variância entre linhagens será h^2 para todos os tipos de delineamentos considerados, exceto RIL e DHL que terão variâncias iguais a $2h^2$. A economia com genotipagem poder ser significativa. Por exemplo, para $h^2 = 0,2$ e $n = 10$, apenas 0,29 genótipos serão necessários no delineamento F-3 quando comparado com o F-2. Para todos os delineamentos, exceto RIL, com n grande, o número de linhagens necessárias será uma função direta da herdabilidade. Para o caso de RIL, o poder é também uma função de r . A taxa de recombinação entre QTL e marcador para RIL tenderá a $r_L = 2r/(1 + 2r)$. Assim, o poder de RIL será proporcional a $1/(1-2r_L)^2$, comparado a $1/(1-2r)^2$ do delineamento F-2.

C.4. Estimativa do poder para populações segregantes

Soller e Genizi (1978) estimaram o poder de detecção de QTL para o delineamento de filhas assumindo uma análise de variância aninhada. Weller et al. (1990) estimaram o poder para os delineamentos de filhas e netas usando o teste de qui-quadrado. Quando o tamanho da amostra é grande, os dois métodos dão

virtualmente os mesmos resultados. Weller et al. (1990) assumiram que a soma dos quadrados dos contrastes entre alelos dentro de famílias paternas teriam uma distribuição central de qui-quadrado sob a hipótese de nulidade, e uma distribuição não-central de qui-quadrado sob a hipótese alternativa. Os seus cálculos foram baseados na pressuposição de que apenas dois alelos do QTL estariam segregando na população. Dessa forma a metade dos touros seriam homozigotos para o QTL e a esperança do contraste dos alelos paternos para estas famílias seria zero.

Assim, considerando-se o erro tipo 1 (α), o poder para se detectar ligação entre QTL e marcador é dado por:

$$P = 1 - B = 1 - p[\chi^2(\text{NC}, \text{NS}) < y] \quad (\text{C.3})$$

onde B é o erro tipo 2, $p[\chi^2(\text{NC}, \text{NS}) < y]$ é a probabilidade do valor do qui-quadrado ser menor do que o valor de y, dado o parâmetro não-central NC e NS graus de liberdade, e y é o valor do qui-quadrado central para um erro tipo 1, igual a α e com NS graus de liberdade. Para o caso da hipótese de nulidade, ou seja, não há ligação entre QTL e marcador, $\text{NC} = 0$. Para uma hipótese alternativa qualquer, a esperança do NC pode ser calculada como:

$$\text{NC} = (\text{NS} * F * A^2) / \text{SE}^2 \quad (\text{C.4})$$

onde F é a fração de touros heterozigotos para o marcador ligado ao QTL, A é o efeito médio da substituição de genes do QTL e SE é o erro padrão do contraste $M_1 - M_2$. F, A e SE são obtidos através das equações:

$$F = 2pq \quad (\text{C.5})$$

$$A = a + (p - q)d \quad (\text{C.6})$$

$$\text{SE}^2 = 4(1 - h^2/4)/\text{ND} \quad (\text{C.7})$$

onde p e q são as frequências dos alelos q_1 e q_2 do QTL, a é a metade da diferença entre os dois homozigotos para o QTL, d é a diferença entre o heterozigoto para QTL e a média dos dois homozigotos, h^2 é a herdabilidade da característica e ND é o número

de filhas. Os termos a e d são medidos em unidades do desvio padrão fenotípico. As equações (C.5) e (C.6) são encontradas em FALCONER (1981).

As mesmas equações para o cálculo do poder do delineamento de filhas podem ser utilizadas para o cálculo do poder do delineamento de netas. Para isto, NS = NGS (número de avós) na equação (C.4), SE = SEG e NC = NCG e são calculados como:

$$SEG^2 = (3/4 * h^2 + (4 - h^2)/ NG)/ NO \quad (C.8)$$

$$NCG = (NGS * F * A^2)/4 SEG \quad (C.9)$$

onde NG é o número de netas por filho (touro) e NO é o número de filhos (tours) de cada avô. Os demais termos são iguais aos já definidos. Assim, o poder do delineamento de netas pode ser calculado pela equação (C.3) substituindo-se NC por NCG e NS por NGS.

Devido à natureza discreta do efeito de touro, a fração de touros heterozigotos (F) pode ser muito diferente do que se espera quando se tem um número pequeno destes. Isto é devido à amostragem ao acaso e, portanto, o poder dos delineamentos devem ser calculados considerando-se a natureza discreta da distribuição dos touros, utilizando-se a distribuição binomial com pequenos ajustes nas equações (C.4) ou (C.9) e (C.3). Nas equações (C.4) ou (C.9), a proporção de touros heterozigotos (F) é calculada como:

$$F_i = \frac{N_i}{NS}, \text{ para } i = 1, 2, \dots, NS, \text{ ou seja, } F_i \text{ varia de } 0 \text{ a } 1;$$

onde N_i é o número de touros heterozigotos e NS o número total de touros sendo considerados, e

$$PF = \sum_{i=1}^{NS} pb_i * P_i \quad (C.10)$$

onde PF é o poder final, pb_i é a probabilidade através da distribuição binomial de que o touro i é heterozigoto e P_i é o poder como calculado em (C.3).

Welller et al. (1990) assumiram completa ligação entre QTL e marcador e consideraram efeitos de substituição de 0,1, 0,2 e 0,3. Quando não há dominância entre

os alelos do QTL, o efeito da substituição é igual à metade da diferença entre as médias dos homozigotos. Para o delineamento de filhas, poder de 70% com erro tipo I de 0,01 é obtido para um QTL com efeito de substituição de 0,2 SDU se 400 filhas de cada 10 touros forem analisadas para uma característica de herdabilidade igual a 0,2. Isto significa ter-se que genotipar 4.000 indivíduos. Em seu paper, Weller et al. (1990) apresenta um quadro com várias combinações de efeito de substituição, número de touros e número de filhas. Quando a frequência dos alelos dos dois QTL é igual, o poder de detecção será máximo. Para o caso de alelos co-dominantes, a frequência dos alelos afetam o poder apenas através do número esperado de touros heterozigotos. Para frequências variando de 0,3 a 0,7, o poder será muito próximo do que se obtém quando a frequência é igual a 0,5, e conseqüentemente o efeito da frequência sobre o poder é muito pequeno.

A situação com o delineamento de netas é similar a do delineamento F-3 considerado anteriormente, em que os filhos são genotipados enquanto as medições são realizadas em suas progênies. Semelhantemente ao delineamento F-3, o poder de delineamento de netas aumenta por indivíduo genotipado porque muitos fenótipos (netas) serão analisados por indivíduo genotipado. Da mesma forma, o poder não é afetado por mais uma geração de recombinação. Ao contrário do delineamento F-3, apenas a metade dos efeitos poligênico e do contraste do QTL são transmitidos para as suas netas. Como no caso das linhagens consangüíneas, aumentando-se o número de netas reduzirá a variância residual, mas não a variância genética entre filhos. Dessa forma a vantagem do delineamento de netas é maior para as características de baixa herdabilidade. Com uma herdabilidade de 0,2 erro tipo I de 0,01, o poder de se detectar um QTL com efeito de substituição de 0,2 SDU será de 74%, se 100 filhos de cada 10 avôs são genotipados e 50 netas por filho são mensuradas para as características quantitativas. Comparando os exemplos do delineamento de netas com o de filhas, observa-se que maior poder é obtido com o delineamento de netas, mesmo quando apenas um quarto dos indivíduos são genotipados em relação ao delineamento de filhas (1.000 vs. 4.000). Weller et al. (1990) apresentam um quadro com o poder de detecção para o delineamento de netas considerando variações em número de avôs, filhos, netas, etc.

Martinez (1998), utilizando-se da metodologia proposta por Weller et al. (1990), calculou os valores máximos do poder estatístico para os delineamentos de netas e filhas variando-se o efeito da substituição de alelos, quando o número de indivíduos a serem genotipados é fixo. Quando apenas 400 indivíduos podem ser genotipados (quadro C3), o poder é maior quando se tem 10 avôs com 40 filhos do que quando se tem 20 avôs com 20 filhos. Em todas as combinações estudadas, à medida que houve aumento do número de filhas por touro, houve diminuição do efeito da estrutura da família, ou seja, mais touros por avô versus o número de avôs com menos touros. À semelhança do delineamento de netas, o poder de delineamento de filhas é afetado pela estrutura das famílias (quadro C4). Quando o número de indivíduos a ser genotipado é de 4.000, o poder para um efeito de substituição de alelos do QTL de 0,2 é mais do que o dobro (56% versus 23%), com uma estrutura de 20 touros com 200 filhas/touro do que com 80 touros com 50 filhas/touro. Isto significa que um planejamento adequado da estrutura de família pode aumentar o poder de detecção sem aumentar os custos com as análises dos genótipos.

Os efeitos da taxa de recombinação entre QTL e marcador, da frequência dos alelos do QTL e da substituição de alelos para características com herdabilidades menores de 0,1 foram estudados por Martinez (1998a) para o delineamento de filhas. Neste estudo a frequência gênica apresentou um grande efeito sobre o poder quando esta variou de 0,5 para 0,1. Observou-se que, quando a frequência gênica é de 0,5, são necessários 4.000 indivíduos (40 touros com 100 filhas/touro) para se obter um poder de 90%, com erro tipo I de 0,01. Todavia, se a frequência gênica for de 0,1, mesmo com 10.000 indivíduos (100 touros com 100 filhas/touro), obtém-se apenas um poder de 24%. Neste caso, para se obter poder equivalente (90%), seriam necessários 80 touros com 400 filhas/touro, ou seja, oito vezes mais indivíduos deveriam ser analisados genética e fenotipicamente.

Outro fator que reduz substancialmente o poder é quando a taxa de recombinação não é igual a zero. Neste trabalho de Martinez (1998a), o poder de 90% foi estimado com 20 touros com 400 filhas/touro quando a taxa de recombinação era zero. Todavia, se a taxa for igual a 0,1, o poder já diminui para 76% e será apenas 51% se a taxa for de 0,2.

Com base nestes resultados de Weiler et al. (1990) e Martinez (1998, 1998a), é possível se tirar as seguintes conclusões sobre o poder de detectar QTL utilizando-se dos delineamentos de filhas e netas:

- a) poder é maior quando se tem poucas famílias grandes do que muitas famílias pequenas, independentemente do delineamento de filhas ou netas;
- b) Com herdabilidade de 0,2, o mesmo poder com o delineamento de filhas pode ser obtido no delineamento de netas com apenas 25% de animais genotipados;
- c) Para um efeito de substituição fixo expresso em desvio padrão fenotípico, o poder com o delineamento de netas diminui com o aumento da herdabilidade;
- d) À semelhança da replicação de progênies em delineamentos de linhagens consangüíneas, aumentando-se o número de netas por filho além de 50, trará incremento muito pequeno no poder estatístico.

O poder diminui nos delineamentos com replicação de progênies com o aumento da herdabilidade se o efeito do QTL é expresso em unidades de desvio padrão fenotípico. Embora o desvio padrão fenotípico seja economicamente relevante, é a variância genética que deve ser explicada pelo QTL. Quando o QTL é expresso relativo ao desvio padrão genético, não há praticamente nenhuma relação entre herdabilidade e poder, se o número de netas for grande.

Quadro C3. Valores máximos do poder¹ para diferentes combinações do número de avôs e touros a serem genotipados para um número fixo de animais a serem analisados no delineamento de netas.

Genótipo	Número total de			A ² :	Poder (% de detecção)		
	Avô	Touros/avô	Filhas/touro		0,1	0,2	0,3
200	5	40	25	2	10	30	
			50	3	15	43	
			100	3	19	52	
	10	20	25	2	6	20	
			50	2	9	31	
			100	2	12	40	
400	10	40	25	3	17	52	
			50	4	26	69	
			100	5	34	78	
	20	20	25	2	10	37	
			50	3	16	54	
			100	3	21	66	
1.000	5	200	25	14	68	91	
			50	20	79	94	
			100	26	85	96	
	10	100	25	8	58	93	
			50	13	74	97	
			100	17	83	98	
2.000	10	200	25	24	90	99	
			50	36	96	99	
			100	46	98	99	
	20	100	25	14	85	99	
			50	23	95	99	
			100	31	98	99	

¹ Considerando-se $\alpha = 0,01$, $p = q = 0,5$, $h^2 = 0,2$ e $r = 0$.

