

ADRIANA MELLO DE ARAÚJO

**PATERNIDADE E DIVERSIDADE GENÉTICA EM  
CAPRINOS NO BRASIL POR MEIO DE  
MICROSSATÉLITES DE DNA**

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

VIÇOSA  
MINAS GERAIS - BRASIL  
2004

**Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV**

T

A663p

Araújo, Adriana Mello de, 1967-  
Paternidade e diversidade genética em caprinos no Brasil  
por meio de microssatélites de DNA / Adriana Mello de  
Araújo. – Viçosa : UFV, 2003.  
56p. : il.

Orientador: Simone E. Facioni Guimarães  
Tese (doutorado) – Universidade Federal de Viçosa

1. Caprino – Genética molecular. 2. Caprino – Melhora-  
mento genético. 3. Caprino – Pedigree. 4. Microssatélites  
(Genética). 5. DNA de satélite. I. Universidade Federal de  
Viçosa. II. Título.

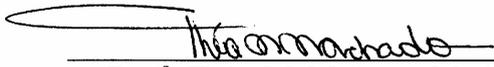
CDD 20.ed. 636.390821

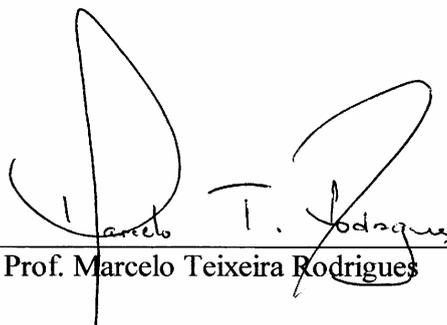
ADRIANA MELLO DE ARAÚJO

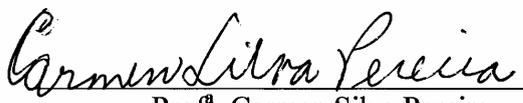
**PATERNIDADE E DIVERSIDADE GENÉTICA EM CAPRINOS NO BRASIL  
POR MEIO DE MICROSSATÉLITES DE DNA.**

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

APROVADA: 26 de janeiro de 2004.

  
Prof.<sup>ª</sup>. Théo M. Medeiros Machado

  
Prof. Marcelo Teixeira Rodrigues

  
Prof.<sup>ª</sup>. Carmen Silva Pereira

  
Prof.<sup>ª</sup>. Cláudia Graça da Fonseca

  
Prof.<sup>ª</sup>. Simone E. Facioni Guimarães  
(Orientadora)

*À Memória de minha mãe, Maria José de Assis Mello.*

*Ao meu marido, Wilson.*

*Ao meu pai, Jorge.*

*Aos meus queridos filhos, Franco e Enzo.*

***Dedico este trabalho.***

## AGRADECIMENTOS

À Universidade Federal de Viçosa, em especial, ao Departamento de Zootecnia e ao Programa de Pós-Graduação em Genética e Melhoramento.

À Embrapa – Empresa Brasileira de Pesquisa Agropecuária, pela liberação de minhas atividades e pela concessão da bolsa de estudo. Ao Centro Nacional de Pesquisa de Caprinos, em especial pela colaboração com amostras da raça naturalizada Moxotó.

À Fapemig – Fundação de Apóio à Pesquisa de Minas Gerais, ao CNPq e à CAPES, pelo financiamento e apoio do projeto.

Ao Capril Sanri, Florestal, Minas Gerais, na pessoa da Sr<sup>a</sup> Maria Pia Guimarães, pela colaboração com os dados.

Ao LABOGENA, INRA, *Jouy-en-Josas*, pela colaboração com a seqüência de *primers* de caprinos.

À professora Simone E. Facioni Guimarães, pela competência, disponibilidade e amizade valiosas. Muito obrigado pela confiança depositada.

À professora Carmen Silva Pereira, pelo empenho à pesquisa científica, pelos ensinamentos, pelas sugestões e críticas, que muito contribuíram para minha formação, e principalmente, pela confiança e amizade ao longo de quase dez anos.

À professora Cleusa Graça da Fonseca, por todo o conhecimento e sabedoria dedicados a este trabalho e a minha formação. Também seguimos por longo tempo em contato.

Ao professor Marcelo Teixeira Rodrigues, pelo empenho na obtenção das amostras e pelas sugestões valiosas na elaboração do projeto e no trabalho final.

À professora Théa Machado, pela dedicação, organização e harmonia com a pesquisa como um todo e com a caprinocultura em especial.

Aos professores Robledo de Almeida Torres e Paulo Sávio Lopes, pela amizade, pelo apoio, pelos ensinamentos e pelos incentivos, que amenizaram a execução deste trabalho.

Aos demais professores do Programa de Pós-Graduação em Genética e Melhoramento, pelos ensinamentos e pela atenção.

Aos colegas do LABTEC, em especial, aos bolsistas e estagiários Virgínia, Patrícia, Paula, George e André, que muito me auxiliaram durante as análises no laboratório.

Meus agradecimentos especiais aos colegas Renata, Fabiano, Lindemberg, Urbano, Jane e Fred, pela colaboração em diferentes etapas do trabalho e do curso.

Aos colegas de curso, Aldrin, Alex, Claudinho, Dani, Danielle Serra, Elizângela, Fernanda, Fausto, Giselle, Guilherme Band, Gustavo, Jaime, Júnior, Leandro, Marcelo, Marcos Vinícius, Peloso, Policarpo, Rachel, Ricardinho, Rodolphinho e Samuel, pelo companheirismo e pelo convívio agradável.

Aos funcionários do Departamento de Zootecnia, pela ajuda e pelas saudações cordiais trocadas nos corredores.

Ao meu marido Wilson, pelo carinho, pela compreensão e pelo incentivo, sempre compartilhando as alegrias e tristezas decorrentes do curso.

Aos meus amados filhos, Franco e Enzo, minha razão de viver. Todas as vitórias ficam mais vibrantes e as derrotas mais toleráveis quando chego em casa e me deparo com estes rostinhos lindos.

## CONTEÚDO

	<b>Página</b>
RESUMO.....	viii
ABSTRACT.....	xi
1. INTRODUÇÃO GERAL .....	1
2. REVISÃO DE LITERATURA .....	4
2.1. Os microssatélites de DNA .....	4
2.2. O mapa genômico de caprinos .....	8
2.3. Marcadores para testes de paternidade.....	10
2.4. Testes de paternidade por meio de microssatélites de DNA em animais domésticos propostos na literatura.....	12
2.5. Estatística para marcadores moleculares e exclusão de paternidade ....	12
2.5.1. Heterozigosidade .....	12
2.5.2. Conteúdo de informação polimórfica.....	15
2.5.3. Probabilidade de exclusão.....	16
2.5.4. Equilíbrio de Hardy-Weinberg.....	18
2.6. Inferências estatísticas para teste de paternidade.....	19
2.6.1. Método da razão de verossimilhança.....	20
2.6.2. Método fracional (Bayesiano) .....	22
2.7. Estudo de variabilidade genética utilizando marcadores de microssatélites.....	23

	<b>Página</b>
2.7.1. Distância genética.....	25
2.7.2. A distância na inferência filogenética .....	29
2.7.3. Variabilidade genética dentro e entre populações .....	33
2.7.4. O <i>Bootstrapping</i> .....	34
2.8. Recursos genéticos no Brasil .....	34
2.9. Diversidade genética com o uso de microssatélites de DNA.....	36
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>38</b>
<b>VERIFICAÇÃO DE PATERNIDADE EM CAPRINOS POR MEIO DE SISTEMA DE MICROSSATÉLITES DE DNA: MÉTODO DE EXCLUSÃO E INFERÊNCIA ESTATÍSTICA UTILIZANDO A RAZÃO DE VEROSSIMILHANÇA .....</b>	<b>46</b>
Resumo .....	46
Abstract .....	47
1. Introdução.....	48
2. Material e Métodos .....	50
2.1. Animais amostrados e extração de DNA.....	50
2.2. Reação de PCR e sistema de genotipagem .....	51
2.3. Análises estatísticas .....	53
3. Resultados .....	57
3.1. Descrição dos <i>loci</i> e probabilidade de exclusão .....	57
3.2. Inferência estatística por meio da razão de verossimilhança (CERVUS) .....	60
4. Discussão.....	63
5. Conclusões .....	66
Referências Bibliográficas.....	67
<b>MARCADORES DE MICROSSATÉLITES NO ESTUDO DE DIVERSIDADE GENÉTICA ENTRE POPULAÇÕES DAS RAÇAS CAPRINAS LEITEIRAS IMPORTADAS SAANEN E ALPINA E DA RAÇA NATURALIZADA MOXOTÓ NO BRASIL .....</b>	<b>70</b>
Resumo .....	70
Abstract .....	71
1. Introdução.....	72
2. Material e Métodos .....	73
2.1. Amostras de animais e extração de DNA.....	73
2.2. Microssatélites utilizados e sistema de genotipagem.....	74
2.3. Análises estatísticas .....	75
3. Resultados .....	76
4. Discussão.....	83

	<b>Página</b>
5. Conclusões .....	86
Agradecimentos .....	86
Referências Bibliográficas .....	86
3. CONCLUSÕES GERAIS .....	89

## RESUMO

ARAÚJO, Adriana Mello, D.S., Universidade Federal de Viçosa, janeiro de 2004. **Paternidade e diversidade genética em caprinos por meio de microssatélite de DNA.** Orientadora: Simone E. Facioni Guimarães. Conselheiros: Paulo Sávio Lopes e Robledo de Almeida Torres.

O registro genealógico confiável é fundamental para possibilitar a implantação de um programa de avaliação genética. Os marcadores de microssatélites constituem uma ferramenta amplamente utilizada atualmente para confirmação de paternidade. No Brasil, a caprinocultura possui um grande potencial para a pecuária, principalmente na agricultura familiar do semi-árido. As principais raças de caprinos especializadas para produção de leite no Brasil são a Alpina e a Saanen, ambas de origem européia. Na Região Nordeste, detentora do maior rebanho caprino do País, existem raças naturalizadas adaptadas às condições adversas do semi-árido, que possuem diversidade genética desconhecida. Estas raças têm sido cruzadas de forma indiscriminada com diversas raças importadas, pondo em risco o seu potencial genético. Os objetivos do trabalho foram estudar a utilização de um sistema de microssatélites para: 1) a verificação de paternidade, utilizando o método de exclusão e o método de inferência estatística da razão de verossimilhança, nos casos em que a mãe é ou não genotipada; e 2) o estudo de diversidade genética entre e dentro as

raças/rebanhos estudados. Foram utilizados, em sistema de genotipagem semi-automático (ABI 310, Perkin Elmer), onze microssatélites anteriormente descritos em bovinos, ovinos ou caprinos. Foram genotipados 292 animais, pertencentes a três rebanhos, sendo que as raças leiteiras Alpina e Saanen foram amostradas em dois rebanhos (UFV e particular) e a raça naturalizada Moxotó foi proveniente do rebanho de conservação da Embrapa Caprinos, Ceará. As análises de frequências alélicas foram realizadas utilizando-se os programas SAS (1998) e CERVUS (Marshall et al., 1998). O teste exato para equilíbrio de Hardy Weinberg, estatística-F e distância genética foram obtidos no programa TFPGA (Miller, 1997). Os *loci* estudados apresentaram herdabilidade e informatividade de moderada à alta. A heterozigosidade esperada média para todos os *loci* foi de 0,717. O conteúdo de informação polimórfica (PIC1 e PIC2) e a probabilidade de exclusão combinada (PE1 e PE2) foram de 0,676 e 0,5420; e 0,999591 e 0,988375, respectivamente, quando se conhecia ou não o genótipo materno. Considerando erros de registros aqueles com mais de três incompatibilidades de genótipo, foram encontrados 27 registros errôneos. Pelo teste de inferência estatística, foram solucionados corretamente 210 dos 276 casos de paternidade (76%). Se fosse adotado apenas o método de exclusão, apenas 160 paternidades seriam assinaladas com certeza (sem incompatibilidades pai alegado-progênie). Desta forma, o método de inferência pode ajudar as verificações de paternidade em larga escala, diminuindo o custo para solucionar os erros de laboratório e as perdas parciais de genotipagem. Quando há conhecimento do genótipo materno, a probabilidade de exclusão foi superior e o programa CERVUS resolveu 95% dos casos. No estudo de diversidade, a heterozigosidade ( $H_E$ ) foi alta nos rebanhos leiteiros, sendo 0,6952 e 0,7043 para a raça Alpina e Saanen, respectivamente. No Moxotó a  $H_E$  obtida foi moderada de 0,4984, provavelmente em decorrência do pequeno número amostrado. O número de alelos variou de cinco (INRA005) a onze (BM3205), com média de 7,0 alelos/*locus* nas raças importadas e 3,5 na Moxotó. A média  $F_{ST}$  (diferenciação entre populações) foi maior entre rebanhos ( $F_{ST} s = 0,0768$ ) do que entre raças ( $F_{ST} p = 0,0263$ ), indicando haver uma similaridade entre raças, dentro do rebanho. Tal similaridade pode ser devido ao

fato de haver algum cruzamento entre raças, dentro do rebanho. A distância genética de Nei ( $D_A$ ) foi maior entre o rebanho Moxotó e os demais ( $D_A > 0,50$ ). Os resultados obtidos estão em coerência com o histórico das raças, demonstrando que este sistema poderá ser adotado em estudos futuros de diversidade em caprinos no Brasil.

## ABSTRACT

ARAÚJO, Adriana Mello, D.S., Universidade Federal de Viçosa, January of 2004. **Paternity and genetic diversity in goats through genotyper of DNA microsatellite**. Adviser: Simone E. Facioni Guimarães. Committee members: Paulo Sávio Lopes and Robledo de Almeida Torres.

The reliable genealogical registration is fundamental to make possible the implantation of a program of genetic evaluation. The microsatellites markers is a tool thoroughly used now for confirmation of paternity. In Brazil, the goats possesses a great potential, mainly in the family agriculture of the semi-arid. The main specialized goat breeds for milk production in Brazil are Alpine and Saanen, both from European origin. In the Northeast area, holder of the largest goat flock of the country, exist naturalized breeds adapted to adverse conditions of the semi-arid, that possess unknown genetic diversity. These breeds have been crossed in an indiscriminate way with several imported animals, endangering their genetic potential. The objectives of the work were to study the use of a microsatellites system for: 1) verification of paternity being used the exclusion method and the method of statistical inference of the likelihood ratio, in the cases where the mother is genotyped or not; and 2) study of genetic diversity among and inside the studied breeds/flocks. Eleven microsatellites were used in genotyping system semiautomatic (ABI 310, Perkin Elmer), previously described in bovine, ovine or goats. The 292

animals were genotyped. The milk breeds Alpine and Saanen were sampled in two flocks (UFV and prived) and the naturalized breed Moxotó from conservation flock of National Goat Research Center of Embrapa, Ceará. The analyses of allelic frequencies were done in SAS (1998) and CERVUS (Marshall et al., 1998) computational programs. The exact test for Hardy Weinberg equilibrium, statistics-F and genetic distance was accomplished in program TFPGA (Miller, 1997). The studied loci presented heterozygosity and informativity from moderate to high. The overall heterozygosity was of .717. The polymorphic information content (PIC1 and PIC2) and the probability of combined exclusion (PE1 and PE2) were .676 and .5420; and .999591 and .988375, respectively when the maternal genotype is or not known. Considering mistakes of registrations those with more than three genotype incompatibilities, 27 erroneous registrations were found. For the test of statistical inference, 210 of the 276 cases of paternity was solved correctly (76%). If just the exclusion method was adopted, only 160 paternities would be marked for save (without incompatibilities alleged father-progeny). By this way, the inference method can help the verifications of paternity in large scale, decreasing the cost to solve the laboratory mistakes and partial losses of genotype. When there is knowledge of the maternal genotype, the exclusion probability was superior and CERVUS solved 95% of the cases. In the diversity study, expected heterozygosity ( $H_E$ ) was high in the milk imported flocks, being .6952 and .7043 for the Alpine and Saanen, respectively. Moxotó obtained a moderate  $H_E$  of .4984, probably due to the small sampled number. The number of alleles ranged of five (INRA005) to eleven (BM3205), with average of 7.0 alleles/locus in the imported breeds and 3.5 in Moxotó. The average  $F_{ST}$  (differentiation among populations) was larger among origin flocks ( $F_{ST} s = .0768$ ) than among breeds ( $F_{ST} p = .0263$ ), indicating a similarity among breeds, inside the flock. Such similarity can due to the fact being some crossing among breeds, inside of the flock. Nei's genetic distance ( $D_A$ ) was larger among the flock Moxotó and the others ( $DA > 0.50$ ). The obtained results are in coherence with the report of the breeds, demonstrating that this system can be adopted in future studies of diversity in goats in Brazil.

## 1. INTRODUÇÃO GERAL

O Brasil possui o décimo rebanho mundial de caprinos e o maior efetivo das Américas, com cerca de 9,5 milhões de cabeças (IBGE, 2001). Entretanto, a atividade não possui expressão econômica proporcional, principalmente na Região Nordeste, detentora de 88% do efetivo nacional. Os órgãos de desenvolvimento desta região apontam a caprinocultura como uma opção sustentável para o avanço da pecuária na região e buscam aumentar a produtividade dos plantéis existentes.

O leite de cabra é considerado pelo mercado consumidor como um produto nobre, de conhecido valor nutritivo, atendendo a dietas especiais de idosos, de crianças e a uma faixa cada vez maior de pessoas que buscam uma alimentação saudável. Entretanto, apesar do grande rebanho caprino, a escassez da oferta têm tornado o produto inacessível à grande parte da população, inibindo o hábito do consumo deste alimento. Além do leite, a carne e a pele se destacam como produtos que elevam a renda do produtor de caprinos. A pele da raça nativa Moxotó é considerada uma das melhores do mundo. A carne de animais precoces tem boa aceitação, embora o consumo ainda seja baixo.

As raças caprinas leiteiras mais expressivas no Brasil são a Saanen e a Alpina. Ambas são consideradas de origem alpina. Segundo Santos (2003), a raça Alpina está muito presente nos rebanhos da Região Nordeste, enquanto o

rebanho Saanen é mais numeroso nas Regiões Sul e Sudeste. Em termos mundiais, a raça Saanen é a maior produtora de leite, podendo chegar a 3.000 kg de leite em 305 dias de lactação.

As raças caprinas naturalizadas (autóctones) apresentam extrema rusticidade e adaptação às condições adversas da caatinga, vegetação que cobre quase todo o Sertão Nordestino. Entretanto, estas estão ameaçadas de extinção devido à sua baixa produtividade e ao desconhecimento de seu potencial futuro, estando hoje restritas a poucos rebanhos de conservação. O cruzamento desordenado entre os animais naturalizados e exóticos contribui para a viabilidade econômica da caprinocultura em curto prazo, mas geram a erosão do patrimônio genético adaptado. A raça Moxotó é, dentre as autóctones, a mais representativa. É reconhecida como uma raça de dupla aptidão, embora sua produção leiteira seja bem reduzida.

O desenvolvimento da genética molecular possibilitou a identificação de relações entre o genótipo progenitor-progênie e o surgimento de metodologias visando a confirmação de paternidade. A importância dos testes de paternidade para o melhoramento animal pode ser ilustrada nas perdas da acurácia de seleção e redução do ganho genético originadas de identificações errôneas na genealogia. O teste de progênie, principal ferramenta para o melhoramento animal, se baseia em avaliações genéticas com base no desempenho de parentes. A acurácia do teste depende, em parte, da precisão das anotações de paternidade. Segundo Ron et al. (1996), erros de paternidade podem atingir até 20% dos registros, reduzindo drasticamente o ganho genético obtido na população pela seleção.

Outro aspecto importante dos testes de paternidade para a indústria animal é a confiabilidade adquirida por espécies/raças que realizam sistematicamente o teste para o registro genealógico. Segundo Penna et al. (1998), parece haver um consenso por parte dos produtores em aceitar erros de genealogia, ignorando-se contudo as conseqüências para o melhoramento genético.

Os primeiros marcadores utilizados para confirmação de paternidade foram os grupos sanguíneos ou marcadores imunológicos. Muitos laboratórios ainda utilizam estes marcadores, pois há amplo conhecimento prévio dos

sistemas e a forma de herança é simples e independente. Avanços na área de biotecnologia e no mapeamento genético revelaram a existência de seqüências repetitivas curtas na molécula de DNA, denominados microssatélites. Estes marcadores fornecem marcas codominantes, apresentam alta variabilidade e são passíveis de automação em suas análises, destacando-se como uma ferramenta importante para testes de paternidade.

Nos últimos trinta anos, o polimorfismo de DNA vem substituindo o estudo de polimorfismos protéicos. Como os microssatélites podem ser amplificados pela reação em cadeia da polimerase (PCR), o teste requer pequena quantidade de DNA e pode ser automatizado e realizado rapidamente, tornando-se a maior ferramenta para verificação de *pedigree*.

Os testes de paternidade são feitos em sistemas de genotipagem semi-automático, em que os *primers* são marcados fluorescentemente, permitindo que se construam painéis onde vários *loci* são analisados simultaneamente, contanto que as fluorescências e os tamanhos dos fragmentos não coincidam. Otimizando custo e tempo, tais sistemas permitem ainda um aumento na precisão dos resultados, pois a análise do tamanho dos fragmentos é obtida através de *softwares* específicos.

Apesar do potencial produtivo dos caprinos, ainda são poucos os estudos envolvendo o melhoramento genético da espécie. Não há no País teste de progênie com caprinos, pois uma das dificuldades enfrentadas na implantação deste é a atual ausência de testes para a confirmação de paternidade.

Este trabalho buscou contribuir para a organização de um teste de paternidade em caprinos, utilizando um sistema multiplex de microssatélites, que possa ser implementado por associações de criadores, juntando esforços para a consolidação de um registro de *pedigree* confiável para um futuro banco de dados para avaliação genética de reprodutores no Brasil. Por outro lado, o conhecimento da variabilidade de alguns *loci* nas raças leiteiras Saanen e Alpina, bem como na nativa Moxotó, poderá contribuir para futuros estudos na área de genética molecular de caprinos no País.

## 2. REVISÃO DE LITERATURA

### 2.1. Os microssatélites de DNA

O DNA de eucariontes é composto de uma fração não-repetitiva, que está presente em cópia única no genoma, sendo a maioria dos genes estruturais constituída de DNA não-repetitivo, com exceção dos genes em tandem que codificam rRNA, tRNA e histonas. Cerca de 50-60% do DNA de mamíferos é composto de DNA não-repetitivo (Lodish et al., 1999). Entretanto, apenas 5% do DNA total representa moléculas funcionais de mRNA, devendo ser ressaltado que o DNA cópia simples restante não possui função bem conhecida e é referido como DNA espaçador.

A fração de DNA repetitivo é composta de seqüências de DNA presentes em mais de uma cópia no genoma, podendo ser subdividido em DNA moderadamente repetido e DNA altamente repetitivo, de acordo com a freqüência de repetição. De acordo com a classificação apresentada por Lodish et al. (1999), a fração moderadamente repetida é composta por elementos móveis, como *transposons*, *LINES (long interspersed elements)* e *SINES (short interspersed elements)*, além de famílias gênicas muito duplicadas. O DNA altamente repetitivo é conhecido como DNA satélite ou DNA com repetição de seqüências simples (SSR).

Os diferentes componentes no genoma eucarionte podem ser isolados através de sua propriedade de reassociação. A forma não-repetitiva de DNA possui uma única seqüência de anelamento, mantendo a temperatura de desnaturação ( $T_m$ ) em condições de reassociação. Por outro lado, o DNA repetitivo não consiste de seqüências exatamente pareadas, mas pode conter erros de pareamento que diminuem o número de pontes de hidrogênio que necessitam ser quebradas na desnaturação, diminuindo, portanto, a  $T_m$  em condições de reassociação.

O DNA repetitivo, portanto, é definido por sua relativamente rápida taxa de reassociação (Lewin, 2000). Os componentes do DNA que reassociam mais rapidamente no genoma eucarionte são seqüências altamente repetitivas e muito curtas, repetidas várias vezes em tandem, formando blocos. Tal conformação cria uma fração com propriedades físicas próprias, que pode ser isolada do restante do DNA através de centrifugação em gradiente de densidade pré-formado. Tal fração de densidade própria é chamada de DNA satélite.

A densidade do duplex de DNA depende do conteúdo de C e G. Em gradiente de densidade de CsCl, frações de DNA contendo mais de 5% de diferença no conteúdo C/G podem ser separadas. Quando o DNA eucarionte é centrifugado em gradiente de densidade, separa-se um pico de amplitude maior, com a densidade correspondente ao conteúdo C/G do DNA genômico, e um pico menor, com valor de densidade menor do que a média, que é o DNA satélite, rico em pareamentos A/T.

Funcionalmente, o DNA satélite não é transcrito nem traduzido. Seqüências repetidas em tandem são especializadas para permitir um alinhamento imperfeito durante o pareamento dos cromossomos, e, portanto, o tamanho dos blocos de repetições tende a ser altamente polimórfico. Estudos com hibridização *in situ* demonstraram que os blocos de DNA altamente repetitivo são localizados em regiões de heterocromatina, principalmente nas regiões dos centrômeros (Burow e Blake, 1998).

Os *Variable Number of Tandem Repeat* (VNTRs) são fragmentos cujo tamanho é extremamente variável, com seqüências curtas, de 5-10 bp, com 15-

500 repetições em tandem, com várias cópias espalhadas pelo genoma. Os VNRTs podem ser microssatélites, quando compreendem repetições de um a seis nucleotídeos, ou minisatélites, se possuem mais do que seis nucleotídeos por repetição. Muitos autores dividem didaticamente o termo VNRT para designar os polimorfismos de minisatélites, reservando aos microssatélites a denominação de STR (*short tandem repeat*).

Os VNRTs são comuns em mamíferos e plantas superiores. Estas seqüências possuem variação nos alelos com diferentes números de unidade de repetição. A causa desta variação é a recombinação entre anelamentos errôneos de seqüências curtas. A taxa de permuta genética em regiões satélites é alta, cerca de  $10^{-4}$  por kb de DNA, que corresponde a cerca de 10 vezes mais que a taxa de recombinação entre homólogos na meiose em seqüências de DNA aleatória. Então, os mini- e microssatélites podem ser pontos (*hotspots*) de maior recombinação meiótica (Lewin, 2000). Além disto, a suposta ausência de pressão de seleção nestas seqüências não-codificadoras contribui para o alto polimorfismo do DNA repetitivo.

