

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



DISSERTAÇÃO DE MESTRADO

**AVALIAÇÃO DE UM PAINEL DE MICROSSATÉLITES PARA
IDENTIFICAÇÃO ANIMAL/AVERIGUAÇÃO DE PATERNIDADE E
ESTRUTURA GENÉTICA DA POPULAÇÃO DE CAVALOS DA RAÇA
MANGALARGA MARCHADOR**

ORIENTADA: Elisângela Alves Silva

ORIENTADORA: Prof^a Dr^a Maria Raquel Santos Carvalho

BELO HORIZONTE – MG

Fevereiro – 2012

Elisângela Alves Silva

**AVALIAÇÃO DE UM PAINEL DE MICROSSATÉLITES PARA
IDENTIFICAÇÃO ANIMAL/AVERIGUAÇÃO DE PATERNIDADE E
ESTRUTURA GENÉTICA DA POPULAÇÃO DE CAVALOS DA RAÇA
MANGALARGA MARCHADOR**

Dissertação apresentada ao programa de Pós-Graduação em Genética do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do grau de mestre em Genética.

Orientadora: Prof^a. Dr^a. Maria Raquel Santos Carvalho

Instituto de Ciências Biológicas

Belo Horizonte - MG

2012

DEDICATÓRIA

Dedico esta dissertação aos meus exemplos de vida, minha mãe e meu marido Carlos que sempre me estimularam a dar este grande passo. Estas duas pessoas com muita sabedoria, discernimento, bom senso e dedicação estiveram ao meu lado me encorajando nas horas difíceis e me aplaudindo nos momentos de glória. Obrigada por serem fonte de inspiração, apoio e ensino diário. Amo vocês.

AGRADECIMENTOS

Em primeiro lugar agradeço a Deus por me amparar nos momentos difíceis, mostrar-me os caminho nas horas incertas e suprir em todas as minhas necessidades

Agradeço à minha orientadora Dra. Maria Raquel, pelo acolhimento, pela paciência que teve comigo e pelos ensinamentos. Suas sugestões foram sempre úteis, bem vindas e acabaram por constituir-se neste trabalho.

Ao laboratório Linhagen, em especial a Kátia, por ter permitido que eu conciliasse o trabalho com o estudo e por ter fornecido o banco de dados da empresa. Ao Roberto pelas boas dicas, à Cecília por compreender os meus horários e a todos os amigos da empresa, que torceram tanto por isso.

À Associação Brasileira dos Criadores do Cavallo Mangalarga Marchador, principalmente ao Ciro, Henrique, Fábio Vilela e Nicolas por disponibilizarem o banco de dados para a realização deste trabalho e pela prontidão em me atender sempre que necessário.

Ao CBio, em especial ao Francislton, que contribuiu muito para a união dos bancos de dados.

Agradeço aos membros da banca por aceitarem ao convite.

Aos amigos que fiz durante minha caminhada no laboratório LGHM, à UFMG pelos conhecimentos adquiridos.

À minha família, marido, mãe, pai e irmãs, obrigada por depositarem em mim a confiança para todas as horas. Sei que vocês se orgulham por eu ter chegado até aqui e eu converto este orgulho na obrigação de cada dia ser mais digna. À minha irmã Flávia que esteve sempre disponível, apesar do pouco tempo que lhe sobra, principalmente na reta final deste trabalho.

Quero agradecer a todas as pessoas que se fizeram presentes, que se preocuparam, que foram solidárias, que torceram por mim. Mas em especial, à Laélia, sempre presente, apoiando, incentivando e humildemente auxiliando sempre que solicitada. Amiga você foi fundamental nessa caminhada. Marlene e Ana Cláudia que sempre estiveram ao meu lado, abrindo mão do seu tempo para me ajudar, Carla e Izinara obrigada pelas dicas. Todas vocês contribuíram muito, jamais esquecerei. Amigas, obrigada por tudo.

Todos vocês foram essenciais para a realização deste trabalho. Muito obrigada a todos!

RESUMO

Introdução: A raça de cavalos Mangalarga Marchador é a mais numerosa no Brasil com mais de 400 mil animais registrados. Originou-se no final do século XIX no sul de Minas Gerais a partir de cruzamentos de animais Alter, Andaluz, Arabe, Crioulo e Quarto de Milha. A principal característica da raça é a marcha, que consiste no passo acelerado. Dois objetivos principais nortearam a realização deste trabalho: avaliar a eficiência do painel de marcadores microssatélites em uso para verificação de parentesco em cavalos MM e caracterizar a estrutura genética da população.

Material e métodos: Foram analisados 6.092 animais nascidos em 2008 e 2009, evitando-se assim sobreposição de geração. Foram analisados dados relativos a 14 marcadores moleculares da classe dos microssatélites (AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG7 e VHL20). Foram estabelecidas frequências alélicas e genotípicas, heterozigosidade observada e esperada, conteúdo de informação polimórfica (PIC), poder de exclusão (PE), índice de paternidade (IP), índice de paternidade combinado (IPC), Equilíbrio de Hardy-Weinberg (EHW), AMOVA e estatísticas F (F_{ST} , F_{IS} e F_{IT}). Além disto, averiguou-se a presença de alelos nulos, exclusão de alelos grandes (*large allele dropout*), erros de genotipagem causados por *stuttering*. Como correção da múltipla testagem, usou-se Bonferroni para um alfa de 0,05. Para estas análises, foram utilizados os softwares: GENEPOP 4.0, Arlequim 3.11, Powerstats 1.2, MicroChecker 2.2.3.

Resultado e Discussão: Foram averiguados 129 alelos, com uma média de 9,2 alelos por marcador, variando de 5 a 20 alelos. A heterozigosidade observada média foi de 0,702 (0,536 a 0,813), a PE global foi de 0,443 (0,221 a 0,624), o PIC médio foi de 0,68 (0,5 a 0,8) e o IP combinado foi 99,9687%. Os valores extremos para todas as estatísticas referidas acima foram sempre do HTG7 (o menor) e do ABS17 (o maior). A análise de alelos nulos pelo Microchecker só mostrou evidências de alelos nulos na população de MG para os marcadores AHT5, ASB2, HMS2, HMS3, HMS6, HTG4, CA425, na população da BA, para os marcadores HMS3 e HTG7 e na população de PE, para os marcadores HTG7 e VHL20. Não foram encontradas evidências de exclusão alélica ou de erros de nomeação gerados por *stuttering*. Para a análise de estrutura de população, a amostra foi considerada em conjunto ou subdividida conforme o estado de nascimento do animal. Na amostra total, os marcadores ASB2, ASB23, CA425, HMS2, HMS6 e VHL20 apresentaram desvios

significativos dos esperados sob condições de EHW. Alguns marcadores apresentaram desvios do EHW em algumas populações (MG, MS e PA). O marcador que mais frequentemente se distanciou do EHW foi o HMS2. Os resultados da AMOVA sugeriram que a maior variação está dentro das populações (99,4) e não entre elas (0,6). O F_{ST} na população em geral considerando-se todos os *loci* foi de 0,006, sugerindo pouca diferenciação entre as populações. O F_{IS} apresentou valores muito próximos de zero, não sendo significativo para a grande maioria dos marcadores analisados. A comparação das populações por F_{ST} par a par, mostrou maior distanciamento entre as populações dos estados de MS-AL, MS-PA e MS-RN. As mais próximas foram SP e MG. Entretanto, a grande maioria das comparações evidenciou F_{ST} muito baixos e altamente significantes. Conclusão: Os resultados das análises estatísticas obtidos para o conjunto de marcadores analisados, como: PIC médio, PE médio e IP combinado sugerem que este painel analisado é eficiente para a raça estudada. Não houve evidências fortes de que a população esteja subestruturada, ou seja, se existe estruturação na população estudada sua base não é geopolítica.

Palavras-chave: cavalo, Mangalarga Marchador, microssatélite, população, paternidade

ABSTRACT

Introduction: Mangalarga Marchador is the most numerous horse breed in Brazil, with more than 400 thousand animals registered in the studbook. It was started by the ending of the XIX century in the south of Minas Gerais State from crosses between Alter, Andaluz, Arab, Creole, and Quarter Horse animals. The breed is defined by a particularly smooth gait and by its rusticity. The present work was developed in order to evaluate the performance of the parentage microsatellite panel in Mangalarga Marchador breed and to ascertain its population genetic structure. *Material and Methods:* Data from 6,092 born in 2008 and 2009 were included in order to avoid generation overlapping. Results from genotyping of 14 microsatellite markers were analyzed (AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG7, and VHL20). Analyses included presence of null alleles, large allele dropout, and allele misassignment due to stuttering, allelic and genotypic frequencies, observed and expected heterozygosity, polymorphic information content (PIC), mean PIC, power of exclusion (PE), paternity index (PI), combined paternity index (CPI), Hardy-Weinberg Equilibrium tests (EHW), AMOVA, and F statistics (F_{ST} , F_{IS} e F_{IT}). P values were Bonferroni corrected for an alpha = 0.05. Softwares used included GENEPOP 4.0, Arlequin 3.11, Powerstats 1.2, MicroChecker 2.2.3. *Results and discussion:* a total of 129 alleles were found, with a mean of 9.2 alleles per marker (range: 5-20). Mean observed heterozygosity was 0.702 (0.536 to 0.813), PE global: 0.9998, a mean PIC: 0.68 (0.5 a 0.8), and a combined PI: 99.9687%. Lower values of number of alleles, observed and expected heterozygosity, PIC, PE, and PI per locus were observed for HTG7 and the higher ones for ABS17. Null alleles were identified for some markers in some populations: MG (AHT5, ASB2, HMS2, HMS3, HMS6, HTG4, and CA425), BA (HMS3 and HTG7), and PE (HTG7 and VHL20). No significant evidences were found for allele dropout or allele misassignment due to *stuttering*. For genetic structure analyses tests were underwent in the complete sample and in the sample divided according the state of birth of the animals. In the complete sample, significant deviations from HWE expectations were found for markers ASB2, ASB23, CA425, HMS2, HMS6, and VHL20. Some markers showed HWE deviations in the subsamples of some states: MG, five markers; MS and PA, one marker each. In these subsamples, HWE

deviations were detected most of the times for HMS2. AMOVA results locate most of the variation within the populations (99.4), with a small fraction of it between them (0.6). F_{ST} for the complete population over all the loci was 0.006, pointing to a slight differentiation between the populations. F_{IS} values were close to null and non-significant for most of the markers tested. Higher pairwise F_{ST} values were found for the populations of the following pairs of states: MS-AL, MS-PA, and MS-RN. Lower pairwise F_{ST} was found for SP-MG. However, most of the pairwise F_{ST} values were very low and highly significant, suggesting that these populations are very similar. Conclusions: results obtained for mean PIC, PE and CPI indicated that the analysed panel is highly efficient for this breed. No strong evidence for population structure was found, meaning that, if there actually is any, it is not geopolitically based.

Keywords: horse, Mangalarga Marchador, microsatellite, population, paternity

Lista de Abreviaturas e Siglas

ABCCMM	Associação Brasileira dos Criadores do Cavalo Mangalarga Marchador
AMOVA	Análise de variância molecular
EHW	Equilíbrio de Hardy-Weinberg
He	Heterozigosidade esperada
Ho	Heterozigosidade observada
IP	Índice de paternidade
ISAG	<i>International Society for Animal Genetics</i>
MM	Mangalarga Machador
PCR	<i>Polimerase Chain Reaction</i>
PE	Probabilidade de exclusão
PIC	<i>Polymorphic information content</i>
STR	<i>Short Tandem Repeats</i>

Lista de figuras

FIGURA 1 – Tríplice apoio, momento em que três membros do MM tocam o solo ao mesmo tempo	14
FIGURA 2 – Caso de inclusão de paternidade para o marcador HMS7.....	18
FIGURA 3 – Caso de exclusão de paternidade para o marcador HMS7	18
FIGURA 4 – Distribuição geográfica da amostra	27

Lista de tabelas e quadros

Tabela 1 – Microsatélites usados, posição de mapa e referência	31
Tabela 2 - Tamanho amostral por estados	Erro! Indicador não definido.
Tabela 3 – Caracterização dos marcadores utilizados	36
Tabela 4 - Frequência alélicas para cada um dos marcadores genéticos estudados.....	37
Tabela 5 - Distribuição dos números de alelos com frequência inferior a 5 e 1% conforme o locus ...	38
Tabela 6 - Estatísticas F para cada locus com respectivos valores de p	40
Tabela 7 - Número médio de alelos em algumas raças.....	42
Quadro 1 - Distribuição dos marcadores que apresentam desvio do EHW por estado	38

SUMÁRIO

Lista de figuras	10
Lista de tabelas e quadros	11
1 INTRODUÇÃO	13
1.1 O cavalo	13
1.2 A raça Mangalarga Marchador	14
1.3 Microsatélites	17
1.4 Microsatélites para identificação animal e verificação de parentesco	17
1.5 Microsatélites para averiguação de estrutura genética de populações	20
1.6 Análises genéticas populacionais	22
1.7 Justificativa	25
1.8 Objetivos	26
1.8.1 Objetivos específicos	26
2 MATERIAIS E MÉTODOS	27
2.1 Amostra	27
2.2 Métodos moleculares	29
2.3 Banco de dados	30
2.4 Análises genéticas populacionais	30
3 RESULTADOS	34
3.1 Avaliação do painel de microsatélites para identificação animal e averiguação de paternidade	34
3.2 Estrutura genética da população	39
4 DISCUSSÃO	41
4.1 Avaliação do painel de microsatélites para identificação animal e averiguação de paternidade	41
4.2 Estrutura genética da população	44
5 CONCLUSÕES	46
REFERÊNCIAS ELETRÔNICAS E BIBLIOGRÁFICAS	47
ANEXOS	56

1 INTRODUÇÃO

1.1 O cavalo

Os cavalos foram domesticados a aproximadamente 6.000 anos atrás e desde então vem desempenhando um papel importante no desenvolvimento de muitas civilizações, atuando em funções básicas na agricultura, lazer, saúde, transporte e economia. Nenhum outro animal doméstico ocupou um lugar tão importante no processo de desenvolvimento social e político como os cavalos, uma vez que, desde a antiguidade fazem parte da cultura, da economia e do lazer do homem (BOWLING e RUVINSKY, 2000).