² A = efeito médio da substituição de alelos (em unidades de desvio padrão fenotípico)

Quadro C4. Valores máximos do poder¹ para diferentes combinações do número de touros e filhas por touro a serem genotipados para um número fixo de animais a serem analisados no delineamento de filhas.

Genótipos	Número total de		A ² :	Poder (% de detecção)		
	Touros	Filhas/touro		0,1	0,2	0,3
1.000	20	50		2	7	25
	40	25		1	4	15
1.500	20	75		2	13	47
	60	25		2	6	21
2.000	20	100		3	21	68
	40	50		2	12	47
	80	25		2	7	28
3.000	40	75		3	24	78
	60	50		3	17	65
4.000	20	200		7	56	95
	40	100		4	39	93
	80	50		3	23	77
6.000	60	100		6	56	98
	80	75		5	46	97

¹ Considerando-se $\alpha = 0,01$, $p = q = 0,5$, $h^2 = 0,2$ e $r = 0$.

² A = efeito médio da substituição de alelos (em unidades de desvio padrão fenotípico)

Referências Bibliográficas

- BOVENHUIS, H. AND WELLER, J.I. (1994) Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers. *Genetics* **137**: 267-280.
- CANTET, R.J.C., AND SMITH, C. (1991) Reduced animal model for marker assisted selection using best linear unbiased prediction. *Gen. Sel. Evol.* **23**: 221-233.
- DARVASI, A., VINREB, A., MINKE, V., WELLER, J.I., AND SOLLER, M. (1993) Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* **134**: 943-951.

- FALCONER, D.S. (1981) *Introduction to Quantitative Genetics*, 2nd Ed. Logman, Inc. New York.
- FERNANDO, R., GROSSMAN, M. (1989) Marker assisted selection using best linear unbiased prediction. *Gen. Sel. Evol.* **21**: 467-477.
- GEORGES, M., NIELSEN, D. MACKINNON, M., MISHRA, A., OKIMOTO, R., PASQUINO, A.T., SARGENT, L.S., SORENSEN, A., STEELE, M.R., ZHAO, X., WOMACK, J.E., AND HOESCHELE, I. (1995) Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* **139**: 907-920.
- HACKETT, C.A. AND WELLER, J.I. (1995) Genetic mapping of quantitative trait loci for traits with ordinal distributions. *Biometrics* **51**: 1252-1263.
- HALEY, C.S. AND KNOTT, S.A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315-324.
- HASEMAN, J.K. AND ELSTON, R.C. (1972) The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.* **2**: 3-19.
- HOESCHELE, I. AND VanRADEN, P.M. (1993) Bayesian analysis of linkage between genetic markers and quantitative trait loci. I. Prior knowledge. *Theor. Appl. Genet.* **85**: 953-960.
- HOESCHELE, I. AND VanRADEN, P.M. (1993a) Bayesian analysis of linkage between genetic markers and quantitative trait loci. II. Combining prior knowledge with experimental evidence. *Theor. Appl. Genet.* **85**: 946-952.
- HOESCHELE, I. (1994) Bayesian QTL mapping via the Gibbs sampler. *Proc. 5th World Cong. Genet. Appl. Livest. Prod. Guelph* **21**: 241-244.
- MACKINNON, M.J. AND WELLER, J.I. (1995) Methodology and accuracy of estimation of quantitative trait loci parameters in a half-sib design using maximum likelihood. *Genetics* **141**: 755-770.
- MARTINEZ, O. AND CURNOW, R.N. (1992) Estimating the locations and the sized of the effects of quantitative trait loci using flanking markers. *Theor. Appl. Genet.* **85**: 480-488.
- MARTINEZ, M.L. (1998) Eficiência dos delineamentos de filhas e netas para detectar ligação entre marcador e loci de característica quantitativa. *R. Bras. Zoot.*, v. 27, n. 4, p. 714-721.

- MARTINEZ, M.L. (1998a) Efeitos da substituição de alelos, frequência gênica e da taxa de recombinação sobre o poder de se detectar ligação entre um marcador genético e loci de característica quantitativa de baixa herdabilidade. *R. Bras. Zoot.*, v. 27, n. 4, p. 722-728.
- NEIMANN-SOERESSEN, A. AND ROBERTSON, A. (1961) The association between blood groups and several production characters in three Danish cattle breeds. *Acta Agr. Scand.* 11: 163-196.
- RON, M., YOFFE, O., EZRA, E., MEDRANO, J.F., AND WELLER, J.I. (1994) Determination of milk protein effects on production traits of Israeli Holsteins. *J. Dairy Sci.* 77: 1106-1113.
- SIMPSON, S.P. (1992) Correction: Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. *Theor. Appl. Genet.* 85: 110-111.
- SOLLER, M., GENIZI, A., AND BRODY, T. (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Appl. Genet.* 47: 35-59.
- SOLLER, M. AND 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* 34: 47-55.
- SOLLER, M. AND BECKMANN, J.S. (1990) Marker-based mapping of quantitative trait loci using replicated progeny. *Theor. Appl. Genet.* 80: 205-208.
- WELLER, J.I. (1986) Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. *Biometrics* 42: 627-640.
- WELLER, J.I., KASHI, Y. AND SOLLER, M. (1990) Power of "daughter" and "granddaughter" designs for genetic mapping of quantitative traits in dairy cattle using genetic markers. *J. Dairy Sci.* 73: 2525-2537.
- WELLER, J.I. (1992) Statistical methodologies for mapping and analysis of quantitative trait loci. In: J.S. Beckmann and T.C. Osborn (Eds.) *Plant Genomes: Methods for Genetic and Physical Mapping*. pp. 181-207. Kluwer Academic Publishers Group, Dordrecht, The Netherlands.

ESTUDO DO DESENVOLVIMENTO MUSCULAR EM AVES

Luiz Lehmann Coutinho

Laboratório de Biotecnologia

ESALQ-USP

1. Introdução

O ser humano pode ter uma vida saudável consumindo apenas produtos de origem vegetal, no entanto, uma dieta contendo produtos animais e em particular a carne, propicia uma alimentação mais saudável e completa. Na dieta norte americana, a carne contribui com 20% da energia, 41% da proteína, 3,5% do cálcio, 26% do fósforo, 29% do ferro, 22% da vitamina A, 28% da tiamina, 24% da riboflavina e 46% da niacina consumida por humanos (Agricultural Statistics, 1974). A carne é um alimento rico em proteínas, vitaminas e minerais. A proteína da carne possui uma composição de aminoácidos mais adequada às exigências humanas, suprimindo, por exemplo, a deficiência de lisina que ocorre em cereais. A presença de minerais como o ferro e vitaminas como a B12, que não estão presentes nos vegetais, reforça a importância da carne como alimento.

Apesar de sua indiscutível qualidade nutricional, o uso de carne na alimentação humana é discutida devido a sua ineficiência energética e protéica. Em frangos de corte, é preciso fornecer 2,4 kg de ração, 14,1 kcal e 1,9 kg de proteína para produzir 1,0 kg de frango, 1,0 kcal de produto final e 1,0 kg de proteína na carcaça, respectivamente. A ineficiência do processo deve-se às perdas em função da digestão, metabolismo e gastos de manutenção do animal. De interesse particular para este trabalho são as perdas relacionadas ao metabolismo. A maneira como os nutrientes absorvidos pelo animal são direcionados para os tecidos muscular e adiposo tem implicações diretas na eficiência do processo produtivo. Como pode ser visto na figura 1, a energia proveniente de 10 kcal de energia metabolizável, se utilizada para a deposição de músculo, resultaria em 3,5 kcal de energia depositada, 0,64 g de proteína e 2,6 g de tecido muscular. Se por outro lado, esta mesma quantidade de energia fosse utilizada para a

deposição de gordura, resultaria em 6,0 kcal de energia depositada, 0,64 g de gordura e 0,7 g de tecido adiposo (Geay, 1984). Portanto, é bastante claro que para uma mesma quantidade de alimento consumido, o ganho de peso será maior se uma maior parte da energia for utilizada para a deposição de tecido muscular em vez de tecido adiposo. Como consequência, a conversão alimentar também será melhor.

Eficiência de deposição dos tecidos muscular e adiposo

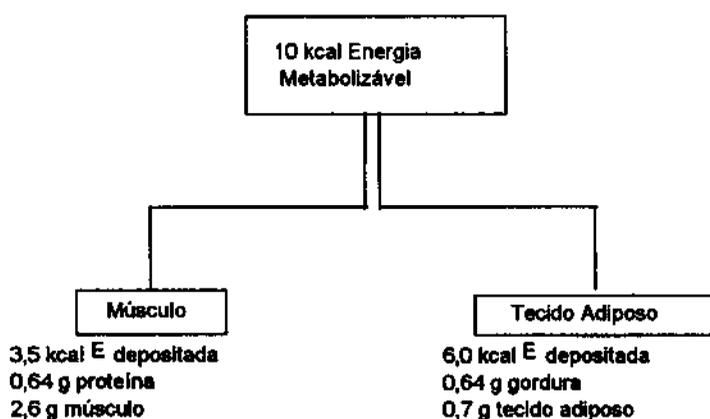


Figura 1: Eficiência da utilização da energia pelos tecidos adiposo e muscular.

É importante ressaltar que a maior utilização de energia e proteína do alimento para a deposição de músculo em vez de gordura não estaria aumentando a eficiência energética do processo, visto que a eficiência para depositar gordura no tecido adiposo é bastante alta. No entanto, o objetivo principal da produção animal é a produção de proteína de alta qualidade. A gordura, apesar de sua importância fisiológica para a sobrevivência do animal, quando depositada em excesso, torna-se um problema. Boa parte da gordura depositada não é utilizada como alimento por seres humanos e, ao contrário, constitui um problema para a saúde. Consumo elevado de gordura está associado com obesidade, câncer e doenças cardiovasculares. Portanto, estudos ou

processos que venham aumentar a deposição da musculatura esquelética em animais estarão contribuindo para aumentar a eficiência do processo produtivo e a qualidade de produtos de origem animal.

As estratégias atualmente utilizadas para aumentar a deposição de músculo concentram-se na manipulação ou seleção dos animais após o nascimento, sendo os eventos pré-natais ignorados. Alguns exemplos destas estratégias são o uso de repartidores de nutrientes (agonistas beta-adrenérgicos), promotores de crescimento e seleção massal. O processo de seleção tem gerado resultados particularmente interessantes. Nesse procedimento animais são selecionados por várias gerações resultando em aumento da frequência dos genes associados com características fenotípicas desejáveis. Tal estratégia é utilizada na produção de frangos de corte, gerando entre 1 a 2% de aumento de ganho de peso ao ano. Uma outra alternativa é a manipulação de genes por meio de técnicas de biologia molecular. Para tal, é necessário o conhecimento profundo dos genes envolvidos na determinação da característica fenotípica de interesse. Neste enfoque, a presente revisão foi conduzida com o objetivo de discutir alguns aspectos dos recentes avanços do conhecimento científico sobre o controle molecular do desenvolvimento da musculatura esquelética em aves.

2. Tecido Muscular

O tecido muscular no animal pode ser classificado em três tipos: musculatura lisa, presente nos vasos sanguíneos e sistema digestivo; musculatura cardíaca, presente no coração; e musculatura esquelética. A musculatura esquelética tem um aspecto estriado, apresenta movimento voluntário e está ligada direta ou indiretamente aos ossos.

No tecido muscular encontra-se tecido conjuntivo, vasos sanguíneos, fibras nervosas, fluidos extracelulares e fibras musculares. A fibra muscular é a unidade fundamental do tecido e representa entre 75 e 92% do volume do músculo. O músculo é formado por fibras musculares agrupadas em feixes, que reunidos formam um grupo muscular. O tecido conjuntivo que envolve cada fibra muscular é chamado de

endomísio, o que envolve um grupo de fibras musculares, perimísio e o que recobre todo o músculo é chamado epimísio.