A maioria dos polimorfismos de microssatélites refere-se aos dinucleotídeos  $(GT)_n:(CA)_n$ ,  $(GA)_n:(CT)_n$ ,  $(CG)_n:(GC)_n$  e  $(AT)_n:(TA)_n$  que são amplamente distribuídos no genoma, sendo que o dinucleotídeo  $(GT)_n$  é a repetição quantitativamente mais abundante em mamíferos (Ellegren, 1993). A primeira análise de microssatélite  $(GT)_n$  foi reportada em ovinos por Crawford et al. (1990).

O trabalho de Saiki et al. (1985) com alelos de  $\beta$ -globulina em humanos mostrou primeiramente o polimorfismo genético devido ao tamanho variável de seqüências repetitivas de DNA. Entretanto, a importância dos marcadores de microssatélites para análise genômica apareceu somente após o desenvolvimento da técnica de reação em cadeia da polimerase (PCR) por Saiki et al. (1988).

Os *loci* de microssatélites ocorrem em todo o genoma de muitos organismos e, portanto, podem ser usados como marcadores para estabelecer grupos de ligação e mapear cromossomos. Se a freqüência alélica for conhecida, os *loci* de microssatélites altamente polimórficos podem ser úteis para

identificação individual em uma população e para determinar a probabilidade de parentesco entre dois indivíduos. Trommelen et al. (1993) aplicaram o perfil de DNA para estudo de identidade e parentesco em bovinos, usando marcadores de mini- e microssatélites, através do método de hibridização e *Southern blot*, com sondas que detectam simultaneamente vários *loci* VNTRs.

Alford et al. (1994) compararam o teste de paternidade em humanos com nove *loci* STR amplificados por PCR, em trios previamente testados usando *loci* minissatélites e *Southern*. Os resultados demonstraram mais de 99% de similaridade entre os dois métodos, validando o teste de paternidade através de *loci* STR amplificados por PCR.

A análise de microssatélites através de PCR possui as seguintes vantagens sobre o método convencional com sondas (Burow e Blake, 1998):

- O tamanho pequeno do fragmento *microssatélite* (geralmente menor do que 350 bp) permite análise de amostras com pequenas quantidades de DNA e/ou DNA degenerado;

- O tamanho pequeno do fragmento *microssatélite* permite a co-amplificação por PCR (multiplex);

- Os alelos dos microssatélites têm tamanhos discretos, permitindo simples interpretação dos resultados; e

- Os testes baseados em PCR são rápidos, com resultados obtidos em 24 horas ou menos;

- Os testes baseados em PCR são padronizados e facilmente automatizados.

Os *loci* de microssatélites podem ser obtidos através de consulta à banco de dados de seqüências de DNA, como o *Genebank* ou o EMBL, o que confere simplicidade escolha dos *loci* a serem utilizados para verificação de paternidade ou para mapeamento de ligação. Em tais bancos de dados, os *primers* flanqueadores da seqüência dos microssatélites são fornecidos junto com a seqüência de nucleotídeos da região do microssatélite.

## 2.2. O mapa genômico de caprinos

Nos últimos anos, um grande esforço tem levado à construção de mapas de ligação de alta resolução em várias espécies. Em bovinos, Bishop et al. (1994) construíram o primeiro mapa de ligação da espécie bovina. Em ovinos, Crawford et al. (1995) identificaram marcadores polimórficos, essencialmente microssatélites, para a construção do primeiro mapa de ligação desta espécie. Tal mapa cobria 2.070 cM, correspondendo a ~75% do genoma total, com marcas nos 26 cromossomos autossomais, espaçadas de 120-30 cM. Três anos depois, um mapa de segunda geração foi publicado, construído com 519 marcas e cobrindo 3.063 cM (DeGartori et al., 1998). O mapa de terceira geração foi desenvolvido com 1.062 *loci* (941 anônimos e 121 genes), oriundo da compilação do trabalho de genotipagem de 15 laboratórios (Maddox et al., 2001), cobrindo 3.400 cM dos autossomais, com intervalo médio de 6,5 cM.

O primeiro mapa genético de caprinos foi desenvolvido por Vaiman et al. (1996), utilizando a estratégia de mapeamento através da identificação de um grande número de microssatélites como uma forma de retirar o mapeamento de caprinos de um estágio incipiente, onde até o final dos anos oitenta só existia descrito o grupo de ligação dos genes da caseína (Vaiman et al., 1996), baseados na estratégia de explorar o curto tempo evolucionário entre as espécies caprina, ovina e bovina - há cerca de 8 e 17 milhões de anos - e, portanto, nas homologies genéticas existentes entre estas espécies de ruminantes. Obtiveram 34% dos *loci* de microssatélites de bovinos fossem polimórficos em caprinos. Ainda no mesmo estudo, dos microssatélites obtidos em ovinos, 43% foram satisfatoriamente amplificados e polimórficos em caprinos.

As principais características comparativas do cariótipo entre as três principais espécies de ruminantes domésticos são (Hayes et al., 1991): 1) a fusão Robertsoniana do centrômero em ovinos, diferindo de caprinos e bovinos. Os cromossomos 1, 2 e 3 de ovinos são equivalentes à fusão dos cromossomos 1 e 3; 2 e 8; 5 e 11 de bovinos ou caprinos; 2) a translocação 9-14 diferenciando as famílias *Bovinae* e *Caprinae*; e 3) o rearranjo intracromossomal do X, em que o

braço Xp em bovinos corresponde ao Xq em caprinos e ovinos.

No mapa de ligação proposto Vaiman et al. (1996), o tamanho estimado do genoma da espécie caprina foi de 3.250 cM. Dos 219 marcadores utilizados, 204 marcas eram de microssatélites, cobrindo >80% do comprimento total do genoma ou 2.300 cM, com espaçamento entre marcadores médio de 16,1 cM. Entre as marcas utilizadas, um conjunto de 55 microssatélites é comum às três maiores espécies de ruminantes domésticos. Além disto, 25 seqüências codificadoras de genes foram localizadas no genoma caprino, sendo que doze genes foram mapeados genética e fisicamente no decorrer deste trabalho, através de FISH (*fluorescent in situ hybridization*) e análise de células somáticas híbridas.

O mapa de segunda geração de caprinos foi gerado através da incorporação de 77 marcadores de microssatélites e 33 genes, gerando 25 localizações consistentes através de FISH. O mapa genético elaborado cobriu 2737 cM, compreendendo >88% do genoma da espécie, inserindo 307 marcas, em 40 grupos de ligações em 29 cromossomos autossomais, o que fornece um intervalo entre marcas de aproximadamente 10 cM (Schibler et al., 1998).

Schibler et al. (1998) realizaram o primeiro estudo de mapeamento comparativo integrado de ruminantes, estabelecido em caprinos. Este mapa constitui um esforço para comparar efetivamente os genomas primatas/ruminantes, ajudando a definir regiões homólogas a *loci* economicamente relevantes em ruminantes, facilitando a ligação de genes candidatos dentro de segmentos limitados de DNA dos cromossomos.

Na análise comparativa de Schibler et al. (1998), dos 255 *loci* Tipo I mapeados em ruminantes, 249 seqüências codificadoras foram avaliadas em humanos e 173 em camundongos. Foram identificados 107 fragmentos conservados entre humano/ruminante e 46 segmentos entre camundongo/ruminante. O tamanho médio do fragmento conservado entre humano/ruminante foi 27,7 cM, implicando no mínimo de 100 rupturas evolutivas, em cerca de 70 milhões de anos. Entre camundongo/ruminante, a estimativa do tamanho do fragmento conservado foi de 7,7 cM, indicando um total de 187 rupturas desde a

divergência entre as espécies, há cerca de 100 milhões de anos atrás.

Atualmente existem no banco de dados da espécie caprina (<http://locus.jouy.inra.fr>) 727 *loci* descritos, sendo 303 genes e 424 referentes a *loci* de microssatélites. No *GoatMap* foram identificados 123 *loci* homólogos a ovinos e 275 homologias com bovinos.

### **2.3. Marcadores para testes de paternidade**

Os marcadores protéicos e os grupos sangüíneos constituíam o padrão de teste de paternidade para diversas espécies, sendo muito usados em humanos, bovinos e eqüinos (Bowling, 2001). Os sistemas imunológicos e protéicos são expressos independentemente de fatores externos, além de possuir uma forma de transmissão hereditária simples (Burow e Blake, 1998). Além destas características favoráveis do ponto de vista do estudo genético, os sistemas conhecidos de grupos sangüíneos pertencem a sistemas genéticos independentes. Estes marcadores possuem ainda a vantagem de possuir fácil execução e requerer pouco trabalho preliminar, quando comparados com os marcadores de DNA (Burow e Blake, 1998). Entretanto, são limitados em número (número de *locus* no genoma e número de alelos por *locus*). Em caprinos, nove sistemas de grupos sangüíneos foram identificados por Vankan e Bell (1993). Grosclaude et al. (1992) fizeram uma revisão sobre os polimorfismos sangüíneos e das proteínas do leite em caprinos.

Os marcadores bioquímicos foram sendo substituídos pelos marcadores moleculares ou de DNA para testes de paternidade e estudos genéticos a partir da década de 1980. Segundo Bowling (2001), para um teste de paternidade ter aplicação de rotina, são características essenciais acurácia, efetividade e baixo custo, além da rapidez de resposta, padronização de referência e fácil transferência de informações entre laboratórios. Ainda segundo Bowling (2001), o processo não pode ser limitado ao sangue fresco e deve ser automatizado.

O primeiro marcador molecular a ser amplamente utilizado para genotipagem foi o RFLP (*restricted fragment length polymorphism*). A técnica

envolvia hibridização de DNA com sondas de oligonucleotídeos marcadas radiotivamente, que consumia longo tempo para sua execução (Usha et al., 1995). O RFLP associado a sondas de minissatélites foi usado para a identificação individual, através da técnica de *DNA fingerprint*, que envolve um padrão *multilocus* complexo de bandas eletroforéticas, o que dificulta o acompanhamento dos alelos segregantes em gerações sucessivas (Bolstein et al., 1980).

Os marcadores RFLP são numerosos, codominantes e herdados segundo as leis de Mendel (Bolstein et al, 1980). Entretanto, Bowling (2001) indica que esta técnica requer muito DNA, consome muito tempo para apresentar os resultados e dificulta aplicações dos resultados obtidos em outro caso, não sendo recomendada na indústria animal.

Dentre as técnicas da biologia molecular, a reação de PCR (*polimerase chain reaction*) permite amplificar a seqüência de DNA específica entre dois oligonucleotídeos conhecidos (*primers*). Sua descrição por Saiki et al. (1988), a qual substituiu o fragmento de Klenow de *E. Coli* pela enzima DNA polimerase termostável (*Taq* Polimerase) tornou a reação de PCR passível de automatização, aumentou a especificidade e sensibilidade da reação, além de possibilitar amplificação de fragmentos maiores, viabilizando, portanto, seu uso em larga escala.

O uso de PCR e eletroforese automatizada, aliada à tecnologia de análise possível com os seqüenciadores de DNA permitiram o processo de automação de genotipagem de *loci* STR (Alford et al., 1994). Os polimorfismos de microssatélites são geralmente pequenos em tamanho, assim, para permitir a obtenção de maior informação, um sistema de análise de alta resolução é necessário. Para análise de fragmentos de DNA, a imagem do gel de eletroforese dos fragmentos amplificados por PCR pode ser obtida pelo padrão de seqüência ou pela análise automática, que captura imagens com câmaras do tipo CCD (*charge coupled device*).

O sistema de teste de paternidade com base na amplificação por PCR possui procedimentos facilmente automatizados e pode ser simplificado através

do carregamento de várias amplificações na mesma canaleta do gel de eletroforese, ou co-amplificando a reação de PCR (multiplex). Os sistemas de reação de PCR por multiplex permitem que vários sistemas de microssatélites sejam eficientemente analisados em conjunto. O conjunto de *primers* utilizado no sistema multiplex deve permitir a amplificação de regiões de microssatélites contendo seqüências de tamanho diferentes e não sobrepostas.

#### **2.4. Testes de paternidade por meio de microssatélites de DNA em animais domésticos propostos na literatura**

Em bovinos, existem vários estudos sobre marcadores de microssatélites para testes de paternidade. Usha et al (1995), com apenas cinco sistemas de microssatélites, obtiveram probabilidade de exclusão de 0,99. Glawatzki-Mullis et al. (1995), utilizaram um sistema semi-automático de genotipagem com seis *loci* microssatélites, e verificou-se o seu desempenho em casos de paternidade não resolvidos pela tipagem sanguínea convencional. Heyen et al. (1997) obtiveram conteúdo de informação polimórfica médio de 0,59 e probabilidade de exclusão do pai incorreto  $>0,99999$ , quando utilizaram 22 sistemas de microssatélites, em multiplex. Em eqüinos, Marklund et al. (1994) obtiveram probabilidade de exclusão variando de 0,96 a 0,99 com nove sistemas de microssatélites. Em caprinos, Luikart et al. (1999) desenvolveram um painel com 22 microssatélites em dois multiplex para o teste de paternidade obtendo uma probabilidade de exclusão  $>0,99999$ .

#### **2.5. Estatística para marcadores moleculares e exclusão de paternidade**

##### **2.5.1. Heterozigosidade**

A heterozigosidade de um *locus* é definida como a probabilidade de que um indivíduo seja heterozigoto naquele *locus*, em uma população. A heterozigosidade pode ser considerada uma medida de variabilidade genética.

Segundo Weir (1996) a frequência de heterozigotos é importante, pois estes carregam diferentes alelos e representam a existência de variação.

A heterozigosidade observada ( $H_O$ ) é a contagem de heterozigotos  $A_u A_v$ ,  $u \neq v$ , no *locus*  $l$ , com uma amostra de  $n$ .

$$H_{Ol} = \sum_u \sum_{v \neq u} \frac{n_{uv}}{n}$$

Se  $m$  *loci* são tipados, então a  $H_O$  média é dada por

$$H_O = \frac{1}{m} \sum_{l=1}^m H_{Ol}$$

Para uma população em equilíbrio de Hardy-Weinberg, a heterozigosidade esperada ( $H_E$ ) pode ser considerada em termos de frequência genotípica dos homozigotos ( $p_{ii}$ ) ou frequência alélica ( $p_i$ ).

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

Quando uma população natural é usada em análise genômica, o número de alelos e a frequência alélica não são conhecidos. Neste caso, a frequência alélica é estimada por

$$\hat{p}_i = \frac{k_i}{N},$$

em que  $k_i$  é o número de observações do alelo  $i$  ( $i = 1, \dots, m$ ), e  $N$  é o número total de informações ( $2n$  alelos).

A estimativa de heterozigosidade, com base em  $k$  indivíduos ( $2n$  alelos) é dada por (Ott, 1992).

$$\hat{H}_u = \frac{n}{n-1} \left( 1 - \sum_{i=1}^m p_i^2 \right)$$

Assim ao adotar  $H_E$ , ao invés de  $\hat{H}_u$ , geralmente se está superestimando a heterozigidade. O viés possui maior impacto quando o número de indivíduos é menor.

Segundo Ott (1992), o polimorfismo para um marcador genético é definido como uma variação detectável e herdável em um *locus*. Um marcador é considerado polimórfico se o mais abundante dos alelos na população possui uma frequência menor do que, por exemplo, 95%. A escolha do nível de 95% implica que a probabilidade do *locus* ser heterozigoto na população é de 10%. A heterozigidade é maximizada quando as frequências alélicas são iguais ( $p_i = p_j = 0,50$ ). Ainda segundo Ott (1992), um *locus* é considerado altamente polimórfico se sua heterozigidade for maior do que 70%, o que implica frequência menor do que 0,55 do mais freqüente alelo.

A variância de  $H_O$  e  $H_E$  foram definidas segundo Weir (1996)

$$VarH_O = \frac{1}{m^2} \sum_l \frac{H_l(1-H_l)}{n_l},$$

em que  $n_l$  é o número de animais genotipados no *locus*  $l$ ;  $m$  é o número de *loci* estudados e  $H_l$  é a heterozigidade observada no *locus*  $l$ .

$$VarH_E = \frac{2}{m^2} \sum_l \frac{(1+f_l) \left[ \sum_u p_{lu}^3 - (\sum_u p_{lu}^2)^2 \right]}{n_l},$$

em que  $p_{lu}$  é a frequência do alelo  $u$  no *locus*  $l$ ; e  $f_l$  é o coeficiente  $F_{IS}$  de Wright.

## 2.5.2. Conteúdo de informação polimórfica

Na genética humana surgiu o PIC (*polimorphic information content*) para quantificar o valor da informação do polimorfismo de um *locus* marcador (Bolstein et al., 1980). Segundo os autores, informatividade representa a probabilidade de uma dada progênie de um progenitor carregando um alelo *i* permitir a dedução do genótipo parental. O PIC pode ser deduzido através do somatório das freqüências de acasalamento multiplicado pela probabilidade da progênie ser informativa, de acordo com Tabela 1.

Tabela 1 – Tipos de acasalamento e probabilidade da progênie ser informativa (adaptada de Bolstein et al., 1980)

Genótipo do progenitor conhecido	Genótipo do progenitor desconhecido (ao acaso)	Probabilidade de a progênie ser informativa		Freqüência do acasalamento
		I	II*	
$A_i A_j$	$A_k A_l$	1.0	1.0	$2 p_i p_j (1 - p_i - p_j)^2$
$A_i A_j$	$A_k A_i$ ou $A_k A_j$	1.0	$\frac{3}{4}$	$2 p_i p_j (1 - p_i - p_j) (2p_i + 2p_j)$
$A_i A_j$	Homozigotos	1.0	$\frac{1}{2}$	$2 p_i p_j \sum p_i^2$
$A_i A_j$	$A_i A_j$	$\frac{1}{2}$	$\frac{1}{2}$	$2 p_i p_j \times 2 p_i p_j$
Homozigotos	Todos os genótipos	0	0	$p_i^2$

I) No caso de ser conhecido o genótipo da mãe, da progênie e do suposto pai; e II) No caso de ser conhecido o genótipo da progênie e do suposto pai apenas (\*adaptado de Da e Lewin, 1995).

Na prática, em dúvidas no pedigree, podem ocorrer duas situações: os genótipos do filho e do suposto progenitor são conhecidos, mas o genótipo do progenitor confirmado é desconhecido; o genótipo do filho, do progenitor confirmado e do alegado são conhecidos. Da e Lewin (1995) adaptaram o conteúdo de informatividade para populações em acasalamento ao acaso, quando o genótipo de ambos os progenitores é desconhecido (Tabela 1-II).

Assim, conforme definido por Bolstein et al. (1980),

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

Conforme Da e Lewin (1995), PIC é a heterozigosidade adicionada da frequência esperada de genótipos informativos na progênie. A frequência esperada de genótipos informativos na progênie (I) corresponde ao somatório de todos os pais heterozigotos de uma amostra aleatória de progenitores (pai ou mãe), quando os dois progenitores possuem o genótipo conhecido, levando a fórmula apresentada anteriormente por Bolstein et al. (1980).

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

No caso de só haver o conhecimento do genótipo do suposto progenitor, então

$$I = 1 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_j (p_i + p_j).$$

Isto leva a uma nova fórmula para definir PIC, no caso de genotipagem de apenas um progenitor:

$$PIC = 1 - \sum p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_j (p_i + p_j)$$

Quando o número de alelos é grande, o conteúdo de informação polimórfica aproximadamente iguala-se a heterozigosidade, o que é desejável.

### 2.5.3. Probabilidade de exclusão

A probabilidade de exclusão é definida como a probabilidade de um pai alegado ser excluído como pai potencial, assumindo que o pai alegado é um indivíduo ao acaso (Heyen et al., 1997). A fórmula de exclusão de paternidade é baseada na probabilidade teórica de exclusão dos pais apontados incorretamente. As probabilidades de inclusão e exclusão são complementares, partes da unidade. A incidência do evento  $AA \times AA = a^5$  é o produto das frequências dos genótipos do acasalamento ( $a^2 \times a^2$ ), multiplicado pela frequência ( $a$ ) do alelo  $A$  na população. A probabilidade de exclusão, neste mesmo caso, seria a probabilidade do acasalamento  $AA \times AA$  ( $a^2 \times a^2$ ), multiplicado pela frequência ( $b$ ) do alelo  $B$ , que exclui o progenitor.

A probabilidade de exclusão, considerando dois alelos  $PE_2$  é dada pelo somatório da probabilidade de exclusão de cada genótipo, apresentada na Tabela 2.

$$PE_2 = a^4b + a^3b^2 + a^2b^3 + ab^4 + a^3b^2 + a^2b^3$$

Quando se consideram vários *loci* para determinar a paternidade, deve-se considerar que o observado de cada probabilidade de exclusão é a probabilidade de inclusão  $PI = (1 - PE)$ , pois uma exclusão é conclusiva. Assim, a probabilidade  $PI$  é a que não é excluída em nenhum dos *loci*,  $(1 - PE_{n1})(1 - PE_{n2}) \dots (1 - PE_{nk})$ .

Se o número de alelos ( $n$ ) for o mesmo para todos os *loci*  $k$ , então a probabilidade de exclusão combinada ( $PEC$ ) é dada por

$$PEC = 1 - (1 - PE_n)^K$$

Em termos de frequência genotípica, a fórmula pode ser deduzida em

$$PE_n = \sum_{i=1}^n (1 - p_i)^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n (p_i p_j)^2 [4 - 3(p_i + p_j)]$$

onde para calcular a probabilidade de usar estes *loci* para detectar parentescos falsos, deve-se substituir as frequências alélicas obtidas na fórmula.

Tabela 2 – Probabilidade de exclusão usando um par de alelos co-dominantes, A e B, com frequências  $a$  e  $b$  (adaptada de Jamieson, 1994)

Genótipos Mãe x Pai	Progênie esperada			Contribuições		Pais excluídos (Frequência do acasalamento vs. Frequência alélica)
	AA	AB	BB	Verdadeiro	Falso	
AA x AA	1	0	0	A	B	$a^2$ x $a^2$ (b)
AA x AB	0,5	0,5	0	A e B		$b^2$ (a)
AA x BB	0	1	0	B	A	$b^2$ (a)
AB x AA	0,5	0,5	0	A e 0,5 B	0,5 B	$2 ab$ x $a^2$ (0,5b)
AB x AB	0,25	0,5	0,25	A e B		$b^2$ (0,5a)
AB x BB	0	0,5	0,5	0,5 A e B	0,5 A	$b^2$ (0,5a)
BB x AA	0	1	0	A	B	$b^2$ x $a^2$ (b)
BB x AB	0	0,5	0,5	A e B		$b^2$ (a)
BB x BB	0	0	1	B	A	$b^2$ (a)

#### 2.5.4. Equilíbrio de Hardy-Weinberg

Segundo o trabalho de Jamieson (1994), o parentesco (endogamia) pode introduzir ambigüidade nos testes de identidade ou pedigree. As técnicas de verificação de pedigree assumem o equilíbrio de Hardy-Weinberg (HW), o que é pouco provável nos rebanhos comerciais. No cálculo de probabilidade de exclusão (PE) para testes de paternidade, é apropriado usar os dados de frequência alélica e genotípica dos alelos marcadores na população ou rebanho em estudo, para possibilitar a correção para afastamentos do equilíbrio de HW (Usha et al., 1995).

Segundo a definição de Falconer (1989), uma população com frequência alélica e genotípica constante, geração após geração, é dita em equilíbrio de HW. A frequência no equilíbrio de Hardy-Weinberg origina-se das leis de segregação de Mendel. A condição de acasalamento ao acaso e ausência de seleção para aquele *locus* é requerida para a manutenção do equilíbrio nas gerações sucessivas. Quando se considera mais de um *locus*, a manutenção do equilíbrio de HW depende da recombinação entre ambos os *loci*. Se a recombinação é grande, próxima de 1, então os *loci* não estão ligados e a seleção para um *locus* não retira do equilíbrio o outro *locus*. Entretanto, se os *loci* estão ligados, a seleção para um dos *loci* retira ambos do equilíbrio.

Os testes mais comuns para equilíbrio de HW são os testes de qui-quadrado *godness-of-fit* para as proporções esperadas e observadas de heterozigotos. Entretanto este teste produz resultados imprecisos quando há um ou mais genótipos raros. Uma solução é o agrupamento de categorias, a fim de minimizar o viés introduzido. Mesmo assim, o qui-quadrado só é recomendado para estudos preliminares (Weir, 1996).

O teste exato ou de probabilidade (Haldane, 1954, citado em Weir, 1996) para proporções de equilíbrio de HW baseia-se na rejeição da hipótese caso o conjunto de dados possua valor baixo de probabilidade. O teste exato é bilateral e o excesso ou falta de heterozigotos podem levar a valores de  $p$  pequenos, rejeitando a hipótese de equilíbrio. O problema do teste consiste em identificar os

valores de significância quando há muitos arranjos de genótipo que possam ocorrer (maior número de alelos).

O teste exato foi implementado por Guo e Thompson (1992) quando há mais de dois alelos por *locus*. Este utiliza o método de Monte Carlo (poucos dados) ou Cadeia de Markov (muitos dados) para obter amostras aleatórias a partir dos dados originais (permutação). O método calcula um erro padrão da estimativa de  $p$ , que deve ser menor que o nível de significância. O valor de  $p$  é dado pela probabilidade agregada de todos os arranjos genotípicos amostrados, com a mesma frequência de alelos.

## 2.6. Inferências estatísticas para teste de paternidade

Os métodos estatísticos para análise de paternidade em populações naturais, ou onde os dados são levantados em larga escala, podem diferir do enfoque utilizado na genética humana, onde apenas a confirmação de compatibilidade do genótipo do trio pai alegado-mãe-progênie é o objetivo (Marshall et al., 1998). O método de exclusão é amplamente utilizado em humanos e é adotado para populações animais. Neste método, o compartilhamento de alelos entre a progênie e o suposto pai estabelece a inclusão de paternidade e o macho que possui o genótipo incompatível com a progênie em questão é rejeitado como pai. O único pai que permaneça não excluído é tomado como o verdadeiro pai. Este método não leva em consideração os erros de tipagem que possam ocorrer no laboratório e não admite que haja *loci* não tipados.

