

**UNIVERSIDADE ESTADUAL PAULISTA  
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CAMPUS DE JABOTICABAL**

**CARACTERIZAÇÃO GENÔMICA DE EQUINOS DAS  
LINHAGENS DE TRABALHO E DE CORRIDA DA RAÇA  
QUARTO DE MILHA**

**Cíntia Maria Marchiori  
Zootecnista**

**2018**

**UNIVERSIDADE ESTADUAL PAULISTA  
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CAMPUS DE JABOTICABAL**

**CARACTERIZAÇÃO GENÔMICA DE EQUINOS DAS  
LINHAGENS DE TRABALHO E DE CORRIDA DA RAÇA  
QUARTO DE MILHA**

**Cíntia Maria Marchiori**

**Orientador: Prof. Dr. Rogério Abdallah Curi**

**Coorientador: Dr. Guilherme Costa Venturini**

Dissertação de mestrado apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Genética e Melhoramento Animal

**2018**

M318c Marchiori, Cíntia Maria  
Caracterização genômica de equinos das linhagens de trabalho e de corrida da raça Quarto de Milha / Cíntia Maria Marchiori. -- Jaboticabal, 2018  
iii, 57 p. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2018  
Orientador: Rogério Abdallah Curi  
Coorientador: Guilherme Costa Venturini  
Banca examinadora: Lilian Madi Ravazzi, Luis Artur Loyola Chardulo  
Bibliografia

1. *BeadChip*. 2. Cavalo. 3. Marcadores moleculares. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.082:636.1

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Diretoria Técnica de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

**CERTIFICADO DE APROVAÇÃO**

**TÍTULO DA DISSERTAÇÃO:** CARACTERIZAÇÃO GENÔMICA DE EQUINOS DAS LINHAGENS DE TRABALHO E DE CORRIDA DA RAÇA QUARTO DE MILHA

**AUTORA:** CÍNTIA MARIA MARCHIORI

**ORIENTADOR:** ROGERIO ABDALLAH CURI

**COORIENTADOR:** GUILHERME COSTA VENTURINI

Aprovada como parte das exigências para obtenção do Título de Mestra em GENÉTICA E MELHORAMENTO ANIMAL, pela Comissão Examinadora:



Prof. Dr. ROGERIO ABDALLAH CURI  
Departamento de Melhoramento e Nutrição Animal / FMVZ / UNESP - Botucatu



Profa. Dra. LILTAN MADI RAVAZZI  
Departamento de Biologia / UNESP- Câmpus de São José do Rio Preto



Prof. Dr. LUIS ARTUR LOYOLA CHARDULO  
Departamento de Melhoramento e Nutrição Animal / FMVZ/USP - Botucatu/SP

Jaboticabal, 20 de abril de 2018

## **DADOS CURRICULARES DO AUTOR**

CÍNTIA MARIA MARCHIORI – filha de José Antônio Marchiori (in memorian) e Maria José Vaccari Marchiori, nasceu em Barra Bonita – SP, em 8 de Abril de 1991. Iniciou o curso de Zootecnia na Faculdade de Ciências Agrárias e Tecnológicas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Dracena -SP, em Agosto de 2009, transferindo no início de 2011 para a Faculdade de Medicina Veterinária e Zootecnia da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Campus de Botucatu –SP, obtendo o título de Zootecnista em dezembro de 2014. Em março de 2016, iniciou o Curso de Mestrado do Programa de Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias – UNESP – Campus de Jaboticabal, sob orientação do Prof. Dr. Rogério Abdallah Curi e coorientação do Dr. Guilherme Costa Venturini. Foi bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) durante o Curso de Mestrado.

## AGRADECIMENTOS

Agradeço primeiramente a **Deus**, por me abençoar todos os dias com pessoas tão especiais, sem as quais certamente não teria dado conta!

A meu pai **José Antonio Marchiori**, que é a estrela mais brilhante do céu e me manda forças todos os dias, se fazendo presente em todos os momentos em meu coração.

A minha mãe **Maria José Vaccari Marchiori**, meu infinito agradecimento. Graças ao seu incentivo e apoio ao estudo e crescimento profissional hoje estou onde estou. Sem ela não seria possível.

Ao meu pai do coração **Aparecido Donizeti Malafatti**, as minhas irmãs **Analu e Anaisa Marchiori** e toda família, principalmente avós **Geraldo Vaccari, Ana Mendes Marchiori e Antônio Marchiori Filho**.

À todos meu amigos da pós-graduação: Rafael de Matteis, Malane Muniz, Marina Elizabeth, Raiza Silveira, Peterson Pacheco, Beatriz Pressi, Alejandra T. Ospina, Felipe Kriek, Verônica Mota, Beatriz L. Vellini, Letícia Custódio, Inaê Regatieri, Betsabeia Milani e Verônica Melo.

Ao grande amigo **Guilherme Luis Pereira**, meu eterno sentimento de gratidão e admiração à você por toda ajuda durante o mestrado.

Ao meu orientador, Prof. Dr. **Rogério Abdallah Curi**, por todos os ensinamentos, paciência, apoio e atenção. Sua contribuição para minha formação profissional e pessoal é imensurável. Obrigada por sempre acreditar em meu trabalho.

Ao meu coorientador **Guilherme Costa Venturini**, por todos os ensinamentos e ajuda.

A **CAPES**, pela concessão da bolsa de estudo.

Agradeço ao **Programa de Pós-Graduação em Genética e Melhoramento Animal** Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal e a todos que de alguma maneira contribuíram para o meu crescimento pessoal e profissional e me ajudaram neste período, meus sinceros agradecimentos.

## SUMÁRIO

	Página
RESUMO .....	ii
ABSTRACT .....	iii
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS .....	1
1. Introdução .....	1
2. Objetivos .....	3
3. Revisão bibliográfica .....	3
3.1. Os equinos no Brasil e no mundo .....	3
3.2. A raça Quarto de Milha .....	5
3.3. Polimorfismos de nucleotídeo único (SNP).....	10
3.4. Estudos relacionados à estrutura de populações.....	12
4. Bibliografia citada .....	17
CAPÍTULO 2 – CARACTERIZAÇÃO GENÔMICA DE EQUINOS DAS LINHAGENS DE TRABALHO E DE CORRIDA DA RAÇA QUARTO DE MILHA.....	28
1. Introdução .....	29
2. Material e métodos .....	30
2.1. Animais .....	30
2.2. Coleta de sangue e extração do DNA.....	31
2.3. Genotipagem e controle de qualidade .....	31
2.4. Imputação de genótipos na linhagem de corrida da raça Quarto de Milha .	33
2.5. Desequilíbrio de ligação .....	33
2.6. Estimação do tamanho efetivo das populações .....	34
2.7. Estrutura populacional .....	34
3. Resultados e discussão.....	36
3.1. Desequilíbrio de ligação .....	36
3.2. Estimação do tamanho efetivo das populações .....	44
3.3. Estrutura populacional .....	46
4. Conclusão .....	51
4. Bibliografia citada .....	52