A fibra muscular pode ter de 10 a 100 μm de diâmetro, vários centímetros de comprimento e pode possuir entre 1.000 e 2.000 miofibrilas que se estendem por todo o comprimento da célula. Estas miofibrilas, por sua vez, são formadas por miofilamentos. Estes, junto com outras 20 proteínas, são responsáveis pelo mecanismo de contração muscular. As principais proteínas presentes são actina e miosina (responsáveis pela função contrátil) e tropomiosina, troponina, alfa-actina e beta-actina (com funções regulatórias).

O aumento no diâmetro (hipertrofia) ou no número (hiperplasia) das fibras musculares resulta em aumento na massa muscular. Em animais em crescimento, a hipertrofia das fibras é o principal processo envolvido no aumento da musculatura esquelética, visto que praticamente não ocorre aumento no número de fibras após o nascimento de mamíferos e aves. O exercício físico pode aumentar o diâmetro de fibras musculares, no entanto, este procedimento exige atividade física intensa, requer dieta rica em energia e proteína e é bastante ineficiente na conversão de ração em proteína muscular.

2.1. Origem do músculo

Os somitos (blocos segmentares de mesoderme paraxial) são as estruturas que dão origem à musculatura esquelética. Durante o desenvolvimento embrionário os somitos são formados ao longo do tubo neural. Inicialmente, o somito é composto de células colunares arranjadas radialmente ao redor de um pequeno lúmen. A parte ventro-medial do somito dá origem ao esclerótomo, enquanto que as células dorso-laterais dão origem ao dermomiótomo (Kahen et al., 1988). Em uma etapa subsequente, células da porção dorso-medial dos somitos migram sob o dermomiótomo, dando origem ao miótomo e, por sua vez, estas células diferenciam-se, formando as primeiras células musculares do embrião. Células individuais mononucleadas vão originar estruturas da musculatura axial e do tronco, enquanto uma

segunda população de células migratórias deverá colonizar outras regiões do embrião, tais como os membros em desenvolvimento e a parede ventral do corpo (Cossu, 1996).

Ordahl e Le Douarin (1992) demonstraram em codornas que as células somíticas próximas ao tubo neural dão origem à musculatura das costas, enquanto que as células mais afastadas migram para outras regiões e dão origem à musculatura do peito, membros e outras.

2.2. Controle molecular da formação de músculo

Os fatores miogênicos pertencem a um grupo de genes (MyoD, Myf-5, MRF4 e miogenina) envolvidos nos processos de determinação e diferenciação de células musculares. Quando expressos em células fibroblásticas, estes genes têm a capacidade de transformá-las em células musculares (Weintraub et al., 1991).

A hipótese atual sobre o desenvolvimento muscular é que a determinação miogênica é controlada pela ativação de um ou mais *loci* miogênicos de determinação, provavelmente por um mecanismo de hipometilação. A ativação destes "loci" miogênicos causa profundas mudanças na expressão gênica pela indução de um número de genes músculo-específicos capazes de direcionar as células mesodérmicas, levando-as indiretamente ou diretamente à ativação da miogênese (Olson et al., 1995).

Estudos de hibridização *in situ* em roedores demonstraram que Myf-5 é ativado antes de qualquer outro gene específico na formação da musculatura esquelética e é, portanto, considerado um gene marcador do desenvolvimento muscular. Após a sua ativação, miogenina e MyoD são ativados (Ott et al., 1991). MRF4 também é ativado em embriões, mas apenas de maneira transitória. O MRF4 parece ser mais importante em animais adultos, pois está presente em quantidades bem superiores aos outros fatores miogênicos (Hinterberger et al., 1991).

Em experimentos conduzidos com animais transgênicos, nos quais fatores miogênicos foram expressos em vários tecidos, observou-se a ativação de proteínas da musculatura esquelética no coração e no cérebro (Miner et al., 1992; Santerre et al., 1993).

Para documentar ainda mais a importância destes fatores, animais mutantes na expressão de miogenina foram obtidos. Nestes animais, a formação da musculatura esquelética foi bastante prejudicada (Hasty et al., 1993). Em resumo, estes genes são responsáveis pela determinação e diferenciação de células multipotentes em células musculares.

O padrão de ativação dos genes em aves durante o desenvolvimento embrionário é diferente do que ocorre em mamíferos. Em codornas, o primeiro gene a ser ativado é qmf1 (MyoD), seguido por qmf3 (Myf-5) e qmf2 (miogenina) (Pownall & Emerson, 1992; Coutinho et al., 1993).

O processo de indução de células somíticas para a formação de músculo parece ser dependente de sinais oriundos do tubo neural. Estudos conduzidos por Rong et al., (1992) mostraram que somitos imaturos, quando afastados do tubo neural por meio de microcirurgia, não formam células musculares. Quando os somitos são reaproximados do tubo neural, observa-se a formação de células musculares. Da mesma forma, estudos preliminares conduzidos por Pownall et al. (1996) revelaram que, quando afastados do tubo neural, os somitos não ativam qmf1 e não formam células musculares. No entanto, se uma seção do tubo neural for enxertada junto ao somito, ocorre a ativação de qmf1. Estes estudos revelam que a ativação do processo miogênico pode ser manipulada através do afastamento e aproximação dos somitos do tubo neural.

Estudos conduzidos por Langman e Nelson (1968) relataram que as células do miótomo não se dividem e que o processo de multiplicação celular ocorre no dermatomo. É interessante observar que em cultura de tecidos, a expressão de MyoD inibe a proliferação celular, no entanto tal relação não foi investigada em células somíticas de embriões.

3. Correlação entre expressão de fatores miogênicos e desenvolvimento muscular.

Estes estudos envolveram o uso de codornas selecionadas, por 87 gerações, para aumento no ganho de peso. Os animais selecionados, denominados linhagem P, apresentavam a quatro semanas de idade, peso 3 vezes superior ao da população não selecionada (controle). Estudos conduzidos por Fowler et al. (1980) revelaram que a linhagem P apresentava aos 10 dias após a eclosão, aumento de 3,5 vezes no peso da musculatura peitoral. Este músculo apresenta também aumento de 3 vezes no conteúdo de DNA, sem diferença na proporção proteína/DNA. Este estudo revelou ainda aumento de 1,6 vezes no número de fibras musculares no músculo semimembranoso. Em estudos conduzidos durante a fase embrionária, quando inicia-se a formação de músculo, Coutinho et al. (1993) observaram atraso na formação de somitos e na expressão de qmf1, qmf2, qmf3 e miosina na linhagem P. Nesta linhagem observou-se também um atraso na ativação de qmf1 nos membros. Em estudos preliminares, conduzidos com animais selecionados para diminuição de ganho de peso, foi observado adiantamento na formação de somitos e na expressão de qmf1. Estes resultados indicam uma relação bastante interessante entre a expressão dos fatores miogênicos e a capacidade de deposição de musculatura esquelética. Como estes fatores são responsáveis pela diferenciação de células multipotentes em células musculares, é bastante provável que a manipulação dos mesmos modifique a deposição de músculo em animais.

4. Referências Bibliográficas

- COSSU, G. How is myogenesis initiated in the embryo? *Trends Gen.*, v.12, n.6, p. 218-23. 1996.
- COUTINHO, L.L.; MORRIS, J.; MARKS, H.L.; BUHR, J.R.; IVARIE, R. Delayed somite formation in a quail line exhibiting myofiber hyperplasia is accompanied by delayed expression of myogenic regulatory factors and myosin heavy chain. *Development*, v.117, p.563-569, 1993.

- FOWLER, S. P.; CAMPION, D.R.; MARKS, H.L.; REAGAN, J.O. Analysis of skeletal muscle response to selection for rapid growth in Japanese quail (*Coturnix coturnix japonica*). **Growth**, v.44, p.235-252, 1980.
- GEAY, Y.. Energy and protein utilization in growing cattle. **J. Anim. Sci.**, v.58, p.766-778, 1984.
- HASTY, P.; BRADLEY, A.; MORRIS, J.H.; EDMONDSON, D.G.; VENUTI, J.M.; OLSON, E.N.; KLEIN, W.H. Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. **Nature**, v.5;364, n.6437, p.501-506, 1993.
- HINTERBERGER, T.J.; SASSOON, D.A.; RHODES, S.J.; KONIECZNY, S.F. Expression of the muscle regulatory factor MRF4 during somite and skeletal myofiber development. **Dev. Biol.**, v.147, n.1, p.144-56, 1991.
- KAHEN, K. J.; CHRIST, H. J.; HINRICHSEN, K.; POELMANN, R. E. The onset of myotome formation in the chick. **Anat Embryol**, v.117, p.191-201. 1988.
- LANGMAN, J.; NELSON, G.R. A radioautograph study of the development of the somite in the chick embryo. **J. Embryology exp Morpho.**, v.19, n.2, p.217-226, 1968.
- MINER, J. H.; MILLER, J. B.; WOLD, B. J. Skeletal muscle phenotypes initiated by ectopic MyoD in transgenic mouse heart. **Development**, v.114, n.4, p.853-60, 1992.
- OLSON E.N.; PERRY, M.; SCHULZ R.A. Regulation of muscle differentiation by the MEF2 family of MADS box transcription factors. **Dev. Biol.**, v.172, p.2-14, 1995.
- ORDAHL, C. P.; LE DOUARIN, N. M. Two myogenic lineages in the development somite. **Development**, v.14, p.339-352, 1992.
- OTT, M. O.; BOBER, E.; LYONS, G.; ARNOLD, H.; BUCKINGHAM, M. Early expression of the myogenic regulatory gene, Myf-5, in precursor cells of skeletal muscle in the mouse embryo. **Development**, v.111, n.4, p.1097-107, 1991.
- POWNALL, M.E.; EMERSON, C.J. Sequential activation of three myogenic regulatory genes during somite morphogenesis in quail embryos. **Dev. Biol.**, v.151, n.1, p.67-79, 1992.
- POWNALL M.E.; STRUNK K.E.; EMERSON C.P. JR. Notochord signals control the transcriptional cascade of myogenic bHLH genes in somites of quail embryos. **Development**, v.122, n.5, p.1475-1488, 1996.
- RONG, P.M.; TEILLET, C.; LE DOUARIN, N.M. The neural tube /nothocord complex is necessary for vertebral but not limb and body wall striated muscle differentiation. **Development**, v.115, p.657-672, 1992.

SANTERRE R.F.; BALES K.R.; JANNEY M.J.; HANNON K.; FISHER L.F.; BAILEY C.S.; MORRIS J.; IVARIE R.; SMITH C.K. 2D. Expression of bovine myf5 induces ectopic skeletal muscle formation in transgenic mice. **Mol. Cell Biol.**, v.13, n.10, p.6044-6051, 1993

WEINTRAUB, H.; DWARKI, V. J.; VERMA I.; DAVIS, R.; HOLLENBERG, S.; SNIDER, L.; LASSAR, A.; TAPSCOTT, S. J. Muscle-specific transcriptional activation by MyoD. **Genes Dev.**, v.5, n.8, p.1377-86, 1991.

RT-PCR NO ESTUDO DA EXPRESSÃO GÊNICA

Lúcia Elvira Alvares

A RT-PCR é uma técnica extremamente sensível que permite detectar a atividade de genes específicos, em pequenas amostras de tecidos ou até mesmo em uma única célula. Esta técnica envolve a síntese de um transcrito de DNA complementar (DNAc) a partir de moléculas de RNA, utilizando-se a enzima Transcriptase Reversa. O DNAc sintetizado é então amplificado exponencialmente pela reação em cadeia da polimerase (PCR).