Exclusões de paternidade são incontestáveis, enquanto a prova da paternidade depende de inferências estatísticas baseadas na ausência de uma exclusão. Isto forma a base para o cálculo da seguinte probabilidade de paternidade  $PP = \frac{1}{(1-PE)}$ . Entretanto, para pares específicos mãe-filho, o alelo paterno obrigatório pode ser freqüente, enquanto em outras situações ele pode ser raro. Assim, a ausência de exclusão na última situação será mais significativa que na primeira. Isto pode ser quantificado pela razão de verossimilhança entre

paternidade e não paternidade, chamada Índice de Paternidade (IP), é uma maneira de quantificar a evidência encontrada em favor da paternidade biológica do possível pai.

### 2.6.1. Método da razão de verossimilhança

Meagher (1986) desenvolveu o método da razão de verossimilhança para inferência de paternidade de dados de aloenzimas de lírios (*Chamaelirium luteum*). O macho assinalado como pai era aquele que possuía o maior LOD score. O termo LOD score definido por Meagher (1986) refere-se ao produto de razão de verossimilhança de vários *loci* marcadores tomados simultaneamente e ao logaritmo natural do valor obtido. Geralmente este valor é positivo se o pai alegado é mais provável de ser o pai verdadeiro, do que outro indivíduo tomado ao acaso na população.

A análise de verossimilhança toma os dados de genótipo da mãe, da progênie e do pai alegado ( $g_m$ ,  $g_p$  e  $g_a$ , respectivamente), e avaliam a probabilidade da hipótese principal (H1=pai alegado é o verdadeiro pai), relativa à hipótese alternativa (H2=pai alegado é um indivíduo ao acaso). A razão de verossimilhança para cada *locus* quando o genótipo da mãe é conhecido (Meagher, 1986) é dada por:

$$L(H1, H2 | g_m, g_p, g_a) = \frac{T(g_p | g_m, g_a)}{T(g_p | g_m)},$$

em que  $T(g_p | g_m, g_a)$ , é a probabilidade do genótipo da progênie, dado o genótipo materno e do pai alegado, na segregação Mendeliana.

Nos casos em que o genótipo materno é desconhecido, a razão de verossimilhança é dada por:

$$L(H1, H2 | g_p, g_a) = \frac{T(g_p, g_a)}{P(g_p)},$$

em que  $P(g_p)$  é a frequência do genótipo da progênie.

Este método foi aplicado por Marshall et al. (1998) no programa CERVUS para inferência de paternidade em populações naturais. Para acessar a confiança de tal paternidade, Marshall et al. (1998) propuseram um teste  $\Delta$ , que é a diferença entre os logaritmos naturais das razões de verossimilhança (LOD score) dos pais mais prováveis, admitindo uma distribuição de  $\Delta$  através de simulação. O valor de  $\Delta$  é definido como  $\text{LOD}_1 - \text{LOD}_2$ , se  $n \geq 2$ ;  $\text{LOD}_1$  se  $n = 1$  e zero se  $n = 0$ , sendo  $n$  o número de machos candidatos com LOD score maior do que zero.

Os valores críticos  $\Delta$  para um intervalo de confiança de 95% (restrito) e 80% (relaxado) são estimados pela sobreposição das distribuições simuladas com base nos dados fornecidos. Alguns parâmetros devem ser especificados para tornar a simulação mais realista. São eles: número de machos candidatos, proporção de machos candidatos amostrados, proporção de *loci* tipados e taxa de erro.

Segundo Marshall et al. (1998) os dois maiores fatores que influenciam o sucesso da inferência de paternidade são a disponibilidade de dados maternos e a confiança requerida. A vantagem de genotipar a mãe está associada a um grande prêmio líquido, pois seriam necessários 50% mais *loci* para obter uma taxa de sucesso (casos resolvidos) com similar nível de confiança. Os resultados mostraram também que existe claramente uma tendência contrária entre o número de paternidades resolvidas e a confiança estabelecida.

Este método prevê que mais de um pai alegado possa não ser excluído pelo teste, e mesmo assim ser possível distinguir estatisticamente qual o verdadeiro pai. Ele oferece as seguintes vantagens sobre o método de exclusão: múltiplos machos não-excluídos podem ser estatisticamente distinguidos, machos não amostrados podem ser assinalados como pai, é robusto aos erros de tipagem de laboratório, permite que haja *loci* não tipados em alguns indivíduos, a confiança estatística do teste é calculada por meio de simulação e pode se adequar às condições dos dados disponíveis. O método é categórico (sim/não) e uma proporção de paternidades permanece não resolvida.

Outra grande vantagem da inferência de paternidade sobre o método de exclusão é que as ocorrências de mutações e alelos nulos (ou não-amplificados) não levam a falsas exclusões de paternidade, pois eles são incorporados aos erros de tipagem (Marshall et al., 1998).

### 2.6.2. Método fracional (Bayesiano)

Em estudos de ecologia, a análise de paternidade é usada para estimação ou detecção do fluxo gênico entre populações, ou para a análise do sucesso reprodutivo de diferentes grupos. Para este tipo de estudo, não é necessário um resultado categórico e o método conhecido como fracional foi desenvolvido (Devlin et al., 1988; Nielsen et al. 2001) e está disponível no programa *PATRI-Paternity Inference Using Genetic Data* (Signorovitch e Nielsen, 2002). Neste procedimento, a estrutura de dados proporciona qual a probabilidade de cada pai a um par mãe-filho, como demonstrado na Tabela 3.

A probabilidade de paternidade *a posteriori* ( $P_{ik}$ ) é estimada pela ponderação da contribuição reprodutiva de cada pai candidato com a probabilidade de paternidade relativa aos outros candidatos.

Tabela 3 – Estrutura de dados para análises fracional de paternidade

Mãe e Filho	Pai 1	Pai 2	-	Pai K	Total
MF <sub>1</sub>	P <sub>11</sub>	P <sub>12</sub>	-	P <sub>1K</sub>	1
MF <sub>2</sub>	P <sub>21</sub>	P <sub>22</sub>	-	P <sub>2K</sub>	1
-	-	-	-	-	-
MF <sub>N</sub>	P <sub>N1</sub>	P <sub>N2</sub>	-	P <sub>NK</sub>	1
Média	$\lambda_1$	$\lambda_2$	-	$\lambda_K$	1

Dado um conjunto de dados de progênie ( $F_i: i = 1...N$ ), filhos das mães ( $M_i: i = 1...N$ ) e um conjunto de pais candidatos ( $Pai_k: k = 1...K$ ),

$$X_{ik} = Pr(\text{genótipo } F_i | \text{genótipo } M_i \text{ e } Pai_k);$$

$\lambda_k = N \cdot K / N$ , que corresponde a contribuição reprodutiva; e

$P_{ik} = \frac{\lambda_k X_{ik}}{\Delta_i}$ , onde  $\Delta_i$  é o  $\sum \sum \lambda_k X_{ik}$  para todos outros candidatos.

Esta aproximação pode ser considerada um procedimento Bayesiano, na qual a todos os pais é dada uma igual probabilidade de paternidade *a priori* (1/N), um valor subjetivo da chance de paternidade anterior ao teste, dado essencial para o cálculo de acordo com as regras do teorema.

Não existem casos não solucionados na análise – a toda a progênie é assinalada paternidade, múltiplos pais podem ser indicados para uma progênie. Este método é robusto a estrutura de família que possa existir dentro da população, permite a existência de machos não amostrados e múltiplos machos não-excluídos. Este método é considerado o melhor para quando existem poucas informações da população.

## 2.7. Estudo de variabilidade genética utilizando marcadores de microsatélites

Segundo Kumar (2000), os marcadores de microsatélites apresentam um alto grau de polimorfismo, são bem documentados em algumas espécies (bovinos, suínos, ovinos, eqüinos, aves e peixes), podendo ser usados para a caracterização de rebanhos, além de estabelecer as relações filogenéticas entre as diversas raças. Ainda segundo o autor, a FAO recomenda o uso de, no mínimo, 25 *loci* de microsatélites para estudo de distância genética, utilizando os seguintes critérios para a escolha de *loci* de microsatélites apropriados:

1) Marcador de domínio público, previamente reportado em publicações científicas.

2) Não-ligados.

3) Exibir herança Mendeliana.

4) Cada *microsatélite* escolhido deve conter no mínimo quatro alelos.

5) *Microsatélite* que possui homologia em várias espécies relacionadas são preferíveis (ex. bovinos, ovinos e caprinos).

6) Possível de ser usado em qualquer seqüenciador automático, em multiplex e de fácil reprodutibilidade.

A diversidade genética pode ser definida em diferentes níveis. Observando o fenótipo dos animais, uma grande parte da diversidade é oculta ou encoberta, pois a diversidade total é composta de elementos geneticamente expressos e não-expressos ou neutros. A diversidade genética, por definição, é aquela contida nos cromossomos de cada célula (Eding e Laval, 1999).

A variabilidade entre populações pode ser acessada através de ferramentas matemáticas, que traduzem as diferenças em medidas de distância entre um par de populações. Dada a distinção entre diversidade genética e fenotípica, a distância também pode ser designada por genética ou fenotípica, dependendo da metodologia adotada na observação da variabilidade. Segundo a definição da FAO (1998), com relação a diversidade genética, o maior número de alelos deve ser preservado, o que implica num esforço de preservar a diversidade total.

Segundo Eding e Laval (1999), a razão para este esforço de conservação na Europa é manter a flexibilidade para atender à mudanças de demanda do mercado e do sistema produtivo. Um parâmetro importante para as estratégias de preservação de recursos genéticos é a singularidade genética da raça, no sentido de não diferente a raça é de todas as outras. Deste ponto de vista, é importante computar as distâncias genéticas entre os pares de raças e construir árvores filogenéticas visando escolher populações que cobrem a maior proporção de variabilidade genética, buscando-se assim otimizar os recursos e preservar o máximo de variabilidade genética (Nei e Takezaki, 1994). Para medir a singularidade, deve-se estudar a diversidade do conjunto de raças, o que pode ser matematicamente traduzido em medidas de distância genética, como será discutido a seguir.

Os métodos para cálculo de distância genética utilizam a diferença das frequências alélicas em diferentes populações. Em princípio, qualquer seqüência de DNA que demonstre polimorfismo pode ser utilizada. Os marcadores genéticos são extremamente importantes para a caracterização das raças/tipos,

pois não são afetados pelo ambiente, como as características fenotípicas. Dentre os marcadores utilizados para estudos de dissimilaridade entre populações, têm-se o tipo sangüíneo e as aloenzimas – ambos marcadores protéicos – e os marcadores de DNA do tipo STR (*single tandem repeat*), RFLP (*restriction fragment length polymorphism*) e RAPD (*random amplified length polymorphism*).

Takezaki e Nei (1996) ressaltaram a importância dos microssatélites de DNA para evidenciar relações de evolução de populações muito próximas. Porém, alguns problemas surgem com o uso deste marcador, como a irregularidade do padrão de mutação nestes *loci* e a diferença de polimorfismo entre populações, suscitando cautela na extrapolação de resultados para populações distintas das relacionadas.

### **2.7.1. Distância genética**

As distâncias genéticas quantificam o grau de diferenciação entre dois grupos populacionais. Sob determinada ótica, podem ser consideradas como mecanismos para reduzir informação, uma vez que transformam todas as informações disponíveis sobre a relação entre duas populações em um único número (Weir, 1996). Se não há diferença, então a distância entre as duas populações é zero. Contudo, se as populações não possuem alelos em comum em todos os *loci*, então a distância deve possuir valor máximo.

A medida de diversidade genética é usualmente feita pela estimativa de distância genética. Na teoria de genética de populações existem quatro forças evolutivas que ocasionam a diferença genética entre as populações: a deriva genética, a seleção, a mutação e a migração. A diversidade genética observada é determinada principalmente pela deriva genética e mutação, uma vez que a migração é ignorada após a separação das populações e assume-se que os *loci* em estudo são seletivamente neutros. Em estudos envolvendo um período relativamente curto de tempo (divergência de raças), então o efeito de mutação pode ser ignorado e assume-se o modelo evolutivo de deriva genética (Reynolds

et al., 1983). Porém para os estudos de divergência genética envolvendo períodos mais longos (relação entre espécies) a mutação não deve ser excluída e adota-se o modelo de evolução deriva-mutação.

As distâncias evolutivas, mais usualmente usadas, são formuladas para estimarem valores com significado evolutivo. As duas fórmulas mais conhecidas são:

a) distância de Nei, que estima o número de substituições gênicas que ocorrem em genes de duas populações desde sua divergência a partir de um ancestral comum; e

b) distância de Reynolds, que estima o tempo decorrido desde a separação, admitindo que as populações se diferenciam devido à deriva genética.

### **- Propriedades matemáticas da distância**

Uma função deve ter algumas propriedades para ser considerada como distância. Primeiro, a distância de uma população X e ela mesma deve ser zero e a distância entre dois pontos X e Y deve ser positiva

$$d(X, Y) \geq 0, d(X, X) = 0.$$

Segundo, a distância entre duas populações X e Y deve ser simétrica

$$d(X, Y) = d(Y, X).$$

Se a distância satisfaz estas duas condições, ela é chamada de distância semi-métrica. Se a distância também satisfaz à igualdade triangular

$$d(X, Y) \leq [d(X, Z) + d(Y, Z)],$$

então ela é chamada de distância métrica. A distância geométrica ou Euclidiana satisfaz os três axiomas e pode ser definida em termos das coordenadas  $x_i$  e  $y_i$  dos pontos X e Y no espaço n-dimensional como

$$d(X,Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \text{ ou na forma matricial, } d(X,Y) = \sqrt{(x-y)'(x-y)}$$

As distâncias geométricas incluem uma série de métodos que resumem dados genéticos como uma função do ângulo entre os vetores de frequência gênica das populações comparadas. Devido à sua formulação em relações trigonométricas, as distâncias geométricas são métricas, o que implica que seus valores podem ser compatíveis com valores de evolução. A distância geométrica é medida pelo ângulo formado entre a junção da origem com os pontos representando as populações (Weir, 1996).

É importante notar que as distâncias de Nei e Reynolds, descritas a seguir, não satisfazem a desigualdade triangular, sendo, portanto, semimétricas. Segundo Nei e Takezaki (1994), a **metricidade** é necessária para representar populações como pontos no espaço Euclidiano e medir as distâncias em termos geométricos. Porém, ainda segundo o autor, em estudos evolutivos ou na construção de árvores filogenéticas, a metricidade não é necessária.

Outra característica importante acerca das distâncias genéticas é a **linearidade** com o tempo, pois a grande relevância do estudo de distância genética para estudos de evolução é a capacidade destas fornecerem estimativas do tempo de separação entre duas populações. Na ausência de linearidade, a distância genética deixa de dar informações sobre a profundidade temporal - conseqüentemente, sendo de pouca utilidade para os estudos evolutivos. A escolha dos marcadores a serem utilizados para estimativas de frequência gênica é de grande importância para obter linearidade, assim como a medida de distância escolhida para o estudo.

O número de mutações que ocorrem entre duas seqüências de DNA equivale a duas vezes a taxa de mutação, vezes o tempo de divergência do ancestral comum ( $2 mt$ ). Desta forma, para que o número inferido de mutações seja proporcional ao tempo, é necessário que **m** seja constante em todas as linhagens. De outra forma não é possível utilizar o número de mutações para inferência de tempo. As fórmulas de distância genética devem estimar a quantia 2

$mt$ , que corresponde ao número cumulativo de substituições que ocorreram desde a separação das duas populações (Eding e Laval, 1999).

### - Fórmulas para o cálculo de distâncias genéticas

Para duas populações  $X$  e  $Y$  das quais se dispõe de dados de frequência alélicas  $x$  e  $y$  para  $m$  loci, cada um com  $n$  alelos, Nei (1972) propôs as seguintes etapas:

Inicialmente se calcula a probabilidade de dois alelos escolhidos ao acaso, um para cada população, serem idênticos. Para um único locus, com  $n$  alelos, esse valor é dado por

$$j_{XY} = \sum_{i=1}^n x_i y_i$$

A probabilidade de alelos idênticos serem sorteados a partir de uma única população são dados para  $X$  e  $Y$  por

$$j_x = \sum_{i=1}^n x_i^2 \quad \text{e} \quad j_y = \sum_{i=1}^n y_i^2$$

A distância de Nei ( $D_A$ ) mede o número mínimo de diferenças por locus. A probabilidade dos alelos serem diferentes é  $1 - \sum(x_i y_i)$ . Ao se considerar os loci analisados, tem-se

$$D_A = 1 - \frac{1}{m} \sum_j \sum_i^{n_j} \sqrt{x_{ij} y_{ij}},$$

em que  $x_{ij}$  e  $y_{ij}$  são as frequências dos alelos no estado  $i$  no locus  $j$ .

Subtraindo a diferença intrapopulacional ( $d_x$  e  $d_y$ ), obtém-se a diferença mínima líquida por locus, que é a distância mínima de Nei ( $D_M$ )

$$D_M = d_{xy} - (d_x - d_y)/2 = 1 - \sum x_i y_i - [(1 - \sum x_i^2) + (1 - \sum y_i^2)]/2 = J_x + J_y / 2 - J_{xy}$$

Onde, em uma análise envolvendo vários *loci*, obtém-se  $J_x$ ,  $J_y$  e  $J_{xy}$  que são as médias aritméticas dos valores de  $j_x$ ,  $j_y$  e  $j_{xy}$ .

A identidade genética,  $I$  é definida entre duas populações  $X$  e  $Y$  como a proporção de alelos que são semelhantes entre as populações e dentro da população.

$$I = \left( \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2 \sum_i y_i^2}} \right)$$

Logo, a medida de distância padrão  $D_S$  de Nei (1972) é dada por

$$D_S = -\ln I$$

Os pressupostos necessários para as distâncias de Nei são:

- a) Os alelos que surgem por mutação são seletivamente neutros;
- b) Os tamanhos efetivos das populações ( $N_e$ ) são iguais;
- c) Ambas as populações são panmíticas (acasalamento ao acaso);
- d) Cada mutação dá origem a um novo alelo, distinto dos preexistentes (modelo infinito de alelos).

Tais pressupostos são dificilmente encontrados nas populações de animais domésticos estudados.

Os efeitos de mutação e deriva se equilibram no tempo, de modo que cada população mantém uma diversidade interna ( $J_x$  e  $J_y$ ) constante. Entretanto, o valor de  $J_{xy}$  varia com o tempo, pois as populações adquirem alelos distintos.

### **2.7.2. A distância na inferência filogenética**

Segundo Barker (1994), não há consenso na análise de dados para cálculo de distância genética entre raças em dados de frequência gênica nos animais

domésticos. A distância de Nei (1972) têm sido a mais utilizada nos estudos com populações naturais em genética evolutiva. Alternativamente, a distância baseada em coancestralidade (Reynolds et al., 1983) pode ser mais apropriada para evoluções a curto prazo, como a divergência entre as raças domésticas. Entretanto, ainda segundo Barker (1994), embora diferindo em suas propriedades matemáticas e bases biológicas, a correlação obtida entre as várias medidas de distância tem sido alta, quando aplicadas a populações dentro de espécies na mesma localidade.

Os principais aspectos que dificultam o uso de distância genética na inferência filogenética são a ocorrência de fluxo gênico, que não é prevista em nenhum dos modelos, a variação dos tamanhos efetivos das populações e variações de taxas de mutação. Ou seja, o ponto fraco dos estudos de distância genética é sua dependência de modelos evolutivos que provavelmente são discrepantes dos processos que de fato operam na natureza (Eding e Laval, 1999).

Analisando os pressupostos, tem-se que a distância de Nei parece ser mais adequada para separação de espécies do que de populações, onde a maior parte da variação de frequências gênicas resulta de deriva, quando se considera o modelo de alelos infinitos (*infinite allele model* -IAM).

Recentemente os *loci* de *microsatélites* de DNA estão sendo freqüentemente utilizados em análises filogenéticas. Estes *loci*, contudo, estão sujeitos a um padrão de mutação que se aproxima do modelo de mutação por passo (*stepwise mutation model*-SMM), diferentemente dos marcadores clássicos (bioquímicos), que seguiriam o modelo de alelos infinitos (IAM). O SMM é um modelo de mutação no qual um alelo com  $i$  repetições muda para  $(i - 1)$  ou  $(i + 1)$  repetições com a probabilidade  $\frac{\mu}{2}$  (Takezaki e Nei,1996).

Vários outros trabalhos têm buscado resolver o problema de linearidade de distância. No estudo de simulação de Nei e Takezaki (1994), foi verificado que  $D_A$  é a medida de distância mais eficiente para obter a topologia da árvore evolutiva, enquanto  $D_S$  e  $(\delta\mu)^2$  são melhores para estimar o tamanho dos ramos evolutivos, pois apresentam maior linearidade com o tempo.

Felsenstein (1984) argumenta que as distâncias genéticas perdem o sentido de distância simplesmente se forem vistas como esperanças estatísticas de quantidades de mudança, ao invés de valores de mudança real. Isto, por sua vez, justifica o critério de ajuste entre a árvore inferida e a distância observada. Contudo, do ponto de vista estatístico, a fraqueza dos métodos de inferência que utilizam a distância está na falta de aditividade e de independência dos erros amostrais das medidas.

### **- Filogenia e árvores evolutivas**

Árvores são representações gráficas ou mapas da matriz de distância. Sob certas suposições, esta representação pode ser tomada por filogenia, em estudos evolucionários. Existem vários métodos de inferência para a construção de árvores, que podem ser agrupados basicamente em três classes:

- a) Métodos de matriz de distância, que constroem o dendograma a partir de uma matriz de distância;
- b) Métodos da Parcimônia, que analisam distribuições de caracteres e buscam topologia que pode ser explicada por um número mínimo de “passos”;
- c) Métodos da Máxima Verossimilhança, que buscam a topologia com maior probabilidade de ter gerado os dados observados, admitindo determinado modelo.

Os métodos que utilizam as matrizes de distância são mais usados em estudos que buscam simplesmente identificar similaridade entre populações de uma mesma espécie. O Método da Parcimônia e o da Máxima Verossimilhança utilizam caracteres discretos (matriz binária de dados), onde a árvore é construída considerando as relações entre as seqüências de DNA em cada sítio de nucleotídeo (Nei e Takezaki, 1994). Tais métodos são mais complexos e, portanto, utilizados em trabalhos que visam estudos evolutivos propriamente ditos.

## - Análises baseadas em matrizes de distância (UPGMA)

Este método se baseia no agrupamento sucessivo de pares de táxons de acordo com o nível de similaridade. O método de “ligação média” ou UPGMA (*Unweighted Pair Group Method With Arithmetic Mean*) define a distância entre agrupamentos como a distância média entre os pares. O algoritmo opera da seguinte forma:

1) O par de táxon mais semelhante (com menor distância) é unido e um nó é posicionado exatamente na metade da distância que separa os dois;

2) A distância entre o par de táxons unidos e os demais é calculado pela média aritmética das distâncias entre os dois táxons unidos e o táxon ao qual ele está sendo comparado. Novamente a menor distância na nova matriz determina quais táxons devem ser unidos e um novo nó é posicionado como anteriormente descrito.

3) Esse procedimento é repetido até que todos os táxons tenham sido incorporados ao dendograma.

Esse método gera árvores com simetria dos ramos ao redor de um nó (árvore ultramétrica) e a condição biológica que gera tal simetria é a ocorrência de taxas evolutivas constantes. O método UPGMA não permite obter topologia final que não seja ultramétrica e seu emprego em situações onde é sabido que a taxa de evolução varia entre as linhagens poderá resultar em uma inferência incorreta.

O algoritmo UPGMA é um método de agrupamento hierárquico, onde não há uma preocupação com o número ótimo de grupos formados. Dentre os métodos hierárquicos, pode-se citar, além do método de ligação média (*average linkage*), ponderado ou não, o método do Encadeamento Simples (*single linkage*); o método do Vizinho mais distante ou Encadeamento completo (*complete linkage*) e o método centróide, também ponderado ou não (Khattree e Naik, 2000).

Em todos os métodos o princípio é o mesmo, sendo que o critério de agrupamento (h) é modificado de acordo com o método utilizado. No método de

ligação média (UPGMA), conforme já visto, a distância ( $h$ ) entre dois agrupamentos é definida como a média das distâncias computadas entre  $x_i$  e  $y_i$ , pertencentes a  $X$  e  $Y$ , respectivamente. Ou seja

$$h(X, Y) = \frac{1}{n_x n_y} \sum_x \sum_y d(x_i, y_i).$$

### 2.7.3. Variabilidade genética dentro e entre populações

A variação existente pode ser quantificada pela estatística  $F$  de Wright (1951), que introduziu os parâmetros  $F_{ST}$ ,  $F_{IT}$  e  $F_{IS}$ . Estes parâmetros fornecem uma maneira de sumarizar a estrutura da população (Weir e Cockerham, 1984). Os três parâmetros se relacionam por

$$F_{IS} = (F_{IT} - F_{ST}) / (1 - F_{ST}),$$

sendo  $F_{IT}$ , a correlação de genes dentro de indivíduos (consangüinidade global),  $F_{ST}$ , a correlação de genes entre indivíduos dentro da população (coancestralidade) e  $F_{IS}$ , a correlação de genes dentro de indivíduos dentro da população (consangüinidade dentro de população). Segundo Weir (1996),  $F_{ST}$  quantifica o grau de divergência entre populações. Frequências alélicas iguais em todas as populações irão causar  $F_{ST} = 0$ .

Os parâmetros da estatística  $F$  não são afetados pela amostragem – número de alelos por *locus*, número de indivíduos amostrados por população ou número de populações amostradas. Eles dependem, contudo, da amostragem genética, ou seja, do tamanho da população e de sua história. Na teoria da genética de populações, a formação das novas gerações irá diferir devido à amostragem genética, mesmo quando mantidas nas mesmas condições.