## CARACTERIZAÇÃO GENÔMICA DE EQUINOS DAS LINHAGENS DE TRABALHO E DE CORRIDA DA RAÇA QUARTO DE MILHA

**RESUMO** - Na raça Quarto de Milha de equinos (*Equus caballus*), diferentes objetivos de seleção levaram à formação de linhagens com distintas habilidades, entre as quais a de trabalho e a de corrida. A linhagem de trabalho destina-se às provas de caráter funcional, explorando habilidades como agilidade e obediência, características consideradas de grande importância no manejo do gado a campo. Os animais de corrida apresentam melhor desempenho em pistas de curtas distâncias do que qualquer outra raça de equinos. O objetivo deste estudo foi caracterizar, por meio da genotipagem de SNP em larga escala, o desequilíbrio de ligação (LD), calculado por  $r^2$ , nas linhagens de trabalho e de corrida de cavalos Quarto de Milha criados no Brasil. Também foi investigado o tamanho efetivo ( $N_e$ ) das duas populações, bem como as suas estruturas e relações. Para tanto, foram utilizados 428 equinos Quarto de Milha, de ambos os sexos, registrados na associação Brasileira de criadores da raça (ABQM), sendo 68 da linhagem de trabalho e 360 da de corrida. Estes animais tiveram o sangue colhido no Jockey Club de Sorocaba (Sorocaba/SP) e em propriedades rurais localizadas em dezenas de cidades do interior do Estado de São Paulo. Cento e oitenta e oito animais, 68 da linhagem de trabalho e 120 da de corrida foram genotipados, no ano de 2011, utilizando arranjo de 54.602 SNP (54K). As demais amostras da linhagem de corrida ( $n = 240$ ) foram genotipadas, no ano de 2015, com arranjo de 65.157 SNP (65K). Após a realização de processo de imputação, o número de SNP informativos na linhagem de corrida foi de 55.114. Desta forma, para as análises genéticas subsequentes realizadas nas linhagens de trabalho e de corrida foram utilizados 43.117 e 55.114 SNP, respectivamente. O  $r^2$  genômico médio entre pares de marcadores foi 0,22 para a linhagem de trabalho e 0,27 para a de corrida. Na linhagem de trabalho o  $r^2$  foi inferior a 0,2 entre 100 e 150 Kb, enquanto na de corrida foram encontrados valores de  $r^2$  menores que 0,2 entre 300 e 350 Kb. A estimativa dos valores de  $N_e$  foi de 60 e 50 animais na última geração para as linhagens de trabalho e de corrida, respectivamente. A maior extensão do LD e o menor  $N_e$  na linhagem de corrida podem ser explicados pela sua estrutura populacional mais fechada em decorrência do maior rigor no registro genealógico dos animais, além da grande influência do Puro-Sangue Inglês na sua formação. O estudo da estrutura da população indicou clara distinção entre as duas linhagens da raça. Na linhagem de corrida, contudo, foram evidenciadas subestruturas populacionais decorrentes da formação de famílias que descendem de diferentes e importantes reprodutores.

**Palavras-chave:** *BeadChip*, cavalo, marcadores moleculares, polimorfismos de DNA, tamanho efetivo



## GENOMIC CHARACTERIZATION OF EQUINES OF CUTTING AND RACING LINES OF QUARTER HORSE BREED

**ABSTRACT** - In the Quarter Horse breed (*Equus caballus*), different selection objectives led to the formation of lineages with different abilities, including cutting and racing. The cutting line intended for functional tests, exploring skills such as agility and obedience, characteristics considered of great importance in the management of cattle. Racing animals perform better on runways short distances than any other horse breed. The objective of this study was to characterize, by means of large-scale SNP genotyping, the linkage disequilibrium (LD), calculated by  $r^2$ , in the cutting and racing lines of Quarter Horses created in Brazil. The effective size ( $N_e$ ) of the two populations was investigated, as well as their structures and relationships. For that, 428 Quarter Horses of both sexes, registered in the Brazilian Association of Breeders (ABQM) were used, of which 68 were from the cutting line and 360 from the racing. These animals had blood collected at the Jockey Club of Sorocaba (Sorocaba/SP) and on rural properties located in dozens of cities in the interior of the São Paulo State. One hundred and eighty-eight animals, 68 of the cutting line and 120 of the racing were genotyped in the year 2011, using a 54,602 SNP array (54K). The other racing line samples ( $n = 240$ ) were genotyped in 2015, with an array of 65,157 SNP (65K). After the imputation process, the number of informative SNP in the racing line was 55,114. Thus, for the subsequent genetic analyzes performed on the cutting and racing lines, 43,117 and 55,114 SNP were used, respectively. The mean genomic  $r^2$  between pairs of markers was 0.22 for the cutting line and 0.27 for the racing. In the cutting line the  $r^2$  was less than 0.2 between 100 and 150 Kb, while in the racing  $r^2$  values were lower than 0.2 between 300 and 350 Kb. The estimates of  $N_e$  values were 60 and 50 animals in the last generation for the cutting and racing lines, respectively. The longer extent of the LD and the smaller  $N_e$  in the racing line can be explained by its more closed population structure due to the greater rigor in the genealogical register of the animals, further on the great influence of the Thoroughbred in its formation. The population structure study indicated a clear distinction between the two lineages of the breed. However, in the racing line, population substructures were evidenced due to the formation of families that descend from different and important stallions.

**Keywords:** BeadChip, horse, molecular markers, DNA polymorphisms, effective size

## **CAPÍTULO 1 – Considerações Gerais**

### **1. INTRODUÇÃO**

A raça de cavalos Quarto de Milha surgiu por volta de 1600, nos Estados Unidos, quando colonizadores europeus trouxeram os primeiros equinos de origem árabe e turca, constituindo assim a primeira raça a ser formada nas Américas (ABQM, 2016). O Quarto de Milha tornou-se, desde sua origem, ponto de referência mundial entre as demais raças e espécies equinas ao apresentar qualidades notáveis, destacando-se a docilidade, velocidade e, principalmente, a versatilidade em executar funções atléticas nas mais variadas modalidades (ABQM, 2016).

A seleção para diferentes propostas deu origem à distintas linhagens da raça Quarto de Milha (EVANS, 1996), entre as quais a de corrida e a de trabalho. Considerados os cavalos mais velozes do mundo, equinos da linhagem de corrida da raça Quarto de Milha destacam-se pelo melhor desempenho em provas de curtas distâncias do que qualquer outra raça (ABQM, 2016), sendo capazes de atingir, a partir de uma posição estática, velocidade de até 88 km/h, percorrendo assim  $\frac{1}{4}$  de milha (aproximadamente 402 metros) em menos de 21 segundos (AMERICA'S HORSE DAILY, 2008). Em equinos da linhagem de trabalho da raça Quarto de Milha, agilidade e obediência são características consideradas de grande importância tendo em vista o manejo de rebanhos de bovinos a campo (MEIRA et al., 2014b). Em animais de trabalho estas habilidades são exploradas em provas de caráter funcional.