De uma forma geral, a RT-PCR tem vantagens significativas sobre os ensaios mais tradicionais com RNA. Uma delas é a sua grande sensibilidade, que teoricamente permite amplificar cópias únicas de um RNAm particular, mesmo em amostras de RNA altamente complexas. Para fins de comparação, enquanto que para um ensaio de RT-PCR é preciso dispor de apenas 1 µg de RNA total, para um ensaio de "Northern blotting" são necessárias 30 µg de RNA total. Outra vantagem importante, é que as condições de reação podem ser facilmente otimizadas, até obter-se um máximo de especificidade. Isto permite, por exemplo, que alvos que difiram em apenas um único nucleotídeo possam ser diferenciados. Além disto, a RT-PCR, mais do que qualquer outra metodologia disponível, é ideal para análises rápidas de um grande número de amostras. Em razão de todas estas vantagens, questões biológicas que antes dificilmente poderiam ser resolvidas pelas técnicas convencionais de análise de RNA podem ser solucionadas pela RT-PCR.

Apesar do grande impacto da técnica de RT-PCR em muitas áreas de pesquisa, a aplicação desta metodologia para fins de quantificação de ácidos nucléicos tem apresentado dificuldades. A natureza exponencial da reação de PCR faz com que, na prática, pequenas variações na eficiência de amplificação tomem-se variações grandes na amplificação do produto. Até mesmo diferenças mínimas de temperatura que eventualmente possam ocorrer entre os poços do termociclador podem causar variações significativas entre amostras idênticas. Por exemplo, usando o mesmo protocolo de amplificação, uma única amostra preparada a partir de um mesmo "mix de

PCR", pode produzir resultados que diferem em até seis vezes quanto à quantidade de produto amplificado nas replicatas individuais (Gilliland et al., 1990).

Chelly et al. (1988) desenvolveram um método para análises semi-quantitativas dos níveis de RNAs específicos pela técnica de RT-PCR, utilizando um RNA endógeno como controle interno. O controle interno é co-amplificado com o RNA de interesse, agindo como um indicador das variações entre as diferentes amostras nas reações de síntese de DNA e na PCR. Este tipo de controle permite monitorar a extensão da degradação e recuperação das amostras de RNA. Após normalização com base no controle interno, os níveis de RNA de cada amostra individual podem ser comparados diretamente. Transcritos codificando enzimas metabólicas, proteínas ribossomais ou fatores de alongação da tradução, têm sido particularmente úteis como controles internos, uma vez que eles tendem a ser expressos de forma ubíqua.

Uma limitação da RT-PCR semi-quantitativa é que a abundância do RNA controle deve ser semelhante à do RNA de interesse, de forma que os amplicons tenham cinéticas de amplificação semelhantes, permitindo a sobreposição da fase exponencial de amplificação. Além disto, para que esta técnica tenha cunho quantitativo, é fundamental que as amostras sejam testadas dentro da fase exponencial da reação de PCR. Isto porque, ao atingir o platô, a eficiência da reação de PCR diminui e as concentrações relativas dos amplicons controle e de interesse começam a variar. Este efeito é resultado de vários fatores, como a saturação do substrato da polimerase, a desnaturação incompleta das seqüências molde e a competição de amplicons que passam a auto-anelar e competir com os primers. Além disto, se o controle interno e o RNA de interesse são amplificados por pares de primers diferentes e divergem significativamente em seqüência, tamanho ou estrutura secundária, é necessário controlar também as diferenças entre as eficiências de amplificação (Kevin et al., 1993).

Wang et al. (1989) elaboraram uma abordagem alternativa aos métodos de análise semi-quantitativa por RT-PCR. Utilizando como controle interno da reação de PCR um RNA heterólogo, estes autores conseguiram quantificar, em termos absolutos, doze tipos diferentes de RNAs. Este procedimento é hoje conhecido como RT-PCR quantitativo ou QC-RT-PCR ("competitive quantitative - RT-PCR"), uma vez que um

controle interno contendo os mesmos sítios de ligação aos primers que um RNAm endógeno, compete com este último pela amplificação por PCR. Desta forma, espera-se que os RNAs endógeno e controle sejam amplificados com a mesma eficiência, uma vez que cada um deles compete igualmente pela amplificação, a despeito de qualquer alteração nas condições de reação. Posteriormente, as moléculas competidoras podem ser separadas das moléculas endógenas por tamanho ou digestão com enzimas de restrição.

Protocolos cada vez mais simples, muitas vezes utilizando o PCR como principal ferramenta, têm sido descritos para a construção de moléculas competidoras idênticas ao RNAm endógeno, exceto pela presença de um sítio de restrição ou de uma pequena inserção ou deleção. Moléculas competidores feitas de RNA são preferíveis às de DNA, em razão das variações de eficiência na síntese do DNAc. Embora a RT-PCR venha sendo aplicado satisfatoriamente usando moléculas competidoras heterólogas, que têm apenas os sítios de ligação com os primers em comum com o RNAm endógeno, a redução das diferenças entre as moléculas controle e endógena podem teoricamente aumentar a acuidade da quantificação. Isto porque diferenças de tamanho ou seqüência entre o RNA endógeno e competidor podem resultar em diferenças na estrutura secundária e, conseqüentemente, na eficiência da amplificação, gerando desvios nos resultados. Por outro lado, quando um controle interno difere do alvo de interesse apenas por um sítio de restrição, a formação de heteroduplex pode ocorrer quando atinge-se a fase de platô do PCR, mascarando a proporção real entre as moléculas controle e endógena (Kevin et al., 1993).

Os experimentos de QC-RT-PCR são geralmente bastante trabalhosos, uma vez que são necessárias de três a cinco reações de PCR para cada amostra a ser analisada. Nestas reações, procede-se a co-amplificação de uma quantidade constante de DNAc com várias diluições seriais do competidor. Após quantificação, é feita uma curva, plotando-se a proporção logarítmica entre os amplicons endógeno e competidor contra a quantidade de competidor (log da concentração) adicionada às reações. Desta forma, é possível determinar o ponto de equimolaridade (log endógeno/competidor = zero), que corresponde ao ponto em que o RNA controle e endógeno são equimolares e que tinham, portanto, as mesmas concentrações iniciais na reação (log zero = 1).

Embora a realização de várias titrações em múltiplas amostras de RNA seja trabalhosa, esta técnica trouxe um avanço significativo, permitindo que quantificações absolutas possam ser feitas mesmo após a fase exponencial da reação (Tsai e Wiltbank, 1996).

Procedimentos mais simples têm sido descritos, como por exemplo a construção de uma única curva padrão para cada ensaio, com a qual as amostras analisadas são comparadas diretamente (Tsai e Wiltbank, 1996). A curva padrão é produzida com diferentes quantidades do RNA endógeno, que é co-amplificado com uma quantidade constante do RNA competidor. O número de transcritos na amostra problema é determinado por RT-PCR desta amostra com a mesma quantidade de competidor usada para a construção da curva padrão. Comparando-se a proporção entre os produtos da amostra (endógeno/ competidor) com a curva-padrão é possível determinar o número de cópias de RNAm presentes nesta amostra. Uma vez que não é preciso realizar para cada amostra desconhecida múltiplas reações com diferentes concentrações do competidor, é possível desta forma reduzir as quantidades de DNAC utilizadas em cada experimento e, conseqüentemente, ampliar o número de RNAm que podem ser estudados a partir de cada amostra de DNAC.

Referências Bibliográficas

- CHELLY, J., KAPLAN, J.C., MAIRE, P., GAUTRON, S. e KAHN, A. Transcription of the dystrophin gene in human muscle and nonmuscle tissue. *Nature*, v.333, p.858-860, 1988.
- GILLILAND, G., PERRIN, S., BLANCHARD, K. E BUNN, H.F. Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction. *Proc.Natl.Acad.Sci.USA*, v.87, p.2725-2729, 1990.
- KEVIN, P.F., LEONARD, M.W. e ENGEL, J.D. Quantitation of RNA using the polymerase chain reaction. *TIG*, v.9, n.11, p.380-385, 1993.
- TSAI, S. e WILTBANK, M.C. Quantification of mRNA using competitive RT-PCR with standard curve methodology. *Biotechniques*, v.21, p.862-866, 1996.
- WANG, A.M., DOYLE, M.V. e MARK, D.F. Quantitation of mRNA by polymerase chain reaction. *Proc.Natl.Acad.Sci.USA*, v.86, p.9717-9721, 1989.

TÉCNICAS EMPREGADAS NO ESTUDO DA EXPRESSÃO GÊNICA

Jane Eyre Gabriel

Um grande desafio da biologia celular é o entendimento do funcionamento das células ao nível molecular. Os avanços nos estudos nesta área estão diretamente associados à sistematização e otimização de técnicas experimentais, possibilitando a investigação detalhada da complexidade dos mecanismos moleculares envolvidos na expressão de um determinado gene (Pratt, 1987).

Nas últimas décadas, um conjunto de técnicas, conhecidas coletivamente como *tecnologia do DNA recombinante*, tem proporcionado imensos avanços nos estudos do controle da expressão gênica, pois com a utilização desta metodologia é possível investigar a estrutura, localização e função de diversos componentes macromoleculares envolvidos nos diferentes mecanismos celulares. Esta metodologia compreende: (a) a clivagem específica do DNA por enzimas nucleases de restrição, que facilitam o isolamento e manipulação de genes individuais; (b) sequenciamento rápido de todos os nucleotídeos de um fragmento purificado de DNA, tornando possível a determinação dos limites precisos de um gene e da seqüência de aminoácidos que ele codifica; (c) hibridização de ácidos nucléicos, possibilitando a procura de seqüências específicas de DNA ou RNA com grande sensibilidade e exatidão, baseando-se na habilidade de ligação entre seqüências complementares de ácidos nucléicos; (d) clonagem de DNA, pelo qual um fragmento específico de DNA é integrado em um elemento genético auto-replicante (plasmídeo ou vírus) que introduzido em uma bactéria pode originar muitos bilhões de cópias idênticas; (e) engenharia genética, pelo qual seqüências de DNA são alteradas, produzindo versões modificadas dos genes, que são inseridas em células ou organismos (Alberts et al., 1989; Lewin, 1997).

O impacto que esta metodologia tem provocado nos setores industrial, farmacêutico e agropecuário é indiscutível e, assim, a otimização de tais sistemas experimentais é de fundamental importância para assegurar a análise precisa dos mecanismos de controle envolvidos na expressão de um determinado gene em diferentes organismos.

Tecnologia de Amplificação dos Ácidos Nucléicos

Recentemente, o DNA de uma determinada região do genoma de qualquer organismo pode ser amplificado em milhões de cópias, graças ao desenvolvimento de uma metodologia, designada PCR, que permite a amplificação em cadeia utilizando uma DNA polimerase (Mullis & Faloona, 1987). Desde sua concepção, esta tecnologia tem causado uma verdadeira revolução na Biologia, tanto na pesquisa visando o entendimento de processos biológicos fundamentais como nas áreas aplicadas envolvendo diagnósticos e melhoramento genético de plantas e animais domésticos (Ferreira & Grattapaglia, 1995). Desta forma, a técnica de PCR tem sido uma ferramenta fundamental nos estudos de expressão gênica.

A técnica de PCR envolve a síntese enzimática *in vitro* de milhares de cópias de um segmento específico de DNA na presença de uma enzima DNA polimerase. A reação de PCR baseia-se no anelamento e extensão enzimática de um par de oligonucleotídeos (pequenas moléculas de DNA de fita simples) empregados como iniciadores (*primers*), que delimitam a seqüência de DNA fita dupla como alvo da amplificação. Estes *primers* são sintetizados artificialmente de maneira que suas seqüências de nucleotídeos sejam complementares a seqüências específicas que flanqueiam a região alvo.

Reações de Hibridização dos Ácidos Nucléicos

Quando uma solução aquosa de DNA é aquecida a 100°C ou exposta a pH muito alto (pH>13), os pares de bases complementares, que normalmente mantêm as duas fitas da dupla hélice unidas, são rompidos, resultando em duas fitas simples. Este processo, chamado desnaturação do DNA, por muitos anos foi considerado um processo irreversível. Entretanto, foi demonstrado que as fitas simples complementares de DNA podem ser dispostas novamente em dupla-hélice por um processo denominado renaturação ou hibridização do DNA, quando tais fitas têm uma seqüência

complementar de nucleotídeos (Alberts et al., 1989; Lewin, 1997). Assim, uma fita simples de DNA pode ser empregada como uma *sonda molecular*, para determinar a concentração de qualquer seqüência de RNA ou DNA em uma mistura de outras seqüências.