No Brasil, a criação de cavalos iniciou-se no período colonial e a história de sua origem está diretamente relacionada com o desbravamento e ocupação do território. Atualmente, podemos encontrar uma variedade de raças com inúmeras aptidões. Apesar da diferença entre as raças ter sido principalmente geográfica, outros fatores contribuíram para essa diferenciação, como as influências do clima, do meio ambiente, das diferentes condições de vida que levaram a diferentes características, qualidades e proporções (CARNEIRO *et al*, 1991; BOWLING e RUVINSKY, 2000; SILVA FILHO *et al.*, 2007;).

Os primeiros exemplares de cavalos chegaram a Minas Gerais vindos de São Paulo e Vale Vertentes do Rio São Francisco, mas não havia àquela época preocupação em manter ou melhorar a qualidade das raças equinas existentes, fato bem diferente da realidade de hoje, onde inúmeros trabalhos têm sido realizados no intuito de melhorar, caracterizar e confirmar pedigree desses animais (GANGARAMBHAI, 2004; OLIVEIRA, 2001; SILVA FILHO *et al*, 2007; DEASSIS *et al*, 2009; CHOI *et al*, 2008). Segundo o último censo feito pelo IBGE¹, em 2010, o Brasil possuía um rebanho de 5.514.253 equinos.

¹ Disponível em: <<http://www.sidra.ibge.gov.br/bda/pecua/default.asp?>>. Acesso em dezembro de 2011

1.2 A raça Mangalarga Marchador

Mangalarga Marchador (MM) é a raça de cavalos mais numerosa no Brasil, com mais de 400 mil animais registrados e teve início no final do século XIX, no sul de Minas Gerais, como resultado dos cruzamentos entre éguas selecionadas dessa região e garanhões vindos da coudelaria de Alter do Chão de Portugal, sendo estes animais das raças Alter, Andaluz, Árabe, Crioulo e Quarto de Milha. (CARNEIRO *et al.*, 1991; COSTA *et al.*, 2004; DEASSIS *et al.*, 2009). Atualmente esta raça está distribuída em todo território nacional, sendo que os estados de Minas Gerais, Rio de Janeiro, São Paulo e Bahia detêm 80% de todos os animais registrados, estando a maior parte na região Sudeste. Apesar do grande número de criadores e proprietários, a maioria deles é detentora de pequenos plantéis (COSTA *et al.*, 2004).

A função principal do MM é a marcha, que se diferencia de outras encontradas nos demais marchadores do mundo. A marcha é o passo acelerado, que transporta o cavaleiro de maneira cômoda, pois não transmite a ele os impactos decorrentes da pisada. Caracteriza-se pelo movimento que o animal faz ao descrever no ar um semicírculo com os membros anteriores, utilizando os posteriores como uma alavanca para ter impulso. Isso permite que o apoio do animal seja alternado nos sentidos diagonal e lateral, sempre suavizados por um tempo intermediário, o tríplice apoio (Fig. 1), momento em que três membros do animal tocam o solo ao mesmo tempo (ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DO CAVALO MANGALARGA MARCHADOR-ABCCMM, 2010)².

² Disponível em: <<http://www.abccmm.org.br>>. Acesso em maio de 2010.



Figura 1 - Tríplex apoio, momento em que três membros do MM tocam o solo ao mesmo tempo.

<http://www.damarcha.com.br>

Este andamento característico do MM é acompanhado de outras importantes características como:

- o temperamento ativo e dócil, podendo ser montado por pessoas de qualquer faixa etária e nível de equitação;
- a resistência, apresentando capacidade de percorrer longas distâncias e enfrentar desafios naturais;
- a inteligência, pois seu adestramento é fácil e rápido em relação a outras raças de sela;
- opção de se criar somente em regime de pasto, diminuindo seu custo de produção e manutenção, bem como facilitando seu manejo;
- a rusticidade pode ser observada na facilidade de adaptação a quaisquer terrenos e climas como o tropical, temperado ou frio.

Todas essas características foram sendo selecionadas e preservadas ao longo da formação dessa raça (CARNEIRO *et al*, 1991).

Com o intuito de unir os criadores desses animais, buscando o aprimoramento da raça, preservação de suas características básicas e o desenvolvimento de uma política de valorização da marcha, em 16 de julho de 1949, durante a I Exposição de Equídeos, no Parque da Gameleira, foi constituída

oficialmente a Associação dos Criadores de Cavalos Marchadores da Raça Mangalarga (ACCMRM) com sede em Belo Horizonte. Essa denominação prevaleceu até 1967, quando foi acrescido o adjetivo Nacional a seu nome, e só bem mais tarde ficou conhecida como Associação Brasileira dos Criadores do Cavalo Mangalarga Marchador (CARNEIRO *et al.*, 1991)

Os objetivos dos criadores são as exposições, os concursos de marcha, o enduro, a lida com o gado, as provas funcionais, as cavalgadas e o lazer em geral (ABCCMM, 2010). Anualmente, acontece em Belo Horizonte o maior evento envolvendo esta raça, a Exposição Nacional do Mangalarga Marchador. Este evento conta com criadores de quase todo o Brasil. Em 2011, foram realizados mais de 250 eventos nos diversos estados do país e a Exposição Nacional movimentou cerca de 10 milhões em negócios ocorridos nos leilões e venda de animais (ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DO CAVALO MANGALARGA MARCHADOR-ABCCMM, 2010³; ENGLER *et al.*, 2010).

A ABCCMM é responsável pelo registro genealógico da raça. Alguns procedimentos são necessários para que um cavalo Mangalarga Marchador receba seu registro definitivo. Assim, entre o nascimento e os 36 meses de idade, potros e potras recebem o Registro Provisório. Após os três anos de idade, cavalos e éguas são submetidos a uma avaliação pelos técnicos da ABCCMM, que determinam se os animais estão dentro do padrão da raça e em condições de receberem o Registro Definitivo. (ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DO CAVALO MANGALARGA MARCHADOR-2010).

Juntamente com o surgimento das diversas Associações de Criadores, foram criados em cada uma delas, os setores de Registro Genealógico, aos quais caberia emitir as genealogias de cada animal registrado, de modo que se documentasse a pureza das raças e estabelecer padrões de seleção. Devido a diversos erros de parentesco, que ocorriam nesta ocasião e passavam despercebidos, ficou clara a importância da realização de exames que pudessem garantir as informações contidas na genealogia. Inicialmente estes testes foram baseados em tipagem sanguínea ou polimorfismos protéicos e atualmente, são baseados na genotipagem de microssatélites (BERNOCO *et al.*, 1997; OLIVEIRA e KUABARA, 2001).

³ Disponível em: <<http://revistadigital.abccmm.org.br>>. Acesso em maio de 2010.

Desde então, foi gerada uma grande quantidade de dados que permitirá investigar a estrutura genética e a evolução destas populações.

1.3 Microssatélites

Os microssatélites, também conhecidos como STR (“*Short Tandem Repeats*” ou repetições curtas em *tandem*), são elementos repetitivos com motivos de dois a seis nucleotídeos, abundantes no genoma eucarioto (CHOI *et al.*, 2008). Eles combinam várias propriedades desejáveis. Além de altos níveis de polimorfismo, apresentam dispersão uniforme, codominância e possibilidade de genotipagem por métodos rápidos e simples (TAUTZ, 1989; ELLEGREN, 1992).

De acordo com sua estrutura, os microssatélites podem ser divididos em: perfeitos, que contem um único motivo de repetição n vezes (ex.: CACACACACACACA); imperfeitos, contêm uma sequência não repetitiva intercalada entre as repetições (ex.: CACACACACATGCTCACACACA) e compostos que estão constituídos por mais de um seguimento repetitivo diferente (ex.: CACACACACACAGAGAGAGAGA). Em equinos, os microssatélites padronizados pela *International Society of Animal Genetics* (ISAG) compreendem repetições de duas bases, ou seja, dinucleotídeos (MENEZES, 2005)

Esses marcadores têm sido aplicados em diversos estudos na área animal, entre eles: estimativas de distâncias genéticas, testes de paternidade; estudos populacionais e no melhoramento genético. (MADALENA, 1999; RODRIGUES, *et al.*, 2002; GANGARAMBHAI, 2004; MENEZES, 2005; CHOI *et al.*, 2008; MELO *et al.*, 2008).

1.4 Microssatélites para identificação animal e verificação de parentesco

Os testes de paternidade são realizados através da comparação do perfil genético dos animais envolvidos, os quais são baseados em microssatélites e constituem a identificação genética individual, inalterável com a idade ou condições fisiológicas, ou seja, em qualquer momento é possível verificar a identidade de um

animal, pelo confronto do resultado obtido de uma nova amostra com o obtido anteriormente (OOM, 2010).

A eficácia do teste não depende apenas do número de marcadores genéticos utilizados, mas também do poder informativo de cada marcador. Desse modo, quanto mais informativos forem os microssatélites individualmente, melhor a eficiência na exclusão de falso parentesco do conjunto (RODRIGUES *et al.*, 2002; DEASSIS *et al.*, 2009; SERENO *et al.*, 2008).

Nas **FIG. 2 e 3**, encontra-se um exemplo de verificação de parentesco e padrão de herança mendeliano usando o *locus* HMS7. Na **FIG. 2**, é apresentado um caso de inclusão para o marcador analisado, uma vez que o primeiro alelo do animal testado é compartilhado com a égua testada e o segundo com o garanhão. Em contrapartida, na **FIG. 3** é mostrado um caso em que o primeiro alelo do animal testado é compartilhado com a égua, mas o segundo alelo não é compartilhado com o garanhão, sugerindo assim um caso de exclusão paterna.

Em casos onde se encontra incompatibilidade em apenas um único sistema, é necessário que se tenha disponível outros marcadores bem caracterizados para resolução deste problema. Bowling *et al.*, em 1997, demonstraram que pode-se evitar as falsas exclusões, decorrentes de casos de mutação ou alelos nulos, usando-se como critério de exclusão a incompatibilidade genética em pelo menos dois *loci*.

Segundo Bowling *et al.*, 1997, a escolha dos microssatélites de um painel para verificação de parentesco inclui aspectos relacionados a:

- informatividade do sistema, como conteúdo de informação polimórfica (PIC), índice de paternidade (IP) e probabilidade de exclusão (PE);
- facilidade técnica, como uma faixa de tamanho alélico adequado;
- bom desempenho dos *primers* em reações de PCR multiplex;
- não ter evidências de alelos nulos;
- baixa taxa de mutação;
- segregação independente.

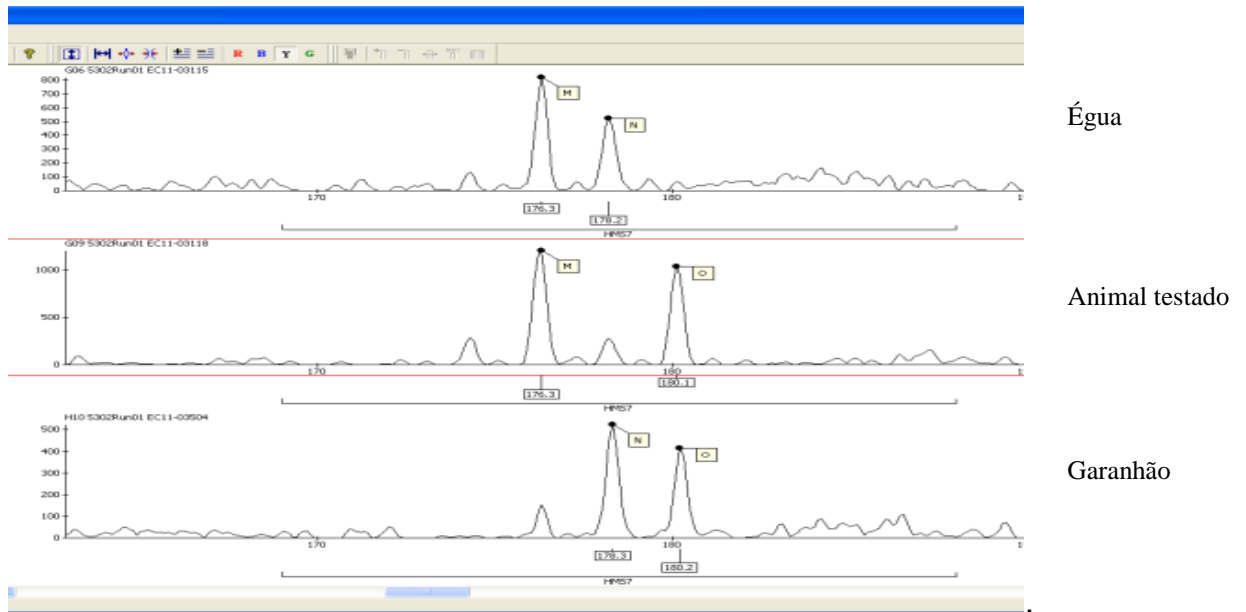


Figura 2 - Caso de inclusão de paternidade para o marcador HMS7
 Nota: Separação por eletroforese capilar em MegaBace 1000 (GE Healthcare).
 Observe que o alelo M do animal testado é compartilhado com a égua suposta mãe e o segundo, o alelo O, com o garanhão, suposto pai.