Em relação às demais espécies de interesse zootécnico, as pesquisas na área de melhoramento genético envolvendo equinos são escassas. Entretanto, na última década, com a disponibilidade de novas ferramentas para a seleção e o melhoramento genético, essa situação vem se modificando (PEREIRA, 2017). Com os avanços nas tecnologias, a identificação dos marcadores moleculares possibilitou a discriminação genotípica, permitindo assim o estudo da variação genética em nível de DNA. É considerado marcador molecular qualquer fenótipo molecular oriundo de um gene expresso ou de um segmento específico de DNA que corresponde a regiões expressas ou não do genoma (FERREIRA; GRATTAPLAGIA, 1998). Os marcadores moleculares podem ser utilizados como ferramenta para estudos de diversidade genética entre indivíduos, dentro e entre populações ou espécies relacionadas (SOUZA et al., 2008). Os arranjos para genotipagem de marcadores de DNA em larga

escala constituem plataforma poderosa, possibilitando a identificação de regiões genômicas modificadas pela seleção e a identificação de regiões cromossômicas, genes e polimorfismos de nucleotídeo único (*Single Nucleotide Polymorphisms* – SNP) que contribuem para características de interesse nas principais raças de equinos criadas atualmente no mundo.

Boa parte destas novas aplicações requerem o conhecimento ou acesso ao desequilíbrio de ligação (*linkage disequilibrium* – LD) existente entre os marcadores em dada população de interesse. O LD entre marcadores moleculares descreve a correlação entre os genótipos de dois marcadores, expressando o grau de associação não aleatória entre seus alelos (PORTO-NETO et al., 2014). A quantificação do nível de LD é importante passo para a seleção genômica (MEUWISSEN e GODDARD, 2000) e para o melhor entendimento da arquitetura genômica e da estrutura histórica da população (HAYES et al., 2009). Assim, o sucesso do mapeamento das regiões em LD depende de um balanço apropriado entre a extensão deste desequilíbrio e a densidade de marcadores (SARGOLZAEI et al., 2008). Mapas comparativos de LD permitem explorar o grau de diversidade entre raças que apresentam diferentes atributos biológicos e detectar regiões genômicas que foram sujeitas a diferentes pressões seletivas (MCKAY et al., 2007).

O conhecimento da estrutura de uma população é importante em programas de melhoramento genético animal e um dos principais parâmetros utilizados nestes estudos é o tamanho efetivo ( $N_e$ ), definido como o conjunto de indivíduos de uma população que se reproduzem e conseguem deixar descendentes. Vozzi et al. (2006) ressaltam a importância do  $N_e$  da população no acompanhamento da variabilidade genética. O contínuo monitoramento da estrutura populacional se faz necessário para evitar a perda da variabilidade genética da população (CARJAVAL, 2017). Estimativas de  $N_e$  baseadas em genealogia são dependentes da integralidade e correção dos *pedigrees* disponíveis, o que, muitas vezes, não ocorre. Estratégia promissora e independente da integralidade do *pedigree* é estimar o tamanho populacional efetivo com base em dados genômicos. Esses métodos são baseados em padrões de LD de milhares de SNP distribuídos ao longo do genoma (HAYES et al., 2003).

Marcadores moleculares podem também ser utilizados para diferenciar populações. A identificação de um conjunto de marcadores que possa efetivamente ser utilizado para a inferência da população, é de grande interesse e pode atribuir indivíduos à populações, estudar zonas híbridas, e identificar os migrantes e

indivíduos miscigenados (FALUSH; STEPHENS; PRITCHARD, 2003; FALUSH; STEPHENS; PRITCHARD, 2007).

## **2. OBJETIVOS**

Considerando as diferentes pressões de seleção exercidas, as diferentes práticas de acasalamentos endogâmicos, os diferentes graus de introdução de animais de outras raças e as diferentes aptidões existentes entre as linhagens de trabalho e de corrida de equinos Quarto de Milha criados no Brasil, este trabalho teve como objetivos a caracterização do desequilíbrio de ligação e a análise do tamanho efetivo em ambos os grupos, bem como a investigação da estrutura populacional da raça por meio da genotipagem de SNP em larga escala.

## **3. REVISÃO BIBLIOGRÁFICA**

### **3.1. Os equinos no Brasil e no mundo**

Existem aproximadamente 58,9 milhões de cavalos no mundo, maior parte vivendo na América, Ásia e alguns países da Europa (FAO, 2014). Os Estados Unidos é o país com maior número de cabeças (10.260.000), seguido do México (6.355.000) China (6.029.780) e Brasil (5.450.601) (FAO, 2014). Deve-se destacar a redução da população de equinos na Ásia, principalmente na China, de 8.916.154 de cabeças em 2000 para 6.029.780 em 2014 (GLIPHA, 2014), a qual encontra-se associada à migração interna da população humana, com menor utilização dos equídeos no transporte e na agricultura e ao maior consumo de carne equina (ALMEIDA; SILVA, 2010). Por outro lado, nos Estados Unidos houve aumento expressivo da população de equinos, de 5.240.000 cabeças em 2000 para 10.260.000 cabeças em 2014 (GLIPHA, 2014), em parte devido a restrições legais internas ao abate e à exportação de carne de equídeos (ALMEIDA; SILVA, 2010). No Brasil, no período de 2000 a 2014, o número de equinos permaneceu praticamente estável, com pequena queda (GLIPHA, 2014), e as raças predominantes são, em ordem, Mangalarga Marchador, Nordeste, Quarto de Milha e Crioula (MAPA, 2016).

O complexo do agronegócio do cavalo no Brasil é bastante expressivo, movimentando valor econômico superior a R\$ 16 bilhões ao ano e ocupando diretamente 607.329 pessoas e indiretamente 2.429.316 (considerando que cada ocupação direta proporciona outras quatro ocupações indiretas), sendo, assim, gerados mais de três milhões de empregos diretos e indiretos. O levantamento do efetivo de equinos no Brasil, mostra o total de aproximadamente cinco milhões de animais, sendo 3.900.000 no manejo de animais de produção e 1.100.000 destinados às atividades esportivas e lazer (MAPA, 2016).

A utilização de animais no manejo de rebanhos bovinos vem sendo o seguimento de maior destaque dentro do agronegócio do cavalo no Brasil, representando mais de 53% da movimentação monetária e 71,3% dos empregos diretos gerados. Os valores do custo de manutenção, o valor anual da tropa e a renda associada à mão-de-obra dos equinos destinados ao manejo resultam em 8,58 bilhões do montante total do complexo do agronegócio equino (MAPA, 2016). Esta intensa relação tem feito os cavalos acompanharem os bovinos em seu deslocamento para as regiões Centro-Oeste e Norte do País. Entretanto, Minas Gerais ainda é o principal estado brasileiro criador de equinos (IBGE, 2015).

Embora a utilização do cavalo no manejo de bovinos seja marcante no contexto da equinocultura nacional, o seguimento que envolve os esportes equestres tem crescido acentuadamente nos últimos anos, acompanhando tendência mundial. Neste sentido, no período de 1999 a 2004 o número de eventos envolvendo as diversas modalidades equestres (salto, adestramento, corridas, enduro, concurso completo de equitação, etc.) cresceu 315%, ou seja, aumento médio de 15,3% ao ano (LIMA; SHIROTA; BARROS, 2006). Animais Quarto de Milha e Crioulo são as raças que prevalecem no segmento das diversas atividades esportivas e a movimentação financeira nessa área totaliza R\$ 5,84 bilhões, empregando diretamente cerca de 125.700 pessoas (MAPA, 2016). Os equinos desse segmento são animais que movimentam com maior intensidade desde a indústria de medicamentos e ferragens até cosméticos e acessórios. Dessa forma, destacam-se no complexo do agronegócio equino os fornecedores de rações e demais produtos e serviços para a criação, as indústrias de medicamentos veterinários, as indústrias de selarias, como também eventos e turismo (como escolas de hipismo e provas como a vaquejada), entre outros.