#### **2.7.4. O *Bootstrapping***

O *Bootstrap* é uma técnica estatística de reamostragem que permite um ganho de informação acerca dos parâmetros de distribuição em um conjunto de dados. Seu uso principal é para avaliar a magnitude dos erros amostrais sobre inferências (Weir, 1996). Em sua implementação, a matriz de dados original é usada para gerar novas matrizes, gerando uma maior robustez. Quando o conjunto de dados consiste de  $n$  observações, o *bootstrap* é uma amostra de  $n$  sorteios aleatórios. Todas as observações do conjunto original possuem igual chance de serem sorteadas. O valor gerado do novo *bootstrap* é a média do parâmetro amostrado. Este processo é repetido um grande número de vezes ( $\geq 1.000$  repetições), para criar um grande número de “novos” dados. Ele pode ser usado para estimar médias e desvios padrões de um estimador calculado a partir dos dados em questão. No caso da construção de árvores, o *bootstrap* pode ser usado nos dados de frequência alélica dos múltiplos *loci* amostrados. Assim, os dados obtidos pelo *bootstrap* simulam os diferentes conjuntos de dados que poderiam ser amostrados se mais dados tivessem sido coletados.

O número obtido pelo *bootstrap* é freqüentemente usado como uma estimativa de confiabilidade da inferência, uma vez que busca quantificar até que ponto os grupos inferidos a partir da matriz de dados são consequência de artefatos amostrais. Entretanto, altos valores de *bootstrap* podem ser obtidos para inferências incorretas, nas quais os dados disponíveis e o método empregado levam todas as amostragens a inferirem árvores erradas sistematicamente (Weir, 1996).

#### **2.8. Recursos genéticos no Brasil**

Recentemente a Embrapa Cenargen tem se dedicado a caracterizar geneticamente as raças envolvidas no projeto de preservação, através do Laboratório de Genética Animal (LGA). São objetivos do LGA, além da caracterização das raças, a formação de um banco de DNA, que em conjunto com

o Banco de Germoplasma (BGA), seria responsável pelo armazenamento de material genético de animais autóctones (Egito et al., 2001). Dos trabalhos realizados pelo LGA, apresentaram resultados na literatura a caracterização de bovinos nativos do Brasil (Serrano et al., 2000) e bubalinos (Albuquerque et al., 2000).

O Instituto de Zootecnia (IZ-SP), em parceria com a Embrapa, tem desenvolvido a caracterização racial de bovinos, ovinos e bubalinos por meio de polimorfismo protéico, obtendo dados de distância genética, heterozigosidade e variabilidade genética (Lara et al., 1998; Lara et al., 1999).

As ações de pesquisa com conservação *in situ* de caprinos estão situadas na Embrapa Caprinos (raça Moxotó), na Embrapa Meio-Norte (raça Marota), na Empresa de Pesquisa Agropecuária da Bahia (raça Repartida), Empresa de Pesquisa Agropecuária da Paraíba (raça Canindé) e Universidade Federal de Pernambuco (raça Moxotó).

Todos os projetos estão veiculados ao projeto de “Conservação e utilização de Recursos Genéticos” (1983), liderado pela Embrapa Cenargen e que têm um convênio internacional com a FAO (*Food and Agricultural Organization*). Seus principais objetivos são (a) identificação de populações em estado avançado de erosão genética; (b) caracterização fenotípica e genética do germoplasma; (c) avaliação do potencial produtivo destes rebanhos nativos (Mariante et al., 2000).

Na Espanha, onde existe uma forte tendência de projetos de preservação de raças nacionais, faz-se um esforço para a caracterização e valorização de rebanhos autóctones (Delgado et al., 1999). Segundo os autores, a caracterização racial por meio de padrões genéticos, leva ao reconhecimento oficial das raças e subsidia uma base para o trabalho das associações específicas de raças, levando a abertura de livros genealógicos. Na Colômbia, considerado um dos países com maior diversidade biológica, existe um programa para caracterização de raças nativas bovinas, ovinas e suínas (Olivera-Angel et al., 1999).

No capítulo de Gandini e Oldenbroek (1999) estão colocadas algumas questões interessantes sobre a escolha da estratégia de conservação. Segundo os autores, até recentemente, as técnicas *ex situ* eram consideradas como a estratégia

mais indicada, devido principalmente aos baixos custos de manutenção quando comparado com os sistemas *in situ*. Hoje, a conservação *in situ*, que mantém a população dentro de seus sistemas de produção, é prioridade para a FAO. Entretanto deve se considerar a integração efetiva entre as técnicas, na qual a conservação *ex situ* é complementar à *in situ*.

Para a escolha da estratégia de conservação, os autores indicam cinco itens a serem analisados: 1) risco de perda da raça em longo prazo, considerando o tamanho da população e a manutenção socio-econômica da área por ela habitada; 2) custos; 3) oportunidade de evolução/adaptação da raça e melhor conhecimento da raça; 4) eficiência de atingir os objetivos da conservação; e 5) eficiência na conservação de valores culturais e ecológicos. Na ponderação dos itens, devem ser considerados os recursos disponíveis, interesses e prioridades do projeto.

Dentro da técnica *ex situ*, pode-se adotar estratégia de criopreservação de sêmen ou embriões, ou banco de células somáticas ou DNA. O sêmen é o material genético eleito para a criação de rebanhos sintéticos, introgressão gênica e suporte ao manejo genético de programas *in situ*.

Segundo Serrano et al. (2000), até o momento a caracterização de rebanhos nativos no Brasil foi feita exclusivamente baseada em características morfológicas e produtivas, as quais são muito influenciadas pelo ambiente. Inclusive, alguns rebanhos nativos, com diferentes nomes e habitando regiões diferentes, possuem fenótipo semelhante, se questionado pertencerem ou não ao mesmo grupo racial ou ecótipo. Além disto, os resultados advindos de pesquisas com marcadores moleculares podem ajudar a selecionar animais para comporem o banco de germoplasma no núcleo *ex situ*, de modo a aproveitar ao máximo o recurso, explorando a variabilidade e evitando as duplicatas de amostras.

## **2.9. Diversidade genética com o uso de microssatélites de DNA**

Os estudos de diversidade com caprinos no Brasil são ainda muito restritos. Machado et al. (2000) utilizaram caracteres morfológicos no estudo de diversidade de raças SRD e mediterrâneas. Igarashi et al. (2000a) estudaram

diversidade em caprinos de raças naturalizadas e importadas através de polimorfismo de proteínas e Igarashi et al (2000b) estudaram diversidade também em rebanhos importados e naturalizados, utilizando dois *loci* de microssatélites (INRA011 e INRA040) em adição a seis proteínas de eritrócitos, obtendo dados de distância e dendogramas semelhantes.

Na literatura mundial existe um considerável número de artigos com dados de diversidade em caprinos. Yang et al. (1999) estudaram diversidade em caprinos de raças autóctones da China por meio de marcadores de microssatélites, obtendo resultados em concordância com o histórico das populações. Saitbekova et al. (1999) verificaram a diversidade entre rebanhos suíços domésticos e cabras Ibex e Bezoar. As raças domésticas apresentaram maior heterozigosidade que os tipos selvagens. Barker et al. (2001) estudaram a divergência genética entre cabras asiáticas nativas utilizando polimorfismos bioquímicos e microssatélites, revelando o isolamento geográfico como maior efeito de diferenciação entre populações. Maudet et al. (2002) estudaram várias metodologias estatísticas no manejo de conservação de biodiversidade, utilizando *Capra ibex*. Foi demonstrada a baixa variação genética existente nesta espécie de cabra selvagem dos Alpes. Os microssatélites possibilitaram a identificação de gargalos populacionais, a simulação de esquemas de migração e a identificação precisa de populações.

Com o estudo de polimorfismo de microssatélites de DNA, Moioli et al. (2001) e Van Hooft et al. (2000) obtiveram a diversidade genética em búfalos africanos (*S. coffer*) de diferentes países africanos e mediterrâneos. Bonnet et al. (2002) estudaram um sistema de microssatélite para estudo de diversidade em veados selvagens e criados em fazendas, propondo um painel com onze *loci* para estudos posteriores de diversidade. Em suínos, Laval et al. (2000) estudaram diversidade de 11 raças européias, constatando um grande índice de fixação dentro de raça ( $F_{ST}$ ). Diez-Tascón et al. (2000) estudaram diversidade de ovinos Merina em rebanhos próximos (com menos de 400 anos de separação), mostrando ser possível separar populações próximas através de estudos com microssatélites.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABI PRISM™GeneScan® *Reference Guide*. Perkin Elmer, Applied Biosystems. 2000.

ALBUQUERQUE, M.M. S.; MARQUES, J. R. F.; GASPAROTTO, C. R.; MARTÍNEZ, A. N.; EGUIO, A. A. Preliminary evaluation of genetic distance among four buffalo populations conserved on Marajó Island. In: GLOBAL CONFERENCE ON CONSERVATION OF DOMESTIC ANIMAL GENETIC RESOURCES, 5., 2000, Brasília. *Proceedings...* Ed. Mariante, A.S., McManus, C., Salomão, A.N. Brasília: Embrapa. 2000. CD-Rom.

ALFORD, R. L.; HAMMOND H. A.; COTO, I.; CASKEY, C. T. Rapid and efficient resolution of parentage by amplification of short tandem repeats. *American Journal of Human Genetics*, v. 55, p. 190-195, 1994.

AREVALO, E.; HOLDER, D. S.; DERR, J. N. et al. Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP 1, 2, 3, 4 and 5 loci. *Animal Genetics*, v. 25, p. 202, 1994.

BARKER, J. S. F. A global protocol for determining genetic distances among domestic livestock breeds. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5. ed., v. 21. *Proceedings...* Guelph, 1994. p. 501-508.

BARKER, J. S. F.; TAN, S. G.; MOORE, S. S.; MUKHERJEE, T. K.; MATHESON, J-L., SELVARAJ, O. S. Genetic variation within and relationship among populations of Asian goats (*Capra hircus*). *Journal of Animal Breeding Genetics*, v. 118, p. 213-233, 2001.

BISHOP, M. D.; KAPPES, S. M.; KEELE, J. W. et al. A genetic linkage map for cattle. *Genetics*, v. 136, p. 619-639, 1994.

BOLSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics*, v. 32, p. 314-331, 1980.

BONNET, A.; THÉVENON, S.; MAUDET, F.; MAILLARD, J. C. Efficiency of semi-automated multiplex PCRs with 11 microsatellite markers for genetic studies of deer populations. *Animal Genetics*, v. 33, p. 343-350, 2002.

BOWLING, A. T. Historical development and application of molecular genetic test for horse identification and parentage control. *Livestock Production Science*, v. 72, p. 111-116, 2001.

BUROW, M. D.; BLAKE, T. K. Molecular tools for the study of complex traits. In: PATERSON, A. H. (Ed.) *Molecular dissection of complex traits*, p. 13-29. New York: CRC Press, 1998.

CRAWFORD, A. M.; BUCHANAN, F. C.; SWARBRICK, P. A. Ovine dinucleotide repeat polymorphism at the MAF18 locus. *Animal Genetics*, v. 21, p. 433-434, 1990.

CRAWFORD, A. M.; DODDS, K. G.; PIERSON, C. A. et al. An autosomal genetic linkage map of the sheep genome. *Genetics*, v. 140, p. 703-724, 1995.

CURI, R. A.; LOPES, C. R. Evaluation of nine microsatellite loci and misidentification paternity frequency in a population of Gyr breed bovines. *Brazilian Journal of Veterinary Research and Animal Science*, v. 39, p. 129-135, 2002.

DA, Y.; LEWIN H. A. Linkage information content and efficiency of full-sib and half-sib designs for gene mapping. *Theoretical and Applied Genetics*, v. 90, p. 699-706, 1995.

DEGARTORI, M. J.; FREKING, B. A.; CUTHBERTSON, R. P.; KAPPES, S. M.; KEELE, J. N.; LEYMASTER, K. A.; DODDS, K. G.; CRAWFORD, A. M.; BEATTIE, C. W. A second-generation linkage map of the sheep genome. *Mammalian Genome*, v. 9, p. 204-209, 1998.

DELGADO, J. V.; BRABA, C.; CAMACHO, M. E.; SERENO, F. T. P. S.; MARTÍNEZ, A. Caracterización genética de los animales domésticos en España. In: SIMPÓSIO DE RECURSOS GENÉTICOS PARA AMÉRICA LATINA E CARIBE, 2., 1999, Brasília. *Anais...* Ed. Mariante, A.S., Bustamante, P.G. Brasília: Embrapa. 1999. CD-Rom.

DEVLIN, B.; ROEDER, K.; ELLSTRAND, N. C. Fractional paternity assignment: theoretical development and comparison to other methods. *Theoretical and Applied Genetics*, v. 76, p. 369-380, 1988.

DIEZ-TASCÓN, C.; LITTLEJOHN, R. P.; ALMEIDA, P. A. R.; CRAWFORD, A. M. Genetic variation within the Merino sheep breed: analysis of closely related populations using microsatellites. *Animal Genetics*, v. 31, p. 243-251, 2000.

EDING, J. H.; LAVAL, G. Measuring genetic uniqueness in livestock. In: *Genebanks and the conservation of farm animal genetic resources*. Lelystad: Ed. J.K. Oldenbroek. ID.DLO, 1999. p. 33-58.

EGITO, A. A.; ALBUQUERQUE, M. S.; MARIANTE, A. S.; MCMANUS, C.; MARQUES, J. R. F.; ABREU, U. P. G.; SERRANO, G. M. S. Caracterização genética de raças naturalizadas. In: SIMPÓSIO DE RECURSOS GENÉTICOS PARA AMÉRICA LATINA E CARIBE, 3., 2001, Londrina. *Anais...* Ed. Rilde, C.R.; Cirino, V.M. Londrina: IAPAR. p.121-126, 2001.

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

FALCONER, D. S. *Introduction to quantitative genetics*. 3. ed. New York: Longman Scientific & Technical, 1989.

FAO. *Secondary Guidelines for development of national farm animal genetic resources management plans: Management of small populations at site*. Rome: FAO, 1998. 215 p.

FELSENSTEIN, J. Distance methods for inferring phylogenies: a justification. *Evolution*, v. 38, p. 16-24, 1984.

GANDINI, G. C.; OLDENBROEK, J. K. Choosing the conservation strategy. In: *Genebanks and the conservation of farm animal genetic resources*. Lelystad: Ed. J.K. Oldenbroek. ID.DLO, 1999. p. 11-31.

GLAWATZKI-MULLIS, M-L.; GAILLARD, C.; WIGGER, G.; FRIES, R. Microsatellite-based parentage control in cattle. *Animal Genetics*, v. 26, p. 7-12, 1995.

GOLDSTEIN, D. B.; LINARES, A. R.; CAVALLI-SFORZA, L. L.; FELDMAN, M. W. An evolution of genetic distances for use with microsatellite loci. *Genetics*, v. 139, p. 463-471, 1995.

GROSCLAUDE, F.; TUCKER, E. M. Blood and milk protein polymorphism in goats. In: INTERNATIONAL CONFERENCE ON GOATS, 5. *Proceedings...* New Delhi: 1992. p. 71-84.

HAYES, H.; PETIT, E.; DUTRILLAUX, B. Comparison of RGB-banded karyotypes of cattle, sheep and goats. *Cytogenet Cell Genet.*, v. 57, p. 51-55, 1991.

HEYEN, D. W.; BEEVER, J. E.; DA, Y., EVERET, R.E.; GREEN, C.; BATES, S. R. E.; ZIEGLE, J. S. Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semi-automated parentage testing. *Animal Genetics*, v. 28, p. 21-27, 1997.

IGBE. *Censo Agropecuário de 2001*. Disponível em: <<http://www.ibge.org.br>> Acesso em: 20 jan. 2004.

IGARASHI, M. L. S. P.; MACHADO, T. M.; CASTRO, S. R.; FERRO, J. A.; CONTEL, E. P. B. Genetic characterization of goats herds of the Brazil. In: GLOBAL CONFERENCE ON CONSERVATION OF DOMESTIC ANIMAL GENETIC RESOURCES, 5., 2000, Brasília. *Proceedings...* Ed. MARIANTE, A. S.; McMANUS, C.; SALOMÃO, A. N. Brasília: Embrapa. 2000a. CD-Rom.

IGARASHI, M. L. S. P.; MACHADO, T. M.; FERRO, J. A.; CONTEL, E. P. B. Structure and genetic relationship among Brazilian naturalized and imported goat breeds. *Biochemical Genetics*, v. 38, p. 353-365, 2000b.

JAMIESON, A. The effectiveness of using co-dominant polymorphic allelic series for (1) checking pedigrees and (2) distinguishing full-sib pair members. *Animal Genetics*, v. 25, suppl. 1, p. 37-44, 1994.

JAMIELSON, A.; TAYLOR, ST. C. S. Comparisons of three probability formulae for parentage exclusion. *Animal Genetics*, v. 28, p. 397-400, 1997.

KHATTREE, R.; NAIK, D. Cluster Analysis. In: *Multivariate data reduction and discrimination. With SAS<sup>®</sup> Software*. Cary, NC: SAS Institute Inc., 2000. p.347-442.

KEMP, S. J.; HISLINDA, O. et al. A panel of polymorphic bovine, ovine and caprine microsatellite markers. *Animal Genetics*, v. 26, p. 299-306.1995.

KUMAR, D. DNA markers for the differentiation of farm animal breeds. In: SAHAI, R.; VIJH, R. K. (Ed.) *Domestic Animal Diversity: Conservation and sustainable development*. Karnal: SI Publications. 2000. P.305-312.

LARA, M. A. C.; CUNHA, E. A.; BUFARAH, G.; CONTEL, E. P. B. Situação da caracterização genética animal no Instituto de Zootecnia. In: SIMPÓSIO DE RECURSOS GENÉTICOS PARA AMÉRICA LATINA E CARIBE, 2., 1999, Brasília. *Anais...* Ed. MARIANTE, A. S.; BUSTAMENATE, P. G. Brasília: Embrapa. CD-Rom.

LARA, M. A. C.; HEICHERT, R. H.; FIGUEIREDO, L. A.; CONTEL, E. P. B. Caracterização genética de bovinos Caracu através de polimorfismos protéicos. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 35., 1998, Botucatu: SBZ. *Anais...* Ed. Moura, A. S. A. M. T.; WECHSLER, F. S.; LEVEZZO, W. 4 vol. 1998b. p.568-571.

LAVAL, G.; IANNUCELLI, N.; LEGAULT, C. et al. Genetic diversity of eleven European pig breeds. *Genetic, Selection, Evolution*, v. 32, p.187-203, 2000.

LEWIN, B. *Genes VII*. Oxford University Press, 2000. 952p.

LODISH, H.; BERK, A.; ZIPURSKY, S.L.; MATSUDAIRA, P.; BALTIMORE, D.; DARNELL, J. *Molecular Cell Biology*. W.H. Freeman and Company, 4. ed., 1999. 1084 p.

LUIKART, G.; BIJU-DUVAL, M-P.; ERTUGRUL, O.; ZAGDSUREN, Y.; MAUDET, C.; TABERLET, P. Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). *Animal Genetics*, v. 30, p. 431-438, 1999.

MACHADO, T. M. M.; CHAKIR, M.; LAUVERGNE, J. J. Genetic distances and taxonomic trees between goats of Ceará State (Brazil) and goats of the Mediterranean region (Europe and Africa). *Genetics and Molecular Biology*, v. 23, n. 1, p. 121-125, 2000.

MADDOX, J. F.; DAVIES, K. P.; CRAWFORD, A. M. et al. An enhanced linkage map of the sheep genome comprising more than 1000 loci. *Genome Research*, v. 11, p. 1275-1298, 2001.

MARIANTE, A. S.; CASTRO, S. T. R.; WETZEL, M. M. V. S. Conservation of animal genetic resources: structure of the Brazilian Network. In: GLOBAL CONFERENCE ON CONSERVATION OF DOMESTIC ANIMAL GENETIC RESOURCES, 5., 2000, Brasília. *Proceedings...* Ed. MARIANTE, A. S.; McMANUS, C.; SALOMÃO, A. N. Brasília: Embrapa. 2000. CD-Rom.

MARKLUND, S.; ELLEGREN, H.; ERIKSSON, S.; SANDBERG, K.; ANDERSSON, L. Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. *Animal Genetics*, v. 25, p. 19-23, 1994.

MARSHALL, T. C.; SLATE, J.; KRUK, L. E. B.; PEMBERTON, J. M. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, v. 7, p. 639-655, 1998.

- MAUDET, C.; MILLER, C.; BASSANO, B. et al. Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine Ibex (*Capra ibex ibex*). *Molecular Ecology*, v. 11, p. 421-436, 2002.
- MEAGHER, T. R. Analysis of paternity within a natural population of *Chamaelirium luteum*. I. Identification of most-likely male parent. *The American Naturalist*, v. 128, p. 199-215, 1986.
- MOIOLI, B.; GEORGOUDIS, A.; NAPOLITANO, F. et al. Genetic diversity between Italian, Greek and Egyptian buffalo populations. *Livestock Production Science*, v. 70, p. 203-211, 2001.
- NEI, M. Genetic distance between populations. *The American Naturalist*, v. 106, p. 283-292, 1972.
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, v. 89, p. 583-590, 1978.
- NEI, M.; TAKEZAKI, N. Estimation of genetic distances and phylogenetic trees from DNA analysis. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5<sup>th</sup>, v. 21. *Proceedings...* Guelph, 1994. p.405-412, 1994.
- NIELSEN, R.; MATTILA, D.K.; CLAPHAM, P.J.; PALSBOLL, P.J.S. Statistical approaches to paternity analysis in natural populations and application to North Atlantic Humpback Whale. *Genetics*, v. 157, p. 1673-1682, 2001.
- OLIVERA-ANGEL, M.; SIERRA, R.; BERDUGO, J. et al. Caracterización de recursos genéticos autóctonos en Colombia. In: SIMPÓSIO DE RECURSOS GENÉTICOS PARA AMÉRICA LATINA E CARIBE, 2., 1999, Brasília. *Anais...* Ed. MARIANTE, A.S.; BUSTAMANTE, P. G. Brasília: Embrapa. 1999. CD-Rom.
- OTT, J. Strategies for characterizing highly polymorphic markers in human gene mapping. *American Journal of Human Genetics*, v. 51, p. 283-290, 1992.
- PENNA, V. M.; MADALENA, F. E.; OLIVEIRA, D. A.; LIMA, R. M. G. Testes de paternidade: investimento em credibilidade. In: ANAIS DO SIMPÓSIO NACIONAL DE MELHORAMENTO ANIMAL, 2., 1998, Uberaba. *Anais...* Uberaba: UFV, 1998, p.35-40.
- PÉPIN, L.; AMIGUES, Y.; LÉPINGLE, A; BERTHIER, J-L.; BENSALID, A; VAIMAN, D. *Sequence conservation of microsatellites between cattle (B. taurus), goat (C. hircus), and related species. Examples of use in parentage testing and phylogeny analysis.* INRA. Laboratoire de Genetique biochimique et de cytogenetique. Jouy-en-Josas. 1995.

REYNOLDS, J.; WEIR, B. S.; COCKERHAM, C. C. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, v. 105, p. 767-779, 1983.

RON, M. ; BLANC, Y. ; BAND, M.; EZRA, E.; WELLER, J. I. Misidentification rate in Israeli dairy cattle population and its implications for genetic improvement. *Journal of Dairy Science*, v. 79, p. 676-681, 1996.

SAIKI, R. K.; GELFAND, D. H.; STOFFEL, S.; SCHARF, S.; HIGUCHI, R.; HORN, G. T.; MULLIS, K. B.; EHRlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, v. 239, p. 487-491, 1988.

SAIKI, R.K., SCHARF, S., FALOONA, F. et al. Enzymatic amplification of beta globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, v. 230, p. 1350-1354, 1985.

SAITBEKOVA, N.; GAILLARD, C.; OBEXER-RUFF, G.; DOLF, G. Genetic diversity in Swiss goat breeds based on microsatellite analysis. *Animal Genetics*, v. 30, p. 36-41, 1999.

SAMBROOK, J.; FRITSCH, E. F.; MANIATS, T. *Molecular cloning: a laboratory manual*. 2. ed. New York: Cold Spring Harbor Lab. Press. 1989.

SANTOS, R. *A cabra e a ovelha no Brasil*. Uberaba, MG: Ed. Agropecuária Tropical Ltda. 2003. 479p.

SCHIBLER, L.; VAIMAN, D.; OUSTRY, A.; GIRAUD-DELVILLE, C.; CRIBIU, E. P. Comparative gene mapping: a fine-scale survey of chromosome rearrangements between ruminants and humans. *Genome Research*, v. 8, p. 901-915, 1998.

SERRANO, G. M.; MCMANUS, C.; EGITO, A. A. Use of RAPD molecular markers in the genetic characterization of native brazilian cattle breeds. In: GLOBAL CONFERENCE ON CONSERVATION OF DOMESTIC ANIMAL GENETIC RESOURCES, 5., 2000, Brasília. *Proceedings...* Ed. MARIANTE, A. S.; McMANUS, C.; SALOMÃO, A. N. Brasília: Embrapa. 2000. CD-Rom.

SIGNOROVICH, J.; NIELSEN, R. PATRI-Paternity Inference Using Genetic Data. *Bioinformatics*, v. 18, p. 341-342, 2002.

SILVA, R. R. *Agrobuisiness da caprinocultura leiteira no Brasil*. Salvador: Bureau, 1998. 74p. il

SLATE, J.; MARSHALL, T.; PEMBERTON, J. A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology*, v. 9, p. 801-808, 2000.

TAKEZAKI, N.; NEI, M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, v. 144, p. 389-399, 1996.

TROMMELEN, G. J. M.; DEN DAAS, N. H. G.; VIJG, J.; UITTERLINDEN, A. G. Identity and paternity testing of cattle: Application of a Deoxyribonucleic Acid profiling protocol. *Journal of Dairy Science*, v. 76, p. 1403-1411, 1993.

USHA, A. P.; SIMPSON, S. P.; WILLIAMS, J. L. Probability of random sire exclusion using microsatellite markers for parentage verification. *Animal Genetics*, v. 26, p. 155-161, 1995.

VAIMAN, D.; SCHIBLER, L.; BOURGEOIS, F.; OUSTRY, A.; AMIGUES, Y.; CRIBIU, E. P. A genetic linkage map of the male goat genome. *Genetics*, v. 144, p. 279-305, 1996.

VANKAN, D. M.; BELL, K. Caprine blood groups. 2. The C, G, H, I, J, K, L, N and Q systems. Biochemical. *Genetics*, v. 31, p. 19-28, 1993.

VANKAN, D. M.; FADDY, M. J. Estimation of the efficacy and reliability of paternity assignments from microsatellite analysis of multiple-sire matings. *Animal Genetics*, v. 30, p. 355-261, 1999.