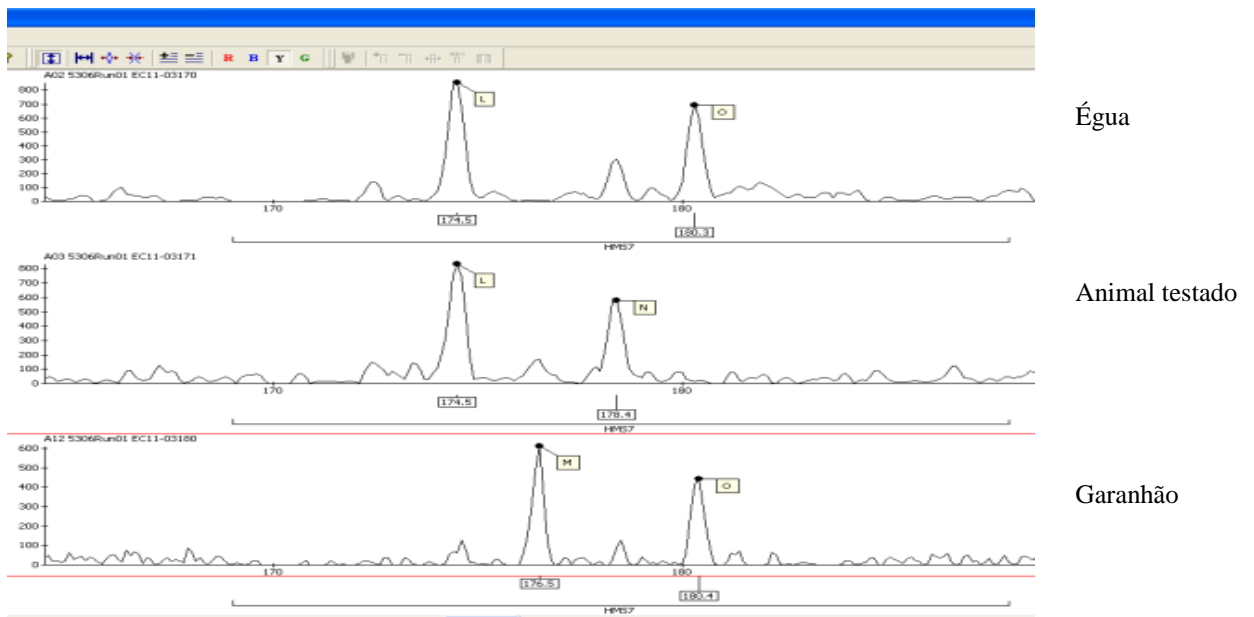


Figura 3 - Caso de exclusão de paternidade para o marcador HMS 7
 Nota: Separação por eletroforese capilar em MegaBace 1000 (GE Healthcare).
 Observe que o alelo L do animal testado é compartilhado com a égua, suposta mãe e o segundo, o alelo N, não é compartilhado com o garanhão, suposto pai.

O principal objetivo dos testes de verificação de parentesco é assistir os criadores, identificando a paternidade e maternidade correta, ou seja, confirmar a relação entre os indivíduos. Casos que envolvam a troca de animais podem ser detectados e comunicados (LEE e CHO, 2006; OLIVEIRA e KUABARA, 2001).

A genealogia de equinos tem sido controlada na maioria dos países através da genotipagem de microssatélites. Em um encontro da *International Society for Animal Genetics* (ISAG) em 1998, ficou decidido um painel mínimo para uso internacional em testes de paternidade. Nove microssatélites foram escolhidos e padronizados com base em um sistema de nomenclatura alfabética dos alelos. A partir disso, a cada dois anos, testes regulares de comparação são realizados sob os cuidados da ISAG. O objetivo destes testes é o controle de qualidade e a manutenção de um sistema internacional de nomenclatura alélica, que possa servir para a identificação individual em casos de exportação, por exemplo.

1.5 Microssatélites para averiguação de estrutura genética de populações

Os microssatélites apresentam algumas vantagens tais como a necessidade de pouca quantidade de amostra biológica, o fato de não se restringir a um único tipo de amostra (podendo usar, inclusive, material de animais mortos), a presença de novas tecnologias que permitem uma análise fácil, rápida e eficiente. Essas são algumas das razões pelas quais estes marcadores têm sido amplamente utilizados na caracterização da estrutura genética de várias populações animais (ELLEGREN, 2004).

A avaliação da diversidade genética em populações foi muito facilitada pelo uso dos microssatélites, devido ao seu alto conteúdo informacional (LUÍS *et al.*, 2007; MELO *et al.*, 2008). A variabilidade genética é a matéria prima da evolução, servindo como base nos processos de adaptação dos seres vivos quando submetidos aos diferentes desafios ambientais como competição, predação, patologias, fatores abióticos e outros. A diversidade dos recursos genéticos tem sido alvo de interesse dos cientistas e daqueles conscientes dos efeitos negativos produzidos pela perda da diversidade genética, como a diminuição da resistência

dos animais a doenças e mudanças nas condições climáticas (SILVA FILHO *et al.*, 2007; GANGARAMBHAI, 2004).

A seleção artificial é um fator que pode reduzir a diversidade genética através da redução do tamanho efetivo da população. Como os marcadores moleculares fornecem informações valiosas para o monitoramento da seleção artificial, eles podem ser usados na identificação de animais ou conjunto de animais que devem ser mantidos em processo de reprodução apenas para evitar a perda da diversidade genética. Além disso, esta informação pode ser utilizada para evitar o cruzamento entre animais altamente endogâmicos (DEASSIS *et al.*, 2009).

1.6 Análises genéticas populacionais

As análises genéticas populacionais vão focar dois aspectos, um relacionado a validação dos sistemas de microssatélites recomendados pela ISAG na raça, uma vez que a padronização internacional foi baseada no puro sangue inglês, e outra para caracterização da estrutura genética da população. Estas análises são fundamentais para a resolução de casos forenses ou de paternidade, além de prover as informações necessárias a programas de conservação da diversidade genética da raça (HUSTON, 1998).

Os parâmetros investigados são heterozigosidade observada e esperada, conteúdo de informação polimórfica (PIC), poder de exclusão (PE), índice típico de paternidade (IP), equilíbrio de Hardy-Weinberg (EHW), análise de variância molecular (AMOVA) e estatísticas F (F_{ST} , F_{IT} e F_{IS}), os quais são amplamente usados na literatura e serão discutidos abaixo.

A heterozigosidade para um marcador é a capacidade de um indivíduo ser heterozigoto para aquele *locus*, ou seja, carregar dois alelos diferentes. Esta é uma medida simples da variação na população e depende do número e da frequência dos alelos na mesma população (Ellegren, 2004). Um marcador é considerado altamente informativo em uma população quando apresenta heterozigosidade maior que 70% (Ott, 1992).

O poder de exclusão indica a capacidade de um sistema de testes excluir uma paternidade ou maternidade falsamente atribuída, ou seja, é a probabilidade de se ter uma exclusão precisa de parentesco, quando um indivíduo e seus pais são analisados (HUSTON, 1998; BIERMAN, 2010).

O índice de paternidade indica quantas vezes é mais provável que um indivíduo, que está sendo testado, seja realmente o pai biológico, do que um indivíduo selecionado aleatoriamente. Geralmente, um IP menor que 1 é um indicativo de inexistência de parentesco entre os envolvidos (HUSTON, 1998).

O conteúdo de informação polimórfica é um indicador da qualidade do marcador para estudos genéticos e uma medida do grau de polimorfismo do marcador. Ele depende apenas do número de alelos e sua frequência para cada marcador (GUO e ELSTON 1999). De acordo com a classificação de Botstein *et al.* (1980), marcadores com valores de PIC superiores a 0,5 são considerados muito informativos, valores entre 0,25 e 0,50 mediamente informativos e inferiores a 0,25

pouco informativos e, segundo Bierman (2010), marcadores com PIC abaixo de 0,5 não são bons para teste de paternidade, uma vez que são pouco informativos.

Um *locus* é dito polimórfico quando o alelo mais comum tem frequência inferior ou igual a 0,99. *Loci* altamente variáveis são utilizados na análise de DNA devido a sua capacidade de diferenciar indivíduos. Uma amostra populacional, idealmente aleatória e representativa é usada para se determinar a frequência de cada alelo e de cada genótipo de um determinado *locus* para essa população, uma vez que os alelos podem ter frequências diferentes em populações diferentes (MENEZES *et al.*, 2005; HUSTON, 1998).

Os pressupostos do Equilíbrio de Hardy-Weinberg (EHW) são que a população considerada seja grande, com acasalamento aleatório, sem atuação da seleção, mutação, migração ou deriva genética. Se uma população segue estes pressupostos, as frequências genóticas correspondem ao estimado considerando-se as frequências alélicas. O princípio do equilíbrio de Hardy-Weinberg estipula que após uma geração de cruzamento, as frequências genóticas para um *locus* com dois alelos A e a, devem seguir a seguinte fórmula:

$$(p+q)^2 = p^2+2pq+q^2$$

Onde: p^2 é a frequência esperada do genótipo AA

$2pq$ é a frequência esperada do genótipo Aa

q^2 é a frequência esperada do genótipo aa

Assim, se as frequências de homozigotos observadas se igualam ao quadrado de suas frequências gênicas, a população está em equilíbrio de Hardy-Weinberg e se elas não se igualam, a população não está em equilíbrio (RIDLEY, 2006).

Uma população pode se desviar do EHW pela ação desses fatores externos e este desvio pode ser determinado através de testes simples de χ^2 comparando-se a heterozigosidade observada (H_o) e a esperada (H_e) (RIDLEY, 2006).

A diversidade genética pode ser avaliada pela análise de variância molecular (AMOVA) e pelas estatísticas F (F_{ST} , F_{IS} e F_{IT}). A AMOVA fornece um quadro geral para a análise da estrutura genética de populações e é baseada na variância das

freqüências alélicas entre populações. A partição da variância pode ser medida através das estatísticas F (MICHALAKIS e EXCOFFIER, 1995).

F é o estimador do grau de diferenciação entre populações. Os F de Wright descrevem a variação da estrutura genética dentro e entre as populações e seguem a seguinte fórmula:

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

O Fst é o parâmetro de interesse na análise da diferenciação genética das subpopulações e além de ser uma estatística descritiva, está diretamente relacionado a variância da freqüência alélica entre as populações e inversamente relacionado com o grau de semelhança entre os indivíduos dentro das populações. Assim, Fst baixo significa que as freqüências dentro de cada população são muito semelhantes e Fst alto, significa que as freqüências alélicas são muito diferentes. Resumidamente, o Fst mede o grau de diferenciação genética entre populações devido às diferenças de freqüências entre elas (WRIGHT, 1951; NEI, 1973; HOLSINGER e WEIR, 2009).

O Fis ou coeficiente de endogamia intrapopulacional é estimado com base na heterozigosidade observada e esperada. Ele equivale à média do desvio das freqüências genotípicas esperadas sob o pressuposto de Equilíbrio de Hardy-Weinberg dentro das populações. Assim denominado como uma medida do grau de endogamia dentro das populações, seus valores podem ser positivos indicando uma deficiência de heterozigotos ou negativos indicando um excesso de heterozigotos (HOLSINGER e WEIR, 2009; OLIVEIRA, 2007; WEIR, 2002).

Fit é a medida de heterozigosidade de um indivíduo em relação ao total da população, ou seja, equivale ao desvio das freqüências esperadas sob o pressuposto de Equilíbrio de Hardy-Weinberg em relação à população inteira (WEIR, 2002; HOLSINGER e WEIR, 2009).

Em conjunto, estas estatísticas descrevem a distribuição da diversidade genética na população e permitem inferir sobre a existência de estruturação na população.

1.7 Justificativa

A raça Mangalarga Marchador é a principal raça de cavalos em termos de número de indivíduos no Brasil. A manutenção de seu valor econômico depende da preservação do seu padrão e de sua identidade genética, o que se busca através da exigência de teste de paternidade por ocasião do registro definitivo dos animais. Para o registro definitivo, é exigido o teste de paternidade animal baseado em DNA e a qualidade desses testes depende da estimativa de parâmetros como, por exemplo, frequências alélicas para a população em estudo. Assim, o estudo da informatividade dos microssatélites utilizados em testes de paternidade para essa raça torna-se pertinente e marcadores com baixa informatividade podem ser detectados e substituídos por outros mais informativos no intuito de tornar o exame mais preciso.

Este é o primeiro estudo avaliando o desempenho do painel de marcadores moleculares recomendado pela ISAG na raça Mangalarga Marchador. Algumas vezes, este número de microssatélites analisados não é suficiente para esclarecer a relação de parentesco, fazendo-se necessária a utilização de outros marcadores. A caracterização desses marcadores para a raça Mangalarga Marchador é importante, uma vez que além de existirem poucos estudos baseados em microssatélites para essa raça, estes trabalhos tem um tamanho amostral pequeno e analisam poucos marcadores (SILVA FILHO *et al.*, 2007; DEASSIS *et al.*, 2009).

Além disto, a disponibilidade de um banco de dados tão amplo, permitirá conhecer melhor a estrutura genética da população de cavalos Mangalarga Marchador das regiões brasileiras estudadas. Os resultados obtidos gerarão as estatísticas básicas usadas nos testes de verificação de parentesco, informações sobre a diversidade genética desses animais e como ela está distribuída. Além disso, podem ajudar no manejo através da determinação do grau de diferenciação genética entre populações e ainda contribuir com outros estudos que tenham como base a raça Mangalarga Marchador.