Além da utilização do cavalo no manejo das fazendas e nas provas equestres, a tendência de sua utilização para lazer vem aumentando significativamente ao longo dos anos. O cavalo também é utilizado na equoterapia, modalidade disponível há milhares de anos e, agora, reconhecida como de grande eficácia para o tratamento de inúmeros males físicos, psíquicos e comportamentais em humanos. Segundo dados fornecidos pela ANDE-BRASIL (Associação Nacional de Equoterapia), há cerca de 280 centros de equoterapia distribuídos por todo território nacional brasileiro, Sul e Sudeste destacam-se como as principais áreas (ANDE-BRASIL, 2015), movimentando anualmente valores acima de 43 milhões e ocupando aproximadamente 2.500 pessoas (LIMA; SHIROTA; BARROS, 2006).

### **3.2. A raça Quarto de Milha**

A raça Quarto de Milha foi a primeira a ser introduzida na América do Norte, a partir do século XVII, com a chegada de equinos de origem árabe e turca, trazidos por colonizadores europeus. O seu desenvolvimento deveu-se à ocupação do oeste Norte Americano, em razão da necessidade de cavalos robustos e versáteis, com aptidão à sela e tração, visto a dificuldade de manutenção de plantel variado de animais para atender à diferentes necessidades. Equinos Quarto de Milha possuem a versatilidade, docilidade, rusticidade e inteligência como principais características. A agilidade e solidez da constituição são contribuições da influência do cavalo espanhol e a velocidade e conformação são contribuições da influência do cavalo Puro-Sangue Inglês na formação da raça (ABQM, 2017).

Em 1940, fundou-se, nos Estados Unidos da América, a *American Quarter Horse Association* (AQHA), primeira envolvendo a raça. Atualmente, é considerada a maior associação de criadores do mundo, com cerca de 500 mil sócios e mais de cinco milhões de cavalos registrados, divididos em 43 países, representando 52% dos equinos em todo o mundo (ABQM, 2016).

A raça Quarto de Milha foi introduzida no Brasil em 1955 e a Associação Brasileira de Criadores de Cavalo Quarto de Milha (ABQM) foi criada em 1969, a qual conta atualmente com plantel de 514.316 equinos registrados, 104.238 criadores, proprietários e associados cadastrados, distribuídos por todos os estados brasileiros. A criação da raça apresenta impacto relevante no agronegócio nacional do cavalo, visto que seus haras ocupam área de aproximadamente um milhão de hectares,

estimados em mais de R\$ 20,0 bilhões (dados de até 13 de janeiro de 2017; ABQM, 2017). O segmento de maior destaque dentro do agronegócio do cavalo no Brasil vem sendo a utilização de animais no manejo de rebanhos bovinos, representando aproximadamente 53% da movimentação monetária e 71% dos empregos diretos gerados (MAPA, 2016).

A mão-de-obra empregada diretamente é significativa, oferecendo mais de 310 mil empregos diretos (média de três funcionários para cada propriedade), sem contar com veterinários, agrônomos, zootecnistas, ferradores, centros de treinamento, centros de reprodução, leiloeiros, leiloeiras, transportadores de cavalos, fabricantes de equipamentos e indústria de ração e produtos veterinários, entre outros. Nos últimos cinco anos o Stud Book da ABQM registrou mais de 130 mil potros. Também neste mesmo período, a raça movimentou em leilões por todo o país em torno de R\$ 1 bilhão com a comercialização de aproximadamente 27 mil animais, pela média de R\$ 37,0 mil/animal (ABQM, 2017).

A seleção nesta raça para diferentes propostas levou à formação de grupos ou linhagens (EVANS, 1996), entre as quais a de corrida, de trabalho e de conformação. A linhagem de trabalho destina-se às provas de caráter funcional, explorando habilidades como agilidade e obediência, características consideradas de grande importância no manejo do gado a campo. Dentro desta linhagem há grande interesse na produção de cavalos com *cow sense* superior (ELLERSIECK; LOCK; VOGT, 1985). *Cow sense*, ou habilidade de trabalhar com o rebanho bovino, pode ser medido pela capacidade do cavalo em cercar e apartar do rebanho um animal escolhido (HINTZ, 1980), com pouca ou nenhuma assistência do cavaleiro (ELLERSIECK; LOCK; VOGT, 1985).

O cavalo Quarto de Milha da linhagem de trabalho apresenta a agilidade e estabilidade emocional, possui condição atlética, temperamento dócil e rápida aprendizagem. Estes animais caracterizam-se por flexibilidade que permite movimentos com velocidade em qualquer direção: são capazes de girar, colocando seu peso sobre os membros pélvicos e mover-se primeiramente com o focinho. Isso ajuda o cavalo a girar rapidamente enquanto permanece focado no bovino. Nas competições esportivas de caráter funcional leva-se em conta o desempenho dos cavalos a partir de uma série de fatores tais como: coragem, apelo visual, trabalho com o rebanho, controle do bovino, grau de dificuldade da execução, tempo de

trabalho, e rédeas soltas. Com base nisso, pontos são adicionados ou subtraídos (NCHA, 2013).

Nos Estados Unidos, a primeira competição registrada de cavalos de apartação foi realizada em Haskell, Texas, em 1898. O interesse por este tipo de competição aumentou tanto que, em 1946, foi criada a *National Cutting Horse Association* (NCHA) a fim de padronizar as regras da competição e criar uma modalidade esportiva (ELLERSIECK; LOCK; VOGT, 1985). Atualmente, a competição de cavalos de apartação é esporte equestre de rápido crescimento e altos investimentos. Um exemplo claro disso é que, nos EUA, em 1984, a premiação girava em torno de US\$ 237.000, enquanto que recentemente se compete em torneios da NCHA *Futurity* por mais de 3,7 milhões de dólares (NCHA, 2013). Ao longo de uma prova de apartação, que apresenta tempo limite de 2 minutos e 30 segundos, o cavalo e cavaleiro devem se mover calmamente para dentro do rebanho, apartar o bovino escolhido e dirigi-lo ao centro da arena, devendo mantê-lo afastado dos outros bovinos. O cavaleiro deve impedir que o boi, já apartado, retorne ao rebanho (ABQM, 2016).

Outra modalidade esportiva de grande destaque nos animais Quarto de Milha da linhagem de trabalho é a prova de rédeas. Durante esta prova, o cavaleiro deve executar um dos 13 percursos existentes pré-estabelecidos e o cavalo deve ser voluntariamente guiado com pouca ou nenhuma resistência. Manobras prescritas de esbarros, *spins* (giros sobre si de 360 graus), *rollbacks* (esbarro com mudança de direção em 180 graus saindo ao galope), mudança de mão e círculos ao galope abrangem esses percursos. Um cavaleiro, ao desempenhar a função de rédeas em um cavalo, também domina todos os movimentos desse animal (ABQM, 2016).