VAN HOOFT, W. F.; GROEN, A. F.; PRINS, H. H. T. Microsatellite analysis of genetic diversity in African buffalo (*Syncerus caffer*) populations throughout Africa. *Molecular Ecology*, v. 9, p. 2017-2025, 2000.

WEBER, J. L.; MAY, P. E. Abundant class of human DNA polymorphism which can be typed using the Polymerase Chain Reaction. *American Journal of Human Genetics*, v. 44, p. 388-396, 1989.

WEIR, B. S.; COCKERHAM, C. Estimating F-statistics for analysis of population structure. *Evolution*, v. 36, p. 1358-1370, 1984.

WEIR, G.S. *Genetic Data Analysis II*. 2. ed. Sinauer Associates Inc., Sunderland MA. 1996.

YANG, L., ZHAO, S.H., LI, K., PENG, Z.Z., MONTGOMERY, G.W. Determination of genetic relationships among five indigenous Chinese goat breeds with six microsatellite markers. *Animal Genetics*, v. 30, p. 452-455, 1999.

## VERIFICAÇÃO DE PATERNIDADE EM CAPRINOS POR MEIO DE SISTEMA DE MICROSSATÉLITES DE DNA: MÉTODO DE EXCLUSÃO E INFERÊNCIA ESTATÍSTICA UTILIZANDO A RAZÃO DE VEROSSIMILHANÇA

**Resumo:** Um sistema de microssatélites para teste de verificação de paternidade em caprinos no Brasil foi proposto, composto de onze *loci* anteriormente descritos em bovinos, ovinos ou caprinos. O objetivo do trabalho foi estudar a utilização deste sistema de microssatélites para a verificação de paternidade em dois casos: 1) quando além do genótipo do suposto pai e da progênie, o genótipo da mãe é conhecido; e 2) apenas o genótipo do suposto pai e da progênie são conhecidos, utilizando-se o método de exclusão e o método de inferência estatística da razão de verossimilhança (Marshall et al., 1998). Foram genotipados 292 animais, pertencentes a três rebanhos, sendo que as raças leiteiras importadas Alpina e Saanen foram amostradas em dois rebanhos (UFV, Viçosa, MG e particular, Florestal, MG) e a raça naturalizada Moxotó foi amostrada do rebanho de conservação da Embrapa Caprinos, Ceará. As análises foram realizadas nos programas SAS (1998) e CERVUS (Marshall et al., 1998). O teste exato para equilíbrio de Hardy Weinberg foi realizado utilizando o programa TFGPA (Miller, 1997). Os *loci* estudados apresentaram heterozigosidade e informatividade de moderada a alta. A heterozigosidade esperada média para todos os *loci* foi de 0,717. O conteúdo de informação polimórfica (PIC1 e PIC2) e a probabilidade de exclusão combinada (PE1 e PE2) foram de 0,676 e 0,5420; e 0,999591 e 0,988375, respectivamente, quando se conhece e não se conhece o genótipo materno. Considerando erros de registros aqueles com mais de três incompatibilidades de genótipo, foram encontrados 27 registros errôneos. Pelo teste de inferência estatística, solucionaram-se corretamente 210 dos 276 casos de paternidade (76%). Se fosse adotado apenas o método de exclusão, apenas 160 paternidades seriam assinaladas com certeza (sem incompatibilidades pai alegado-progênie). Desta forma, o método de inferência pode ajudar as verificações de paternidade em larga escala, diminuindo o custo para solucionar os erros de laboratório e perdas parciais de genotipagem. Quando há conhecimento do genótipo materno, a probabilidade de exclusão foi superior e o programa CERVUS resolveu 95% dos casos.

**Palavras-Chave:** polimorfismo, DNA, Cervus, LOD score, confiança estatística.

## VERIFICATION OF PATERNITY IN GOAT THROUGH SYSTEMS OF DNA MICROSATELLITES: METHOD OF EXCLUSION AND STATISTICAL INFERENCE USING THE LIKELIHOOD RATIO

**Abstract:** A microsatellite system for verification of paternity in goats in Brazil was proposed, composed for eleven *loci* previously described in bovine, ovine or goat. The objective of the work was to study the use of the microsatellites for these verification of paternity in two cases: 1) when besides the assumption father's genotype and of the progeny, the mother's genotype is known; and 2) just the assumption father's genotype and of the progeny are known, being used the exclusion method and the method of statistical inference likelihood-based (Marshall et al., 1998). 292 animals were genotyped, belonging to three flocks. The milking breeds Alpine and Saanen were amostras in two flocks (UFV, Viçosa, MG State and matter, Florestal, MG State) and the naturalized breed Moxotó of the conservation flock of National Goat Research Center (Embrapa), Ceará State. The analyses were done in computational programs SAS (1998) and CERVUS (Marshall et al, 1998). The exact test for Hardy Weinberg equilibrium was accomplished in TFGA (Miller, 1997). The studied loci presented moderate to high heterozygosity and informativity. The average heterozygosity for all loci was of .717. The polymorphic information content (PIC1 and PIC2) and the probability of combined exclusion (PE1 and PE2) were of .676 and .542; and .999591 and .988375, respectively when the maternal genotype was and not known. Considering mistakes of registrations those with more than three genotype incompatibilities, they were found 27 erroneous registrations. For the test of statistical inference, it was solved 210 of the 276 cases of paternity (76%) were solved correctly. If it was just adopted the exclusion method, only 160 paternities would be marked for sure (without incompatibilities alleged father - progeny). This way, the inference method can help the verifications of paternity in large scale, decreasing the cost to solve the laboratory mistakes and partial losses of genotyping. When there is knowledge of the maternal genotype, the exclusion probability was superior and CERVUS solved 95% of the cases.

**Keywords:** polymorphism, DNA, CERVUS, LOD score, statistical confidence.

## 1. Introdução

O caprino é a espécie doméstica mais vastamente distribuída no mundo e desempenha um papel muito importante para a economia de países em desenvolvimento da Ásia e África. O Brasil detém o décimo maior plantel de caprinos do Mundo, com cerca de 9,5 milhões de cabeças (IBGE, 2001). Os caprinos encontram-se concentrados na região Nordeste, contribuindo para a sustentabilidade da agricultura familiar no semi-árido. Sua exploração é diversificada, sendo criados de modo ultra-extensivo na caatinga até em sistemas intensivos nas bacias leiteiras da região Sudeste do país.

O material genético existente no país também é bastante diversificado. Nas regiões de atividade leiteira, predominam as raças importadas da Europa, como a Saanen e a Alpina. Esta última, bastante expressiva também no Nordeste, onde está presente em cruzamentos com os caprinos adaptados (SRD) ou dos grupos naturalizados (Moxotó, Canindé, Marota e Repartida) (Machado, 2000a).

Não existe nenhum programa de melhoramento genético nacional em andamento para esta espécie. As poucas iniciativas de criar um banco de dados se devem a esforços isolados de criadores e instituições de pesquisa. Os registros genealógico são realizados pela Associação Nacional de Criadores de Caprinos (ABCC) e suas filiadas.

A importância dos testes de paternidade para o melhoramento genético dos animais domésticos pode ser ilustrada nas perdas da acurácia de seleção e redução do ganho genético originadas de identificações errôneas na genealogia (Carneiro et al., 1999). Todas as metodologias existentes para avaliação genética são baseadas em informações de parentesco e, portanto, subordinadas a um registro confiável de paternidade.

Os métodos estatísticos para análise de paternidade em populações naturais, ou onde os dados são levantados em larga escala, podem diferir do enfoque utilizado nos laboratórios comerciais, onde apenas a confirmação de compatibilidade do genótipo do trio pai alegado-mãe-progênie é o objetivo, sendo conhecido como método de exclusão.

O compartilhamento de alelos entre a progênie e o suposto pai estabelece a inclusão de paternidade e o macho que possui o genótipo incompatível com a progênie em questão é rejeitado como pai. O único macho que permaneça não excluído é tomado como o verdadeiro pai. A fórmula de exclusão de paternidade baseia-se nas frequências genótípicas da população, deduzidas pelas frequências gênicas estimadas, assumindo equilíbrio de Hardy-Weinberg na população (Jamielson e Taylor, 1997).

O método de exclusão não leva em consideração os erros de tipagem ou mutações, que levam a incompatibilidades de genótipo pai-progênie. Outro problema parece surgir quando mais de um pai candidato não é excluído pelo teste, sendo necessário um método estatístico para indicar paternidade (Slate et al., 2000).

A razão de verossimilhança representa quanto mais provável é o pai candidato que um pai ao acaso de ter passado aqueles alelos à progênie. Para pares específicos mãe-filho, o alelo paterno obrigatório pode ser freqüente, enquanto em outras situações ele pode ser raro. Assim, a ausência de exclusão na última situação será mais significativa que na primeira. Isto pode ser quantificado pela razão de verossimilhança entre paternidade e não paternidade, e fornece uma maneira de quantificar a evidência encontrada em favor da paternidade biológica do possível pai.

A aplicação do método da razão de verossimilhança por Marshall et al. (1998), utilizando o programa CERVUS para teste de paternidade em populações naturais trouxe uma inferência estatística ao teste de paternidade. Para acessar a confiança de paternidade, Marshall et al. (1998) propuseram um teste delta, que é a diferença entre os logaritmos naturais das razões de verossimilhança (LOD score) dos pais mais prováveis, assumindo uma distribuição de delta por meio de simulação. Este método prevê que mais de um pai alegado possa não ser excluído pelo teste, e mesmo assim ser capaz de distinguir estatisticamente qual o verdadeiro pai.

Este trabalho tem o objetivo de estudar um sistema de microssatélites e sua utilização para a verificação de paternidade em dois casos: (1) quando além

do genótipo do suposto pai e da progênie, o genótipo da mãe é conhecido (2) apenas o genótipo do suposto pai e da progênie são conhecidos, utilizando-se o método de exclusão (Jamielson e Taylor, 1997) e o método de inferência estatística da razão de verossimilhança (Marshall et al.,1998).

## **2. Material e Métodos**

### **2.1. Animais amostrados e extração de DNA**

Foram utilizados no estudo 292 caprinos, pertencentes a três rebanhos. Do rebanho da Universidade Federal de Viçosa (UFV), Minas Gerais, foram amostrados 121 animais da raça Alpina e 94 da raça Saanen; de um rebanho particular situado em Florestal, Minas Gerais, foram amostrados 36 animais da raça Alpina e 19 da raça Saanen; e do rebanho de preservação da Embrapa Caprinos, Ceará, foram amostrados 22 animais da raça naturalizada Moxotó. Os animais utilizados possuem registro de paternidade conhecido, a ser confirmado pelo teste. Alguns possuíam mãe conhecida que foram genotipadas no rebanho. As amostras foram distribuídas dentro das famílias dos reprodutores existentes dentro de cada rebanho, de forma a conseguir um grupo de progênie de cada macho do plantel.

O sangue para extração de material genético das amostras foi colhido na veia jugular pelo sistema de vácuo. Depois de colhido, o sangue foi acondicionado em geladeira (4 °C) por 12 horas antes de iniciar a extração de DNA, segundo protocolo do fenol:clorofórmio (Sambrook et al.,1989).

As amostras de DNA do rebanho Moxotó foram obtidas de folículos pilosos extraídos da região da cauda, visando facilitar o transporte do material para análise. A extração foi realizada segundo o protocolo do CTAB (brometo de cetiltrimetilamônio), seguindo os seguintes passos: preparar três microtubos por amostra de DNA: 1) 500 µl de CTAB a 2%, com 10-30 folículos pilosos; 2) 500 µl de álcool isoamínico; e 3) vazio para adição posterior de isopropanol. Macerar os folículos no CTAB e vortexar. Levar ao banho-maria a 65°C durante

60 minutos, vortexando a cada 20 minutos. Centrifugar a 14.000 rpm durante 2 minutos. Remover o sobrenadante para o segundo conjunto de microtubo e homogeneizar. Centrifugar a 14.000 rpm durante 15 minutos, Remover o sobrenadante e transferi-lo para o terceiro conjunto de microtubo. Descansar por 30 minutos a 4°C. Centrifugar 14.000 rpm durante 15 minutos. Descartar o sobrenadante e lavar o *pellet* em 1 ml de etanol 75%. Centrifugar a 14.000 rpm durante 2 minutos. Descartar o sobrenadante e lavar o *pellet* em 1 ml de etanol 100%. Centrifugar a 14.000 rpm durante 2 minutos. Descartar o sobrenadante e deixar secar por 20 minutos. Adicionar 50 a 100 µl de tampão Tris-EDTA.

## 2.2. Reação de PCR e sistema de genotipagem

Foram utilizados onze *primers* de microssatélites, descritos na Tabela 1. Os *loci* inicialmente testados foram BETACAP, INRA005, ILSTS087, INRA006, INRA063 e INRABERN172 (em colaboração com o LABOGENA, INRA); foram adicionados os *loci* ILSTS005, ILSTS011, SRCRSP05, OARFCB48 e BM3205 para aumentar a confiabilidade do painel. O *locus* BETACAP é pertencente ao *cluster* gênico da caseína (Pépin et al., 1995). O *loci* ILSTS0087, INRA005, INRA006, INRA063, INRABERN172, ILSTS005, ILSTS011 e BM3205 foram descritos em caprinos por homologia em bovinos e o *locus* OARFCB48 por homologia em ovinos (Bishop et al., 1994; Crawford et al., 1995; Vaiman et. al., 1996). O SRCRSP005 foi descrito por Arevalo et al. (1994) em caprinos.

Cada 20 µl de reação de PCR continha 25 ηg de DNA genômico, 20 mM de Tris-HCl (pH 8.3), 50 mM de KCl, 0,2 mM de dNTPs, 1,25 mM a 2,5 mM de MgCl<sub>2</sub> e 1 U *Taq polimerase*. As concentrações dos *primers* foram ajustadas para uma boa sinalização na leitura dos picos de fluorescência, ficando entre 0,12 µM e 0,20 µM. Após o período de ativação de 94° C/3 min., foram realizados 27 ciclos de 94° C/1 min, T<sub>a</sub> °C/1 min, 72° C/1 min, seguidos por uma extensão final de 72° C/20 min.

Tabela 1 – *Locus*, cromossomo de sua localização (CHA), *primers* utilizados e dados na literatura para cada *locus* estudado em caprinos

<i>Locus</i>	CHA	<i>Primer</i>	<i>Citação</i>
BETACAP	6	F: ACGTGGCTGAGAAGCTTGGGTAAATG R: TCTAGAAGGCAAGAATGGACAGACAA	Pépin et al. (1995)
INRA005	12	F: AAATATTAGCCAACCTGAAAACCTGGG R: C TTCAGGCATACCCTACACCACATG	Bishop et al. (1994) e Vaiman et al. (1996)
ILSTS0087		F: CTGCCTCTTTTCTTGAGAGC R: AGCAGACATGATGACTCAGC	Bishop et al. (1994)
INRA006	3	F: CTGAGCTGGGGTGGGAGCTATAAATA R: AGGAATATCTGTATCAACCGCAGTC	Bishop et al. (1994) e Vaiman et al. (1996)
INRABERN172	26	F: GGCCTTGCTAGCCTCTGCAAAC R: CCAGGGCAGTAAAATGCATAACTG	Vaiman et al. (1996)
ILSTS005	10	F: GAAGCAATGAAATCTATAGCC R: TGTTCTGTGAGTTTGTAAAGC	Bishop et al. (1994) e Vaiman et al. (1996)
ILSTS011	14	F: GCTTGCTACATGGAAAGTGC R: CTAAAATGCAGAGCCCTACC	Bishop et al. (1994) e Vaiman et al. (1996)
SRCRSP05	21	F: GGACTCTACCAACTGAGCTACAAG R: TGAAATGAAGCTAAAGCAATGC	Arevalo et al. (1994)
INRA063	18	F: ATTTGCAGAAGCTAAATCTAACC R: AAACCACAGAAATGCTTGGAAG	Bishop et al. (1994) e Vaiman et al. (1996)
OARFCB48	17	F: GACTCTAGAGGATCGCAAAGAACCAG R: GAGTTA GTACAAGGATGACAAGAGGCAC	Crawford et al. (1995) e Vaiman et al. (1996)
BM3205	1	F: TCTTGCTTCCTTCCAAATCTC R: TGCCCTTATTTAACAGTCTGC	Bishop et al. (1994) e Vaiman et al. (1996)

Os *primers* reversos foram marcados fluorescentemente pelos corantes Tet (verde), 6-Fam (azul) e Hex (amarelo), de forma que o tamanho dos fragmentos dos diferentes *loci* não se sobrepusessem, permitindo a utilização da eletroforese em multiplex. Para cada amostra, os fragmentos amplificados no PCR foram misturados e preparados para a eletroforese, utilizando um mix contendo 1,0 µl da amostra amplificada e 0,25µl do padrão interno de tamanho (TAMRA 500), desnaturados em 24 µl de formamida deionizada (HI-DI), incubados por 94° C por 5 min e levados ao gelo.

Os fragmentos amplificados, depois de preparados, foram aplicados para eletroforese em capilar a 9 V, 60° C, por 22 minutos e analisados no *Genetic Analyzer* ABI 310, da *Applied Biosystem*<sup>TM</sup>, utilizando o *software Genescan*®. As reações de PCR foram feitas em multiplex, quando permitido (Tabela 2). A concentração dos *primers* foi ajustada para permitir um bom sinal de fluorescência para todas as cores.

A aplicação das amostras para eletroforese capilar no ABI 310 foi carregada em apenas dois multiplexes por amostra. A combinação de corantes (*dye*) e de tamanho dos fragmentos analisados é dada na Tabela 2. Para permitir que a eletroforese seja multiplex no sistema automático, é necessário que *loci* marcados com a mesma fluorescência não possuam fragmentos de tamanhos sobrepostos.

### **2.3. Análises estatísticas**

Os dados foram analisados no *Statistical Analysis Software* (SAS,1998), obtendo-se as frequências gênicas de cada alelo por contagem direta, a heterozigidade esperada sob equilíbrio de Hardy-Weinberg ( $H_E$ ) e a variância da estimativa, conforme descrito por Ott (1992). O conteúdo de informação polimórfica (PIC) de cada *locus* foi calculado conforme descrito por Bolstein et al. (1980), assumindo que o genótipo materno e da progênie são conhecidos (PIC1) e segundo Da e Lewin (1995), quando somente o genótipo da progênie é conhecido (PIC2). O teste exato para o equilíbrio de Hardy-Weinberg foi

Tabela 2 – *Locus*, mix de PCR, temperatura de anelamento em °C ( $T_a$ ), concentração de  $Mg^{2+}$  em mM ( $Mg^{2+}$ ), concentração dos *primers* em  $\mu$ M (*Pr*, marcação fluorescente (*dye*), número de alelos (N), tamanho dos alelos encontrados (pb) e a frequência observada para cada *locus* nas raças Alpina (FA), Saanen (FS) e Moxotó (FM) e geral (FG)

<i>Locus</i>	Mix	$T_a$	$Mg^{2+}$	<i>Pr</i>	<i>Dye</i>	N	(pb)	FA	FS	FM	FG
BETACAP	1	55	1,25	0,16	Tet	7	160	0,334	0,345	-	0,318
							162	0,073	0,133	0,361	0,115
							164	0,064	0,040	0,222	0,064
							166	-	-	0,333	0,021
							168	0,013	0,013	-	0,012
							170	0,516	0,465	0,083	0,469
							174	-	0,004	-	0,002
INRA005	1	55	1,25	0,16	6-Fam	5	113	0,003	0,004	-	0,003
							115	0,494	0,571	0,800	0,545
							117	0,102	0,252	0,125	0,162
							119	0,366	0,102	0,075	0,243
							121	0,035	0,071	-	0,047
ILSTS0087	1	55	1,25	0,12	6-Fam	9	135	0,010	-	-	0,005
							137	0,006	0,018	-	0,010
							139	0,019	0,004	0,237	0,028
							141	0,041	0,195	0,079	0,104
							143	0,067	0,004	0,289	0,057
							145	0,083	0,066	0,105	0,078
							147	0,357	0,323	0,289	0,339
							149	0,363	0,389	-	0,349
							153	0,054	-	-	0,029
INRA006	2	55	2,50	0,16	Hex	10	106	0,029	0,031	0,200	0,041
							108	0,013	0,053	0,100	0,034
							112	0,272	0,195	0,125	0,224
							114	0,066	0,173	-	0,102
							116	0,140	0,088	-	0,110
							118	0,226	0,323	0,575	0,288
							120	0,143	0,075	-	0,107
							122	0,124	0,035	-	0,081
							124	0,003	0,018	-	0,009
							126	-	0,009	-	0,003
INRABERN172	1	55	1,25	0,20	Hex	8	136	-	0,075	-	0,029
							140	0,003	0,009	0,053	0,009
							142	0,131	0,075	0,684	0,145
							144	0,369	0,425	0,105	0,374
							146	0,140	0,031	-	0,088
							148	0,197	0,137	-	0,161
							150	0,156	0,234	0,132	0,185
							152	0,003	0,013	0,026	0,009
ILSTS005	3	56	2,0	0,16	Tet	7	175	0,023	0,071	0,075	0,045
							177	-	0,022	-	0,009
							179	0,623	0,554	0,875	0,615
							181	0,266	0,326	0,050	0,274
							183	-	0,004	-	0,002
							185	-	0,008	-	0,003
							187	0,084	0,013	-	0,051

Continua...

Tabela 2, Cont.

<i>Locus</i>	<i>Mix</i>	<i>T<sub>a</sub></i>	<i>Mg<sup>2±</sup></i>	<i>Pr</i>	<i>Dye</i>	<i>N</i>	<i>(pb)</i>	<i>FA</i>	<i>FS</i>	<i>FM</i>	<i>FG</i>
ILSTS011	3	56	2,0	0,16	6-Fam	10	262	-	0,004	-	0,002
							264	0,010	0,053	-	0,026
							266	0,094	0,097	0,278	0,107
							268	0,260	0,053	-	0,161
							270	0,003	0,004	-	0,003
							272	0,006	0,013	-	0,009
							274	0,399	0,407	-	0,377
							276	0,110	0,296	0,583	0,214
							278	0,114	0,071	0,139	0,098
							280	0,003	-	-	0,002
SRCRSP05	3	56	2,0	0,16	Hex	7	157	0,003	0,044	0,219	0,032
							159	0,026	0,084	-	0,048
							163	0,221	0,035	-	0,134
							165	0,153	0,270	-	0,191
							167	0,562	0,451	0,219	0,498
							169	0,032	0,097	-	0,056
							173	0,003	0,012	0,562	0,041
INRA063	3	56	2,0	0,20	6-Fam	7	162	0,032	-	-	0,017
							164	0,093	0,027	-	0,061
							166	0,045	0,203	0,158	0,115
							168	0,400	0,310	0,632	0,380
							170	0,281	0,460	0,210	0,347
							172	0,145	-	-	0,078
							207	0,003	-	-	0,002
OARFCB48	4	58	3,0	0,20	Tet	7	152	0,074	0,069	-	0,071
							154	0,011	0,037	0,100	0,024
							156	0,182	0,259	-	0,212
							158	0,230	0,250	-	0,234
							160	0,137	0,222	0,900	0,189
							162	0,352	0,162	-	0,262
							164	0,015	-	-	0,008
BM3205	5	50	1,25	0,20	Hex	11	215	0,077	0,054	-	0,067
							217	0,008	0,086	-	0,044
							219	0,008	0,014	-	0,010
							223	0,339	0,364	-	0,350
							225	0,115	0,150	-	0,131
							227	0,350	0,200	-	0,281
							229	0,004	0,073	-	0,035
							231	0,092	0,023	-	0,060
							233	0,004	-	-	0,002
							235	0,004	0,014	-	0,008
							237	-	0,023	-	0,010

realizado pelo programa TFPGA v 1.3 (Miller, 1997), utilizando o método de Guo e Thompson (1992) da Cadeia de Markov (10 ciclos, 5.000 repetições, 1.000 dememorizações).

A probabilidade de exclusão foi calculada segundo Jamieson (1994), para os casos em que um progenitor é confirmado (exclusão de paternidade), sendo conhecidos os genótipos do pai alegado, da mãe confirmada e da progênie (PE1); e segundo Jamieson e Taylor (1997), quando apenas os genótipos do pai e da progênie são conhecidos (PE2). A probabilidade de exclusão combinada para todos os *loci* analisados foi calculada pelo produto da probabilidade de inclusão em cada *loci*  $PI = 1 - PE$ , pois uma exclusão é conclusiva. Assim, a probabilidade PI é a probabilidade daquele pai não ser excluído em nenhum dos *loci*.

A análise de verossimilhança toma os dados de genótipo da mãe, da progênie e do pai alegado ( $g_m$ ,  $g_p$  e  $g_a$ , respectivamente), e avalia a probabilidade da hipótese principal ( $H_1 =$  pai alegado é o verdadeiro pai), relativa à hipótese alternativa ( $H_2 =$  pai alegado é um indivíduo ao acaso). A razão de verossimilhança para cada *locus* quando o genótipo da mãe é conhecido (Meagher, 1986) é dada por:

$$L(H1, H2 | g_m, g_p, g_a) = \frac{T(g_p | g_m, g_a)}{T(g_p | g_m)},$$

em que  $T(g_p | g_m, g_a)$ , é a probabilidade do genótipo da progênie, dado o genótipo materno e do pai alegado, na segregação Mendeliana.

Nos casos em que o genótipo materno é desconhecido, a razão de verossimilhança é dada por:

$$L(H1, H2 | g_p, g_a) = \frac{T(g_p, g_a)}{P(g_p)},$$

em que  $P(g_p)$  é a frequência do genótipo da progênie.

Marshall et al. (1998) propuseram um teste estatístico no programa CERVUS v 2.0, que assinala a paternidade a partir de um delta crítico calculado através de simulação. O delta é definido como a diferença entre LOD score do pai alegado mais-provável, com o próximo mais-provável (LOD1-LOD2), para todos os pais alegados com LOD positivo.

O termo LOD score definido por Meagher (1986) refere-se ao produto da razão de verossimilhança de vários *loci* marcadores tomados simultaneamente e ao logaritmo natural do valor obtido. Geralmente este valor é positivo se o pai alegado é mais provável de ser o pai verdadeiro, do que outro indivíduo tomado ao acaso na população.