Os fragmentos dos ácidos nucléicos podem ser identificados por transferência de RNA total ou DNA genômico para membranas de nitrocelulose, nas quais as reações de hibridização são realizadas empregando sondas moleculares específicas ao gene de interesse. Tal procedimento experimental pode ser realizado com DNA ou RNA, sendo conhecido como "Southern" e "Northern blotting", respectivamente. Um procedimento análogo que é empregado para transferir proteínas é conhecido como "Western blotting".

Tais técnicas envolvendo hibridização de sondas moleculares de DNAs ou RNAs possibilitam: (a) a investigação das variações na expressão de um determinado gene e (b) a verificação de se tais variações ocorrem ao nível transcricional ou pós-transcricional. Atualmente, os métodos de hibridização são fundamentais nos estudos da estrutura e expressão gênicas (Alberts et al., 1989; Lewin, 1997).

Como as demais macromoléculas, os ácidos nucléicos ocupam posições precisas nas células e tecidos, e informações importantes quanto à localização destas moléculas nas células são perdidas quando tais moléculas são extraídas por homogeneização. Por esta razão, técnicas têm sido desenvolvidas, onde sondas de ácidos nucléicos e anticorpos marcados são empregados para localizar seqüências específicas de DNA ou RNA, sendo este procedimento experimental denominado hibridização *in situ*. Os métodos de hibridização *in situ* têm sido desenvolvidos com o objetivo de reconhecer o DNA nos cromossomos, mas também revelar a distribuição de moléculas específicas de RNA nas células em diferentes tecidos. Neste caso, o padrão diferenciado de expressão dos genes pode ser investigado, proporcionando informações importantes sobre os mecanismos envolvidos na expressão gênica durante o crescimento e desenvolvimento.

Referências Bibliográficas

ALBERTS, B. et al. **The molecular biology of the cell**. 2. ed. New York: Garland Publishing Inc, 1989. 1219p.

FERREIRA, M.C. & GRATTAPAGLIA, D. **Introdução ao uso de marcadores RAPD e RFLP na análise genética**. Brasília: EMBRAPA, CENARGEN, 1995. 220p.

LEWIN, B. **Genes VI**. New York, Oxford University Press, 1997. 1260p.

MULLIS, K. & FALOONA, F. 1987 Specific synthesis of DNA *in vitro* via a polymerase catalysed chain reaction. **Methods of Enzymology**, v.55, p.335-50.

PRATT, J.M. Coupled transcription-translation in prokaryotic cell-free systems. *In*: Hames, B.D. & Higgins, S.J. **Transcription and translation: a practical approach**. Oxford: IRL Press Limited, 1987. 350p.

ANIMAIS TRANSGÊNICOS

Luciana de A. Ribeiro

CPG – Genética e Melhoramento, USP - ESALQ

Durante séculos, produtores têm realizado seleção artificial em várias raças e linhagens de animais domésticos, objetivando aumentar a frequência dos genes com efeitos favoráveis. No entanto, quando o objetivo é a obtenção de mudanças mais drásticas no potencial genético, como mudança da base alimentar (pasto x grãos) ou nos requerimentos de mercado (redução de gordura), produtores têm lançado mão de estratégias de substituição de raças ou cruzamentos para transferir genes de uma população para outra, dentro de uma mesma espécie (Cundiff et al., 1993). Na década passada, foram desenvolvidas técnicas para transferir genes específicos com efeitos desejáveis não somente de uma raça para outra, mas de uma espécie para outra (Pursel & Rexroad, 1993). Estas técnicas de transgênese foram desenvolvidas para se obter mais rapidamente animais com características economicamente importantes do que quando utilizava-se acasalamento seletivo (Wall, 1996). Esta tecnologia tem tido enorme impacto nos componentes da eficiência de produção e, sobretudo na agregação de maior valor econômico aos produtos de origem animal.

Animal transgênico pode ser definido como aquele que contém moléculas de DNA exógeno em seu genoma, introduzidas por intervenção humana intencional, objetivando a expressão de novas características (Wall, 1996). O animal que integra o DNA exógeno em seu genoma é chamado de transgênico e o gene transferido de transgene (Pursel & Rexroad, 1993). Entretanto, a integração por si só não garante a expressão do transgene, e, uma outra definição, seria aquele animal que expressa o transgene e que quando acasalado com animais normais, produz progênie que herdaram este gene de forma mendeliana (Gordon & Ruddle, 1981).

O primeiro experimento com transgênese foi realizado em células da linhagem germinativa de camundongos em 1974. O genoma inteiro do vírus de símios foi microinjetado na cavidade blastocélica de embriões em estágio inicial do desenvolvimento. Entretanto, a integração do DNA viral só foi detectada em estudos

subsequentes, onde embriões de camundongos foram microinjetados com o retrovírus da leucemia de Moloney, gerando a primeira linhagem de camundongos transgênicos, em 1977 (Jaenisch, 1988). A partir desta data, uma variedade de protocolos têm sido desenvolvidos visando alterar o genótipo de animais de maneira estável.

Os primeiros resultados positivos com a expressão de DNA exógeno foram também obtidos em camundongos, no início da década de 80 (Gordon & Ruddle, 1981; Palmiter et al., 1982 e 1983). Camundongos gigantes gerados a partir da introdução do transgene (promotor da metalotioneína de camundongos e gene do hormônio do crescimento humano) em embriões de uma única célula, demonstraram que a integração foi estável e a expressão foi correta nos tecidos do animal adulto (Palmiter et al., 1983). Estes resultados incentivaram a especulação sobre a aplicação das técnicas de transgênese para aumentar o crescimento em espécies domésticas.

Os primeiros coelhos, ovelhas e porcos transgênicos foram obtidos em meados da década de 80 (Hammer et al., 1985) e os primeiros bovinos e caprinos no início dos anos 90 (Pursel & Rexroad, 1993). Entretanto, a eficiência de transformação foi bem menor que a obtida em camundongos.

Animais transgênicos têm sido amplamente utilizados como modelos biológicos experimentais para pesquisas dos mecanismos do desenvolvimento e regulação gênica, estudos de doenças genéticas humanas (Jaenisch, 1988) e de diversos tipos de câncer, através da inserção de diferentes oncogenes (Lewin, 1997).

Um dos objetivos imediatos da transgênese em animais é desenvolver um sistema alternativo de produção de proteínas recombinantes de interesse farmacológico humano (Jaenisch, 1988; Pursel & Rexroad, 1993; Wall, 1996). Uma outra aplicação, para um futuro próximo, é a produção de órgãos de animais (frequentemente o porco) para transplante em humanos (xenotransplante), produzindo animais transgênicos expressando fatores que evitem a rejeição (Lanza et al., 1997).

Vários métodos têm sido utilizados para a introdução de genes em animais, tais como: 1) microinjeção de DNA em pronúcleo, 2) infecção por retrovírus, 3) células embrionárias indiferenciadas ("stem cells"), 4) biobalística, e 5) espermatozóides como vetores.

1. Microinjeção de DNA em Pronúcleo

Microinjeção de DNA em pronúcleos de ovos fertilizados tem sido o método mais amplamente utilizado para a produção de animais transgênicos (Jaenisch, 1988). A maior vantagem deste método é a eficiência em gerar linhas transgênicas que expressam o transgene de maneira correta. Entretanto, esta técnica é limitada, não podendo ser utilizada em embriões em estágio mais avançado do desenvolvimento (Jaenisch, 1988). Outra desvantagem da injeção direta de DNA é a introdução de várias cópias, originando animais com expressão variável do transgene.

2. Infecção por Retrovírus

Ao contrário do DNA microinjetado, retrovírus integram-se por um mecanismo precisamente definido no genoma da célula hospedeira. Somente uma única cópia do vírus é inserida num determinado sítio do cromossomo e nenhum rearranjo no genoma é induzido, exceto uma pequena duplicação da seqüência hospedeira no sítio de integração (Jaenisch, 1988; Menck, 1998). O uso desta técnica para inserir genes em aves tem recebido muito mais atenção, devido ao fato que, após a postura, há cerca de 60.000 células na blastoderme, inviabilizando a microinjeção com DNA (Pursel & Rexroad, 1993). A principal vantagem do uso de vetores retrovirais para transferir genes em animais é a facilidade de se introduzir vírus em embriões em vários estágios do desenvolvimento. No entanto, o tamanho do DNA a ser introduzido é limitado (menos de 6 Kb) e, geralmente, pode apresentar problemas de expressão do gene devido à alta instabilidade de tais vetores.

3. Células Embrionárias Indiferenciadas ("stem cells")

As células embrionárias indiferenciadas ("stem cells") têm sido mais comumente utilizadas em camundongos, pois é a única espécie que possui linhagens estabelecidas de células indiferenciadas estáveis. Estas células são estabelecidas "in vitro" a partir de embriões em estágio de blastocisto, conservam seu cariótipo em cultura e são

pluripotentes. Quando injetadas em um blastocisto hospedeiro, estas células podem colonizar o embrião e contribuir para a formação da linhagem germinativa, originando um animal quimérico para o gene exógeno (Jaenisch, 1988). A vantagem deste procedimento é que um genótipo particular pode ser selecionado "in vitro", antes da introdução das células indiferenciadas no embrião. Por esta razão, esta é a única técnica que permite o uso de regiões específicas no transgene para uma possível inserção sítio-específica por recombinação homóloga (Pursel & Rexroad, 1993). Nas espécies domésticas, a utilização desta técnica é limitada até os dias de hoje, pelo fato dos animais obtidos serem quiméricos. Isto implica em um aumento no tempo necessário para a obtenção do animal fundador homozigoto (pelo menos uma geração a mais), sendo, principalmente para bovinos, um importante obstáculo a ser considerado (Menck, 1998).

4. Biobalística

A biobalística (ou biolística) foi desenvolvida em 1987 como uma alternativa para a introdução de DNA em plantas, principalmente para aquelas espécies resistentes à infecção por *Agrobacterium* (Sanford et al., 1993). Devido à sua praticidade, esta técnica passou a ser utilizada também em animais e tem sido muito empregada para estudos em pesquisas básica e aplicada. O princípio desta técnica consiste em impulsionar, em alta velocidade, micropartículas com o DNA exógeno adsorvido para o interior de células ou tecidos vivos. Diversos organismos já têm sido transfectados via biolística, tais como: bactérias, fungos, protozoários, insetos, peixes, anfíbios, camundongos, coelhos, macaco, galinha, entre outros. Recentemente, esta técnica tem sido utilizada em animais intactos vivos, ao invés de cultura de células e tecidos. Este procedimento tem mostrado ser muito eficiente na eliciação de resposta imune (imunização genética), produzindo anticorpos específicos contra o antígeno codificado pelo gene introduzido no bombardeamento (Tang et al., 1992).

5. Espermatozóides como Vetores

Espermatozóides podem ser utilizados como vetores para a introdução de DNA exógeno no interior de ovócitos, no momento da fecundação (Lavitrano et al., 1989). Isto foi demonstrado em camundongos, quando após a incubação de espermatozóides em um meio contendo o DNA exógeno, observou-se a presença deste DNA nos tecidos de alguns camundongos recém-nascidos (Menck, 1998). Até hoje nenhum outro laboratório conseguiu sucesso com a utilização desta técnica (Pursel & Rexroad, 1993).

Em suma, dependendo do objetivo do trabalho e da espécie utilizada, deve-se escolher qual o melhor método a ser aplicado. Há necessidade, também, independente dos resultados obtidos até o momento, de se realizar novos trabalhos e de se otimizar as técnicas mencionadas acima.