A linhagem de corrida da raça Quarto de Milha explora a aptidão dos animais quanto à velocidade em pistas retas e de curta distância. Destacadamente equinos desta raça/linhagem tem melhor desempenho em corridas de curtas distâncias do que qualquer outra (ABQM, 2016), sendo os mais velozes cavalos do mundo e um dos mais velozes dentre todos os animais, podendo alcançar velocidade de até 88 km/h e percorrer, a partir de uma posição estática,  $\frac{1}{4}$  de milha (402 metros) em menos de 21 segundos (AMERICA'S HORSE DAILY, 2008). Apesar do efetivo de animais ser relativamente menor na linhagem de corrida do que nas demais, sua importância econômica é substancial, não somente por gerar renda por meio de premiações e apostas (US\$ 2,5 milhões – LIMA; SHIROTA; BARROS, 2006), mas também pelo elevado custo gerado na manutenção destes animais dentro desta modalidade



esportiva (entre R\$ 800,00 e R\$ 1.400,00, em média mensal, excluindo-se medicamentos e procedimentos veterinários).

O principal atributo de seleção utilizado pelos criadores do cavalo Quarto de Milha de corrida é a pontuação conhecida como Índice de Velocidade (IV). Este índice é obtido durante a campanha de um animal com o intuito de classificar o seu desempenho em diferentes condições (distâncias, hipódromo, clima, país) (EVANS, 1996). Cada hipódromo tem sua própria tabela de IV, que é elaborada a partir da média das três vitórias mais rápidas (três melhores tempos) para cada um dos três últimos anos consecutivos, em cada distância, sendo que o valor da média desses nove tempos equivalerá ao IV igual a 100 (JCS, 2002). Em razão das diferenças entre as pistas de corrida, os tempos variam ligeiramente e como um meio de comparar cavalos nas diferentes pistas, desenvolveu-se o sistema IV.

Os pontos de IV são inteiros e variam de acordo com o tempo, ao nível de centésimos de segundo, seguindo ajustes de acordo com a distância percorrida. Cada fração de segundo é igual a um ponto no IV. Como exemplo, para distâncias de 365 metros (m), 402 m e 503 m, a cada quatro centésimos de segundos, a mais ou a menos, no tempo de um animal, em relação ao tempo médio obtido para se estabelecer o índice de velocidade igual a 100, diminui-se ou cresce-se um ponto neste índice.

Equinos da raça Quarto de Milha praticam diversos tipos de exercícios físicos, variando de exercícios predominantemente aeróbios até exercícios predominantemente anaeróbios. Animais Quarto de Milha da linhagem de corrida apresentam metabolismo muscular predominantemente anaeróbio durante as corridas de curtas distâncias, diferentemente do ocorrido nas modalidades desempenhadas pela linhagem de trabalho. Durante exercícios de arena, como apartação e rédeas, os animais apresentam períodos curtos de exercício anaeróbio com maiores períodos de atividade aeróbia (FREEMAN, 2013), assim como ocorre durante longas corridas em animais da raça Puro-Sangue Inglês (PSI).

As diferenças morfológicas entre os animais das linhagens de corrida e de trabalho da raça Quarto de Milha são, em geral, mais evidentes, embora existam diferenças fisiológicas, como o já mencionado metabolismo energético muscular. Os cavalos de corrida da raça Quarto de Milha são mais altos que os de trabalho devido à influência de animais Puro-Sangue Inglês, utilizados para melhorar a sua velocidade (MEIRA et al., 2013). Entretanto, a seleção direcionada a diferentes objetivos (corrida

ou trabalho), também explica as mudanças significativas nas características corporais dos equinos Quarto de Milha. Em trabalho realizado por Meira et al. (2013), mostrou-se diferenças significativas com relação às características morfológicas entre as linhagens, onde animais da linhagem de corrida apresentaram maiores peso, altura, comprimentos e perímetros corporais (canela, quartela, guarupa, cabeça, pescoço, casco) em relação aos de trabalho. Dessa forma, animais da linhagem de trabalho possuem comprimento corporal mais curto e compacto, enquanto animais da linhagem de corrida são mais compridos e com pernas mais longas, sendo portanto, animais mais altos. Os autores sugerem que as diferenças fenotípicas observadas entre as duas linhagens apresentam evidente componente genético. Neste mesmo trabalho foram realizadas ainda análises de *pedigree* que sugeriram a ocorrência de maior prática de acasalamentos endogâmicos na linhagem de corrida em relação à de trabalho.

No mundo todo, as pesquisas na área de melhoramento genético são proporcionalmente menores em equinos quando comparadas às outras espécies de interesse zootécnico. Dada a grandeza do número de equinos no Brasil, essa escassez torna-se ainda maior (BELTRÁN, 2014). Embora algumas pesquisas publicadas envolvam a área de melhoramento genético equino (SANTOS, 2006; CORRÊA; MOTA, 2007; PRADO; MOTA, 2008), ainda não há programas consistentes de seleção para as diferentes raças equinas criadas no Brasil. Visando obter avanços na área, pequenos grupos de pesquisadores ligados à institutos de pesquisas e universidades trabalham especialmente aspectos quantitativos de caracteres de interesse econômico em raças nacionais e importadas ou aspectos conservacionais em raças nacionais. Pesquisas envolvendo Genética e Biologia Molecular começaram a ser realizadas a pouco tempo e em algumas raças equinas específicas no Brasil (MEIRA et al., 2014a,b,c; MANSO FILHO, 2015; REGATIERI et al., 2016; PEREIRA et al., 2016a,b; PEREIRA, 2017; FONSECA et al., 2017; LIMA, 2017).

Com a utilização de modelos estatísticos mais adequados para a obtenção de valores genéticos individuais, programas efetivos de seleção e melhoramento genético equino são praticados na Europa (ARNASON, 2013), especialmente em equinos destinados a esportes olímpicos, tais como: *cross contry*, adestramento e salto (PEREIRA, 2017).

Dentre todas as dificuldades inerentes à espécie equina, a principal causa da complicação na aplicação da seleção baseada em estimação de valores genéticos é

a baixa herdabilidade do desempenho atlético e os valores de correlações genéticas existentes entre as características de maior interesse nos cavalos, especialmente características associadas a reprodução e desempenho em provas funcionais e corridas (WALLIN; STRANDBERG; PHILIPSSON, 2003; BOKOR et al., 2005). Contudo, a utilização de novas ferramentas disponíveis para a seleção e o melhoramento genético, como os marcadores de DNA, é argumento a ser considerado visto a baixa herdabilidade do desempenho atlético em equinos (PEREIRA, 2017).

### **3.3. Polimorfismos de nucleotídeo único (SNP)**

Marcador molecular é toda e qualquer variação oriunda de um gene expresso ou de um segmento específico de DNA, correspondente a regiões gênicas ou inter gênicas do genoma. Ao se verificar que esses marcadores segregam de acordo com as leis mendelianas para características monogênicas, ou apresentam distribuições compatíveis com as esperadas para características poligênicas, um marcador molecular é também definido como marcador genético (FERREIRA; GRATTAPAGLIA, 1998). Do ponto de vista molecular ocorrem três tipos principais de variações na molécula de DNA, as regiões repetitivas (minissatélites e microssatélites), as inserções e deleções (*insertion/deletion polymorphisms* - InDels) e as alterações de uma base (*Single Nucleotide Polymorphisms* – SNP). À medida que as sequências de nucleotídeos dos genomas foram sendo desvendadas, característica observada foi o grande número de variações de uma base encontradas ao se comparar segmentos correspondentes do mesmo genoma, ocorrendo, aproximadamente, a cada 600 bases (PEREIRA, 2017).