O valor crítico de delta, com determinado intervalo de confiança é obtido por simulação de distribuição de delta (10.000 rodadas) usando as frequências alélicas observadas dos *loci* tipados, com e sem informação do genótipo materno. O valor de delta é registrado, junto com o *status* (verdadeiro/falso) do pai mais-provável. Assim, o delta crítico com, por exemplo, 95% de confiança é aquele valor de delta para o qual 95% dos pais alegados com delta acima daquele valor sejam obtidos por pais verdadeiros (Figura 1).

Os valores estabelecidos pelo estudo para o cálculo de simulação de delta, foram 10% para erro de tipagem e 95% para a proporção de *loci* tipados. Foram considerados sete machos candidatos por rebanho, sendo 95% tipados. Não foram considerados os possíveis parentescos existentes entre os pais candidatos.

### **3. Resultados**

#### **3.1. Descrição dos *loci* e probabilidade de exclusão**

As frequências alélicas obtidas em cada *locus* estão demonstradas na Tabela 2. A Heterozigosidade Esperada ( $H_E$ ), Conteúdo de Informação Polimórfica ( $PIC_1$  e  $PIC_2$ ) e Probabilidade de Exclusão ( $PE_1$  e  $PE_2$ ) para os *loci* analisados estão na Tabela 3. A heterozigosidade esperada média ( $H_E$ ) foi de 0,717, variando de 0,825 (INRA006) a 0,542 (ILSTS005). O número médio

Tabela 3 – Número de observações (N), heterozigosidade esperada e erro padrão da estimativa ( $H_E \pm SE$ ), conteúdo de informação polimórfica (PIC1 e PIC2) e probabilidade de exclusão (PE<sub>1</sub> e PE<sub>2</sub>) para os *loci* estudados

<i>Loci</i>	N	$H_E \pm SE$	PIC1	PIC2	PE1	PE2
BETACAP	288	0,663±0,009	0,606	0,462	0,409	0,247
INRA005	289	0,617±0,011	0,560	0,414	0,362	0,203
ILSTS0087	290	0,742±0,008	0,702	0,568	0,524	0,347
INRA006	290	0,825±0,005	0,802	0,689	0,656	0,484
INRABERNI72	289	0,772±0,007	0,739	0,610	0,568	0,387
ILSTS005	286	0,542±0,012	0,480	0,338	0,294	0,154
ILSTS011	285	0,765±0,007	0,732	0,600	0,558	0,377
SRCRSP05	283	0,690±0,011	0,656	0,514	0,477	0,294
INRA063	287	0,713±0,007	0,665	0,525	0,475	0,303
OARFCB48	248	0,792±0,004	0,757	0,631	0,584	0,406
BM3205	240	0,771±0,008	0,738	0,610	0,570	0,392
Todos os <i>loci</i> *	292	0,717±0,004	0,676	0,542	0,999591	0,988375

\* Médias de  $H_E$ , PIC1 e PIC2. Probabilidade combinada de PE1 e PE2.

de alelos encontrados foi oito, variando de cinco (INRA005) a onze (BM3205). Conforme esperado, o PIC foi maior para os *loci* de maior heterozigosidade (INRA006), e quando o genótipo materno é conhecido.

A probabilidade de exclusão combinada obtida foi de 0,99959 e 0,98837 (PE1 e PE2), respectivamente (Tabela 3). Existe com este sistema de microssatélites a chance de inclusão errada de aproximadamente 1:2400 e 1:85, respectivamente.

Ao se considerar as raças separadamente (Tabela 4), a probabilidade de exclusão foi de 0,9992 (PE1) e 0,9817 (PE2), para a raça Alpina e 0,9993 (PE1) e 0,9834 (PE2) para a raça Saanen. A raça Moxotó possui uma amostragem pequena, de um só rebanho, mesmo assim os *loci* mostraram-se informativos e podem ser utilizados em estudos futuros, sendo 0,9731 (PE1) e 0,8248 (PE2). A percentagem de *loci* tipados para cada raça foi de 0,965; 0,993; e 0,791, respectivamente para a Alpina, Saanen e Moxotó. O número médio de alelos por *locus* foi 7,18 para a Alpina e Saanen e 3,5 para a Moxotó.

Tabela 4 – Número de observações (N), número médio de alelos por *loci* (n), heterozigosidade esperada ( $H_E \pm SE$ ), conteúdo de informação polimórfica (PIC1), probabilidade de exclusão (PE1 e PE2) para a raça Alpina, Saanen e Moxotó

<i>Locus</i>	N	n	$H_E \pm SE$	PIC1	PE1	PE2
<b>Alpina</b>						
BETACAP	157	5	0,614 $\pm$ 0,013	0,546	0,348	0,201
ILSTS0087	157	9	0,727 $\pm$ 0,011	0,682	0,501	0,326
INRA005	157	5	0,613 $\pm$ 0,011	0,536	0,335	0,195
INRA006	157	9	0,824 $\pm$ 0,005	0,798	0,647	0,472
INRABERNI72	157	7	0,766 $\pm$ 0,008	0,729	0,550	0,369
ILSTS005	154	4	0,558 $\pm$ 0,017	0,531	0,144	0,466
ILSTS011	154	9	0,741 $\pm$ 0,010	0,701	0,519	0,340
SRCRSP05	154	7	0,613 $\pm$ 0,016	0,562	0,369	0,205
INRA063	155	7	0,731 $\pm$ 0,010	0,687	0,501	0,324
OARFCB48	135	7	0,769 $\pm$ 0,009	0,731	0,553	0,373
BM3205	130	10	0,738 $\pm$ 0,010	0,693	0,509	0,333
Todos os <i>Loci</i>	151,8	7,18	0,697 $\pm$ 0,003	0,648	0,9992	0,9817
<b>Saanen</b>						
BETACAP	113	6	0,648 $\pm$ 0,013	0,581	0,379	0,226
ILSTS0087	113	7	0,705 $\pm$ 0,010	0,648	0,447	0,280
INRA005	113	5	0,598 $\pm$ 0,019	0,542	0,349	0,191
INRA006	113	10	0,813 $\pm$ 0,010	0,786	0,635	0,460
INRABERNI72	113	8	0,736 $\pm$ 0,014	0,698	0,520	0,339
ILSTS005	113	7	0,584 $\pm$ 0,016	0,512	0,318	0,179
ILSTS011	113	9	0,729 $\pm$ 0,013	0,686	0,505	0,329
SRCRSP05	113	7	0,707 $\pm$ 0,015	0,662	0,478	0,302
INRA063	113	4	0,653 $\pm$ 0,011	0,584	0,374	0,221
OARFCB48	108	6	0,792 $\pm$ 0,006	0,755	0,582	0,403
BM3205	110	10	0,792 $\pm$ 0,012	0,763	0,605	0,426
Todos os <i>Loci</i>	112,2	7,18	0,705 $\pm$ 0,003	0,656	0,9993	0,9834
<b>Moxotó</b>						
BETACAP	18	4	0,722 $\pm$ 0,021	0,645	0,439	0,273
ILSTS0087	19	5	0,780 $\pm$ 0,018	0,718	0,531	0,057
INRA005	20	3	0,347 $\pm$ 0,065	0,311	0,174	0,354
INRA006	20	4	0,619 $\pm$ 0,049	0,558	0,366	0,200
INRABERNI72	19	5	0,514 $\pm$ 0,066	0,470	0,299	0,137
ILSTS005	20	3	0,232 $\pm$ 0,063	0,214	0,116	0,026
ILSTS011	18	3	0,579 $\pm$ 0,044	0,495	0,296	0,159
SRCRSP05	16	3	0,607 $\pm$ 0,045	0,523	0,318	0,173
INRA063	19	3	0,546 $\pm$ 0,052	0,474	0,283	0,141
OARFCB48	5	2	0,200 $\pm$ 0,134	0,164	0,082	0,016
BM3205	-	-	-	-	-	-
Todos os <i>Loci</i>	17,4	3,5	0,515 $\pm$ 0,038	0,457	0,9731	0,8248

Foram verificados desvios do equilíbrio HW significativo ( $p < 0,01$ ) para os *loci* INRA063 e OARFCB48 em todas as raças, para o *locus* INRABERN172 na raça Alpina, *locus* SRCRSP05 na raça Saanen e *locus* BETACAP na raça Moxotó, estando os demais em proporções de equilíbrio.

No *pedigree* disponível, obtiveram-se 27 genótipos paternos registrados não compatíveis com o genótipo do filho, em pelo menos três dos *loci* analisados, sendo considerados para o estudo como erros de paternidade por exclusão. Os erros de registro de paternidade ocorreram a uma taxa de 10,4% na Alpina-UFV, 13,5% na Saanen-UFV, 10% na Moxotó e 3,9% no rebanho particular. De acordo com as normas internacionais para testes de paternidade, não se recomenda que a exclusão do pai alegado seja baseada em um *locus* individualmente. Segundo Luikart et al. (1999), este cuidado deve ser tomado principalmente para os *loci* que estão fora da proporção de Hardy-Weinberg, como o INRA063 e OARFCB48.

Considerando que a ocorrência de uma ou duas incompatibilidades entre o genótipo pai alegado-filho poderiam ser registradas como paternidade correta, verificaram-se ainda 64 incompatibilidades de pareamento entre os genótipos pai-filho. A grande margem de incompatibilidade foi devida principalmente aos *loci* INRA063 e OARFCB48, responsáveis por aproximadamente 20% das incompatibilidades cada um.

### **3.2. Inferência estatística por meio da razão de verossimilhança (CERVUS)**

Foram efetuados testes de verificação de paternidade para 276 progênes, sendo 43 com informação do genótipo materno em pelo menos oito *loci*. Os limites críticos de delta gerados pela simulação, com 10.000 rodadas, são apresentados na Tabela 5, com nível de confiança restrito (95%) e relaxado (80%). Na Figura 1 é demonstrada a distribuição de delta para os casos onde o genótipo materno é conhecido (Figura 1(a)), e para onde se desconhece ambos os progenitores (Figura 1(b)).

Tabela 5 – Limites de delta estabelecidos por meio de simulação ( $r = 10.000$ ) pelo programa CERVUS, para os casos onde a mãe possui genótipo conhecido (Caso 1) e quando não possui genótipo conhecido (Caso 2). Entre parênteses a porcentagem de casos de paternidade resolvidos nos dados apresentados

	Caso 1 ( n = 43)	Caso 2 (n = 233)
Confiança 95%	0,71 (63%)	0,92 (30%)
Confiança 80%	0,10 (95%)	0,30 (73%)

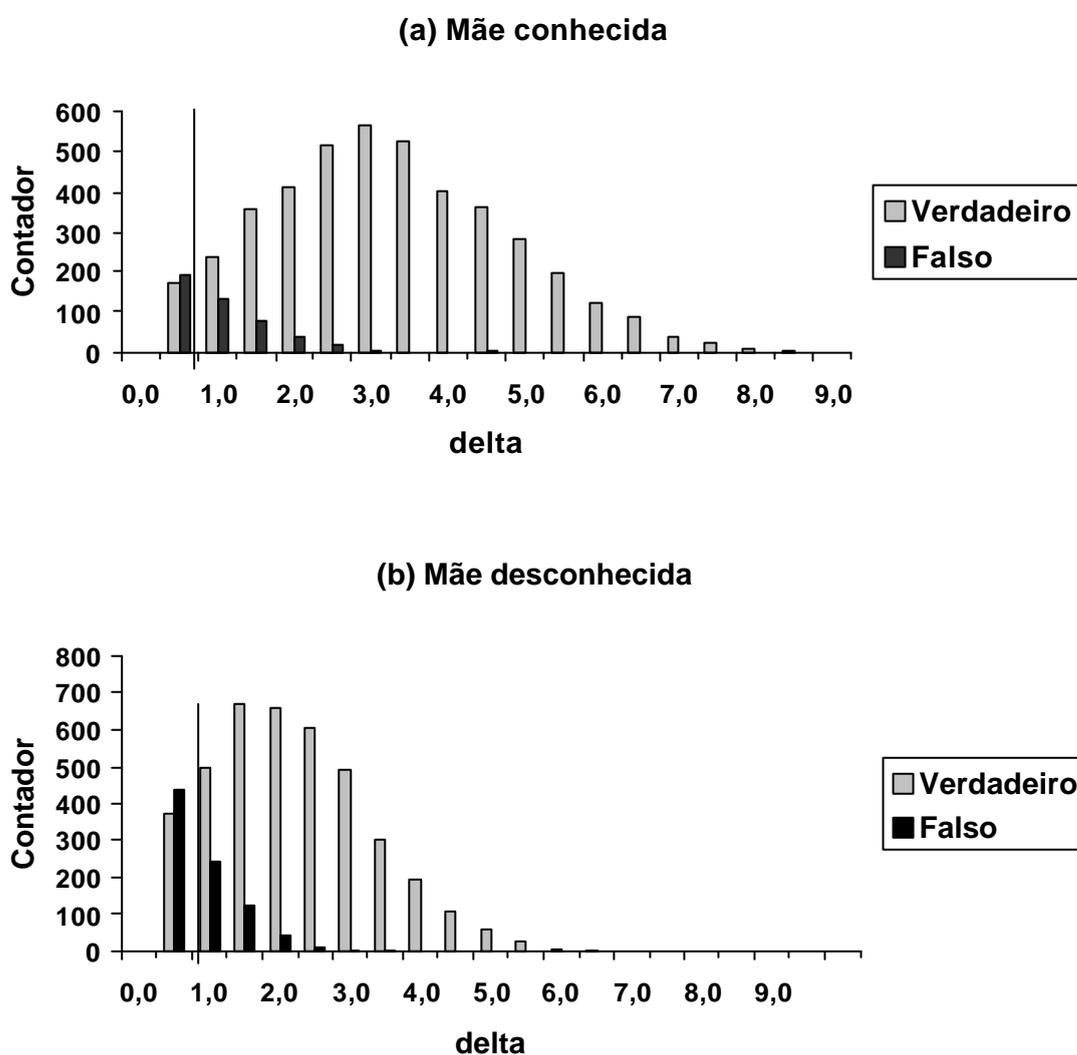


Figura 1 – Histograma dos escores de delta gerados por simulação usando as frequências alélicas de caprinos. a) Quando a mãe tem genótipo conhecido; e b) quando a mãe não é conhecida. O valor crítico de delta (linha vertical) é calculado pela sobreposição das duas distribuições pelo programa CERVUS.

O programa CERVUS conseguiu solucionar 41 (95%) das paternidades com 80% de confiança quando o genótipo da mãe foi informado e 169 (73%) dos casos, quando não foi informado, totalizando 210 casos solucionados. Dos 66 casos (24% do total de progênes) que permaneceram não resolvidos, 23 referiam-se a progênie com pai registrado conhecido não genotipado (fora da amostra de pais candidatos).

Para comparar a eficiência do método da razão de verossimilhança com o método de exclusão, foi observado o número de pais candidatos que apresentaram maior delta (LOD1-LOD2) no programa CERVUS e que não correspondiam aos machos apontados como verdadeiros pais nos registros genealógicos e que não foram cortados pelo método da exclusão. Constataram-se apenas 12 ocorrências (~4%), sendo cinco na Alpina-UFV, quatro na Alpina-particular, uma na Saanen-UFV e duas na Saanen-particular. Em todos estes casos os pais com o segundo maior delta correspondiam ao verdadeiro pai (registrado e não-excluído).

Na raça Moxotó, duas progênes não tiveram pais mais prováveis por meio do programa dentre os candidatos genotipados, apresentando delta zero. Tais progênes tiveram também o pai alegado excluído pelo método de exclusão, gerando erro de paternidade registrada. Uma progênie foi genotipada em menos de três *loci* e não foi testada pelo programa. O *locus* BM3205 não foi analisado para esta raça, por possuir poucas observações.

Os erros de registro de *pedigree* detectados pelo método de exclusão foram detectados também pelo programa CERVUS. Neste método, além de excluir o pai falso, um pai candidato da população é apontado como o mais provável para a progênie, podendo ser aceito ou não como verdadeiro, dependendo do delta calculado. Vê-se na Figura 2 que, com base na exclusão de paternidade simplesmente, apenas 160 progênes teriam pai assinalado (~60%). No restante dos casos ou o pai alegado seria rejeitado como verdadeiro, ou então teria de confirmar sua genotipagem no laboratório. Na Figura 2, as 64 incompatibilidades registradas correspondem à soma das incompatibilidades dos casos resolvidos pelo programa CERVUS.

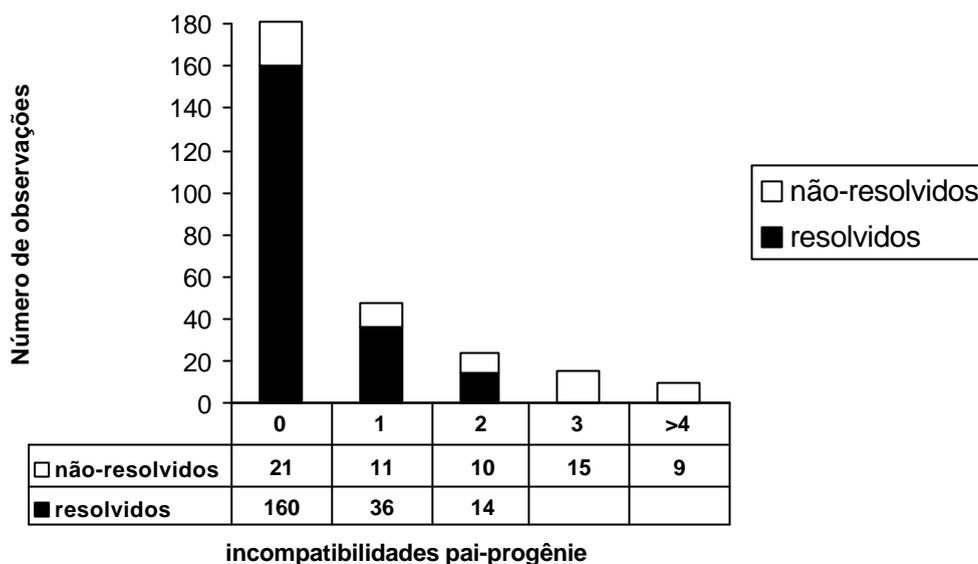


Figura 2 – Histograma do número de incompatibilidades pai alegado-progênie, nos casos resolvidos e não-resolvidos pelo programa CERVUS em caprinos.

A frequência de ocorrência de alelos nulos calculadas pelo programa CERVUS foi maior para os *loci* INRA063 e OARFCB48, os mesmos *loci* que foram responsáveis pelo maior número de incompatibilidades, sendo a frequência de ocorrência de respectivamente 0,1571 e 0,080.

Ocorreram ainda 15 incompatibilidades entre os genótipos mãe-filho, sendo seis provavelmente em decorrência de registro errado de maternidade, pois ocorreram em mais de três *loci*. Quando excluídos estes registros, a taxa de erro de genotipagem calculada pelo programa CERVUS decresceu de 10% para 6%.

#### 4. Discussão

Usha et al. (1995), com apenas cinco sistemas de microssatélites, obtiveram probabilidade de exclusão de 0,99 em bovinos. Também em bovinos, Heyen et al. (1997) obtiveram conteúdo de informação polimórfica médio de 0,59 e probabilidade de exclusão do pai incorreto  $>0,99999$ , quando foram utilizados 22 microssatélites, em multiplex. No Brasil, Curi e Lopes (2002) obtiveram probabilidade de exclusão combinada de 0,9789 com nove *loci* de

microssatélites em bovinos. Em caprinos, Luikart et al. (1999) desenvolveram um painel com 22 microssatélites em dois multiplex para o teste de paternidade obtendo uma probabilidade de exclusão  $>0,99999$ .

Segundo Vankan e Faddy (1999), ao se utilizar o teste de paternidade em touros múltiplos na Austrália, a razão custo:benefício dos testes de paternidade através de análise de DNA deve ser considerada na sua adoção pelos criadores e testes com poder de exclusão menor que 99% não são viáveis. O custo estimado do teste foi de cerca de US\$5 por microssatélite, de forma que deve-se ponderar o número de marcadores no sistema proposto. O aumento da acurácia do teste com o número de marcadores não é linear, e a probabilidade de exclusão de 100% é inatingível (Penna et al., 1998).

Em caprinos, existem poucas informações disponíveis sobre *loci* de microssatélites. Grande parte dos *loci* adotados por este e outros trabalhos vem do conhecimento existente em bovinos e ovinos, e da capacidade de amplificação interespecífica dos mesmos (Vaiman et al., 1996; Schibler et al., 1998). Isto reflete a história evolutiva destas espécies, que descenderam de uma mesma família (*Bovidae*) que se separou há cerca de 18 milhões de anos atrás. No Goatmap (<http://locus.jouy.inra.fr>) foram identificados 123 *loci* homólogos à ovinos e 275 homologias com bovinos. O INRA, na França, colaborou com a seqüência de cinco *primers* que foram inicialmente usados. Neste trabalho, a probabilidade de exclusão foi de cerca de 96%, o que levou a inclusão de mais seis marcas retiradas do GoatMap.

Os *loci* escolhidos são independentes (localizados em cromossomos distintos) e com herança Mendeliana simples, condição essencial para a escolha de *loci* marcadores. Dentro do definido como marcador polimórfico por Ott (1992), os *loci* BETACAP, INRA005, ILSTS005 e SRCRSP05 são moderadamente polimórficos ( $H_E > 0,70$ ). Ainda segundo Ott (1992), podem ser considerados polimórficos os marcadores que apresentem o alelo de maior freqüência, com freqüência observada menor que 0,50. Com a retirada de cada um dos quatro *loci* citados anteriormente, a PE2 abaixou para 0,9846, 0,9854, 0,9862 e 0,9835, respectivamente; e com a retirada de todos abaixou para 0,9676.

O afastamento do equilíbrio de HW verificado neste estudo para alguns *loci* pode ser devido à amostragem (efeito Wahlund), devendo atingir o equilíbrio quando utilizados mais rebanhos (Luikart et al., 1999). Foi verificada nos *loci* INRA063 e OARFCB48 a possibilidade de ocorrência de alelos nulos, uma vez que ocorreu um déficit de heterozigotos observados. Os alelos nulos ou não-amplificados podem ser responsáveis também pelo afastamento do equilíbrio de HW.

A taxa de erro de paternidade encontrada (cerca de 10%) está próxima à encontrada em raças européias de bovinos leiteiros (Ron et al., 1996). No Brasil, Rosa et al. (1997) obtiveram taxa de erro de aproximadamente 15% em bovinos Nelore. Curi e Lopes (2002) observaram erros de 27,5% em famílias da raça Gir registradas na ABCZ (Associação Brasileira de Criadores de Zebu)

O poder de exclusão do painel apresentado foi similar ao proposto por Luikart et al. (1999) para a raça Saanen na Suíça. O sistema de microsatélites montado pode servir para a verificação de paternidade em caprinos no Brasil. A partir deste, modificações podem ser implementadas, principalmente visando incorporar outros *loci*, a fim de melhorar a probabilidade de exclusão, principalmente quando o genótipo materno não é conhecido e quando a disputa de paternidade for entre animais aparentados. Para ambas as metodologias apresentadas o conhecimento do genótipo materno aumentou a acurácia do teste.

O método que utiliza a razão de verossimilhança é robusto a erros de tipagem (Marshall et al., 1998; Slate et al., 2000) e facilita a genotipagem em larga escala, como, por exemplo, de rebanhos inteiros. Vankan e Faddy (1999) discutem o teste de paternidade baseado em microsatélites para grupos de acasalamento por monta natural com múltiplos-touros, utilizado em gado de corte. O teste através de DNA poderia resolver as paternidades, favorecendo as análises dos dados provenientes de lotes com touros múltiplos (Carneiro et al., 1999).

Para utilizar com confiança o método de exclusão de paternidade, é necessário refazer as análises laboratoriais de todas as amostras onde as incompatibilidades foram observadas, para não incidir falsas exclusões. Dado ao

alto preço das análises, tal fato poderia inviabilizar levantamentos em larga escala no Brasil.

A taxa de erro laboratorial encontrada foi alta, provavelmente em decorrência do grande volume de amostras processadas e do caráter de treinamento de pessoal na rotina das análises. O programa CERVUS calcula a taxa de erro de acordo com as incompatibilidades de genótipo mãe-progênie. Porém, poucas mães foram genotipadas e ocorreram alguns erros de registro de maternidade. A taxa de erro admitida por Marshall et al. (1998) foi de cerca de 1%. A frequência de mutação e a ocorrência de alelos nulos ou não-amplificados podem também comprometer a taxa de erro.

Os resultados apresentados pela inferência estatística geradas pelo programa CERVUS foram bem consistentes com o método de exclusão. A principal vantagem é que o método confere uma confiabilidade estatística às análises, predeterminada pelo usuário, e o conjunto de dados é solucionado proporcionalmente a qualidade dos dados fornecidos pelo laboratório. Outra vantagem é que o método indica um pai provável dentro dos candidatos genotipados, e não simplesmente exclui a paternidade do pai alegado.

## **5. Conclusões**

Os onze *loci* de microssatélites estudados foram capazes de compor um painel para o teste de verificação de paternidade com uma confiabilidade próxima da obtida na literatura para o método de exclusão. O uso da metodologia proposta por Marshall et al. (1998) de inferência estatística por meio da razão de verossimilhança pode ser aplicado quando a genotipagem é feita em larga escala (rebanho), reduzindo o custo com repetições de genotipagem em casos de incompatibilidades entre os genótipos pai alegado-progênie. O conhecimento do genótipo materno aumenta muito a acurácia do teste, devendo-se sempre que possível proceder a genotipagem materna.

## Referências Bibliográficas

AREVALO, E.; HOLDER, D. S.; DERR, J. N. et al. Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP 1, 2, 3, 4 and 5 loci. *Animal Genetics*, v. 25, p. 202, 1994.

BISHOP, M. D.; KAPPES, S. M.; KEELE, J. W. et al. A genetic linkage map for cattle. *Genetics*, v. 136, p. 619-639, 1994.

BOLSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics*, v. 32, p. 314-331, 1980.

CARNEIRO, P. L. S.; EUCLYDES, R. F.; SILVA, M. A. et al. Efeito de erros de pedigree na seleção. *Revista Brasileira de Zootecnia*, v. 28, p. 269-274, 1999.

CRAWFORD, A. M.; DODDS, K. G.; PIERSON, C. A. et al. An autossomal genetic linkage map of the sheep genome. *Genetics*, v. 140, p. 703-724, 1995.