Referências Bibliográficas

- CUNDIFF, L. V.; BISHOP, M. D.; JOHNSON, R. K. Challenges and opportunities for integrating genetically modified animals into traditional animal breeding plans: Genetically modified livestock: progress, prospects and issues. **Journal of Animal Science**, v. 71, Supplement 3, p. 20-25, 1993.
- GORDON, J. W.; RUDDLE, F. H. Integration and stable germ line transmission of genes injected into mouse pronuclei. **Science**, v. 214, p. 1244-1246. 1981.
- HAMMER, R. E.; PURSEL, V. G.; REXROAD, C. E.; WALL, R. J.; BOLT, D. J.; EBERT, K. M.; PALMITER, R. D.; BRINSTER, R. L. Production of transgenic rabbits, sheep and pigs by microinjection. **Nature**, v. 315, p. 680-683. 1985.
- JAENISCH, R. Transgenic Animals. **Science**, v. 240, p. 1468-1474, 1988.
- LANZA, R. P.; COOPER, D. K. C.; CHICK, W. L. Xenotransplantation. **Scientific American**, p. 40-45, jul. 1997.
- LAVITRANO, M.; CAMAIONI, A.; FAZIO, V. M.; DOLCI, S.; FORACE, M. G.; SPADAFORA, C. Sperm cells as vectors for introducing foreign DNA into eggs: Genetic transformation of mice. **Cell**, v. 57, p. 717. 1989.
- LEWIN, B. **Genes VI**. Oxford: Oxford University Press and Cell Press, 1997. 126p.

- MENCK, M. C. **Seleção de embriões bovinos transgênicos utilizando-se o cDNA da luciferase de vagalume como gene repórter.** Ribeirão Preto, 1998. Tese (doutorado). Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.
- PALMITER, R. D.; BRINSTER, R. L.; HAMMER, R. E.; TRUMBAUER, M. E.; ROSENFELD, M. G.; BIRNBERG, N. C.; EVANS, R. M. Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. **Nature**, v. 300, p. 611-615. 1982.
- PALMITER, R. D.; NORSTEDT, G.; GELINAS, R. E.; HAMMER, R. E.; BRINSTER, R. L. Metallothionein-human GH fusion genes stimulate growth of mice. **Science**, v. 222, p. 809-814. 1983.
- PURSEL, V. G.; REXROAD, C. E. Status of research with transgenic farm animals. **Journal of Animal Science**. Genetically modified livestock: progress, prospects and issues. v. 71, p. 10-19, 1993. Supplement, 3.
- SANFORD, J. C.; SMITH, F. D.; RUSSEL, J. A. Optimizing the biolistic process for different biological applications. **Methods in Enzymology**, v. 217, p. 483-509. 1993.
- TANG, D. C.; DE VIT, M.; JOHNSTON, S. A. Genetic immunization is a simple method for eliciting an immune response. **Nature**, v. 356, p. 152-154. 1992.
- WALL, R. J. Transgenic Livestock: progress and prospects for the future. **Theriogenology**, v. 45, p. 57-68, 1996.

Manual de Laboratório

*Técnicas de Biologia
Molecular Aplicadas à
Produção Animal*

EXTRAÇÃO DE DNA

INTRODUÇÃO

Diversos métodos são disponíveis para a obtenção de DNA genômico de mamíferos. A escolha do método depende basicamente do tecido a ser utilizado como fonte desse ácido nucléico, da finalidade do material e da disponibilidade de reagentes no laboratório.

De um modo geral, os métodos compreendem as seguintes etapas:

1. Lise das células;
2. Remoção das proteínas e de restos celulares;
3. Precipitação dos ácidos nucleicos;
4. Remoção enzimática do RNA.

A seguir descrevemos dois métodos de extração de DNA.

O primeiro foi adaptado de Olerup & Zetterquist (1992)¹ e tem sido utilizado em nosso laboratório para amostras de sangue.

O segundo, próprio para extração a partir de amostras de sêmen, foi adaptado de Zadworny & Kuhnlein (1996)², substituindo-se o fenol, utilizado por aqueles autores para extração das proteínas, por precipitação dessas com cloreto de sódio.

Alguns cuidados especiais devem ser observados durante o trabalho com ácidos nucleicos:

1. As secreções e microrganismos da pele podem conter enzimas que degradam os ácidos nucleicos. Para prevenir a contaminação das amostras é importante que se utilizem luvas descartáveis durante a manipulação das amostras, dos reagentes e de acessórios de laboratórios (ponteiras, tubos, pipetas, etc.).
2. O DNA é adsorvido pelo vidro quando em presença de alguns sais como o iodeto de Sódio. Por isso, é recomendável que se utilize tubos e acessórios de polipropileno para as fases em que o DNA se encontra livre na solução.

¹ Olerup, O. & Zetterquist, H. (1992) *Tissue Antigens*, 39: 225 - 235

² Zadworny, D. & Kuhnlein, U. (1996) *Theor. Appl. Gent.*, 80: 631-634

PROTOCOLO DE EXTRAÇÃO DE DNA DE SANGUE FRESCO

Amostras : _____, _____, _____, _____, _____, _____, _____, _____,
_____, _____, _____, _____, _____, _____, _____, _____.

A. Obtenção de Leucócitos

- 1- Colher 5 ml de sangue em tubos Vacutainer com EDTA potássico (50µl EDTA (K₃) a 15%). Equilibrar os tubos completando com solução salina.
- 2- (opcional) Centrifugar 5' a 390 g e descartar o plasma, usando um pipetador, com cuidado para não descartar as células brancas.
- 3- Lisar os glóbulos vermelhos com 10 ml do tampão de hemólise (**solução A**), e homogeneizar, misturando bem por inversão ou passando os tubos no vórtex.
- 4- Centrifugar 10' a 693 g a 4°C.
- 5- Dispensar o sobrenadante. Se o pellet for pequeno, recentrifugar.
- 6- Ressuspender o pellet em 5 ml do tampão de hemólise. Para isto, tampar os tubos com parafilm e misturar bem por inversão e, se preciso, usar o vórtex. Centrifugar a 2000 rpm por 10' a 4°C.
- 7- Repetir a lavagem até obter somente as células brancas.
- 8- Ressuspender o pellet em 500 µl do tampão de hemólise e passar para tubos eppendorfs. Centrifugar por 1' a 15.993 g e descartar o sobrenadante.

obs: O pellet pode ser conservado no freezer, por vários meses.

B. Extração do DNA:

- 9) Ressuspender o pellet em 500µl de **solução B** (Proteinase K), vórtexando.
- 10) Incubar à 55°C até dissolver o pellet (4-6 horas ou overnight). De vez em quando vortexar as amostras para que o pellet fique bem dissolvido.
- 11) Ajustar o volume para 760 µl com TE. Dividir em 2 tubos e acrescentar 120 µl de NaCl 5M para 380 µl de solução.
- 12) Centrifugar 15' à 15.993 g.
- 13) Transferir o sobrenadante para um novo tubo (devidamente marcado).
- 14) Adicionar 1ml de etanol absoluto gelado e inverter o tubo várias vezes. Centrifugar por 15' a 15.993g.
- 15) Lavar com 500 µl de etanol 70% gelado. Centrifugar por 5'a 15.933 g.
- 16) Ressuspender em 250 µl de TE + RNase (10 µg por ml de amostra) e incubar por 1 hora.
- 17) Quantificar em espectrofotômetro:

Os ácidos nucleicos absorvem luz no comprimento de onda de 260 nm. Para fazer a leitura no espectrofotômetro, normalmente se utiliza uma diluição em água.

Para estimar a concentração de DNA utiliza-se a seguinte relação: $1 \text{ OD}_{260} = 50 \mu\text{g}$ DNA dupla-hélice. Dessa forma a concentração de DNA na amostra, é obtida pelo seguinte cálculo:

$$[\text{DNA}] = \text{Valor da leitura em O.D.} \times 50 \times \text{Fator de diluição}$$

As proteínas absorvem luz no comprimento de onda de 280 nm. Sendo assim, a relação A_{260}/A_{280} fornece um parâmetro de avaliação da qualidade das preparações de ácidos nucleicos. Valores inferiores a 1,8 resultam de contaminação com proteína.

Amostra	OD_{260}	OD_{280}	Razão	Concentração

PREPARO DAS SOLUÇÕES

Solução A

Para 1 amostra (25ml) → 250 μl de Tris.Cl pH 7,6 1M (final 10mM)
 250 μl de MgCl_2 0,5M (final 5mM)
 50 μl de NaCl 5M (final 10mM)
 Completar com H_2O pura

Para n amostras (25.n) _____ ml → 250.n: _____ de Tris.Cl pH 7,6 1M
 250.n: _____ de MgCl_2 0,5M
 50.n: _____ de NaCl 5M
 Completar com H_2O pura

Solução B

Para 1 amostra (500 μl) X 5 μl de Tris.Cl pH 8,0 1 M (final de 10mM)
 10 μl de NaCl 5 M (final de 100mM)
 10 μl de EDTA pH 8,0 0.5 M (final 10mM)
 12,5 μl de SDS 20% (final 0,5%)
 2 μl de Proteinase K na concentração 20mg/ml
 460,5 μl H_2O pura

Para n amostras (500.n) _____ 5.n : _____ de Tris.Cl pH 8,0 1 M
10.n : _____ de NaCl 5 M
10.n : _____ de EDTA pH 8,0 0.5 M
12,5.n : _____ de SDS 20%
2.n : _____ μ l de prot.K 20mg/ml
460,5.n : _____ de H₂O

TE

Tris HCl pH 7,6 (final 10mM)

EDTA pH 8,0 (final 1mM)

Para 1 amostra (510 μ l) - Para n amostras (510.n) _____ de solução.

SOLUÇÃO DE RNase (10mg/ml)

Dissolver a enzima Ribonuclease A livre de DNase (Sigma R- 6513) em solução Tris.HCl 10mM (pH 7.5), NaCl 15mM.

Armazenar a -20°C.

SOLUÇÃO DE PROTEINASE K (20mg/ml)

Dissolver 100 mg da enzima em 5 ml de solução Tris-HCl 10mM (pH 7.5), Cloreto de Cálcio 20mM, Glicerol 50%.

Armazenar a -20°C.

EXTRAÇÃO DE DNA DE SÊMEN

- 1) Descongelar 1 paleta em cada eppendorf de 1,5 ml;
- 2) Centrifugar 8 min a 15.993 g .
- 3) Lavar o pellet 4 vezes em 1 ml de PBS 1X (a cada lavagem passar no vortex e centrifugar novamente).
- 4) Ressuspender em 100 µl de PBS - pode-se parar nesse ponto e armazenar no freezer.
- 5) Adicionar 400 µl de Solução de lise. Incubar 30 min a 50°C e dissolver o pellet.
- 6) Adicionar Proteinase-K a 200 µg/ml (100 µg = 5 µl da solução estoque de 20mg/ml). Passar no vortex até dissolver. Incubar por 16 horas a 50°C .

Precipitação com sal:

- 7) Dividir as amostras em 2 tubos, 250 µl em cada tubo.
- 8) Adicionar 80 µl de NaCl 5M por tubo. Agitar vigorosamente por inversão. Centrifugar a 15.993 g por 5 min.
- 9) Transferir o sobrenadante para tubos limpos. Acrescentar 1ml de etanol absoluto a cada tubo e agitar por inversão. Centrifugar a 15.993 g por 5 min.
- 10) Desprezar o sobrenadante. Lavar o pellet com etanol 70% - simplesmente acrescenta-se 100 µl de etanol 70% a cada tubo, não passar no vortex.
- 11) Centrifugar a 15.993 g por 5 min. Desprezar o sobrenadante. Secar o pellet e ressuspender em 100 µl de TE + RNase.
- 12) Incubar por 1 hora.
- 13) Leitura no espectrofotômetro.

Preparo das soluções

Solução de PBS (1X):

KCl	2,7 mM
KH ₂ PO ₄	1,5 mM
NaCl	137 mM
Na ₂ HPO ₄	8 mM
Ajustar o pH para	7.0

Solução de lise:

	2%
2-mercaptoetanol	
Tris.Cl (pH 8.0)	10 mM
NaCl	100 mM
EDTA (pH 8.0)	10 mM
SDS	0.5%

Anotações:

Observações:

DIGESTÃO DO DNA - β -LACTOGLOBULINA

Data: ____/____/____

Amostras : _____, _____, _____, _____, _____, _____, _____,
_____, _____, _____, _____, _____, _____, _____,
_____, _____, _____, _____, _____.