As substituições mais frequentemente observadas no DNA envolvem bases nitrogenadas de mesma característica estrutural, ou seja, são trocas entre duas purinas (A/G ou G/A) ou duas pirimidinas (C/T ou T/C) e são denominadas transições. As transversões são substituições de uma purina por uma pirimidina ou o contrário. Essas alterações podem ser provocadas por erros de incorporação de bases durante a replicação do DNA ou em outros casos, são causadas por agentes ambientais (físicos, químicos ou biológicos). Caso essas alterações ocorram em células germinativas e sejam transmitidas às gerações seguintes passam a ser denominadas mutações herdáveis. Caso fixem-se na população em uma frequência mínima de 1% ou mais, passam a ser denominadas de polimorfismos (KWOK; GU, 1999). Neste sentido, as mutações de uma base passam a ser denominadas SNP.

Os SNP podem ocorrer em regiões gênicas codificadoras ou com função regulatória, bem como em espaços intergênicos, muitas vezes sem função determinada. Em regiões codificadoras, quando resultam em substituição de aminoácido na sequência proteica, são denominados não sinônimos, podendo a substituição ser conservativa ou não conservativa em função das características dos aminoácidos envolvidos na troca. Nesses casos, pode haver modificações estruturais e funcionais na proteína. Embora SNP sinonímicos não alterem a sequência proteica, eles podem modificar a estrutura e a estabilidade do RNA mensageiro, e, conseqüentemente afetar a quantidade de proteína produzida. Esta também pode ser afetada quando ocorrem alterações nas regiões não traduzidas do RNA mensageiro (5' UTR e 3' UTR). Além disso, polimorfismos gênicos podem promover processamentos alternativos, geração ou supressão de códons de terminação, alteração nos códons de iniciação da tradução e alterações no padrão de expressão de genes quando a troca de bases ocorre em sequências regulatórias (GUIMARÃES; COSTA, 2002).

Polimorfismos em regiões de intron ganharam importância pelo fato de também poderem ser possíveis responsáveis diretos por alterações fenotípicas. RNAs não codificantes transcritos a partir de regiões de introns (micro-RNAs) podem estar envolvidos em diferentes processos biológicos tais como os controles transcricional e pós-transcricional da expressão gênica (NAKAYA et al., 2007). Estudos mostram cada vez mais que as regiões intergênicas apresentam importância no processo de regulação gênica, mesmo que ainda as funções da maior parte dessas regiões sejam desconhecidas. Neste sentido, polimorfismos situados nestas regiões também podem estar associados à variações fenotípicas (PEREIRA, 2017).

Estudos em humanos e em espécies de interesse zootécnico mostraram a ocorrência de milhões SNP distribuídos ao longo do genoma de um indivíduo (*Human Genome Project Information, The SNP Consortium LTD, Bovine Genome Sequencing and Analysis Consortium, EquCab2.0 SNP Collection*). Além dos marcadores SNP serem abundantes, as formas como se originam permitem sua distribuição homogênea pelo genoma (CAETANO, 2009).

De acordo com Chowdhary e Raudsepp (2008), entre os maiores destaques provenientes da análise do genoma de equinos está o seu sequenciamento completo (EquCab2.0) e, a partir deste, a identificação de 1.162.753 de SNP que ocorrem entre diversas raças (WADE et al., 2009). Avanços tecnológicos na área de Genética

Molecular, incluindo os métodos de sequenciamento de próxima geração, permitiram a identificação em massa destes polimorfismos de DNA distribuídos ao longo dos cromossomos de indivíduos da mesma espécie, culminando no desenvolvimento dos arranjos de SNP, também chamados *chips* de genotipagem de SNP em larga escala (BELTRÁN, 2014).

Projetado para permitir a identificação de regiões genômicas modificadas pela seleção (assinaturas de seleção) e a identificação de SNP e genes que contribuem para características de interesse nas principais raças de equinos criadas atualmente no mundo (estudos de ampla associação do genoma – GWAS), o *EquineSNP50 BeadChip* da empresa Illumina (Illumina Inc., EUA), com cerca de 54.000 SNP distribuídos uniformemente ao longo do genoma, constituiu poderosa plataforma para a seleção e o melhoramento genético da espécie, habilitando pesquisadores da área a conduzir vasta gama de experimentos em que a aplicação da genotipagem de polimorfismos de DNA é necessária. Já em sua segunda geração, o novo SNP chip equino (*EquineSNP70 BeadChip*; Illumina Inc., EUA) possui aproximadamente 65 mil SNP, dos quais 19 mil são novos marcadores e 45 mil foram validados no *EquineSNP50 BeadChip* (PEREIRA et al., 2017).

Em um primeiro momento, esses *chips* serviram com sucesso, principalmente, à identificação de SNP, regiões genômicas e genes relacionados/associados à importantes doenças e síndromes que acometem determinadas raças de equinos tais como lordose (COOK; GALLAGHER; BAILEY, 2010), osteocondrose (TEYSSÉDRE et al., 2012), neuropatia laringeal recorrente (DUPUIS et al., 2011), nanismo (EBERTH; SWERCZAK; BAILEY, 2009) e síndrome do potro lavanda (BROOKS et al., 2010). Mais recentemente, características complexas relacionadas à desempenho em provas esportivas e aptidões específicas têm sido alvo de pesquisas por meio de *chips* de SNP para equinos (HILL et al., 2010; BINNS; BOEHLER; LAMBERT, 2010; SCHRÖDER et al., 2011; PETERSEN et al., 2013, MEIRA et al., 2014b; FARIA, 2016; REGATIERI et al., 2016; FONSECA et al., 2017, PEREIRA, 2017).

### **3.4. Estudos relacionados à estrutura de populações**

O estudo simultâneo de milhares de polimorfismos espalhados ao longo do genoma tem possibilitado também o estudo da estrutura genética de diferentes populações em várias espécies de animais domésticos (KIJAS et al., 2009; McKAY et al., 2008; GIBBS; TAYLOR; VAN TASSELL, 2009), estimar o grau de diversidade

dentro e divergência genética entre populações (ZENGER et al., 2006), determinar a relação entre a perda de alelos e o aumento da endogamia devido à seleção (MUIR et al., 2008), e, recentemente, identificar e localizar regiões do genoma sujeitas à seleção (PEREIRA, 2017; CARJAVAL, 2017; PETERSEN et al., 2013; QANBARI; SIMIANER, 2014; HESTAND et al., 2015; ISO-TOURU et al., 2016).

Boa parte destas aplicações requerem o conhecimento ou acesso ao desequilíbrio de ligação (*linkage disequilibrium* – LD) existente entre os marcadores moleculares em dada população da espécie de interesse. O LD é a associação não aleatória entre alelos de diferentes locos do mesmo cromossomo, ou seja, é quando a frequência de um determinado haplótipo é maior do que seria esperado (ARDLIE; KRUGLYAK; SEIELSTAD, 2002; KHATKAR et al., 2006; GODDARD; HAYES et al., 2009). O LD entre marcadores moleculares descreve a correlação entre os genótipos de dois marcadores, expressando o grau de associação não aleatória entre seus alelos (PORTO-NETO et al., 2014). Em geral, esta associação ou ligação é consequência da proximidade física existente entre dois locos (ARDLIE; KRUGLYAK; SEIELSTAD, 2002; BOLORMAA et al., 2011), mas é influenciada por eventos como mutação, deriva, seleção, recombinação e tamanho efetivo da população (GU et al., 2007).