CURI, R. A.; LOPES, C. R. Evaluation of nine microsatellite loci and misidentification paternity frequency in a population of Gyr breed bovines. *Brazilian Journal of Veterinary Research Animal Science*, v. 39, p. 129-135, 2002.

DA, Y.; LEWIN H. A. Linkage information content and efficiency of full-sib and half-sib designs for gene mapping. *Theoretical and Applied Genetics*, v. 90, p. 699-706, 1995.

GUO, S. W.; THOMPSON, E. A. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, v. 48, p. 361-372, 1992.

HEYEN, D. W.; BEEVER, J. E.; DA, Y.; EVERET, R. E.; GREEN, C.; BATES, S. R. E.; ZIEGLE, J. S. Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semi-automated parentage testing. *Animal Genetics*, v. 28, p. 21-27, 1997.

IGBE. *Censo Agropecuário de 2001*. Disponível em: <<http://www.ibge.org.br>> Acesso em: 20 jan. 2004.

JAMIESON, A. The effectiveness of using co-dominant polymorphic allelic series for (1) checking pedigrees and (2) distinguishing full-sib pair members. *Animal Genetics*, v. 25, suppl. 1, p. 37-44, 1994.

JAMIELSON, A.; TAYLOR, ST. C. S. Comparisons of three probability formulae for parentage exclusion. *Animal Genetics*, v. 28, p. 397-400, 1997.

LUIKART, G.; BIJU-DUVAL, M-P.; ERTUGRUL, O.; ZAGDSUREN, Y.; MAUDET, C.; TABERLET, P. Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). *Animal Genetics*, v. 30, p. 431-438, 1999.

MACHADO, T. M. M. Caprins autochtones du Brésil: L'identification, la standardisation et la savergarde. In: INTERNATIONAL CONFERENCE ON GOATS, 7. Tours, 2000. *Proceedings...* INRA, 2v, 1.049 p. v. 2, p.941-943, 2000.

MARSHALL, T. C.; SLATE, J.; KRUK, L. E. B.; PEMBERTON, J. M. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, v. 7, p. 639-655, 1998.

MEAGHER, T. R. Analysis of paternity within a natural population of *Chamaelirium luteum*. I. Identification of most-likely male parent. *The American Naturalist*, v. 128, p. 199-215, 1986.

MILLER, M. P. *TFPGA – Tools for populations genetics analyses*. V 1.3 A Window program for the analysis of allozyme and molecular population genetic data. 1997.

OTT, J. Strategies for characterizing highly polymorphic markers in human gene mapping. *American Journal of Human Genetics*, v. 51, p. 283-290, 1992.

PENNA, V. M.; MADALENA, F. E.; OLIVEIRA, D. A.; LIMA, R. M. G. Testes de paternidade: investimento em credibilidade. In: ANAIS DO SIMPÓSIO NACIONAL DE MELHORAMENTO ANIMAL, 2., 1998, Uberaba. *Anais...* Uberaba: UFV, 1998, p.35-40.

PÉPIN, L.; AMIGUES, Y.; LÉPINGLE, A; BERTHIER, J-L., BENSALD, A; VAIMAN, D. *Sequence conservation of microsatellites between cattle (B. taurus), goat (C. hircus), and related species. Examples of use in parentage testing and phylogeny analysis.* (1995). INRA. Laboratoire de Genetique biochimique et de cytogenetique. Jouy-en-Josas.

ROSA, A. J. M.; PACKER, I. U. ; REGITANO, L. C. A. Caracterização da raça Nelore e teste de paternidade por marcadores moleculares. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34., 1997, Juiz de Fora. *Anais...* Juiz de Fora :Embrapa, 1997. v. 3, p. 62-64.

RON, M.; BLANC, Y.; BAND, M.; EZRA, E.; WELLER, J. I. Misidentification rate in Israeli dairy cattle population and its implications for genetic improvement. *Journal of Dairy Science*, v. 79, p. 676-681, 1996.

SAMBROOK, J.; FRITSCH, E. F.; MANIATS, T. *Molecular cloning: a laboratory manual*. 2. ed. New York: Cold Spring Harbor Lab. Press. 1989.

SAS (1998). *User's Guide. Statistical Analysis System*. 1998.

SCHIBLER, L.; VAIMAN, D.; OUSTRY, A.; GIRAUD-DELVILLE, C.; CRIBIU, E. P. Comparative gene mapping: a fine-scale survey of chromosome rearrangements between ruminants and humans. *Genome Research*, v. 8, p. 901-915, 1998.

SLATE, J.; MARSHALL, T.; PEMBERTON, J. A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology*, v. 9, p. 801-808, 2000.

USHA, A. P.; SIMPSON, S. P.; WILLIAMS, J. L. Probability of random sire exclusion using microsatellite markers for parentage verification. *Animal Genetics*, v. 26, p. 155-161, 1995.

VAIMAN, D.; SCHIBLER, L.; BOURGEOIS, F.; OUSTRY, A.; AMIGUES, Y.; CRIBIU, E. P. A genetic linkage map of the male goat genome. *Genetics*, v. 144, p. 279-305, 1996.

VANKAN, D. M.; FADDY, M. J. Estimation of the efficacy and reliability of paternity assignments from microsatellite analysis of multiple-sire matings. *Animal Genetics*, v. 30, p. 355-261, 1999.

WEIR, G. S. *Genetic Data Analysis II*. 2. ed. Sunderland MA.: Sinauer Associates Inc., 1996.

## MARCADORES DE MICROSSATÉLITES NO ESTUDO DE DIVERSIDADE GENÉTICA ENTRE POPULAÇÕES DAS RAÇAS CAPRINAS LEITEIRAS IMPORTADAS SAANEN E ALPINA E DA RAÇA NATURALIZADA MOXOTÓ NO BRASIL

**Resumo:** Na região Nordeste, detentora do maior rebanho caprino do País, existem raças naturalizadas adaptadas às condições adversas do semi-árido, que possuem diversidade genética desconhecida. Estas raças sofreram um processo de cruzamento indiscriminado com diversas raças importadas, pondo em risco seu potencial genético. O Brasil possui um rebanho expressivo de caprinos. Além das raças naturalizadas, o país possui um grande plantel de animais importados, destacando-se as raças leiteiras de origem européia, como a Alpina e a Saanen. Entretanto, não existem estudos de diversidade com o uso de sistemas de microssatélites de DNA. Este trabalho propõe um sistema composto por onze microssatélites para o estudo da diversidade genética entre as raças leiteiras importadas Alpina e Saanen e a raça naturalizada Moxotó. Foram genotipados 292 animais pertencentes aos rebanhos da UFV, um rebanho particular (raças importadas); e do rebanho de conservação da raça Moxotó na Embrapa Caprinos, Ceará. A heterozigosidade ( $H_E$ ) foi alta nos rebanhos leiteiros, sendo 0,6952 e 0,7043 para a raça Alpina e Saanen, respectivamente. Na raça Moxotó a  $H_E$  obtida foi moderada de 0,4984. O número de alelos variou de cinco (INRA005) a onze (BM3205), com média de 7,0 alelos/*locus* nas raças importadas e 3,5 alelos/*locus* na Moxotó. A média  $F_{ST}$  (diferenciação entre populações) foi maior entre rebanhos ( $F_{ST} s=0,0768$ ) do que entre raças ( $F_{ST} p= 0,0263$ ), indicando haver uma similaridade entre raças, dentro do rebanho. Tal similaridade pode ser devido ao fato de haver algum cruzamento entre raças, dentro do rebanho. A distância genética de Nei ( $D_A$ ) foi maior entre o rebanho Moxotó e os demais ( $D_A>0,50$ ). Os resultados obtidos estão em coerência com o histórico das raças, demonstrando que este sistema poderá ser adotado em estudos futuros de diversidade em caprinos no Brasil.

**Palavras-chave:** distância genética, DNA, heterozigosidade, polimorfismo.

## MICROSATELLITES MARKERS IN THE STUDY OF GENETIC DIVERSITY AMONG POPULATIONS OF IMPORTED MILKING GOATS BREEDS SAANEN AND ALPINE AND THE NATURALIZED BREED MOXOTÓ, IN BRAZIL

**Abstract:** In the Northeast part of Brazil, holder of the largest goat flock of the country, naturalized breeds exist adapted to the adverse conditions of the semi-arid, and possess unknown genetic diversity. These breeds are suffering a process of indiscriminate crossing with several imported breeds, endangering their genetic potential. The Brazil possess an expressive flock of goats. Besides the naturalized breeds, the country possesses a great number of imported animals, standing out the milking breeds of European origin, like the Alpine and Saanen. However, diversity studies don't exist with the use of microsatellites systems. This work proposes a system composed by eleven microsatellites for the study of the genetic diversity among the milking breeds Alpine and Saanen and the naturalized breed Moxotó. Were genotyped 292 animals belonging to the flocks of UFV, a private flock (imported breeds); and of the flock of conservation of the race Moxotó in National Goat Research Center, Embrapa, Ceará State. The heterozygosity ( $H_E$ ) was high in the milkf locks, being .6952 and .7043 for the Alpine and Saanen breeds, respectively. Moxotó breed obtained a moderate  $H_E$  of .4984. The allelic number varied of five (INRA005) to eleven (BM3205), with average of 7.0 allels/locus in the imported breed and 3.5 allels/locus in Moxotó breed. The average  $F_{ST}$  (differentiation among populations) was larger among flocks ( $F_{ST} s=.0768$ ) than among breeds ( $F_{ST} p = .0263$ ), indicating there to be a similarity among breeds, inside flock. Such similarity can feel to the fact of there being some crossing among breeds, inside flock. Nei's genetic distance ( $D_A$ ) was larger among the flock Moxotó and the others ( $D_A > .50$ ). The obtained results are in coherence with the report of the breeds, demonstrating that this system can be adopted in future studies of diversity in goats in Brazil.

**Keywords:** DNA, genetic distance, heterozygosity, polymorphism.

## 1. Introdução

As raças caprinas para a produção de leite mais expressivas no Brasil são de origem européia, importadas intensamente a partir de 1970. Segundo Santos (2003), as raças exóticas Saanen e Alpina são originárias da mesma região dos Alpes (França, Alemanha, Suíça). Ainda segundo o autor, a maioria dos exemplares de caprinos leiteiros introduzidos no Brasil é de origem francesa. A raça Saanen sofreu forte pressão de seleção em seu país de origem (Suíça), sendo na atualidade a maior produtora de leite em termos mundiais. Porém, esta raça é muito sensível ao clima tropical, sendo criada intensivamente principalmente na região Sudeste. A raça Alpina é muito numerosa no país, principalmente nas bacias leiteiras do Nordeste, onde é muito utilizada para cruzamento com as raças naturalizadas, dando origem a diversos grupos, como o da chamada “Parda Sertaneja” (Machado, 2000a).

O Nordeste é a região brasileira que possui o maior plantel de caprinos, aliado quase sempre à agricultura familiar no semi-árido. A maior parte dos caprinos desta região pertencem ao grupo SRD (Sem-raça-definida), que não possui padronização, formado pela mistura de exemplares trazidos na época da colonização. Outro grande grupo refere-se aos animais padronizados, destacando-se os grupos Moxotó, Marota, Canindé e Repartida. Estes caprinos são conhecidos como naturalizados, pois foram introduzidos na época da colonização e sofreram forte seleção natural para se adaptar às condições do semi-árido. Dentre estes grupos, talvez o mais tradicional seja a raça Moxotó, originária de Pernambuco, que possui reconhecimento do Ministério da Agricultura e Livro Genealógico (Machado, 2000a).

Os caprinos Moxotó são bem adaptados às condições edafoclimáticas da região semi-árida, produzindo, entretanto, pouco leite. São suas características principais a rusticidade, a prolificidade e a alta qualidade do couro. Devido ao uso de cruzamento maciço com raças caprinas importadas, seu rebanho está vulnerável, o que poderá causar perda de patrimônio genético para a caprinocultura no futuro. Estudos de diversidade genética nesta e nas demais

raças naturalizadas são urgentes para melhor direcionar os recursos existentes para sua conservação.

A análise de microssatélites tem sido apontada na literatura como uma ferramenta valiosa para estudo de diversidade genética entre populações, sobretudo devido ao alto polimorfismo apresentado por estes marcadores (Takezaki e Nei, 1996; Moiola et al., 2001). Apesar do expressivo rebanho caprino existente, ainda não existe um sistema de microssatélite proposto para raças caprinas no Brasil. Igarashi et al. (2000b) publicaram o primeiro estudo de estrutura genética de rebanhos naturalizados e importados no Brasil, com base em polimorfismos protéicos. Em resumo publicado por Igarashi et al. (2000a) uma subamostra do primeiro trabalho foi submetida ao estudo com dois *loci* de microssatélites, obtendo a mesma topografia anterior, com o uso de seis marcadores bioquímicos. Machado et al. (2000b), estudaram características morfológicas para a construção de filogenia de população caprina tradicional (SRD) do Nordeste brasileiro e de raças da região do Mediterrâneo.

Este trabalho tem por objetivo propor um sistema de *locus* polimórficos de microssatélites que possam ser utilizados posteriormente no estudo de diversidade genética de raças caprinas no Brasil, bem como verificar a diversidade entre as raças importadas Alpina e Saanen e a raça naturalizada Moxotó.

## **2. Material e Métodos**

### **2.1. Amostras de animais e extração de DNA**

Foram genotipados 292 animais, oriundos de três rebanhos. Do rebanho de caprinos da Universidade Federal de Viçosa (UFV), Minas Gerais, foram amostrados 121 caprinos da raça Alpina e 94 da raça Saanen. De um rebanho particular da região Sudeste foram amostrados 36 animais Alpino e 19 animais Saanen. Do rebanho de conservação da Embrapa- Centro Nacional de Pesquisa de Caprinos, localizado em Sobral, Ceará, foram amostrados 22 animais da raça Moxotó.

As amostras de sangue para extração de material genético foram colhidas na veia jugular pelo sistema de vácuo. Depois de colhido, o sangue foi acondicionado em geladeira (4 °C) por 12 horas antes de iniciar a extração de DNA, segundo protocolo do fenol:clorofórmio (Sambrook et al., 1989).

Do rebanho Moxotó, foram colhidos pêlos da região distal da cauda, e a extração foi feita por meio do método CTAB (brometo de cetiltrimetilamônio). A extração de DNA utilizando folículos pilosos pode ser bastante útil quando se desejam amostras de localidades distantes de grandes centros, onde o transporte aéreo rápido não está presente. O volume para transporte é pequeno, resistente e não perecível. Visando a padronização da técnica de extração de DNA de folículos pilosos, foi feito um pré-estudo com pelos de vinte caprinos da UFV.

## **2.2. Microssatélites utilizados e sistema de genotipagem**

Foram utilizados onze *primers* de microssatélites, localizados em cromossomos diferentes. Os *loci* inicialmente testados foram BETACAP, INRA005, ILSTS087, INRA006, INRA063 e INRABERN172 (em colaboração com o LABOGENA, INRA); adicionados os *loci* ILSTS005, ILSTS011, SRCRSP05, OARFCB48 e BM3205. Todos os *primers* utilizados haviam sido descritos anteriormente na construção do mapa de ligação em caprinos (Vaiman et al., 1996). Os *primers* reversos foram marcados fluorescentemente com Tet, 6-Fam e Hex, (Perkin Elmer).

Cada 20 µl de reação de PCR continha 25 ηg de DNA genômico, 20 mM de Tris-HCl (pH 8.3), 50 mM de KCl, 0,2 mM de dNTPs, 1,25 mM a 2,5 mM de MgCl<sub>2</sub> e 1 U *Taq polimerase*. As concentrações dos *primers* foram ajustadas para uma boa sinalização na leitura dos picos de fluorescência, ficando entre 0,12 µM e 0,20 µM. Após o período de ativação de 94 °C/3 min., foram realizados 27 ciclos de 94 °C/1 min, 50-58 °C (Ta)/1 min, 72 °C/1 min, seguidos por uma extensão final de 72 °C/20 min. Foi utilizado um termociclador *MJ Research PTC-100*. Os *primers* BETACAP, INRA005, ILSTS087, e INRABERN172 (mix 1); e INRA063 ILSTS005, ILSTS011 e SRCRSP05 (mix 2) foram

amplificados em PCR multiplex; as amplificações dos *loci* INRA006, OARFCB48 e BM3205 foram feitas em reações únicas.

Os produtos dos PCRs únicos e múltiplos foram misturados e preparados para a eletroforese em multiplex. Cada amostra analisada continha 1,0 µl da amostra amplificada e 0,25 µl do padrão interno de tamanho (TAMRA 500, *Applied Biosystem*), desnaturados em 24 µl de formamida deionizada (HI-DI, *Applied Biosystem*), incubados por 94 °C por 5 min e levados ao gelo. Para a amostra de DNA de cada animal foram realizadas duas corridas, uma com o mix 1, o mix 2 e o *locus* INRA006 e outra com os *loci* OARFCB48 e BM3205.

Os fragmentos amplificados, depois de preparados, foram aplicados para eletroforese em capilar a 9 V, 60 °C, por 22 minutos e analisados no *Genetic Analyzer ABI 310 (Applied Biosystem)*, utilizando o *software Genescan®*.

### 2.3. Análises estatísticas

As frequências alélicas e estimativas de heterozigosidade esperadas sob equilíbrio de Hardy-Weinberg ( $H_E$ ) e observadas ( $H_O$ ) para cada *locus* e média de todos os *loci* foram obtidas utilizando o programa TFPGA (*Tools for Population Genetic Analyses*) v 1.3 (Miller, 1997). As análises foram realizadas tomando rebanho de origem como subpopulação dentro de raça (Alpina-UFV = 1, Alpina-Particular = 2, Saanen-UFV = 3, Saanen-particular = 4 e Moxotó = 5).

As variâncias de  $H_O$  e  $H_E$  foram calculadas segundo Weir (1996) para cada uma das subpopulações estudadas:

$$VarH_O = \frac{1}{m^2} \sum_l \frac{H_l(1-H_l)}{n_l},$$

em que  $n_l$  é o número de animais genotipados no *locus*  $l$ ;  $m$  é o número de *loci* estudados e  $H_l$  é a heterozigosidade observada no *locus*  $l$ .

$$VarH_E = \frac{2}{m^2} \sum_l \frac{(1 + f_l) \left[ \sum_u p_{lu}^3 - (\sum_u p_{lu}^2)^2 \right]}{n_l},$$

em que  $p_{lu}$  é a frequência do alelo  $u$  no *locus*  $l$ ; e  $f_l$  é o coeficiente  $F_{IS}$  de Wright.

Para acessar a diferenciação genética das populações, os índices de fixação de Wright  $F_{IS}$  (consangüinidade ou perda de heteroziguidade dentro da população),  $F_{ST}$  (medida de diferenciação entre populações) e  $F_{IT}$  (perda global de heterozigose) foram computados segundo método de Weir e Cockerham (1984), utilizando um *bootstrapping* com intervalo de confiança de 95% (1000 repetições). Para detectar a variação dentro de raças de diferentes populações, foi realizada análise tendo população como raça e subpopulação como rebanho de origem. O programa computacional TFPGA pressupõe que a população está em equilíbrio de Hardy-Weinberg. O teste exato para múltiplos alelos proposto por Guo e Thompson (1992) para o equilíbrio foi realizado utilizando a cadeia de Markov (10 ciclos, 1000 repetições, 1000 dememorizações).

A distância genética ( $D_A$ ) foi estimada segundo Nei (1978), corrigida para amostras de tamanho pequeno. O método de grupamento com média aritmética não ponderada (UPGMA-*Unweighted Pair Group Method With Arithmetic Mean*), foi utilizado para construção do dendograma. Para dar um intervalo de confiança à árvore construída, o *bootstrap* (1.000 repetições) foi realizado pelo programa TFPGA, fornecendo a porcentagem de replicatas similares aos dados originais.

### 3. Resultados

O pêlo da região da cauda de caprinos mostrou-se 80% mais eficiente do que do dorso do animal para extração de DNA. O protocolo CTAB mostrou eficiência na extração e amplificação pelo PCR de cerca de 91%.

A proporção de *loci* com frequência perdida (não analisados) foi de cerca de 5%, sendo que o *locus* BM3205 não foi genotipado nos animais da raça Moxotó. A quantidade e qualidade do DNA extraído de folículos pilosos não permitem seu armazenamento prolongado. As reações de amplificação com os *loci* OARFCB48 e BM3205 foram feitas posteriormente às demais, e o DNA extraído pelo CTAB degradou mais rapidamente e não amplificou mais nas reações de PCR.

O número de alelos por *loci* variou de cinco (INRA005) a onze (BM3205), sendo o número médio de alelos por *locus* de 7,18 para Alpina e Saanen e 3,5 para Moxotó. A heterozigosidade esperada em equilíbrio de Hardy-Weinberg ( $H_E$ ) para cada *loci* estudado, nas cinco populações, está na Tabela 1. A heterozigosidade esperada ( $H_E$ ) foi de  $0,6952 \pm 0,0071$ ,  $0,7043 \pm 0,0075$  e  $0,4984 \pm 0,0325$ , respectivamente para as raças Alpina, Saanen e Moxotó. O valor de  $H_E$  está, de maneira geral, próximo da heterozigosidade observada (contagem direta), indicando não estar havendo perda global de heterozigosidade (fixação de alelos). Os rebanhos das raças Alpina e Saanen mostraram alto grau de diversidade (heterozigosidade), enquanto a Moxotó revelou menor diversidade nos *loci* estudados.

A  $H_E$  média (Tabela 1) para cada *loci* separadamente variou de 0,825 (INRA006) a 0,542 (ILSTS005). Tomando cada população separadamente, nas populações 4 e 5 os *loci* INRABERN172 e BETACAP apresentaram maior  $H_E$  que os demais. Nas outras populações, a  $H_E$  seguiu a média.

Os alelos compartilhados entre as populações estão demonstrados na Tabela 2. A raça Alpina detém alguns alelos exclusivos, com mais de três observações nos *loci* ILSTS087 (135 e 153 pb) e INRA063 (162,172 e 207 pb). Isto provavelmente se deu devido ao fato de haver maior número de amostras nesta raça, quando comparadas com as demais. A raça Saanen demonstrou apenas um alelo exclusivo com 237 pb (cinco observações) no *locus* BM3205. A raça Moxotó apresentou um fragmento de 166 pb no BETACAP, com uma frequência de observação alta (0,333), podendo ser, portanto, um alelo exclusivo da raça.

Tabela 1 – Tamanho efetivo da amostra (n), heterozigosidade esperada e observada ( $H_E$  e  $H_O$ ) para cada microsatélite nas populações estudadas (1 = Alpina-UFV; 2 = Alpina-Particular; 3 = Saanen-UFV; 4 = Saanen-Particular; 5 = Moxotó) e total para cada *loci*

<i>Locus</i>		1 (n = 116)	2 (n = 35)	3 (n = 93)	4 (n = 19)	5 (n = 18)	Total (n = 288)
BETACAP	$H_E$	0,640	0,489	0,668	0,555	0,693	0,663
	$H_O$	0,661	0,556	0,660	0,526	0,333	0,618
INRA005	$H_E$	0,628	0,527	0,621	0,422	0,339	0,617
	$H_O$	0,628	0,611	0,617	0,526	0,400	0,600
ILSTS087	$H_E$	0,711	0,632	0,688	0,632	0,759	0,742
	$H_O$	0,752	0,694	0,681	0,737	0,684	0,716
INRA006	$H_E$	0,817	0,760	0,810	0,731	0,604	0,825
	$H_O$	0,901	0,861	0,883	0,947	0,650	0,876
INRABERN172	$H_E$	0,764	0,725	0,706	0,796	0,500	0,772
	$H_O$	0,843	0,861	0,766	0,737	0,579	0,796
ILSTS005	$H_E$	0,542	0,469	0,585	0,525	0,226	0,542
	$H_O$	0,542	0,611	0,495	0,474	0,150	0,503
ILSTS011	$H_E$	0,749	0,675	0,710	0,737	0,563	0,765
	$H_O$	0,797	0,694	0,681	0,842	0,722	0,744
SRCRSP05	$H_E$	0,588	0,629	0,712	0,663	0,588	0,690
	$H_O$	0,636	0,583	0,649	0,632	0,500	0,625
INRA063	$H_E$	0,731	0,663	0,665	0,554	0,532	0,713
	$H_O$	0,571	0,444	0,372	0,526	0,263	0,467
OARFCB48	$H_E$	0,777	0,672	0,765	0,661	0,180	0,792
	$H_O$	0,657	0,818	0,483	0,579	0,200	0,601
BM3205	$H_E$	0,732	0,561	0,794	0,690	-	0,771
	$H_O$	0,740	0,667	0,834	0,706	-	0,767
Todos os <i>loci</i>	$H_E$	0,698±0,006	0,618±0,011	0,702±0,006	0,633±0,014	0,498±0,033	0,717±0,004
	$H_O$	0,702±0,004	0,673±0,008	0,648±0,003	0,657±0,010	0,448±0,018	0,665±0,002

Tabela 2 – Número de alelos/*locus* presentes em cada população e compartilhados entre populações (1 = Alpina-UFV; 2 = Alpina-Particular; 3 = Saanen-UFV; 4 = Saanen-Particular; e 5 = Moxotó)

Alelos	A <sup>1</sup>	B	C	D	E	F	G	H	I	J	K	Média
Populações												
1	5	4	9	8	5	4	8	7	7	7	10	6,7
2	4	4	6	8	7	3	6	5	5	6	6	5,5
3	6	5	6	10	7	7	9	7	4	6	9	6,9
4	4	3	5	5	6	3	4	4	3	5	8	4,5
5	4	3	5	4	5	3	3	3	3	2	-	3,5
Compartilha												
1-2	4	3	6	7	5	3	5	5	5	6	6	5,0
1-3	5	4	6	9	5	4	7	7	4	6	8	5,8
1-4	4	3	5	5	4	2	4	4	3	5	2	3,7
1-5	3	3	5	4	3	3	3	3	3	2	-	3,2
2-3	4	4	5	8	6	2	6	5	4	6	6	5,0
2-4	3	3	5	4	5	2	4	3	3	5	5	3,8
2-5	3	3	5	3	5	3	3	1	3	2	-	3,2
3-4	4	3	4	5	5	3	4	4	3	5	6	4,2
3-5	3	3	4	4	4	3	3	3	3	2	-	3,1
4-5	2	3	4	3	4	2	2	2	3	1	-	2,6
Média de compartilha	3,5	3,2	4,1	5,2	4,6	2,7	4,1	3,7	3,4	4,5	5,5	4,0
N. total	7	5	9	10	8	7	10	7	7	7	11	8

<sup>1</sup> *primers* (A) BETACAP, (B) INRA005, (C) ILSTS087, (D) INRA006, (E) INRABERN172, (F) ILSTS005, (G) ILSTS011, (H) SRCRSP05, (I) INRA063, (J) OARFCB48 e (K) BM3205.