1.8 Objetivos

- Avaliar a eficiência do painel de marcadores microssatélites em uso para a raça MM;
- Analisar a estrutura genética da população brasileira de cavalos MM considerando cada estado como uma subpopulação.

1.8.1 Objetivos específicos

- Estimar as frequências alélicas e genotípicas para cada marcador;
- Determinar a heterozigosidade esperada e observada;
- Testar a hipótese de Equilíbrio de Hardy-Weinberg (EHW) para os diversos marcadores;
- Estimar o conteúdo de informação polimórfica (PIC), calcular o poder de exclusão (PE) e o índice de paternidade (PI) de cada marcador;
- Caracterizar a estrutura genética da população através da análise de variância molecular (AMOVA) e das estatísticas de F (F_{ST} , F_{IS} , F_{IT}).

2 MATERIAIS E MÉTODOS

2.1 Amostra

Todos os dados analisados neste trabalho foram obtidos de resultados de testes de verificação de parentesco realizados no laboratório Linhagen Produtos e Serviços em Biotecnologia LTDA.

As amostras de sangue ou crina dos animais foram coletadas por técnicos credenciados pela ABCCMM e chegaram ao laboratório acompanhadas de uma ficha de cadastro, preenchida pelo técnico responsável pela coleta, com as seguintes informações:

- nome do animal, número de registro na ABCCMM, sexo;
- nome e endereço do proprietário;
- assinatura do técnico responsável pela coleta e data da mesma;
- nos casos de verificação de parentesco, nome dos supostos pais.

Todos os procedimentos realizados com as amostras (recepção, triagem, cadastro, processamento e armazenamento) foram feitos de acordo com procedimentos operacionais padrão desenvolvidos pela Linhagen Produtos e Serviços em Biotecnologia LTDA.

Este laboratório possui um banco de dados com os genótipos de mais de 100.000 animais da raça MM. Para a seleção da amostra a ser analisada, primeiro foi feita a união do banco de dados fornecidos pelo Laboratório Linhagen Produtos e Serviços em Biotecnologia LTDA (nome e dados da genotipagem dos animais) com o fornecido pela Associação Brasileira dos Criadores do Cavallo Mangalarga Marchador (nome, local, data de nascimento e sexo dos animais). Do arquivo criado, foram filtrados os animais nascidos em 2008 e 2009, restando assim uma amostra de 12.111 animais, cuja distribuição por cidade está demonstrada na Fig. 4. Esse critério de seleção da amostra foi adotado no sentido de se evitar sobreposição de gerações, uma vez que equinos são usados para reprodução a partir de 2 anos de idade. Além disto, este é o período com o maior número de animais. Um segundo filtro foi feito, onde foram retirados todos os animais que apresentavam ausência de algum dado, finalizando-se assim com uma amostra de 6.092 animais.

Essa amostra está distribuída em todas as regiões do Brasil, sendo que a grande concentração se encontra na região Sudeste, principalmente em Minas Gerais e a menor na região Norte, a qual foi representada por apenas um estado, Pará.



Figura 4 - Distribuição geográfica da amostra

Nota: Os pinos representam as populações, conforme a posição de GPS da cidade de nascimento do animal, informado no registro da ABCCMM. Fonte do mapa: Google Earth.

2.2 Métodos moleculares

Embora o presente trabalho tenha sido desenvolvido com as informações de genótipos já depositados na base de dados, a metodologia usada em sua geração será brevemente descrita, para melhor avaliação do trabalho.

O DNA foi extraído a partir de amostras de sangue ou crina dos animais. A extração baseia-se primeiro em quebrar as membranas das células, solubilizar o DNA e, posteriormente, eliminar através de processos químicos (proteínase K e precipitação salina), as proteínas e outras macromoléculas (adaptado de Miller *et al.*, 1988). Para isso, foram utilizados protocolos específicos para cada tipo de amostra. Este método fornece grandes quantidades de DNA a partir de pequenas amostras de sangue ou crina e permite que este material seja armazenado por longos períodos.

Para a amplificação dos microssatélites foi utilizada a técnica de PCR. A reação de PCR consiste em fazer cópias de DNA “*in vitro*”, usando-se os elementos básicos do processo de replicação do DNA e é baseada na repetição cíclica e automática de três simples etapas: desnaturação da fita dupla do DNA molde a ser amplificado, anelamento dos iniciadores nas fitas desnaturadas e a síntese da fita complementar à seqüência molde de DNA pela enzima *Taq* DNA polimerase. Tanto os protocolos de extração quanto os de amplificação foram adaptados da literatura na Linhagen Produtos e Serviços em Biotecnologia LTDA.

As amostras amplificadas foram genotipadas por eletroforese capilar em sequenciador MegaBACE 1000 (*GE Healthcare*) e analisadas com *software* Fragment Profile v1.2 (*GE Healthcare*). Este sequenciador é um sistema de eletroforese capilar, que detecta marcações fluorescentes nos produtos de PCR. Este sistema reduz tempo e custos, permite utilizar amostras com uma concentração pequena de DNA e aumenta a possibilidade de análises. Pode-se trabalhar com os marcadores em reações multiplex, ou seja, fragmentos com o mesmo tamanho podem ser analisados em uma mesma reação desde que estejam marcados com fluorescências diferentes. O MegaBACE 1000 é capaz de detectar quatro fluorescências.

2.3 Banco de dados

Foram criadas duas planilhas em Excel (*Microsoft Corporation*), uma extraída do banco de dados do laboratório Linhagen Produtos e Serviços em Biotecnologia LTDA, contendo as informações de nome e genótipos dos animais, e outra extraída do banco de dados da ABCCMM, com as informações de nome, local e data de nascimento dos animais, sexo e número de descendentes.

Destas planilhas foi criada uma única tabela unindo todas estas informações, utilizando como parâmetro o nome dos animais, uma vez que este é o dado comum às duas planilhas. A partir disso, pode-se obter qualquer combinação de dados de acordo com o interesse do estudo, facilitando assim o acesso às informações necessárias.

2.4 Análises genéticas populacionais

Os parâmetros populacionais foram calculados a partir de um painel de 14 microssatélites, todos com motivos de repetição do tipo dinucleotídeo, os quais fazem parte de um sistema multiplex, que são usados na rotina do laboratório Linhagen Produtos e Serviços em Biotecnologia LTDA e incluem marcadores recomendados pela ISAG (Tab. 1).

Primeiramente, os dados foram convertidos através do *software* Convert 3.1 (GLAUBITZ, 2004), que permite a conversão dos dados em *input* de diferentes programas utilizados em estudos de genética de população. A partir disso, foram estimadas frequências alélicas e genotípicas, número de alelos por *locus* e total, heterozigosidade observada (H_o) e esperada (H_e) para cada marcador, com o programa GENEPOP 4.0 (RAYMOND e ROUSSET, 1995).

Tabela 1 – Microssatélites usados, posição de mapa e referência

Microssatélite	Localização cromossômica	Referência
AHT04*	24	BINNS <i>et al.</i> , 1995
AHT05*	8	BINNS <i>et al.</i> , 1995
ASB02*	15	BREEN <i>et al.</i> , 1997
ASB17**	2	BREEN <i>et al.</i> , 1997
ASB23**	3	IRVIN <i>et al.</i> , 1998
CA425	28	EGGLESTON-STOTT <i>et al.</i> , 1997
VHL20*	30	VAN HAERINGEN <i>et al.</i> , 1994
HMS01	15	GUERIN <i>et al.</i> , 1994
HMS02**	10	GUERIN <i>et al.</i> , 1994
HMS03*	9	GUERIN <i>et al.</i> , 1994
HMS06*	4	GUERIN <i>et al.</i> , 1994
HMS07*	1	GUERIN <i>et al.</i> , 1994
HTG04*	9	ELLEGREN <i>et al.</i> , 1992
HTG07	4	MARKLUND <i>et al.</i> , 1994

Nota: *marcadores recomendados pela ISAG, **marcadores acrescentados pela ISAG em 2010.

As frequências alélicas, o conteúdo de informação polimórfica (PIC), poder de exclusão (PE) e índice de paternidade (IP) para cada marcador foram estimados utilizando-se o programa Powerstats 1.2 (*Promega Corporation*). Esse programa é de fácil manuseio, não sendo necessário criar-se arquivo de entrada, uma vez que é necessário informar apenas os indivíduos a serem analisados com seus respectivos genótipos por *locus*. Todos estes parâmetros analisados medem a informatividade do marcador.

As fórmulas utilizadas foram:

- Freqüências alélicas (P_i):

$$P_i = \frac{n_i}{\sum n} \quad (1)$$

onde n_i é o número de vezes que o alelo aparece e n é o total de alelos averiguados para o marcador.

- Conteúdo de informação polimórfica (PIC):

$$PIC = 1 - \sum_{i=1}^n \sum_{j=1}^n P_{ij}^2 - \left(\left(\sum_{i=1}^n \sum_{j=1}^n P_{ij}^2 \right)^2 + \sum_{i=1}^n \sum_{j=1}^n P_{ij}^4 \right) \quad (2)$$

onde i e j são as freqüências dos alelos na população n e P_{ij} é a freqüência dos genótipos para um determinado marcador genético.

- Poder de exclusão (PE):

$$PE = h^2(1 - 2hH^2) \quad (3)$$

em que, h é a freqüência de heterozigotos e H a de homozigotos para um determinado *locus*;

- Índice de paternidade (IP):

$$IP = \frac{(H+h)}{2H} \quad (4)$$

onde h é a frequência de heterozigotos e H a frequência de homozigotos para um determinado *locus*.

O IP combinado deste painel de microssatélites foi estimado pela multiplicação dos resultados individuais do IP de cada marcador e a probabilidade de paternidade pela seguinte fórmula:

$$\frac{IP}{IP + 1}$$

Para a análise da estrutura genética da população, a AMOVA, as estatísticas F e o teste do equilíbrio de Hardy-Weinberg, foram conduzidas pelo programa Arlequin (Excoffier *et al.*, 2007).

Para o teste do equilíbrio de Hardy-Weinberg, foi usada a correção de Bonferroni, no qual o valor de alfa de 0,05 foi dividido pelo número de testes (RICE, 1989). Assim, todos os marcadores com valor de $P > 0,0036$, encontram-se em Equilíbrio de Hardy-Weinberg.

Utilizando-se o software Microchecker versão 2.2.3, foi averiguada a existência de alelos nulos, a exclusão alélica e os erros de nomeação alélica devido a *stuttering* na amostra subdividida por estado de nascimento do animal.

3 RESULTADOS

Os dados foram analisados em conjunto e separados por estado da Federação. Os estados com menos de 30 animais foram excluídos da análise. As frequências alélicas, genotípicas, H_o , H_e , PIC, PE e IP foram estimadas para a amostra total. Frequências alélicas, AMOVA, F_{ST} , F_{IS} e F_{IT} foram estimados para a amostra inteira e subdividida por estado, os quais foram tratados como subpopulações.

3.1 Avaliação do painel de microssatélites para identificação animal e averiguação de paternidade

Na Tab. 2, encontram-se os dados relativos ao tamanho amostral (N) das populações de cada estado que foram utilizadas nas análises.

Tabela 2 - Tamanho amostral por estado

Estado	N
Alagoas	67
Bahia	768
Espírito Santo	317
Goiás	66
Minas Gerais	2373
Mato Grosso do Sul	70
Pará	34
Pernambuco	320
Paraná	157
Rio de Janeiro	686
Rio Grande do Norte	45
Rio Grande do Sul	49
Santa Catarina	152
Sergipe	36
São Paulo	952
Total	6092

Todos os 14 microssatélites analisados foram polimórficos para o cavalo Mangalarga Marchador. Um total de 129 alelos foi encontrado para o conjunto dos marcadores analisados, com um número médio de 9,2 alelos por *locus*, variando de 5 para o HTG7 a 20 para o ASB17. A heterozigosidade observada média foi de 0,702, variando de 0,536 para o HTG7 a 0,813 para o ASB17, enquanto a heterozigosidade esperada média foi de 0,714, variando de 0,541 para o HTG7 a 0,821 para o ASB17.

A média do PE foi 0,443 para o conjunto de microssatélites avaliados, variando de 0,221 para o HTG7 a 0,624 para o ASB17 e o PE global foi de 0,9998. O IP combinado foi de 2.391, sendo o menor valor de 1,08 para o HTG7 e maior valor de 2,68 para o ASB17, o que remete a uma Probabilidade de Paternidade de 99,97%. Para o PIC médio foi encontrado um valor de 0,68 e seguindo os resultados dados pelos parâmetros anteriores, o HTG7 com 0,5 e o ASB17 com 0,8 apresentaram, respectivamente, o menor e o maior resultado.

Na Tab. 3, são mostrados os resultados relativos a número de alelos, heterozigosidade observada (H_o) e esperada (H_e), PE, IP e PIC dos marcadores estudados.

As frequências alélicas para os *loci* usados são mostradas na Tab. 4.