As medidas mais utilizadas para avaliar o LD entre marcadores bialélicos são o  $D'$  (LEWONTIN, 1964) e o  $r^2$  (HILL; ROBERTSON, 1968), cada uma com diferentes propriedades estatísticas, porém com valores entre 0 e 1 (BOHMANOVA; SARGOLZAEI; SCHENKEL, 2010). Um valor de  $D' < 1$  evidencia a ocorrência de recombinação entre os dois locos, e  $D' = 1$  (completo desequilíbrio), evidencia a ausência de recombinação. Para marcadores bialélicos,  $D'$  tem o valor de 1 se um alelo em dado loco está completamente associado com um alelo de outro loco, isto é, se um ou mais dos quatro haplótipos possíveis estão ausentes. Os valores  $D'$  são inferiores a 1, se todos os quatro possíveis haplótipos estão presentes (GURGUL et al., 2014). Esta medida é superestimada quando a amostra é pequena e pela presença de alelos raros na população (DU; CLUTTER; LOHUIS, 2007; BOHMANOVA; SARGOLZAEI; SCHENKEL, 2010). O  $r^2$  representa a correlação entre dois locos e demonstrou ser a medida mais adequada para estimar o LD entre marcadores bialélicos, como os SNP, por ser menos sensível à frequência alélica e ao tamanho da amostra (ZHU; ZHAO, 2007; BOHMANOVA; SARGOLZAEI; SCHENKEL, 2010).

Para um par de locos bialélicos,  $r^2$  é igual a 1 se apenas dois haplótipos estão presentes em uma população (GURGUL et al., 2014).

O primeiro passo necessário para determinar o número de marcadores necessários para um mapeamento eficiente de QTL por estudo de associação ampla do genoma (*genome wide association studies* – GWAS) e a aplicação da seleção genômica é a quantificação da extensão do desequilíbrio de ligação ao longo do genoma da espécie de interesse. Sargolzaei et al. (2008) ressaltaram que para a obtenção de sucesso na aplicação destes estudos, há necessidade de se basear na relação entre a extensão deste desequilíbrio da população alvo e a densidade (cobertura) de marcadores utilizados. A densidade dos marcadores deve ser suficientemente alta para garantir que todos os QTL estejam em LD com um marcador ou haplótipo de marcadores.

Desta forma, os mapas de LD são ferramentas fundamentais para explorar a base genética de características economicamente importantes (MEIRA et al., 2014a). Além disso, mapas comparativos de LD permitem explorar o grau de diversidade entre populações (raças/linhagens) que apresentam diferentes atributos biológicos e detectar regiões genômicas que foram submetidas a diferentes pressões seletivas (McKAY et al., 2007). Para o mapeamento mais eficiente de características de interesse no genoma, utilizando poucos marcadores, é necessário LD em grandes extensões, de forma oposta, LD em pequenas extensões implicam na necessidade de utilização de maior número de marcadores (BARNAUD et al., 2010).

A extensão do LD foi caracterizada por meio de  $r^2$  em 14 raças equinas para estimar o poder de mapeamento de associações entre polimorfismos gênicos e características de importância. Em geral, valores de  $r^2$  caíram abaixo de 0,2 entre 100 e 150 Kb de distância. Valores de  $r^2$  continuaram superiores a 0,2 em distâncias maiores à 400 Kb nas raças Puro-Sangue Inglês, *Standardbred* e *French Trotter*. Em comparação às demais raças, a Quarto de Milha destacou-se por apresentar baixos valores de LD, com decaimento de  $r^2$  abaixo de 0,2 entre 50 e 100 Kb de distância (McCUE; MICKELSON, 2013). Meira et al. (2014a) destaca, que provavelmente além do grande tamanho efetivo populacional, a estrutura aberta explica esta pequena extensão do LD na raça.

Para muitas espécies de interesse zootécnico a seleção visando o melhoramento genético tem sido praticada concomitantemente com o controle das taxas de endogamia, mas em outras não. No nível individual, a consanguinidade é o

resultado do acasalamento deliberado de indivíduos aparentados. Ao nível da população, a deriva genética aleatória faz com que ocorra a divisão de uma população em subpopulações com um número menor de pais resultando, portanto, em endogamia (FALCONER, 1987; BARRET; KOHN, 1991; YOUNG; BOYLE; BROWN, 1996). No nível molecular, a deriva aleatória afeta as frequências alélicas, leva à perda de variação genética neutra, e à fixação de alelos favoráveis ou deletérios. Uma abordagem que auxilia na predição dessas perdas é a determinação do tamanho efetivo da população ( $N_e$ ), definido como o conjunto de indivíduos de uma população que se reproduzem e conseguem deixar descendentes (FALCONER, 1987).

Estimativas de  $N_e$  baseadas em genealogia são dependentes da integralidade dos *pedigrees* disponíveis, o que muitas vezes não ocorre. Uma abordagem promissora e independente da integralidade do *pedigree* é estimar o tamanho populacional efetivo com base em dados genômicos. Tais métodos são baseados em padrões de desequilíbrio de ligação (SVED, 1971; HILL, 1981; HAYES et al., 2003) de milhares de SNP distribuídos ao longo do genoma. Esta nova tecnologia é ferramenta útil para estudar a estrutura das populações e permite comparação mais detalhada entre populações do que as abordagens anteriores baseadas em *pedigree*.

Em populações com tamanho efetivo pequeno, um nível substancial de LD deverá abranger grandes distâncias físicas (FARNIR et al., 2000; McRAE et al., 2002; ODANI et al., 2006). Por outro lado, em populações com tamanho efetivo grande, espera-se uma menor extensão do desequilíbrio de ligação médio.

Os marcadores moleculares também podem ser utilizados para diferenciar populações. A identificação de um conjunto de marcadores que possa efetivamente ser utilizado para a inferência da população, oferece subsídios, por exemplo, para a identificação de locos genéticos que tenham sido submetidos à pressões seletivas. Dados de genótipos multi-local podem ser analisados com a finalidade de investigar a estrutura das populações. Seus usos incluem inferir a presença de populações distintas, atribuir indivíduos às populações, identificar marcadores grupo-específicos, estudar zonas híbridas, e identificar os migrantes e indivíduos miscigenados (FALUSH; STEPHENS; PRITCHARD, 2003; FALUSH; STEPHENS; PRITCHARD, 2007). Estas análises podem ser aplicadas à maior parte dos marcadores genéticos normalmente utilizados em análises genéticas animal, incluindo microssatélites, RFLP (*restriction fragment length polymorphism*), AFLP (*amplified fragment length*



*polymorphism*) e SNP (HRISTOV et al., 2014; BARCACCIA et al., 2013; AL ABRI et al., 2017).