	<i>Emix 1X</i>	<i>Emix _____ X</i>	
Água Milli Q	1,45 μ l	_____ μ l	<input type="checkbox"/>
10X Medium Buffer *	1,3 μ l	_____ μ l	<input type="checkbox"/>
Hae III (5u / μ l)	0,25 μ l	_____ μ l	<input type="checkbox"/>

*100 mM Tris- HCl (pH 7.5), 100mM MgCl₂, 500mM NaCl, 10mM DTT

Por reação : 3,0 μ l Emix + 10 μ l DNA amplificado. Incubar à 37°C por 3 horas.

Início : _____

Um segmento de 262 pares de bases do gene da β -lactoglobulina, que contém polimorfismo de restrição para a enzima *HaeIII*, é amplificado. A digestão do DNA amplificado a partir do alelo **A** produz fragmentos de 153 e 109 pares de bases. O alelo **B** caracteriza-se pela presença de um segundo sítio de restrição dentro do fragmento de 153 pares de bases, resultando em três fragmentos com 79, 74 e 109 pares de bases, respectivamente.

Resultados

Animal	Genótipo

Anotações:

ELETROFORESE EM GEL DE AGAROSE

- **gel 1%** → 0,3 g de agarose para 30 ml de gel, que deve ser preparado com tampão Tris-Borato-EDTA 1X e com 1 µl de brometo de etídeo. Para isso, aplicar 5 µl do amplificado + 1 µl de loading buffer. Submetê-lo a uma voltagem constante na faixa de 2 a 5 V/cm (considerando a distância entre os eletrodos) e visualizá-lo sob luz ultra-violeta.
- **gel 3%** → 0,9 g de agarose de baixo ponto de fusão para 30 ml de gel, que deve ser preparado com tampão Tris-Borato-EDTA 1X, além de 1,2 µl de brometo de etídeo*. Utilizado para analisar o resultado das digestões com enzimas de restrição. Aplicar todo o produto da digestão (13 µl) + 2,6 µl de loading buffer. A separação dos fragmentos deve ser realizada em TBE 1X nas condições descritas anteriormente.
* (0.4 µg/ml de gel)

Procedimento:

1. Pesar a agarose;
 2. Colocá-la em um erlenmayer contendo o volume necessário de tampão TBE 1X;
 3. Aquecer no forno de microondas até iniciar a ebulição (aproximadamente 30 segundos em potência média para um gel de 30 ml). Agitar o frasco e retornar ao forno de microondas por mais alguns segundos.
 4. Aguardar a redução da temperatura para aproximadamente 60 ° C e adicionar Brometo de Etídio, na quantidade indicada para cada gel.
 5. Verter no molde previamente nivelado e colocar os pentes.
 6. Após a solidificação, colocar o gel na cuba de eletroforese e cobrir com tampão TBE 1X.
 7. Aplicar as amostras acrescidas de loading buffer e aplicar a corrente de acordo com o tamanho do gel⁴.
-
- O Brometo de etídeo é um potente agente mutagênico. Utilizar luvas e máscara para preparar a solução estoque e manipular os géis.
 - A luz ultra-violeta produz queimaduras severas. Utilize máscara/óculos de proteção adequados.

⁴ Ron et al. (1994). J. Dairy Sci., 77: 1106 - 1113

Loading Buffer 6 x

Bromophenol Blue 0,25%

Xileno Cianol 0,25%

Glicerol 30% em água (estocar a 4°C).

Limites de eficiência da separação de DNA em diferentes concentrações de agarose

Concentração de agarose no gel (%)	Separação de moléculas de DNA (Kb). Limites de eficiência.
0,3	60 – 5,0
0,6	20 – 1,0
0,7	10 – 0,8
0,9	7 – 0,5
1,2	6 – 0,4
1,5	4 – 0,2
2,0	3 – 0,1

Resultados

Animal	Genótipo

Anotações:

O microssatélite CSFM50, localizado no braço longo do cromossomo 2 dos bovinos, é amplificado com os iniciadores descritos por Moore et al. (1994). Seus alelos apresentam tamanho entre 168 e 182 pb.

Resultados

Animal	Genótipo

Anotações:

O microssatélite TEXAN15, localizado no cromossomo 5 dos bovinos a 7,7 Centimorgans do microssatélite IGF1, é amplificado com os iniciadores descritos por Burns *et al.* (1995). Os alelos variam de tamanho entre 203 e 221 pb.

Resultados

Animal	Genótipo

Anotações:



Preparação das placas para gel de poliacrilamida

1. Limpeza das placas

- Deixar as placas de molho no NaOH 1N por 30 minutos;
- Lavar com detergente comum fazendo movimentos circulares ao longo de toda a placa, usando uma esponja macia;
- Enxaguar abundantemente em água corrente esfregando com a esponja;
- Deixar de molho em água destilada;
- Secar;
- Limpar com etanol absoluto, utilizando papel toalha.

2. Preparação da placa maior usando Bind-Silane

- Separar a placa menor da maior;
- Preparar a solução de Bind-Silane, utilizando sempre a mesma luva:

Bind-Silane	3 μ l
Etanol absoluto	1 ml
Ácido acético glacial	5 μ l
- Aplicar a solução de Bind com uma pipeta sobre a placa e espalhar com papel toalha sobre toda a superfície em um sentido único, não esquecendo de deixar um pedaço sem Bind para o pente;
- Secar por 30 minutos;
- Retirar o excesso com etanol absoluto, passando papel toalha no sentido perpendicular ao da etapa anterior.

Obs: Separar as luvas usadas com Bind para serem usadas somente para a preparação da placa maior.

3. Preparação da placa menor com Repel-Silane

- Aplicar 1 ml de Repel puro e espalhar com papel toalha ao longo de toda a superfície da placa;
- Secar por 10 minutos;
- Retirar o excesso com papel toalha umedecido em etanol absoluto.

obs: passar Repel no pente. Não há necessidade de retirar o excesso com etanol.

4. Montagem do sanduíche

GEL DE ACRILAMIDA DESNATURANTE PARA MICROSSATÉLITES

• SOLUÇÕES:

1. 20X modified TBE buffer:			1,2X
Tris base	242,2g (2M)		120mM
Ácido bórico	12,0g (195mM)		12mM
EDTA dissódico	5,8g (15mM)		1mM
Completar o volume para 1 litro			

2. Solução de uréia:

Uréia	612g
Completar o volume para 1 litro	

Aquecer para dissolver. Não deixar passar de 60 ° C

3. Tampão do gel:

20X modified TBE	60 ml	12ml
Solução de uréia	1 l	200 ml (9,5M)

4. Tampão de corrida (1,2X modified TBE):

20X modified TBE	60 ml
Completar o volume para 1 litro com água destilada	

5. Persulfato de Amônia 10%

0,1 g de persulfato + 1 ml de água Eiga

6. Acrilamida : Bisacrilamida

Acrilamida.....	38,67g
N´N´metileno bisacrilamida.....	1,33g
Completar o volume para 100ml	

Se necessário aquecer até 55°C. Armazenar em frasco escuro a 4°C.

- ☠ Cuidado ao manipular esses reagentes; Utilize luvas e máscaras para partículas.
A acrilamida é neurotóxica!

Para cada gel grande (6,67%), utilizar:

Tampão do gel	53,78 ml
Acrilamida-Bisacrilamida 40%	10,84 ml
TEMED	37,52 µl
Persulfato de Amônia 10%	375,2 µl

- Verter o gel na placa inclinada, cuidadosamente para não formar bolhas.
- Colocar o pente, cobrir a extremidade com filme plástico e prender o pente com grampos de metal.

II. PREPARO DAS AMOSTRAS

1. Stop solution

Formamida	95 ml	14,4 ml
EDTA 0,5M pH 7.5 a 8.0	4 ml	0,6 ml
Bromofenol Blue	0,05 g	0,0075 g
Xileno Cianol	0,05 g	0,0075 g

- Adicionar 16 μ l do "stop solution" em 25 μ l de PCR (no próprio tubinho da reação). Aqueça as amostras a 92°C por 10 minutos → Termociclador: programa DENATURA.
- Aplicar 15 μ l quando for usado o pente pequeno.
- Aplicar 6 μ l do padrão 10 pb.

Taxa de migração dos marcadores através do Gel Desnaturante de Poliacrilamida

% Poliacrilamida	Bromophenol Plue	Xileno Cianol
5	35	130
6	26	106
8	19	76
10	12	55
20	8	28

Sambrook *et al.* 2ª edição, Molecular Cloning: Laboratory manual v 2, pag 1355, 1989

GEL DE ACRILAMIDA NÃO DENATURANTE PARA MICROSSATÉLITES

SOLUÇÕES

1. TBE 10 X

Tris (0,89M)	108 g
Acido Bórico (0,89M)	55 g
EDTA 0.5M pH 8.0 (0,02M)	40 ml
Completar para 1 litro.	

2. Persulfato de amônia 10%

- 0,1 g de persulfato + 1ml de água ultra-pura.

Preparação do gel não desnaturante 6,6% (70 ml):

Acril-Bisacrilamida 40%	11,64 ml
TBE 10 x	7,0 ml
Água	50,92 ml
TEMED	40,34 μ l
Persulfato de Amônia 10%	403,64 μ l

I. Preparação das amostras

1. Loading Buffer 6 x

Bromophenol Blue 0,25%

Xileno Cianol 0,25%

Glicerol 30% em água (estocar a 4°C).

- Aplicar : 9 μ l do amplificado + 1,5 μ l de loading
- Aplicar 10 μ l do padrão 10 pb (33,3 ng/ μ l) : 1,6 μ l loading + 8,4 μ l água + 0,3 μ l de ladder.

Propriedades do gel de poliacrilamida não desnaturante

Concentração %	faixa de separação (pb)	Xileno Cianol	Bromophenol
3.5	1000 - 2000	460	100
5.0	80 - 500	260	65
8.0	60 - 400	160	45
12.0	40 - 200	70	20
15.0	25 - 150	60	15
20.0	6 - 100	45	12

ELETROFORESE EM GEL DE POLIACRILAMIDA

1. Retirar os grampos e a fita adesiva;
2. Retirar o pente cuidadosamente (antes, retirar o excesso de acrilamida);
3. Lavar as placas com água destilada e enxugar (as placas devem ser colocadas na pia com os poçinhos para baixo);
4. Colocar o gel na cuba e virar as presilhas sem apertá-las. Apertá-las da esquerda para a direita, de cima para baixo.
5. Certificar-se que o registro do dreno (direita) está fechado;
6. Colocar tampão de corrida na cuba de cima, certificando-se que não vaze, e depois na de baixo.
7. Limpar os poçinhos com seringa e, se preciso, com o "caça-bolhas";
8. Fazer uma pré-corrida a 35-40W (não deixar passar de 40 miliamperes) por uma hora;
9. Aplicar as amostras no gel e correr nas condições previamente descritas.

PROCOLO DE COLORAÇÃO COM PRATA

(Comincinis *et al.*, J. Anim. Breed. Genet. 112, 415-420, 1995)

1. **Pré-tratamento** : 20 minutos em 1,5 l de etanol 10% e ácido acético 10% em água deionizada. Guardar a solução para usá-la no ítem 8 (stop).
2. **Enxágüe e Lavagem** : 3 x 30 segundos em água deionizada.
3. **Oxidação** : 3 minutos em 1,5 l de ácido nítrico 1%.
4. **Lavagem** : 3 x 30 segundos em água deionizada.
5. **Impregnação** : 30 minutos em 1,5 l de solução contendo 1,5g de nitrato de prata e 2,0 ml de formaldeído 37%.
6. **Lavagem** : 30 segundos em água deionizada.
7. **Revelação** : 3-5 minutos em 1,5 l de solução contendo 11,13g de carbonato de sódio anidro, 2,0 ml de formaldeído 37% e 7,5 ml de tiosulfato de sódio pentahidratado 200mg/l.
8. **Stop** : 5 minutos em 1,5 l da solução de pré-tratamento.