Tabela 3 – Caracterização dos marcadores utilizados

Locus	Nº Alelos	Ho	He	PE	IP	PIC
AHT04	9	0,765	0,774	0,535	2,13	0,75
AHT05	8	0,753	0,761	0,514	2,02	0,73
ASB02	11	0,790	0,810	0,581	2,38	0,79
ASB17	20	0,813	0,821	0,624	2,68	0,8
ASB23	10	0,765	0,780	0,536	2,13	0,75
CA425	10	0,657	0,665	0,366	1,46	0,60
HMS1	7	0,652	0,654	0,358	1,44	0,59
HMS2	9	0,628	0,649	0,325	1,34	0,61
HMS3	8	0,747	0,763	0,504	1,97	0,72
HMS6	7	0,570	0,589	0,256	1,16	0,55
HMS7	8	0,755	0,767	0,518	2,04	0,73
HTG4	6	0,637	0,650	0,338	1,38	0,60
HTG7	5	0,536	0,541	0,221	1,08	0,5
VHL20	11	0,761	0,776	0,529	2,09	0,75
Média	9,2	0,702	0,714	0,443	1,81	0,68
Combinado	-	-	-	0,9998	2.391	-

Tabela 4 - Frequência alélicas para cada um dos marcadores genéticos estudados

Alelo	AHT4	AHT5	ASB2	ASB17	ASB23	CA425	HMS1	HMS2	HMS3	HMS6	HMS7	HTG4	HTG7	VHL20
A			0.0023											
B			0.0266	0.0001										
C			0.0452											
F				0.1535										
G				0.0240	0.0313	0.0002								
H	0.1212			0.0046		0.0015		0.2167						
I	0.1343		0.0961	0.0054	0.0634	0.1106	0.0154	0.1070	0.1239					0.2063
J	0.1688	0.2102		0.0001	0.3230	0.0766	0.3814	0.0047			0.0127			0.0042
K	0.3906	0.0253	0.0519	0.0016	0.1808	0.0007	0.0210	0.0539		0.0463	0.0223	0.0353	0.1258	0.0018
L	0.0256	0.0109	0.0009	0.0034	0.1114	0.0099	0.0004	0.5353		0.0184	0.2714	0.3075		0.0794
M	0.0057	0.0863	0.2194	0.1097	0.0001	0.3769	0.4315	0.0410	0.2157	0.1634	0.2052	0.4913	0.1123	0.3875
N	0.0706	0.3803	0.3158	0.0958		0.4184	0.1074		0.0104	0.0006	0.2738	0.0779	0.1154	0.0492
O	0.0823	0.1662	0.1201	0.0252		0.0050		0.0089	0.2429	0.1666	0.2038	0.0089	0.6463	0.0167
P	0.0008	0.0003		0.0013		0.0001		0.0013	0.3359	0.5955	0.0100	0.0791	0.0002	0.0657
Q		0.1205	0.1160	0.1950	0.0002		0.0429		0.0602	0.0092	0.0007			0.1279
R			0.0057	0.2972				0.0313	0.0109					0.0414
S				0.0006	0.0342				0.0002					0.0200
T					0.0023									
U				0.0792	0.2534									
V				0.0002										
W				0.0001										
X				0.0002										
Y				0.0028										

Nota: Em negrito, está marcado o alelo mais freqüente de cada locus.

A distribuição das freqüências alélicas variou bastante entre os *loci*. Alguns marcadores apresentaram um único alelo com freqüência muito alta, o que reduziu sua informatividade. Isto pode ser observado para os marcadores HTG7, HMS2 e HMS6. Outros marcadores apresentam dois alelos com freqüência alta, como observado para CA425, HMS1 e HTG4.

Além disto, foi avaliado o número de alelos raros por *locus* na população em geral (Tab. 5). Todos os marcadores apresentaram alelos com freqüência menor do que 5% e menor do que 1%.

Tabela 5 - Distribuição dos números de alelos com freqüência inferior a 5 e 1% conforme o locus

Marcador genético	Nº Total de alelos	Nº Alelos com freqüência <5%	Nº Alelos com freqüência <1%
AHT4	9	3	2
AHT5	8	3	1
ASB2	11	5	3
ASB17	20	13	11
ASB23	10	5	3
CA425	10	6	6
HMS1	7	4	1
HMS2	9	5	3
HMS3	8	3	1
HMS6	7	4	2
HMS7	8	4	1
HTG4	6	2	1
HTG7	5	1	1
VHL20	11	6	2

3.2 Estrutura genética da população

O Equilíbrio de Hardy-Weinberg foi testado para todos os marcadores na população como um todo e em cada população subdividida por estado de nascimento do animal. Foi observado que alguns marcadores em cinco subpopulações (ES, MG, MS, PA, PE) apresentavam desvios significativos do Equilíbrio de Hardy-Weinberg (Qdo 1).

Quadro 1- Distribuição dos marcadores que apresentam desvio do EHW por Estado

Estado	Marcadores que apresentaram desvio para o EHW
MG	AHT4 - ASB2 - HMS2 - HMS3 - HMS7
MS	CA425
PA	HMS2

Os marcadores moleculares ASB2; ASB23; CA425; HMS2; HMS6 e VHL20 apresentaram desvio do EHW na população como um todo.

A variabilidade genética se distribuiu, conforme avaliado pela AMOVA, em um percentual de variação de 0,6 entre as populações e de 99,4 dentro das populações. Na Tab. 6, são mostrados os índices de fixação (F_{ST}) por *locus*. O F_{ST} médio para todos os *loci* foi de 0,006. O valor médio de F_{IS} para todos os *loci* foi de 0,01244 e de F_{IT} foi de 0,01836.

Tabela 6 - Estatísticas F para cada locus com respectivos valores de p

Marcador	F_{ST}	p- value	F_{IS}	p- value	F_{IT}	p-value
AHT4	0,00508	0,00000	0,00772	0,10851	0,01276	0,03109
AHT5	0,00474	0,00000	0,00715	0,15228	0,01185	0,04921
ASB17	0,00730	0,00000	0,00423	0,23277	0,01150	0,03990
ASB2	0,00572	0,00000	0,01953	0,00069	0,02513	0,00000
ASB23	0,00649	0,00000	0,01437	0,01545	0,02077	0,00158
CA425	0,00378	0,00000	0,00801	0,17099	0,01175	0,09980
HMS1	0,00293	0,00010	0,00132	0,45079	0,00424	0,34168
HMS2	0,00818	0,00000	0,02719	0,00030	0,03515	0,00000
HMS3	0,00558	0,00000	0,01646	0,00851	0,02194	0,00218
HMS6	0,00947	0,00000	0,02431	0,00188	0,03355	0,00010
HMS7	0,00722	0,00000	0,01036	0,07574	0,01751	0,01089
HTG4	0,00725	0,00000	0,01476	0,03594	0,02191	0,00743
HTG7	0,00472	0,00000	0,00478	0,28733	0,00948	0,15129
VHL20	0,00565	0,00000	0,01398	0,01168	0,01955	0,00149
Média	0,00600	0,00000	0,01244	0,11092	0,01836	0,04181

A análise de alelos nulos pelo Microchecker só mostrou evidências de alelos nulos na população de MG para os marcadores AHT5, ASB2, HMS2, HMS3, HMS6, HTG4, CA425, na população da BA, para os marcadores HMS3 e HTG7 e na população de PE, para os marcadores HTG7 e VHL20. Não foram encontradas evidências de exclusão alélica ou de erros de nomeação gerados por *stuttering*.

4 DISCUSSÃO

4.1 Avaliação do painel de microssatélites para identificação animal e averiguação de paternidade

Embora diversos estudos tenham avaliado o painel de marcadores moleculares recomendados pela ISAG para verificação de parentesco/identificação individual em eqüinos em diversas raças, o conjunto de marcadores efetivamente usado varia bastante entre as publicações, uma vez que inclui outros marcadores na análise, o que dificulta a comparação direta do conjunto aqui analisado (BEHL *et al.*, 2006; LUÍS *et al.*, 2007; CHOI *et al.*, 2008; TAKASU *et al.*, 2011)

Outro problema diz respeito aos tamanhos amostrais, que são bastante reduzidos em alguns trabalhos, particularmente naqueles que descrevem raças ameaçadas de extinção.

Em função disto, serão discutidos preferencialmente artigos com maior número de marcadores do painel da ISAG e/ou maiores amostras. Além disto, os presentes resultados serão comparados aos descritos por SILVA FILHO *et al.*, 2007 e de DEASSIS *et al.*, 2009, por serem estudos de estrutura genética de populações com marcadores moleculares na raças MM.

A maioria dos marcadores analisados mostraram um grande número de alelos, sendo a média de 9,2 e os extremos observados para HTG7, com 5 alelos, e para ASB17, com 20 alelos. Na Tab. 7 abaixo, encontram-se os resultados obtidos por outros autores para outras raças, utilizando um painel com pelo menos 10 marcadores em comum com o presente estudo.

Tabela 7 - Número médio de alelos em algumas raças

Raça	No de marcadores	No de indivíduos	No médio de alelos	Ho	PIC	PE	Referências
MM	14	6.092	9,2	0,702	0,68	0,9998	Presente trabalho
Jeju	20-14*	79	7,45	0,685	0,659	0,9999	CHOI <i>et al.</i> , 2008
Thoroughbred	14 -12*	1.285	6,36	0,663	0,652	0,9998	LEE e CHO, 2006
Zanskari	25 -10*	42	5,8	0,61	0,78	-	BEHL <i>et al.</i> , 2006
Sorraia	22 -10*	131	3,3	0,450	0,415	0,999	LUIS <i>et al.</i> , 2007

Nota: * n° de marcadores em comum com o presente estudo

Através dessa comparação, é possível verificar-se que a média do número de alelos por *locus* e a heterozigosidade observada do cavalo MM, foi maior que a encontrada em outras raças, apesar do painel analisado variar bastante. Isso pode ter ocorrido devido ao maior tamanho amostral analisado no presente estudo, que permitiu a identificação de diversos alelos raros. O maior valor de PIC foi encontrado para a raça Zanskari, mas esta analisou um número muito maior de marcadores. Em relação ao PE, o maior valor encontrado foi para cavalos Jeju o que também pode ser devido ao maior número de marcadores analisados para esta raça.

Todos marcadores analisados neste painel em estudo apresentaram alelos raros, sendo que para uma frequência abaixo de 1% a média foi de 2,7 e abaixo de 5% foi de 4,6. O ASB17 foi o marcador que apresentou o maior número de alelos raros. Os demais artigos incluídos nesta comparação não fazem referência a alelos raros.

A diversidade genética, de maneira geral, pode ser medida pela heterozigosidade média. Diferenças na heterozigosidade observada podem ser explicadas pelos marcadores microssatélites utilizados e pelo tamanho amostral, assim como pela composição da amostra em termos de aleatoriedade e representatividade (CHOI *et al.*, 2008). Isso pode ser

observado nos resultados já descritos na literatura para a raça MM, em que valores mais altos de heterozigidade média (0,839) foram encontrados por Silva-Filho *et al.*, (2007) analisando 3 marcadores em 79 animais e valores menores (0,637) foram observados por DeAssis *et al.*, (2009) com um painel com 10 microssatélites em 115 indivíduos. No presente trabalho, analisando 14 marcadores em 6.042 animais foi encontrada uma heterozigidade média de 0,702, um valor intermediário aos dois trabalhos citados anteriormente.

Dentre os 14 marcadores analisados, oito foram altamente informativos com PIC maior que 0,7 e apresentaram valores de PE acima de 0,5, sugerindo que este seja um painel eficiente para testes de paternidade em cavalos Mangalarga Marchador.

De acordo com o IP combinado para este painel, tem-se 99,97% de chance de que o animal acreditado como pai seja realmente o pai biológico.

No presente estudo, o marcador ASB17 deve ser destacado, uma vez que este é o mais informativo do painel analisado, o que pode ser confirmado pelo seu maior valor de PIC (0,8), PE (0,624) e IP (2,68). Com 20 alelos, este é o marcador que possui o maior número de alelos dos 14 analisados e está entre os três novos marcadores recomendados pela ISAG. Por isso, este foi considerado o melhor marcador do painel analisado para a raça MM. Em outras raças, o ASB17 também está entre os mais informativos (LEE e CHO, 2006; CHOI, *et al.*, 2008; LING, *et al.*, 2010).

Em contrapartida, o microssatélite menos informativo desse painel é o HTG7, apresentando os menores valores de PIC (0,5), PE (0,215) e IP (1,06), Além de possuir o menor número de alelos, a frequência destes é muito mal distribuída, sendo que um único alelo representa 65,1% da frequência alélica deste marcador. Assim, para o MM este foi o pior marcador e apesar de ser utilizado por outros laboratórios de paternidade, ele não está entre os recomendados pela ISAG. Dentre os recomendados, o HMS6 é o menos informativo com valores de PIC, PE e IP bem próximos do HTG7.

Pouquíssimos trabalhos foram publicados investigando o desempenho de painéis de marcadores microssatélites na raça MM. O autor Silva Filho e colaboradores (2007) estudaram uma amostra de 79 animais de um plantel do interior do estado de São Paulo, utilizando os marcadores VHL20, HTG4 e HMS7, que são utilizados na rotina de paternidade de cavalos. Os três

marcadores foram considerados como bastante informativos. No presente estudo, encontramos valores aproximados de heterozigosidade e PIC para o VHL20 e HMS7, confirmando a boa informatividade destes, mas para o marcador HTG4 encontramos valores bem menores, indicando uma baixa informatividade. Os alelos mais freqüentes de cada *locus*, não foram os mesmos relatados neste estudo.

Já DeAssis e colaboradores (2009) estudaram um conjunto diferente de microssatélites, em uma amostra aleatória de 115 animais, não aparentados e nascidos de uma mesma estação de monta, de 18 plantéis de Minas Gerais. O único em comum com o presente estudo foi o VHL20. Os valores de heterozigosidade foram semelhantes para este marcador, o PIC foi igual, mas o número de alelos encontrado foi quase o dobro do publicado por DeAssis. Isso pode ser devido ao maior tamanho amostral utilizado no presente estudo e a origem da amostra, permitindo assim encontrar alelos raros.