O *locus* INRA006 foi o mais polimórfico, compartilhando na média 5,2 alelos entre populações, do total de 10 alelos (Figura 1a). O *locus* ILSTS005 foi o menos polimórfico, compartilhando na média 2,7 alelos dos 7 totais entre todas as populações (Figura 1b). Este, juntamente com o INRA005, foram os dois *loci* que apresentaram o mesmo alelo com maior frequência nas três populações. Para outros cinco *loci* (BETACAP, ILSTS087, INRABERN172, ILSTS011 e SRCRSP005) o alelo mais frequente para a raça Alpina, era também para a raça Saanen, diferindo ambos da raça Moxotó. Apenas para o *locus* OARFCB48 o alelo mais frequente foi diferente nas três populações.

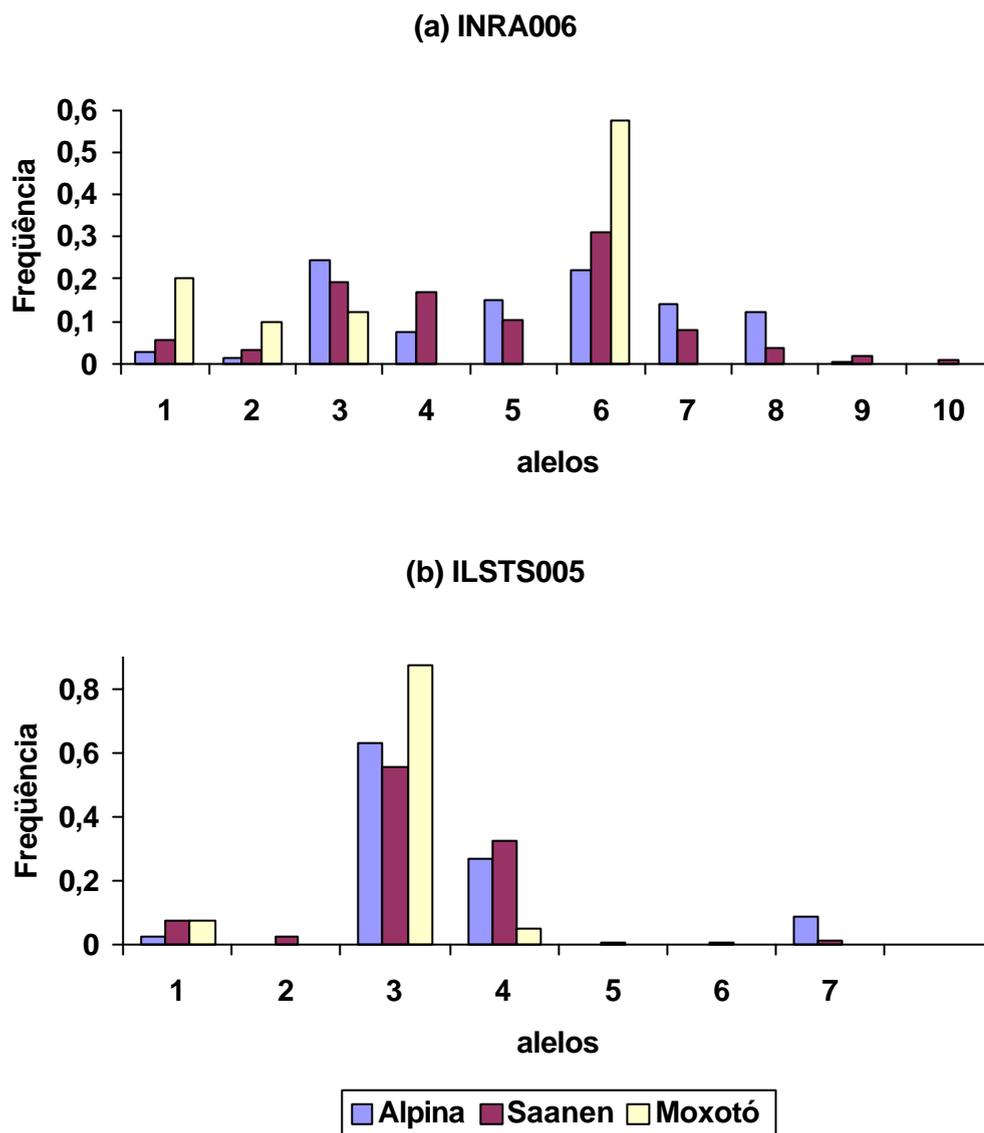


Figura 1 – Frequência alélica nos *loci* INRA006 e ILSTS005 nas três raças caprinas estudadas: a) apresentou melhor compartilhamento de alelos; e b) apresentou pouco compartilhamento de alelos entre as raças.

Os *loci* BETACAP, INRABERN172, SRCRSP05, INRA0063 e OARFCB48 mostraram afastamento significativo do equilíbrio HW nos dados amostrados ( $p < 0,001$ ). Dentro das raças, o *locus* INRABERN172 ( $p < 0,01$ ), apresentou desequilíbrio apenas na raça Alpina da UFV; o SRCRSP05 ( $p < 0,05$ ) apenas na raça Saanen da UFV e o BETACAP ( $p < 0,01$ ) apenas na raça Moxotó.

A estatística F de Wright está demonstrada na Tabela 3. O maior índice de fixação dentro de população ( $F_{IS}$ ) foi observado nos *loci* INRA063 (0,3171) e

Tabela 3 – Índices de fixação dentro das populações ( $F_{IS}$ ) entre as populações  $F_{ST}$  (P) e subpopulações (S) e consangüinidade global ( $F_{IT}$ ) para os *loci* e sobre todos os *loci*. Limites estimados por *bootstrapping* com 95% de confiança

<i>Locus</i>	$F_{IS}$	$F_{ST} S$	$F_{ST} P$	$F_{IT}$
BETACAP	0,0246	0,0835	0,0536	0,1060
INRA005	-0,0264	0,0873	0,0597	0,0632
ILSTS087	-0,0270	0,0788	-0,0213	0,0540
INRA006	-0,1041	0,0576	-0,0117	-0,0405
INRABERN172	-0,0889	0,0850	0,0563	0,0036
ILSTS005	0,0474	0,0420	0,0200	0,0874
ILSTS011	-0,0324	0,0903	0,0705	0,0608
SRCRSP05	0,0310	0,1111	0,0852	0,1386
INRA063	0,3171	0,0679	0,0491	0,3635
OARFCB48	0,1955	0,0810	-0,0038	0,2606
BM3205	-0,0394	0,0550	-0,0543	0,0178
Todos <i>loci</i>	0,0252	0,0768	0,0263	0,1000
<i>Bootstrapping</i>				
Superior	0,1028	0,0873	0,0511	0,1726
Inferior	-0,0379	0,0672	-0,0014	0,0411

OARFCB48 (0,1955), com média sob todos os *loci* de 0,0252. A média de  $F_{ST}$  (diferenciação entre populações) foi de 0,0717 quando consideradas amostras das cinco populações separadas e 0,0263 quando considerada as subpopulações de acordo com o rebanho (UFV ou particular). Isto implica que a maior parte da variabilidade se encontra entre rebanhos de uma mesma raça ( $F_{ST}$  da subpopulação = 0,0768) do que propriamente entre as raças. Através de *bootstrapping*, com intervalo de confiança de 95%, o valor de  $F_{ST}$  difere de zero, indicando haver diferença significativa entre as populações estudadas. O índice de fixação global ( $F_{IT}$ ) foi de 0,100.

A matriz das distâncias genéticas calculada (Nei, 1978) é dada na Tabela 4 e pode ser tomada como uma medida da diversidade entre rebanhos. A similaridade entre o rebanho 1 e 3 (0,9039), respectivamente Alpino e Saanen da UFV, foi maior do que entre 1 e 2 (0,8929), respectivamente Alpino UFV e Particular, e entre 3 e 4 (0,8837) Saanen UFV e Particular. Os animais Saanen do rebanho da UFV são na verdade, puros por cruza com a raça Alpina, como constatado em registros zootécnicos consultados posteriormente.

Tabela 4 – Distância genética de Nei (1978)  $D_A$  entre as populações estudadas (1 = Alpina-UFV; 2 = Alpina-Particular; 3 = Saanen-UFV; 4 = Saanen-Particular; 5 = Moxotó)

População	2	3	4	5
1	0,1133	0,1010	0,1707	0,5456
2		0,1236	0,1477	0,6358
3			0,1220	0,4468
4				0,6139

A população Moxotó demonstrou distância maior do que 0,50 em relação a todas as outras, com exceção do rebanho Saanen-Particular ( $D_A = 0,4468$ ), provavelmente em função do pequeno número de amostras dos dois rebanhos. O *bootstrap* dos nós da árvore gerada na Figura 2 apresentou 0,433 e 0,417 de similaridade nos nós 1 e 2, respectivamente; e 100% nos nós 3 e 4. O número de *loci* suportando cada nó foram 2, 2, 6 e 10, respectivamente.

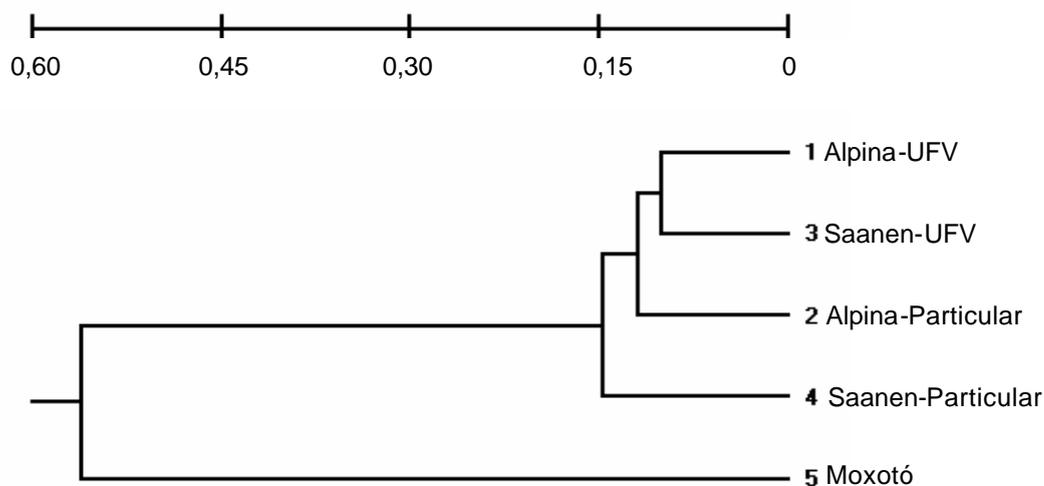


Figura 2 – Dendrograma UPGMA (*unrooted*) construído usando a distância genética ( $D_A$ ).

#### 4. Discussão

A precisão da estimativa de diversidade genética é função do número de *loci* analisados, a heterozigosidade destes *loci* e do número de animais amostrados em cada população (Barker, 1994). A amostragem ocorre hierarquicamente, com os *loci* amostrados dentro dos indivíduos, ou seja, os *loci* amostrados devem ser os mesmos para todas as populações sob investigação. Os animais amostrados devem ser retirados ao acaso da população, com recomendação mínima de N= 25 (FAO, 1998). Os *loci* amostrados devem ser informativos no sentido de possuir polimorfismo suficiente para distinguir os alelos semelhantes em estado. Barker (1994) sugere para estudos de diversidade, *loci* com no mínimo quatro diferentes alelos, para reduzir o erro padrão das estimativas de distância.

Takezaki e Nei (1996) ressaltaram a importância dos microssatélites de DNA para evidenciar relações evolutivas de populações muito próximas. Porém, alguns problemas surgem com o uso deste marcador, como a irregularidade do padrão de mutação e a diferença de polimorfismo entre populações, suscitando cautela na extrapolação de resultados para populações distintas das relacionadas. Diez-Tascón et al. (2000) estudaram diversidade de ovinos Merinos em rebanhos próximos (com menos de 400 anos de separação), mostrando ser possível distinguir estas populações através de estudos com microssatélites.

Os estudos de diversidade em caprinos com marcadores de microssatélites são raramente encontrados na literatura. Maudet et al. (2002) estudaram várias metodologias estatísticas no manejo de conservação de biodiversidade, utilizando *Capra ibex*. Foi demonstrada a baixa variação genética existente nesta espécie de cabra selvagem dos Alpes.

Yang et al. (1999) estudaram diversidade em caprinos de raças autóctones da China através de marcadores de microssatélites, obtendo resultados em concordância com o histórico das populações. A heterozigosidade obtida foi alta em seis dos treze *loci* testados Saitbekova et al. (1999) verificaram a diversidade entre rebanhos caprinos suíços domésticos e cabras Ibex e Bezoar.

As raças domésticas apresentaram maior heterozigosidade que os tipos selvagens, com  $H_E$  média de 0,51 a 0,58 para os rebanhos domésticos e 0,17 e 0,19 para as selvagens, respectivamente. Barker et al. (2001) estudaram a divergência genética entre cabras asiáticas nativas utilizando polimorfismos bioquímicos e microssatélites, revelando o isolamento geográfico como maior efeito de diferenciação entre populações.

No conjunto de *loci* estudado foi observado polimorfismo de alto a moderado ( $H_E > 0,70$ ) nas populações importadas amostradas. A raça Moxotó apresentou menor polimorfismo ( $H_E < 0,50$ ), provavelmente devido ao menor tamanho da amostra e também devido a uma certa redução de variação genética dentro da população.

O baixo valor de  $F_{ST}$  obtido para as raças, inferior ao obtido por rebanho (isolado geograficamente), sugere que os animais do rebanho da UFV podem estar sendo acasalados entre raças, proporcionando uma maior similaridade. Segundo Laval et al. (2000), entre as populações comerciais, a migração possui um efeito maior do que a mutação ou a deriva na redução da diferenciação genética entre populações.

Não houve nos dados analisados um padrão de déficit ou excesso de heterozigosidade observada. Os desvios do equilíbrio de HW em *loci* de microssatélites foram encontrados em várias citações (Luikart et al., 1999; Laval et al., 2000; Barker et al., 2001) e indicam afastamento do acasalamento ao acaso. Tal desvio pode ser devido à ocorrência de alelos nulos ou não-amplificados, vícios de grupamento de alelos, efeito da estrutura de amostragem (efeito Wahlund), seleção contra heterozigotos ou consangüinidade. É difícil distinguir que efeito está afetando o equilíbrio, porém os *loci* que apresentaram o desequilíbrio apenas dentro de uma subpopulação específica podem estar sob efeito Wahlund. Tal efeito tem como causa o agrupamento de indivíduos de diferentes rebanhos na análise, causando um déficit global de heterozigose, mas não um resultado similar nas observações dentro de *locus* e dentro de rebanho (Luikart et al., 1999)

Os *loci* INRA0063 e OARFCB48 mostraram  $F_{IS}$  maior que os demais *locus*, podendo estar havendo fixação de alelos. Entretanto, quando se analisa o  $F_{IS}$  global, não parece haver um acréscimo da endogamia nestas populações. Os valores de  $F_{IT}$  e  $F_{IS}$  estão próximos (Tabela 3), indicando estar havendo um impedimento a acasalamentos endogâmicos dentro dos rebanhos. O valor de  $F_{IS}$  (consangüinidade dentro dos rebanhos) não diferiu de zero, com intervalo de confiança de 95%.

Conforme dados obtidos neste estudo, existe grande diversidade genética entre as raças importadas e a raça naturalizada Moxotó e uma similaridade muito grande entre as raças Alpina e Saanen. O valor de  $F_{ST}$  indica que apenas 2,6% da variação genética existente é devido a diferença entre as raças. As populações da UFV foram mais semelhantes entre si do que as populações do rebanho particular, mostrando que pode estar havendo fluxo gênico entre as duas raças dentro do rebanho da UFV. De maneira geral, as raças Saanen e Alpina diferiram muito pouco entre si, provavelmente devido à origem em comum. Tais resultados corroboram com os obtidos por Igarashi et al. (2000b) e com o histórico das raças. As populações de caprinos leiteiros diferiram mais quanto ao rebanho de origem (subpopulações) do que a raça propriamente (população).

Na construção de árvores filogenéticas, o método UPGMA mostrou topologia correta. No estudo de simulação de Nei e Takezaki (1994), foi verificado que  $D_A$  (Nei, 1978) é a medida de distância mais eficiente para obter a topologia da árvore evolutiva. O UPGMA ou método de “ligação média” se baseia no agrupamento sucessivo de pares de táxons de acordo com o nível de similaridade. O UPGMA define a distância entre grupamentos como a distância média entre os pares. Esse método gera árvores com simetria dos ramos ao redor de um nó e a condição biológica que gera tal simetria é a ocorrência de taxas evolutivas constantes. Porém, em estudos de diversidade dos animais domésticos, o interesse maior está no valor de conservação das populações, mais do que na filogenia exata.

## 5. Conclusões

Em estudo com marcadores de microssatélites da variabilidade das populações caprinas no Brasil, foi demonstrado que as raças importadas Alpina e Saanen possuem identidades genéticas, sendo bem distantes da raça naturalizada Moxotó. Mais estudos sobre a diversidade das raças naturalizadas do nordeste brasileiro deverão ser conduzidos para monitorar a conservação. Os *loci* estudados foram informativos e devem ser mantidos em estudos posteriores de diversidade de caprinos no Brasil.

## Agradecimentos

Ao LABOGENA, INRA, pela colaboração com o desenho dos *primers* BETACAP, INRA005, INRA006, INRA063 e INRABERN172. Ao Capril Sanri, Florestal, MG, pela colaboração com a pesquisa. A Embrapa Caprinos, pelos dados do projeto de Conservação de Germoplasma. Aos bolsistas e estagiários do LABTEC, Dep. de Zootecnia, UFV, pela ajuda nos processos laboratoriais.

## Referências Bibliográficas

BARKER, J. S. F. A global protocol for determining genetic distances among domestic livestock breeds. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5<sup>th</sup>, v. 21. *Proceedings...* Guelph, 1994. p. 501-508, 1994.

BARKER, J. S. F.; TAN, S. G.; MOORE, S. S.; MUKHERJEE, T. K.; MATHESON, J-L.; SELVARAJ, O. S. Genetic variation within and relationship among populations of Asian goats (*Capra hircus*). *Journal of Animal Breeding Genetics*, v. 118, p. 213-233. 2001.

DIEZ-TASCÓN, C.; LITTLEJOHN, R. P.; ALMEIDA, P. A. R.; CRAWFORD, A. M. Genetic variation within the Merino sheep breed: analysis of closely related populations using microsatellites. *Animal Genetics*, v. 31, p. 243-251, 2000.

FAO. *Secondary Guidelines for development of national farm animal genetic resources management plans: Management of small populations at site*. Rome: FAO, 1998. 215 p.

GUO, S. W.; THOMPSON, E. A. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, v. 48, p. 361-372, 1992.

IGARASHI, M. L. S. P.; MACHADO, T. M.; CASTRO, S. R.; FERRO, J. A.; CONTEL, E. P. B. Genetic characterization of goats herds of the Brazil. In: GLOBAL CONFERENCE ON CONSERVATION OF DOMESTIC ANIMAL GENETIC RESOURCES, 5., 2000, Brasília. *Proceedings...* Ed. MARIANTE, A. S.; McMANUS, C.; SALOMÃO, A. N. Brasília: Embrapa. 2000a.CD-Rom.

IGARASHI, M. L. S. P.; MACHADO, T. M.; FERRO, J. A.; CONTEL, E. P. B. Structure and genetic relationship among Brazilian naturalized and imported goat breeds. *Biochemical Genetics*, v. 38, p. 353-365, 2000b.

LAVAL, G.; IANNUCELLI, N.; LEGAULT, C. et al. Genetic diversity of eleven European pig breeds. *Genetic, Selection, Evolution*, v. 32, p. 187-203, 2000.

LUIKART, G.; BIJU-DUVAL, M-P.; ERTUGRUL, O.; ZAGDSUREN, Y.; MAUDET, C.; TABERLET, P. Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). *Animal Genetics*, v. 30, p. 431-438, 1999.

MACHADO, T.M.M. Caprins autochtones du Brésil: L'identification, la standardisation et la sauvegarde. In: INTERNATIONAL CONFERENCE ON GOATS, 7. Tours, 2000. *Proceedings...* INRA, 2000a. 2 v, 1.049 p., v. 2, p. 941-943.

MACHADO, T. M. M.; CHAKIR, M.; LAUVERGNE, J. J. Genetic distances and taxonomic trees between goats of Ceará State (Brazil) and goats of the Mediterranean region (Europe and Africa). *Genetics and Molecular Biology*, v. 23, n. 1, p. 121-125, 2000b.

MAUDET, C.; MILLER, C.; BASSANO, B. et al. Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine Ibex (*Capra ibex ibex*). *Molecular Ecology*, v. 11, p. 421-436, 2002.

MILLER, M.P. *TTPGA - Tools for Populations Genetics Analyses*. V 1.3 A Window program for the analysis of allozyme and molecular population genetic data. 1997.

MOIOLI, B.; GEORGOUDIS, A.; NAPOLITANO, F. et al. Genetic diversity between Italian, Greek and Egyptian buffalo populations. *Livestock Production Science*, v. 70, p. 203-211, 2001.

NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, v. 89, p. 583-590, 1978.

NEI, M.; TAKEZAKI, N. Estimation of genetic distances and phylogenetic trees from DNA analysis. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5<sup>th</sup>, v. 21. *Proceedings...* Guelph, 1994. p. 405-412, 1994.

SAITBEKOVA, N.; GAILLARD, C.; OBEXER-RUFF, G.; DOLF, G. Genetic diversity in swiss goat breeds based on microsatellite analysis. *Animal Genetics*, v. 30, p. 36-41, 1999.

SAMBROOK, J.; FRITSCH, E.F.; MANIATS, T. *Molecular cloning: a laboratory manual*. 2. ed. New York: Cold Spring Harbor Lab. Press, 1989.

SANTOS, R. *A cabra e a ovelha no Brasil*. Uberaba-MG: Ed. Agropecuária Tropical Ltda., 2003. 479p.

TAKEZAKI, N.; NEI, M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, v. 144, p. 389-399, 1996.

VAIMAN, D.; SCHIBLER, L.; BOURGEOIS, F.; OUSTRY, A.; AMIGUES, Y.; CRIBIU, E.P. A genetic linkage map of the male goat genome. *Genetics*, v. 144, p. 279-305, 1996.

YANG, L.; ZHAO, S. H.; LI, K.; PENG, Z. Z.; MONTGOMERY, G. W. Determination of genetic relationships among five indigenous Chinese goat breeds with six microsatellite markers. *Animal Genetics*, v. 30, p. 452-455, 1999.

WEIR, B. S.; COCKERHAM, C. Estimating F-statistics for analysis of population structure. *Evolution*, v. 36, p. 1358-1370, 1984.

WEIR, G. S. *Genetic Data Analysis II*. 2. ed. Sunderland MA: Sinauer Associates Inc., 1996.

### 3. CONCLUSÕES GERAIS

O sistema de microssatélites de DNA em genotipagem semi-automática forneceu uma ferramenta valiosa para o teste de paternidade em caprinos no Brasil. A técnica é rápida, requer pouca quantidade de DNA e é reproduzida em qualquer laboratório. Tais condições tornam o painel composto por onze microssatélites BETACAP, INRA005, ILSTS087, INRA006, INRABERN172, ILSTS005, ILSTS011, SRCRSP05, INRA063, OARFCB48 e BM3205 a primeira proposta para viabilizar a realização de teste de verificação de paternidade em caprinos no Brasil. A probabilidade de exclusão obtida foi próxima aos padrões internacionais requeridos para o teste de paternidade,  $>0,999$ . Os testes de paternidade são importantes para o melhoramento genético da espécie, imprimindo confiabilidade aos registros geneológicos oficiais para que estes forneçam uma base de dados para os métodos estatísticos de avaliação genética através do parentesco.

A metodologia baseada na razão de verossimilhança proposta no programa computacional CERVUS (Marshall et al., 1998) resolveu os casos de paternidade com grande eficiência e pode ser uma opção para quando se deseja realizar a paternidade em larga escala (com muitas amostras simultaneamente). O método é robusto a erros de tipagem que possam ocorrer durante o manuseio das amostras coletadas, permitindo uma verificação acurada de casos de paternidade, mesmo

quando existem incompatibilidades entre os genótipos pai-progênie. As incompatibilidades geram falsas exclusões de paternidade e, pelo método convencional de verificação de paternidade, precisam ser revistas antes do laudo conclusivo. Isto acarreta em aumento de tempo e recursos gastos para a realização de verificações de paternidade.

A genotipagem do trio mãe-progênie-pai candidato melhora a eficiência do teste de paternidade, quando comparada a genotipagem progênie-pai candidato. O custo adicional para a genotipagem materna deve ser avaliado, uma vez que o custo por *locus* de microssatélite adicionado à análise, para torná-la mais precisa, também é alto.

Os marcadores de microssatélites propostos pelo estudo foram também eficientes para o estudo de diversidade genética em caprinos. Foi verificada uma similaridade grande entre os recursos genéticos importados, o que corrobora com o histórico das raças Alpina e Saanen. Por outro lado, a raça naturalizada Moxotó mostrou-se divergente das raças importadas estudadas, demonstrando a necessidade de preservação deste recurso genético adaptado às condições do semi-árido nordestino. Por outro lado, foi verificada uma perda de variabilidade no rebanho Moxotó, indicando que este rebanho está em processo de erosão de sua diversidade. Há necessidade de mais estudos de caracterização dos ecótipos caprinos naturalizados por meio de marcas de microssatélites, para conhecer a diversidade genética existente, visando direcionar melhor os escassos recursos disponíveis para conservação de germoplasma animal no Brasil. Os *loci* de microssatélites propostos neste trabalho poderão ser utilizados para a realização de estudos futuros.