OBS: As soluções devem ser preparadas um pouco antes do início da coloração. O formaldeído e a solução de tiosulfato de sódio pentahidratado devem ser adicionados às respectivas soluções no momento do uso.

Descrição dos primers utilizados.

Marcador	Cromossomo	Seqüência 5'-3' 1	T.A.	Referência
CSN3 ^a	6q26-33	ATC ATT TAT GGC CAT TCC ACC AAA G GCC CAT TTC GCC TTC TCT GTA ACA GA	55 °C	Medrano & Cordova, 1990 ⁵
LGB ^b	11q28	GTC CTT GTG CTG GAC ACC GAC TAC A CAG GAC ACC GGC TCC TGG TAT ATG A	55 °C	Ron et al., 1994 ⁶
GH ^c	19q17-qter	GCT GCT CCT GAG GGC CCT TCG GCG GCG GCA CTT CAT GAC CCT	58 °C	Schlee et al., 1994 ⁷
CSFM50	2	AGT TCT CCT CTT GAT TTC AGT AAC CCT ACT TCC TGC CTT TGT AGC ATA	55 °C	Moore et al. 1994 ⁸
TEXAN15	5	TCG CAA ACA GTC AGA GAC CAC TC TGG ATG AGA AAG AAG AGC AGA GTT G	56 °C	Burns et al., 1995 ⁹
OPX7	?	GAGCGAGGCT	35 °C	Operon
OPX18	?	GACTAGGTGG	35 °C	Operon

¹ A seqüência do *primer forward* para cada marcador é descrita na linha superior. A linha de baixo corresponde ao *primer reverse*.

T.A. = Temperatura de anelamento;

^aK-caseína, ^bB-lactoglobulina, ^cHormônio de crescimento.

⁵ Medrano, J.F. & Cordova, E., 1990, *Biotechnology* 8: 144-146.

⁶ Ron, M., Yoffe, O., Ezra, E., et al., 1994, *J. Dairy Sci* 77: 1106-1113.

⁷ Schlee, P., Graml, R., Schallenberger, E., et al., 1994, *Theor. Appl. Genet.* 88: 497-500.

⁸ Moore, S.S., Byrne, K., Berger, K.T., et al., 1994, *Mamm. Genome* 5: 84-90.

⁹ Burns, B.M., Taylor, J.F., Herring, A.D., et al., 1995, *Anim. Genet.* 26: 201-202.

Extração de RNA com Trizol LS

Protocolo:

1. Remover o tecido a ser submetido à extração.
2. Em tubo de 1 ml, acrescentar:
750 µl Trizol + 250 µl H₂O livre de RNase
3. Incubar 5 min., 15-30 °C.
4. Acrescentar 0,2 ml de clorofórmio.
5. Agitar vigorosamente por 15 segundos.
6. Incubar 15-30 °C por 2 a 15 minutos.
7. Centrifugar a (12.000 g) por 15 min., à temperatura de 2-8 °C.
8. Remover a fase aquosa (superior e translúcida) e transferir para um tubo limpo.
Reservar a fase orgânica para extração de DNA.
9. Adicionar 0,5 ml de álcool isopropílico.
10. Incubar por 10 min., à temperatura de 15-30 °C.
11. Centrifugar a (12.000 g) por 10 min., à temperatura de 2-8 °C.
12. Descartar o sobrenadante.
13. Lavar com 1ml de etanol 75%. Vortex para misturar.
14. Centrifugar a (7.500 g) por 5 min., à temperatura de 2-8 °C.
15. Secar brevemente o RNA (à temperatura ambiente ou à vácuo).
16. Dissolver o "pellet" em H₂O livre de RNase. Incubar por 10 min., à 55-60 °C antes de levar ao gel.

Eletroforese RNA

1. Gel com formaldeído	30 ml
Agarose (1%)	0,3 g
Buffer 5x	6 ml
Formaldeído (2.2 M)	5,4 ml
H ₂ O livre de RNase	11,4 ml

Verificar se o pH formaldeído é maior que 4,0.

2. Tampão de corrida

MOPs	0,1 M
Acetato de sódio	40 mM
EDTA pH 8,0	5 mM

Pesar 20,6 g MOPs. Dissolver em 800 ml de acetato de sódio 500 mM. Ajustar o pH para 7,0 com NaOH 2N. Adicionar 10 ml de EDTA. Guardar protegido da luz.

3. Preparação das amostras

Colocar em um tubo estéril:

RNA (até 30 µg)	4,5 µl
Tampão de corrida 5x	2,0 µl
Formaldeído	3,5 µl
Formamida	10,0 µl
Total	20,0 µl

Incubar as amostras por 15 min. a 65 °C. Levar ao banho gelo. Centrifugar para baixar o conteúdo. Adicionar 2 µl de tampão de amostra.

Marcadores: 28S (6.333 pb)
18S (2.366 pb)
9S Rabbit
mRNA β-globina

RT-PCR (Kit Gibco)

A. Síntese do cDNA fita simples

1. Misturar e centrifugar brevemente cada componente do Kit antes de utilizá-los.

2. Preparar: (ependorf 0,5 ml)

1-5 µg RNA total	x µl
Oligo dT (0,5 µg/ul)	1 µl
H ₂ O DEPC q.s.p.	12 µl

3. Incubar as amostras a 70°C por 10 min.

4. Deixar no gelo por pelo menos 1 min.

5. Preparar a seguinte mistura de reação:

PCR buffer 10x	2 µl
MgCl ₂ 25 mM	2 µl
Mix dNTP 10 mM	1 µl
DTT 0,1 M	2 µl
Total	7 µl

6. Adicionar 7 µl da mistura de reação em cada mistura RNA/primer. Homogeneizar delicadamente e centrifugar brevemente.

7. Incubar 5 min., 42°C

8. Adicionar 1 µl (200u) Superscript IIRT em cada tubo e incubar 42°C por 50 min.

9. Interromper as reações, colocando as amostras a 70°C por 15 min.

10. Retirar e deixar em gelo.

11. Coletar a reação através de breve centrifugação.

12. Adicionar 1 µl de RNase H em cada tubo e incubar a 37°C por 20 min. antes de proceder o protocolo B.

B. Amplificação do cDNA

β-actina

1. Adicionar em um tubo 0,5 ml:

cDNA (10% reação)	2,0 µl
PCR buffer 10x (1x)	5,0 µl
MgCl ₂ 50 mM (1,5mM)	1,5 µl
Mix dNTP 10 mM (20pM)	1,0 µl
Amplification primer 1 (10 mM)	5,0 µl
Amplification primer 2 (10 uM)	5,0 µl
Taq DNA pol (2 U)	0,5 µl
H ₂ O MilliQ	30,0 µl

2. Misturar gentilmente e adicionar 2 gotas de óleo mineral.
3. Analisar 10 µl do produto PCR em gel de agarose.
4. Programa PCR:

1^o Ciclo - 95^oC 5 min.
 55^oC 1 min.
 74^oC 1 min.

2^o - 34^o Ciclo- 95^oC 1 min.
 55^oC 1 min.
 74^oC 1 min.

35^o Ciclo - 95^oC 1 min.
 55^oC 1 min.
 74^oC 10 min.

Obs: Tamanho esperado do produto: 409 pb

PROCOLO EXPERIMENTAL DA TÉCNICA DE HIBRIDIZAÇÃO "IN SITU"

Este protocolo foi inicialmente descrito por Coutinho et al. (1992).

Dia 1. Incubação dos ovos de galinha por aproximadamente 48 horas a 37°C.

Dia 2. Coleta dos embriões e fixação em paraformaldeído 4% por 12 horas, a 4°C.

Dia 3. Remoção das membranas envoltórias e estocagem em etanol 70% a -20°C.

Dia 4. Pré-hibridização e hibridização.

- Desidratação em 1 ml etanol 80% por 2 min.
- Desidratação em 1 ml etanol 95% por 2 min.
- Desidratação em 1 ml etanol 100% por 2 min.
- Lavagens em 1 ml etanol 100% por 5 min.
- Incubação em 1 ml de solução contendo etanol 50%/xileno 50% por 5 min.
- Incubação em 1 ml xileno por 30 min.
- Lavagens rápidas (5 vezes) em 1 ml etanol 100%
- Hidratação em 1 ml etanol 95% por 2 min.
- Hidratação em 1 ml etanol 80% por 2 min.
- Hidratação em 1 ml etanol 70% por 2 min.
- Lavagens (3 vezes) em 1 ml PBT.3 (PBS 1X contendo 0,3% TRITON X-100) por 2 min. cada lavagem
- Digestão dos embriões em 1 ml de tampão proteinase K (100 mM TRIS-HCl pH 8,0; 50 mM EDTA; 30 µg/ml proteinase K) por 7 min. a 37 °C. ¹⁰
- Lavagens (3 vezes) em 1 ml PBT.3 por 2 min. cada lavagem.
- Incubação por 10 min. em 200 µl de solução de hibridização 50% (50% formamida, 4X SSC, 1X Denhardt's solution, 0,5 µg/µl DNA de esperma de salmão desnaturado, 0,25 µg/µl tRNA de levedura e 10% dextran sulfato)/PBT.3 50%.

PRÉ-HIBRIDIZAÇÃO: 3 horas a 63 °C em 100 µl solução de hibridização (50% formamida, 4X SSC, 1X Denhardt's solution, 0,5 µg/µl DNA de esperma de salmão desnaturado, 0,25 µg/µl tRNA de levedura e 10% dextran sulfato).

HIBRIDIZAÇÃO: 42 horas a 63 °C em 100 µl de solução de hibridização contendo a sonda digoxigenina-marcada do RNA "antisense" do gene de interesse em uma concentração final de 0,3 ng/µl.

¹⁰ O tempo de digestão da proteinase K deve ser otimizado em função do tamanho do embrião. Para embriões com 45 horas de incubação 7 min. de digestão são suficientes.

Dia 5. Lavagens

- 100 µl solução de hibridização por 5 min. a 63 °C.
- 100 µl solução de hibridização por 2:30 horas a 63 °C,
- 100 µl solução de hibridização por 16 horas a 63 °C.

Dia 6. Lavagens e Imunodeteccção

- 80 µl solução de hibridização/ 20 µl PBT.3 (80%/20%) por 30 min.
- 60 µl solução de hibridização/ 40 µl PBT.3 (60%/40%) por 30 min.
- 40 µl solução de hibridização/ 60 µl PBT.3 (40%/60%) por 30 min.
- 20 µl solução de hibridização/ 80 µl PBT.3 (20%/80%) por 30 min.
- Lavagens (2 vezes) em 500 µl PBT.3 por 1 hora cada lavagem.

IMUNODETECCÃO:

- Incubação por 1 hora em 500 µl PBT.3, contendo 5% de soro de galinha sob level agitação
- Incubação por 2 horas em 500 µl PBT.3 plus 5% de soro de galinha, contendo 0,25 µl de anti-dig acoplado à fosfatase alcalina (diluição 1:2000)
- Lavagens rápidas (2 vezes) em 500 µl PBT.5 (PBS 1X plus 0,5% TRITON X-100) e estocagem a 4 °C, O/N.

Dia 7. Imunodeteccção

- Lavagens (5 vezes) em 500 µl PBT.5 por 20 min. cada lavagem.
- Lavagens (2 vezes) em 500 µl CT.3 (100 mM TRIS, 150 mM NaCl, 25 mM MgCl₂, pH 9,5, 0,3 % TRITON X-100) por 15 min. cada lavagem.

REAÇÃO DE DESENVOLVIMENTO DA COR: em 1 ml de CT (100 mM TRIS, 150 mM NaCl, 25 mM MgCl₂, pH 9,5) adicionar 4,5 µl NBT e 3,5 µl X-phosphate. Esta reação ocorre em 12 horas e deve ser realizada na ausência de luz.

Coutinho, L.L. *et al.* 1992 Whole mount in situ detection of low abundance transcripts of the myogenic factor qmf1 and myosin heavy chain protein in quail embryos. *Biotechnology*, 13(5):722-4.

Centrífuga	rpm	g
Fanem 204 R	1.500	389,90
	2.000	692,8
	5.000	4330
Eppendorf 5415C	14.000	15.994