4.2 Estrutura genética da população

Esta é a maior amostra de uma raça de eqüinos estudada, quanto à estrutura de população, utilizando-se marcadores microssatélites.

Para uma análise exploratória inicial da estrutura genética da raça MM, a amostra foi dividida conforme o estado de nascimento do animal. Os resultados apresentados pela AMOVA mostraram que a maior variação foi encontrada dentro das populações e não entre elas, sugerindo que as populações estudadas são muito parecidas geneticamente. Os testes de EHW para cada marcador em cada população identificaram alguns marcadores fora do EHW (**Qdo 1-** acima). A população de Minas Gerais foi a que apresentou mais marcadores que desviaram-se do EHW, podendo isso ser devido ao efeito de Wahlund, no qual subpopulações podem ter sido analisadas em conjunto e as diferenças nas freqüências alélicas entre essas subpopulações podem ser observadas pela deficiência de heterozigotos.

O valor do F_{ST} para cada *locus* na população como um todo foi muito baixo e altamente significativo, com uma média para todos os *loci* de 0,006.

Este resultado sugere que o fluxo gênico entre estas populações deve ser alto, levando à homogeneidade.

Esses achados sugerem que se há subestruturação, sua base provavelmente não é geopolítica. Isto faz sentido, uma vez que principalmente na região Sudeste as populações são contíguas, particularmente as de São Paulo e Minas Gerais. Além disto, as técnicas de reprodução assistida, em uso crescente na raça, favorecem a homogeneização, mesmo de populações distantes geograficamente.

Através da comparação das populações dos estados a partir do F_{ST} par a par, identificou-se todos os resultados com valor de p altamente significativo. Os valores de F_{ST} observados para cerca de 1/3 (31/105) dos pares de populações se situam entretanto, abaixo de 0,01. Os valores mais baixos de F_{ST} são observados entre MG-SP, conforme esperado, uma vez que estas são as maiores amostras. Os valores mais altos de F_{ST} par a par foram observados entre MS-AL, MS-PA e MS-RN, que estão entre as menores amostras.

Com o programa Microchecker, só foram encontradas evidências de alelos nulos em alguns *loci* nas populações de três estados, MG, BA e PE. Esse resultado sugere que não existem realmente alelos nulos na população estudada uma vez que essas evidências aconteceram esporadicamente para marcadores diferentes em populações diferentes. Com o mesmo programa, não foram encontradas evidências para exclusão de alelos grandes (*large allele dropout*) ou para erros de identificação alélica causados por *stuttering*.

O tamanho amostral grande, embora vantajoso, criou uma dificuldade para a realização de algumas análises, pela falta de programas que analisem os dados de mais de, geralmente, 1000 indivíduos. Além disto, alguns programas que aceitam o tamanho amostral usado, não geraram resultados em função do tempo computacional muito grande. Assim, este estudo terá continuidade, pois algumas perguntas permaneceram, como por exemplo encontrar outras formas de estruturação geográfica, mas não-geopolítica, ou não-geográfica.

5 CONCLUSÕES

O painel analisado se mostrou eficiente para estudos de verificação de parentesco na raça Mangalarga Marchador, com valores elevados de heterozigosidade observada, PIC médio, PE e IPC. O marcador mais informativo foi o ASB17 com PIC de 0,8; PE de 0,624 e IP de 2,68 e os menos informativos foram o HTG7 (PIC-0,5, PE-0,5, IP-1,18) e o HMS6 (PIC-0,5, PE-0,5, IP-1,18), sendo que HTG7 e HMS6 possuem um único alelo com frequência alélica de 65,4% e 59,5% respectivamente.

Seis marcadores apresentaram desvios das proporções esperadas sob o pressuposto do Equilíbrio de Hardy-Weinberg na população como um todo e nas subpopulações apenas os estados do PA, MS e MG mostraram marcadores que desviaram do equilíbrio, sendo MG com mais desvios.

Conforme a AMOVA a maior diferenciação genética se encontra entre as populações e não dentro delas. Além disso, a maioria das comparações par a par mostrou F_{ST} muito baixos. Os estados de SP e MG apresentaram os menores valores, sendo os dois estados mais próximos geneticamente.

Utilizando-se a subdivisão da população por estado, não se tem evidências de que essa população esteja estruturada, assim, se existe estruturação sua base não é geopolítica.

REFERÊNCIAS ELETRÔNICAS E BIBLIOGRÁFICAS

ALMEIDA, P. A. R. *Diversidade genética e diferenciação das raças portuguesas de ovinos com base em marcadores de DNA – microssatélites: uma perspectiva de conservação*. 2007. 333 f. Tese (Doutorado em Ciência Animal) - Faculdade de Trás-os-Montes e Alto Douro, Vila Real, Filoso, 2007.

ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DO CAVALO MANGALARGA MARCHADOR. Belo Horizonte, 2010. Disponível em <<http://www.abccmm.org.br/>>. Acesso em: 28 maio 2010.

BEHL, R.; BEHL, J.; GUPTA, N.; GUPTA, S. C.; AHLAWAT, S. P. S.; RAGNEKAR, M.; AHMED, Z. Genetic characterization of Zanskari breed of horse. *Journal of Genetics*, v. 85, n. 3, December 2006.

BERNOCO, D.; OLIVEIRA, D. A. A.; KUABARA, M. Y. Análise de parentesco em cavalos de grupos sanguíneos e polimorfismos bioquímicos à tipagem por DNA. *Rev. Bras. Reprod. Anim.*, v. 21, n. 4, p. 38-44, 1997.

BIERMAN, A. *Refinement and validation of a microsatellite based identification and parentage testing panel in horses*. 2010. 500 f. Tese (Doutorado em Psicologia) - Faculdade de Filosofia e Ciências Humanas, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

BINNS, M. M.; UOLMES, N. G; HOLLIMAN, A. M. The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing. *Br. Vet. J.*, n. 151, p. 9-15, 1995.

BOWLING, A. T.; EGGLESTON-STOTT, M. L.; BYRNS, G.; CLARK, R. S.; DILEANIS, S.; WICTUM, E. Validation of microsatellite markers for routine horse parentage testing. *Animal Genetics*, n. 28, p. 247-252, 1997.

BOWLING, A. T.; RUVINSKY, A. *The genetics of the horse*. Edited by BOWLING, A. T. & RUVINSKY, A. CABI Publishing, p. 527, 2000.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Hum. Genet.*, n. 32, p. 314-331, 1980.

BREEN, M.; LINDGREN, G.; BINNS, M. M.; NORMAN, J.; IRVIN, Z.; BELL, K.; SANDBERG, K.; ELLEGREN, H. Genetical and physical assignments of equine microsatellites-first integration of anchored markers in horse genome mapping. *Mamm.Genome*, n. 8, p.267-273, 1997.

CARNEIRO, E.; DUARTE, J. S.; DIAS, M. V. A História do cavalo Mangalarga Marchador. Belo Horizonte: Nova Fronteira, 1991.

CHAUHAN, M.; GUPTA, A. K.; DHILON, S. Genetic diversity and population structure of three Indian horse breeds. *Mol Biol Rep.*, v. 38, p. 3505 – 3511, 2010.

CHOI, S. K.; CHO, C. Y.; YEON, S. H.; CHO, B. W.; CHO, G. J. Genetic characterization and polymorphisms for parentage testing of the Jeju horse using 20 microsatellite loci. *J. Vet. Med. Sci.*, v. 70, n. 10, p. 1111-1115, 2008.

COOGLE, L.; BAILET, E.; REID, R.; Russ, M. Equine dinucleotide repeat polymorphisms at loci LEX002, -003, -004, -005, -007, -008, -009, -010, -011, -013 and -014. *Anim. Genet.*, n. 27, p. 126-127, 1996.

COOGLE, L.; REID, R.; BAILEY, E. Equine dinucleotide repeat loci from LEX025 to LEX033. *Anim. Genet.*, n. 27, p. 289-290, 1996.

COSTA, M. A. P.; BRESSEL, R. M. C.; ALMEIDA, D. B.; OLIVEIRA, P. A.; BASSINI, L. N.; MOREIRA, C. G. A.; MANZKE, V. H. B.; SIEWERDT, F.; MOREIRA, H. L. M. genotyping in the Brazilian Criollo horse stud book: resources and perspectives. *Genet. Mol. Res.*, n. 9, 2010.

COSTA, M.D.; BERGMANN, J.A.G.; RESENDE, A.S.C.; FONSECA, C.G.; FARIA, F.J.C. Análise temporal da endogamia e do tamanho efetivo da população de equinos da raça Mangalarga Marchador. *Arq. Bras. Med. Vet. Zootec*, v. 57, n. 1, 2005.

COSTA, M. D.; BERGMANN, J. A. G.; RESENDE, A. S. C; MARTINS, G. A.; BRETAS, M. S. Caracterização demográfica da raça Mangalarga Marchador. *Arq. Bras. Med. Vet. Zootec*, v. 56, n. 5, p. 687-690, 2004.

DEASSIS, J. B.; DELAAT, D. M.; PEIXOTO, M. G. C. D.; BERGMANN, J. A. G.; FONSECA, C. G.; CARVALHO, M. R. S. Genetic diversity and population structure in Brazilian Mangalarga Marchador horses. *Genet. Mol., Res.*, v. 8, p. 1519-1524, 2009.

DIERKS, C.; LOHRING, K.; LAMPE, V.; WITWER, C.; DROGEMULLER, C.; DISTL, O.; et al. Genome-wide search for markers associated with osteochondrosis in Hanoverian warmblood horses. *Mammalian Genome*, v.18, p. 739-747, 2007.

EGGLESTON-STOTT, M.; L DELVALLE, A.; BAUTISTA, M.; DILEANIS, D.; WICTUM, E.; BOWLING, A. T. Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ412, UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502 and UCDEQ505. *Anim. Genet.*, n. 28, p. 370-371, 1997.

EIZIRIK, E. The effect of habitat fragmentation on the genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (*Panthera onca*). *Molecular Ecology*, n. 19, p. 4906-4921, 2010

ELLEGREN, H.; JOHANSSON, M.; SANDBERG, K.; ANDERSSON, L. Cloning of highly polymorphic microsatellites in the horse. *Anim. Genet.*, n. 23, p. 133-142, 1992.

ELLEGREN, H. Microsatellites: simple sequence with complex evolution. *Nature reviews – Genetics*, v.5, p. 435-445, 2004.

ENGLER, Y. S.; BORGES, M. L.; VENÂNCIO, P.; VENTUROLI, E.; FIGUEIREDO, M. Expectativa de bons negócios. *Mangalarga Marchador*, n. 68, 2010.

EXCOFFIER, L.; HECKEL, G. Computer programs for population genetics data analysis: a survival guide. *Nature Reviews – Genetics*, v. 7, p. 745-758, 2006.

EXCOFFIER, L. G.; LAVAL; SCHNEIDER, S. Arlequin ver. 3.11: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, 47-50, 2007.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotype: application of human mitochondrial DNA restriction data. *Genetics*, v. 131, p. 479-491, 1992.

GANGARAMBHAI, K. P. *Molecular characterization of “Kathiawari” breed of horse using microsatellite markers*. 2004. 75f. Dissertação (Mestrado em Ciência Veterinária) Department of Animal Genetics and Breeding College of Veterinary Science and Animal Husbandry Anand Agricultural University Anand, Anand, 2004.

GLAUBITZ J.C. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes*, v. 4, p. 309-310, 2004.

GUERIN, G.; BERTAUD, M.; AMIGUES, Y. Characterization of seven new horse microsatellites: HMS1, HMS2, HMS3, HMS5, HMS6, HMS7 and HMS8. *Anim. Gene.*, n. 25, p. 62, 1994.

Guo, X.; Elston, R. C. Linkage Information Content of Polymorphic Genetic Markers. *Human Heredity*, n. 49, p. 112-118, 1999.

HAAG, T.; SANTOS, A. S.; SANA, D. A.; MORATO, R. G.; CULLEN JR, L.; CRAWSHAW JR.; DE ANGELO, C.; DI BITETTI, M. S.; SALZANO, F. M.; EIZIRIK, E. The effect of habitat fragmentation on the genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (*Panthera onca*). *Molecular Ecology*, n. 19, p. 4906-4921, 2010.

HOLSINGER, K. E. e WEIR, B. S. Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} . *Nature Reviews Genetics*, n. 10, p. 639-650, 2009.

.

Huston, K. A. Statistical analysis of STR data. *GenePrint*, p.14-15, 1998.

INTERNATIONAL SOCIETY FOR ANIMAL GENETICS. 2009. Disponível em: <<http://www.isag.org.uk/>>. Acesso em: 29 maio 2010.

IRVIN, Z.; GIFFARD, J.; BRANDON, R.; BREEN, M.; BELL, K. Equine dinucleotide repeat polymorphisms at loci ASB21, 23, 25 and 37-43. *Anim. Genet.*, n. 29, p. 67, 1998.

LEE, S.; CHO, G. Parentage testing of Thoroughbred horse in Korea using microsatellite DNA typing. *J. Vet. Sci.*, n. 7, p. 66-67, 2006.

LING, Y. H.; GUAN, W. J.; CHENG, Y. J.; WANG, Y. P.; HAN, J. L.; MANG, L.; ZHAO, Q. J.; HE, X. H.; PU, Y. B.; FU, B. L. Evaluation of the genetic diversity and population structure of Chinese indigenous horse breeds using 27 microsatellite markers. *Animal Genetics*, n. 42, p. 56-65, 2010.

LUÍS, C.; COTHRAN, E. G.; OOM, M. M. Inbreeding and genetic structure in the endangered sorraia horse breed: implications for its conservation and management. *Journal of Heredity*, v. 98, n. 3, p. 232-237, 2007.

MADALENA, F. E. Verificação do parentesco em programas de melhoramento animal. *Cad. Téc. Esc. Vet. UFMG*, n. 27, p. 57-66, 1999.

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. *Anim. Genet.*, n. 25, p. 19-23, 1994.

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

MCMANUS, C.; PAIVA, S.; CORRÊA, P. S.; SEIXAS, L.; MELO, C. B. Estatísticas para descrever Genética de Populações

MELO, D. C.; OLIVEIRA, D. A. A.; SEERIG, A.; CARVALHO, D. C. Aplicações práticas de marcadores microsatélites na caracterização genética e identificação de plantéis de tilápia. *Rev. Bras. Reprod. Anim.*, v. 32, n. 4, p. 220-224, 2008.

MENEZES M. P. R. *Variabilidade e relações genéticas entre raças caprinas brasileiras, ibéricas e canárias*. 2005. 126 f. Tese (Doutorado integrado em Zootecnia) – Universidade Federal da Paraíba, Universidade Federal Rural de Pernambuco e Universidade Federal do Ceará, Paraíba, 2005.

MICHALAKIS, Y.; EXCOFFIER, L. A. generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics Society of America*, 1995.

MILLER, S. A.; DYKES, D. D.; POLESKY, H. F. *A simple salting out procedure for extracting DNA from human nucleated cells*. *Nucl. Acid. Res.* v. 16, p. 1215, 1988.

NEI, M. Analysis of Gene Diversity in Subdivided Populations. *Proc. Nat. Acad. Sci.* v. 70, n. 12, p. 3321-3323, 1973.

OLIVEIRA, J. D.; *Origem, distribuição e relação genética entre populações de*

Capra hircus do Nordeste do Brasil e sua relação com populações do Velho Mundo. 2007. 128 f. Tese (Doutorado em Genética) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2007.

OLIVEIRA, D. A. A.; KUABARA, M. Y. A genealogia, sua importância para o melhoramento e para a produção animal e os métodos de confirmação dos pedigrees. In: PEREIRA, J. C. C. Melhoramento genético aplicado à produção animal. Belo Horizonte, p. 435-444, 2001.

OOM, M. M. Identificação individual e testes de paternidade em raças equinas e caninas. Disponível em: <http://cba.fc.ul.pt/pdf/folheto_caes_equinos.pdf>. Acesso em: 29 maio 2010.

OTT, J. Strategies for characterizing highly polymorphic markers in human gene mapping. *Am. S. Hum. Genet.* n. 51, p. 283-290, 1992.

PENA, S. D. J. O DNA Como (Única) Testemunha em Determinação de Paternidade. *Bioética*, v. 5, n. 2, p. 231-240, 1997.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multi/locus genotype data. *Genetics*, v. 155, p. 945-959, 2000.

RAYMOND. M.; ROUSSET, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity*, n. 86, p,248-249, 1995.

RICE W. W. Analyzing tables of statistical tests. *Evolution* 43, p. 223-225, 1989.

RIDLEY, 2006. *Evolução*. Porto Alegre, 3 ed., Artmed, 2006.

RODRIGUES, S. G.; OLIVEIRA, D. A. A.; TEIXEIRA, C. S.; OLIVEIRA, P. F.; COELHO, E. G. A.; ALVES, C.; VELLOSO, A. P. S.; PEREIRA, J. C. C. Microsatélites BM2113, ILSTS005, ILSTS008, ETH131 e RM88 em testes de verificação de parentesco para bovinos da raça Gir. *Arq. Bras. Med. Vet. Zootec.*, v. 54, n. 3, 2002.

ROMASZKAN, G.; JUNQUEIRA, J. F. D. *O cavalo*. Belo Horizonte: Itatiaia LTDA, 1992.

ROSENBERG, N. A.; PRITCHARD, J. K.; WEBER, J. L.; CANN, H. M.; KIDD, K. K.; ZHIVOTOVSKY, L. A.; FELDMAN, M. W. Genetic structure of human populations. *Science*, v. 298, 2002.

SERENO, F. T. P. S.; SERENO, J. R. B.; VEGA-PLA, J. L.; DELGADO, J. V. DNA testing for parentage verification in a conservation nucleus of Pantaneiro horse. *Genet. Mol. Biol.*, v. 31, n. 1, p. 64-67, 2008.

SILVA FILHO, E.; SCHNEIDER, M. P. C.; SILVA, A. L. C. Variabilidade genética de cavalos baseada em DNA microssatélites. *Revista Trópica – Ciências Agrárias e Biológicas*, v. 1, n. 1, p. 76-87, 2007.

TAKASU, M.; HIRAMATSU, N.; TOZAKI, T.; KAKOI, H.; NAKAGAWA, T.; HASEGAWA, T.; MAEDA, M.; MURASE, T.; MUKOYAMA, H. Genetic characterization of endangered Kiso horse using 31 microsatellite DNAs. *The Journal of Veterinary Medical Science*, 2011.

TAMMELEHT, E.; REMM, J.; KORSTEN, M.; DAVISON, J.; TUMANOV, I.; SAVELJEV, A.; MA"NNIL, P.; KOJOLA, I.; SAARMA, U. Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. *Molecular Ecology*, 2010.

TAUTZ, D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acid Research*, v. 17, p. 6463-6471, 1989.

VAN HAERINGEN, H.; BOWLING, A. T.; SCOTT, M. L.; LENSTRA, J. A.; ZWAAGSTRA, K. A. A highly polymorphic horse microsatellite locus: VHL20. *Anim. Genet.* n.25, p. 207, 1994.

VAN DE GOOR, L. H. P.; VAN HAERINGEN, W. A.; LENSTRA, J. A.
Populations studies of 17 equine STR for forensic and phylogenetic analysis.
Animal Genetics, v. 42, p. 627- 633, 2010.

WEIR, B. S.; HILL, W. G. Estimating F-Statistics. *Annu. Rev. Genet.*, n. 36, 721-
50, 2002.

WRIGHT S. (1951) The genetical structure of populations. *Ann. Eugen.* 15: 323–354.

ANEXOS

Anexo 1 - F_{ST} comparação das populações par a par

Comparisons of pairs of population samples

 List of labels for population samples used below:

Label	Population name
1:	AL
2:	BA
3:	ES
4:	GO
5:	MG
6:	MS
7:	PA
8:	PE
9:	PR
10:	RJ
11:	RN
12:	RS
13:	SC
14:	SE
15:	SP

 Population pairwise F_{ST} s

Distance method: No. of different alleles (FST)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00000														
2	0.01302	0.00000													
3	0.01810	0.00843	0.00000												
4	0.01898	0.00555	0.00809	0.00000											
5	0.01179	0.00404	0.00360	0.00655	0.00000										
6	0.05235	0.03585	0.03202	0.03920	0.02746	0.00000									
7	0.01694	0.01319	0.01939	0.01497	0.01637	0.05434	0.00000								
8	0.01196	0.00412	0.00565	0.00827	0.00294	0.03453	0.01345	0.00000							
9	0.01371	0.01361	0.00798	0.01557	0.00529	0.04011	0.02711	0.00842	0.00000						
10	0.01492	0.00382	0.00341	0.00675	0.00230	0.02900	0.01619	0.00475	0.01037	0.00000					
11	0.02377	0.00961	0.02250	0.01218	0.01823	0.05847	0.01713	0.01500	0.03042	0.01770	0.00000				
12	0.03354	0.01699	0.01217	0.02269	0.01407	0.03602	0.02636	0.01717	0.02622	0.01107	0.03165	0.00000			
13	0.02645	0.01242	0.01435	0.01905	0.00784	0.02696	0.03448	0.01360	0.01581	0.01051	0.03361	0.01822	0.00000		
14	0.01448	0.00920	0.01146	0.01677	0.00628	0.04030	0.02023	0.00547	0.01109	0.01184	0.01890	0.02149	0.01569	0.00000	
15	0.01805	0.00674	0.00346	0.00896	0.00193	0.02181	0.02062	0.00543	0.00889	0.00235	0.02469	0.01133	0.00654	0.01021	0.00000

Anexo - 2 Equilíbrio de Hardy-Weinberg

No. of steps in Markov chain = 1000000
 No. of Dememorisation Steps = 1000000
 Required precision on Probability = 0
 Test association at the Locus level

=====
 == ANALYSES AT THE INTRA-POPULATION LEVEL
 =====

=====
 == Sample : Equino
 =====

=====
 == Hardy-Weinberg equilibrium : (Equino)
 =====

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

Locus #	Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1 (AHT4)	6092	0.76477	0.77382	0.01536	0.00010	1001000
2 (AHT5)	6092	0.75263	0.76089	0.50330	0.00035	1001000
3 (ASB17)	6092	0.81320	0.82138	0.42580	0.00024	1001000
4 (ASB2)	6092	0.79022	0.80960	0.00000	0.00000	1001000
5 (ASB23)	6092	0.76510	0.78026	0.00228	0.00003	1001000
6 (CA425)	6092	0.65742	0.66471	0.00016	0.00001	1001000
7 (HMS1)	6092	0.65200	0.65437	0.73232	0.00037	1001000
8 (HMS2)	6092	0.62771	0.64945	0.00000	0.00000	1001000
9 (HMS3)	6092	0.74672	0.76257	0.00999	0.00008	1001000
10 (HMS6)	6092	0.56976	0.58836	0.00153	0.00004	1001000
11 (HMS7)	6092	0.75476	0.76703	0.00671	0.00007	1001000
12 (HTG4)	6092	0.63723	0.65050	0.08561	0.00025	1001000
13 (HTG7)	6092	0.53595	0.54054	0.67866	0.00035	1001000
14 (VHL20)	6092	0.76133	0.77558	0.00089	0.00003	1001000

Hardy-Weinberg equilibrium test:

```

-----
No. of steps in Markov chain      = 1000000
No. of Dememorisation Steps      = 1000000
Required precision on Probability = 0
Test association at the Locus level
=====

```

== ANALYSES AT THE INTRA-POPULATION LEVEL

== Sample : AL

== Hardy-Weinberg equilibrium : (AL)

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

```

Forecasted chain length      :1000000
Dememorization steps        :1000000
=====

```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	67	0.64179	0.63753	0.28741	0.00047	1001000
2	67	0.71642	0.77275	0.39230	0.00048	1001000
3	67	0.77612	0.77320	0.89467	0.00030	1001000
4	67	0.71642	0.76916	0.30608	0.00047	1001000
5	67	0.79104	0.77735	0.20236	0.00031	1001000
6	67	0.83582	0.70699	0.42515	0.00053	1001000
7	67	0.68657	0.63640	0.98389	0.00012	1001000
8	67	0.61194	0.67288	0.32208	0.00046	1001000
9	67	0.70149	0.73089	0.67876	0.00039	1001000
10	67	0.52239	0.61374	0.05572	0.00025	1001000
11	67	0.76119	0.78308	0.06524	0.00021	1001000
12	67	0.71642	0.71159	0.40529	0.00044	1001000
13	67	0.68657	0.60016	0.22853	0.00040	1001000
14	67	0.88060	0.84031	0.01670	0.00012	1001000

```
=====
== Sample :      BA
=====
```

```
=====
== Hardy-Weinberg equilibrium : (BA)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):
 Forecasted chain length :1000000
 Dememorization steps :1000000

```
-----
```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	768	0.77865	0.79264	0.01895	0.00010	1001000
2	768	0.76562	0.74916	0.10826	0.00026	1001000
3	768	0.77734	0.79129	0.22800	0.00024	1001000
4	768	0.78646	0.78989	0.00591	0.00006	1001000
5	768	0.76562	0.76646	0.90806	0.00024	1001000
6	768	0.67057	0.67789	0.82192	0.00038	1001000
7	768	0.70312	0.67671	0.75311	0.00052	1001000
8	768	0.67839	0.69282	0.04091	0.00017	1001000
9	768	0.72396	0.75982	0.00962	0.00010	1001000
10	768	0.58854	0.60484	0.13769	0.00029	1001000
11	768	0.75260	0.76510	0.22722	0.00037	1001000
12	768	0.62891	0.64200	0.45937	0.00047	1001000
13	768	0.54688	0.57111	0.17197	0.00033	1001000
14	768	0.76172	0.76951	0.64419	0.00032	1001000

```
-----
```

```
=====
== Sample :      ES
=====
```

```
=====
== Hardy-Weinberg equilibrium : (ES)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	317	0.76972	0.75038	0.02014	0.00014	1001000
2	317	0.78549	0.78405	0.30361	0.00039	1001000
3	317	0.82019	0.81465	0.08505	0.00022	1001000
4	317	0.77918	0.79647	0.47166	0.00040	1001000
5	317	0.76656	0.78244	0.94721	0.00020	1001000
6	317	0.69716	0.67559	0.21797	0.00032	1001000
7	317	0.61199	0.66106	0.21852	0.00037	1001000
8	317	0.64353	0.66284	0.75269	0.00037	1001000
9	317	0.71293	0.72314	0.57331	0.00047	1001000
10	317	0.64984	0.61492	0.48133	0.00040	1001000
11	317	0.80442	0.76407	0.00517	0.00006	1001000
12	317	0.70662	0.69099	0.73448	0.00050	1001000
13	317	0.53628	0.52457	0.94724	0.00023	1001000
14	317	0.74132	0.72865	0.41425	0.00046	1001000

```
-----
```

```
=====
== Sample :      GO
=====
```

```
=====
== Hardy-Weinberg equilibrium : (GO)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000
 Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	66	0.83333	0.80500	0.60238	0.00044	1001000
2	66	0.78788	0.74960	0.40201	0.00048	1001000
3	66	0.77273	0.81078	0.03909	0.00019	1001000
4	66	0.80303	0.76475	0.54755	0.00045	1001000
5	66	0.75758	0.78880	0.81146	0.00032	1001000
6	66	0.57576	0.66875	0.12122	0.00029	1001000
7	66	0.66667	0.72033	0.76468	0.00042	1001000
8	66	0.71212	0.65603	0.47166	0.00042	1001000
9	66	0.72727	0.75330	0.12471	0.00027	1001000
10	66	0.63636	0.65765	0.69339	0.00038	1001000
11	66	0.69697	0.71177	0.58838	0.00045	1001000
12	66	0.74242	0.71120	0.26860	0.00045	1001000
13	66	0.46970	0.47594	0.56183	0.00050	1001000
14	66	0.72727	0.78510	0.35760	0.00040	1001000

```
-----
=====
```

== Sample : MG

=====
 == Hardy-Weinberg equilibrium : (MG)
 =====

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	2373	0.75432	0.77360	0.00025	0.00001	1001000
2	2373	0.73873	0.75567	0.76651	0.00038	1001000
3	2373	0.81458	0.82367	0.06012	0.00024	1001000
4	2373	0.79351	0.81580	0.00184	0.00003	1001000
5	2373	0.77876	0.78293	0.22761	0.00032	1001000
6	2373	0.65065	0.66420	0.02221	0.00012	1001000
7	2373	0.62874	0.64376	0.03460	0.00015	1001000
8	2373	0.62874	0.65600	0.00000	0.00000	1001000
9	2373	0.74589	0.76251	0.00204	0.00004	1001000
10	2373	0.58744	0.60614	0.01157	0.00009	1001000
11	2373	0.74926	0.76227	0.00356	0.00006	1001000
12	2373	0.62663	0.64719	0.00828	0.00008	1001000
13	2373	0.52845	0.53763	0.24536	0.00037	1001000
14	2373	0.76485	0.78531	0.04146	0.00015	1001000

```
=====
== Sample :      MS
=====
```

```
=====
== Hardy-Weinberg equilibrium : (MS)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	70	0.72857	0.78900	0.25207	0.00044	1001000
2	70	0.72857	0.75478	0.33221	0.00050	1001000
3	70	0.80000	0.75355	0.84597	0.00035	1001000
4	70	0.72857	0.75653	0.97046	0.00017	1001000
5	70	0.77143	0.78911	0.07235	0.00027	1001000
6	70	0.75714	0.57811	0.00000	0.00000	1001000
7	70	0.57143	0.57831	0.71631	0.00038	1001000
8	70	0.41429	0.44913	0.12735	0.00033	1001000
9	70	0.74286	0.70966	0.61138	0.00050	1001000
10	70	0.32857	0.32117	0.90214	0.00027	1001000
11	70	0.75714	0.75015	0.58009	0.00051	1001000
12	70	0.68571	0.60092	0.13971	0.00034	1001000
13	70	0.38571	0.34584	1.00000	0.00000	1001000
14	70	0.74286	0.69825	0.28987	0.00040	1001000

```
-----
```

```
=====
== Sample :      PA
=====
```

```
=====
== Hardy-Weinberg equilibrium : (PA)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000
 Dememorization steps :1000000

```
-----
```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	34	0.67647	0.75812	0.84913	0.00040	1001000
2	34	0.70588	0.66418	0.32900	0.00047	1001000
3	34	0.91176	0.82572	0.61454	0.00039	1001000
4	34	0.85294	0.82441	0.73132	0.00037	1001000
5	34	0.76471	0.75285	0.16958	0.00034	1001000
6	34	0.61765	0.63784	0.10129	0.00028	1001000
7	34	0.76471	0.73705	0.51791	0.00055	1001000
8	34	1.00000	0.71378	0.00076	0.00003	1001000
9	34	0.70588	0.76514	0.65077	0.00051	1001000
10	34	0.64706	0.62687	0.90988	0.00031	1001000
11	34	0.88235	0.75066	0.45742	0.00043	1001000
12	34	0.70588	0.74759	0.79653	0.00038	1001000
13	34	0.50000	0.52722	0.24678	0.00041	1001000
14	34	0.73529	0.78885	0.60466	0.00043	1001000

```
-----
```



```
=====
== Sample :      PE
=====
```

```
=====
== Hardy-Weinberg equilibrium : (PE)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	320	0.76250	0.74997	0.36216	0.00049	1001000
2	320	0.73125	0.74710	0.55873	0.00052	1001000
3	320	0.83437	0.83375	0.61254	0.00044	1001000
4	320	0.75625	0.80114	0.32766	0.00032	1001000
5	320	0.75313	0.77270	0.10005	0.00025	1001000
6	320	0.60000	0.61853	0.15411	0.00035	1001000
7	320	0.68437	0.66339	0.69042	0.00040	1001000
8	320	0.67812	0.68849	0.47758	0.00048	1001000
9	320	0.80312	0.76874	0.19903	0.00036	1001000
10	320	0.57500	0.61275	0.75246	0.00041	1001000
11	320	0.75625	0.78345	0.00787	0.00009	1001000
12	320	0.71250	0.63719	0.23045	0.00044	1001000
13	320	0.56563	0.60053	0.41437	0.00048	1001000
14	320	0.72813	0.78537	0.06477	0.00027	1001000

```
-----
```

```
=====
== Sample :      PR
=====
```

```
=====
== Hardy-Weinberg equilibrium : (PR)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000
Dememorization steps :1000000

```
-----
```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	157	0.70064	0.71838	0.53048	0.00051	1001000
2	157	0.73248	0.75939	0.50344	0.00043	1001000
3	157	0.77070	0.78912	0.11253	0.00028	1001000
4	157	0.78344	0.79634	0.24035	0.00031	1001000
5	157	0.74522	0.76433	0.09351	0.00026	1001000
6	157	0.69427	0.67483	0.67005	0.00037	1001000
7	157	0.72611	0.65318	0.60524	0.00045	1001000
8	157	0.65605	0.67347	0.68317	0.00043	1001000
9	157	0.73885	0.78067	0.24142	0.00036	1001000
10	157	0.56051	0.59191	0.79737	0.00033	1001000
11	157	0.77707	0.75312	0.92852	0.00022	1001000
12	157	0.57325	0.59970	0.43637	0.00057	1001000
13	157	0.61783	0.55674	0.27944	0.00039	1001000
14	157	0.76433	0.79408	0.03135	0.00016	1001000

```
-----
```

```
=====
== Sample :      RJ
=====
```

```
=====
== Hardy-Weinberg equilibrium : (RJ)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	686	0.78134	0.76926	0.06945	0.00022	1001000
2	686	0.75948	0.77349	0.06222	0.00022	1001000
3	686	0.78134	0.81041	0.72505	0.00030	1001000
4	686	0.78717	0.80683	0.77980	0.00027	1001000
5	686	0.74636	0.76223	0.03193	0.00017	1001000
6	686	0.66910	0.65361	0.02535	0.00015	1001000
7	686	0.69534	0.67254	0.55635	0.00047	1001000
8	686	0.62391	0.63710	0.35169	0.00031	1001000
9	686	0.75073	0.75721	0.64561	0.00052	1001000
10	686	0.55685	0.56898	0.90079	0.00030	1001000
11	686	0.74344	0.75793	0.19853	0.00027	1001000
12	686	0.62828	0.65414	0.29077	0.00048	1001000
13	686	0.50583	0.51538	0.66177	0.00041	1001000
14	686	0.75219	0.75729	0.04958	0.00016	1001000

```
-----
```

```
=====
== Sample :      RN
=====
```

```
=====
== Hardy-Weinberg equilibrium : (RN)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	45	0.75556	0.78377	0.90572	0.00028	1001000
2	45	0.77778	0.73633	0.84903	0.00032	1001000
3	45	0.71111	0.72734	0.05787	0.00016	1001000
4	45	0.84444	0.82622	0.24327	0.00039	1001000
5	45	0.80000	0.74856	0.90067	0.00030	1001000
6	45	0.66667	0.59850	0.13533	0.00032	1001000
7	45	0.57778	0.67491	0.23040	0.00046	1001000
8	45	0.71111	0.72759	0.35528	0.00052	1001000
9	45	0.82222	0.76454	0.73347	0.00040	1001000
10	45	0.66667	0.62871	0.64640	0.00049	1001000
11	45	0.71111	0.75855	0.92679	0.00023	1001000
12	45	0.64444	0.66092	0.23343	0.00036	1001000
13	45	0.55556	0.50836	0.88314	0.00029	1001000
14	45	0.86667	0.81174	0.59347	0.00050	1001000

```
-----
```

```
=====
== Sample :      RS
=====
```

```
=====
== Hardy-Weinberg equilibrium : (RS)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000

Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	49	0.75510	0.73827	0.23344	0.00041	1001000
2	49	0.65306	0.66295	0.71368	0.00051	1001000
3	49	0.87755	0.83295	0.84442	0.00028	1001000
4	49	0.81633	0.80160	0.96792	0.00019	1001000
5	49	0.77551	0.74122	0.29190	0.00045	1001000
6	49	0.46939	0.57732	0.06836	0.00024	1001000
7	49	0.61224	0.58742	0.98709	0.00012	1001000
8	49	0.42857	0.42310	0.60583	0.00044	1001000
9	49	0.75510	0.72985	0.47360	0.00056	1001000
10	49	0.53061	0.53208	0.85021	0.00036	1001000
11	49	0.79592	0.75468	0.17703	0.00038	1001000
12	49	0.67347	0.69577	0.90060	0.00030	1001000
13	49	0.42857	0.39323	0.40206	0.00049	1001000
14	49	0.71429	0.69367	0.84731	0.00039	1001000

```
-----
```

```
=====
== Sample :      SC
=====
```

```
=====
== Hardy-Weinberg equilibrium : (SC)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length :1000000
Dememorization steps :1000000

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	152	0.82237	0.76268	0.13974	0.00032	1001000
2	152	0.78289	0.74666	0.55894	0.00046	1001000
3	152	0.88816	0.84174	0.07124	0.00020	1001000
4	152	0.77632	0.77184	0.32283	0.00041	1001000
5	152	0.75000	0.79805	0.22889	0.00051	1001000
6	152	0.66447	0.69224	0.04621	0.00023	1001000
7	152	0.60526	0.61306	0.36807	0.00037	1001000
8	152	0.51974	0.51880	0.30841	0.00045	1001000
9	152	0.82237	0.79436	0.49887	0.00054	1001000
10	152	0.46711	0.42668	0.46815	0.00035	1001000
11	152	0.69737	0.73975	0.13702	0.00031	1001000
12	152	0.50658	0.56794	0.06272	0.00026	1001000
13	152	0.56579	0.53904	0.53374	0.00047	1001000
14	152	0.77632	0.77347	0.66737	0.00034	1001000

```
=====
== Sample :      SE
=====
```

```
=====
== Hardy-Weinberg equilibrium : (SE)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):
 Forecasted chain length :1000000
 Dememorization steps :1000000

```
-----
Locus  #Genot      Obs.Het.      Exp.Het.      P-value      s.d.      Steps done
-----
```

1	36	0.80556	0.73631	0.75884	0.00041	1001000
2	36	0.77778	0.69484	0.81678	0.00043	1001000
3	36	0.86111	0.83685	0.95728	0.00018	1001000
4	36	0.77778	0.75626	0.18770	0.00043	1001000
5	36	0.72222	0.72692	0.66479	0.00049	1001000
6	36	0.63889	0.65141	0.97243	0.00016	1001000
7	36	0.72222	0.64710	0.21475	0.00039	1001000
8	36	0.63889	0.66823	0.20470	0.00040	1001000
9	36	0.80556	0.77543	0.01045	0.00012	1001000
10	36	0.63889	0.69366	0.19276	0.00042	1001000
11	36	0.75000	0.76956	0.60989	0.00050	1001000
12	36	0.69444	0.61307	0.24374	0.00037	1001000
13	36	0.52778	0.57238	0.38667	0.00051	1001000
14	36	0.75000	0.74374	0.76591	0.00036	1001000

```
-----
```

```
=====
== Sample :      SP
=====
```

```
=====
== Hardy-Weinberg equilibrium : (SP)
=====
```

Reference: *Guo, S. and Thompson, E. 1992.*
Levene H. (1949).

Exact test using a Markov chain (for all Loci):
 Forecasted chain length :1000000
 Dememorization steps :1000000

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	952	0.77731	0.77886	0.46932	0.00041	1001000
2	952	0.77311	0.76791	0.59754	0.00037	1001000
3	952	0.84979	0.82987	0.20628	0.00029	1001000
4	952	0.80882	0.81487	0.11271	0.00028	1001000
5	952	0.75105	0.77906	0.11694	0.00022	1001000
6	952	0.65126	0.66854	0.01889	0.00012	1001000
7	952	0.63655	0.64367	0.92302	0.00026	1001000
8	952	0.58088	0.59682	0.10029	0.00029	1001000
9	952	0.74580	0.75796	0.11089	0.00022	1001000
10	952	0.51786	0.53326	0.05701	0.00021	1001000
11	952	0.76576	0.76653	0.36850	0.00040	1001000
12	952	0.63655	0.64194	0.63670	0.00040	1001000
13	952	0.55042	0.53387	0.02377	0.00015	1001000
14	952	0.76786	0.76047	0.26400	0.00038	1001000