Comumente utilizados e, de forma muito útil, métodos tradicionais de análises multivariadas, como a análise de componentes principais (*principal component analysis* – PCA) e a análise de coordenadas principais (*principal coordinates analysis* – PCoA), permitem acessar a variação entre grupos genéticos e indivíduos a partir de dados genômicos (PRICE et al., 2006; JOMBART, 2008). Contudo, a PCA exige definição a priori de grupos para o estudo de estruturas populacionais, não sendo possível obtê-los a partir dos dados. Além disso, evidenciam não somente a variação entre os grupos (ou seja, a variabilidade genética estruturada), mas também a variação que ocorre dentro de grupos (variabilidade genética aleatória), portanto, a variabilidade geral entre os indivíduos (JOMBART; DEVILLARD; BALLOUX, 2010).

Em contrapartida, a análise discriminante (*discriminant analysis* – DA) define um modelo em que a variação genética é particionada em um componente entre grupos e dentro do grupo e produz variáveis sintéticas (chamadas funções discriminantes) que maximizam o primeiro enquanto minimizam o segundo (JOMBART; PONTIER; DUFOUR, 2009). Em outras palavras, a DA tenta evidenciar a diferenciação genética entre os grupos, ao mesmo tempo que desconsidera a variação dentro do grupo. O método consegue, portanto, a melhor discriminação dos indivíduos em grupos pré-definidos. Curiosamente, esse método também permite a atribuição probabilística de indivíduos para cada grupo, como em métodos de agrupamento bayesiano. Contudo, algumas restrições podem ser atribuídas a DA (BEHARAV; NEVO, 2003), como a necessidade de que o número de variáveis seja menor que o de amostras e também, assim como na PCA, há necessidade da informação a priori dos grupos a serem estudados. Nesse sentido, a PCA pode se tornar ferramenta útil para transformação dos dados originais (alelos) em componentes principais (PC), o que reduz drasticamente a dimensão das variáveis. Embora, a DA também necessite de número prévio de grupos, isso pode ser obtido por algoritmos como o *K-means* (JOMBART; DEVILLARD; BALLOUX, 2010).

Portanto, a DA a partir de PC (*discriminant analysis of principal component* – DAPC) pode ser ferramenta de grande utilidade no estudo da estrutura genética de populações e suas linhagens.

Diversos estudos no Brasil mostraram ganhos genéticos para características de importância em diferentes espécies de produção, (STEFANI, 2016; PIRES et al.,

2015; ANRAIN et al., 2015; MOREIRA et al., 2015), entretanto, diferentemente do ocorrido nestas espécies, não existem programas intensos de melhoramento genético em equinos. O primeiro passo para viabilizar os programas de melhoramento genético consiste no conhecimento dos diferentes fatores que interferem potencialmente na seleção e no progresso genético, como, tamanho efetivo, intervalo de gerações e variabilidade genética.

O monitoramento do ganho genético deve ser uma preocupação para toda e qualquer raça (EUCLIDES FILHO et al., 2000). De acordo com esses pesquisadores, o conhecimento da variabilidade e da evolução genética de uma população tem importância não só para proceder as modificações que se fizerem necessárias, mas também para a avaliação do programa de seleção que vem sendo adotado.

#### 4. BIBLIOGRAFIA CITADA

- ABQM. Associação brasileira dos criadores de cavalos Quarto de Milha. Disponível em: <<http://www.abqm.com.br/>>. Acesso em: 26 out. 2017.
- AL ABRI, M. A.; BORSTEL, U. K. V.; STRECKER, V.; BROOKS, S. A. Application of Genomic Estimation Methods of Inbreeding and Population Structure in an Arabian Horse Herd. **Journal of Heredity**, v. 108, p. 361-368, 2017.
- ALMEIDA, F. Q.; SILVA, V. P. Progresso científico em equideocultura na 1ª década do século XXI. **Rev. Bras. Zootec.**, v. 39, p. 119-129, 2010.
- AMERICA'S HORSE DAILY. All About the Racing American Quarter Horse. 26 de fevereiro de 2008. Disponível em: <[http://americashorsedaily.com/all-about-the-racing-american-quarter-horse/#.Urx8N\\_RDvik](http://americashorsedaily.com/all-about-the-racing-american-quarter-horse/#.Urx8N_RDvik)> Acesso em: 25 set. 2017.
- ANDE-BRASIL. Associação nacional de equoterapia. Disponível em: <<http://www.http://equoterapia.org.br/>>. Acesso em: 09 out. 2017.
- ANRAIN, M.; BERGMANN, J. A. G.; IRGANG, R.; VALENTE, B. D. Parâmetros genéticos para características de prolificidade nas raças Landrace e Large White. **Arq. Bras. Med. Vet. Zootec.**, Belo Horizonte, v. 67, n. 3, p. 846-854, 2015.
- AQHA. AMERICAN QUARTER HORSE ASSOCIATION. Horse-Breeding History. Disponível em: <<https://www.aqha.com/daily/breeding/2016/breeding-archive/horsebreeding-history/>>. Acesso em: 27 out. 2017.

- ARDLIE, K. G.; KRUGLYAK, L.; SEIELSTAD, M. Patterns of linkage disequilibrium in the human genome. **Natural Review Genetics**, v. 3, p. 299-309, 2002.
- ARNASON, T. Bright future for research in horse breeding! **Journal of Animal Breeding and Genetics**, v. 9, p. 130-167, 2013.
- BARCACCIA, G.; FELICETTI, M.; GALLA, G.; CAPOMACCIO, S.; CAPPELLI, K.; ALBERTINI, E.; BUTTAZZONI, L.; PIERAMATI, C.; SILVESTRELLI, M.; SUPPLIZI, A. V. Molecular analysis of genetic diversity, population structure and inbreeding level of the Italian Lipizzan horse. **Livestock Science**, v. 151, p. 124-133, 2013.
- BARNAUD, A.; LAUCOU, V.; THIS, P.; LACOMBE, T.; DOLIGEZ, A. Linkage disequilibrium in natural French grapevine, *Vitis vinifera* L. subsp. *silvestris*. **Heredity**, v. 104, p. 431–437, 2010.
- BARRET, S. C. H.; KOHN, J. R. Genetic and evolutionary consequences of small population sizes in plants: Implications for conservation. In: Falk, D.A. and Holsinger, K.A. (Eds.). **Genetics and Conservation of Rare Plants**. Oxford: University Press, New York, 1991. p. 3-30.
- BEHARAV, A.; NEVO, E. Predictive validity of discriminant analysis for genetic data. **Genetica**, v. 119, p. 259–267, 2003.
- BELTRÁN, N. A. R. **Assinaturas de seleção na linhagem de trabalho de equinos quarto de milha**. 2014. 59 f. Tese (Doutorado corrente em Genética e Melhoramento Animal) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2014.
- BINNS, M. M.; BOEHLER, D. A.; LAMBERT, D. H. Identification of the myostatin locus (MSTN) as having a major effect on optimum racing distance in the Thoroughbred horse in the USA. **Animal Genetics**, v. 41, supl. 2, p. 154-158, 2010.
- BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. Characteristics of linkage disequilibrium in North American Holsteins. **BMC Genomics**, v. 11, p. 421, 2010.
- BOKOR, A.; BLOUIN, C.; LANGLOIS, B.; STEFLER, J. Genetic parameters of racing merit of Thoroughbred horses in steeplechase races. **Italian Journal of Animal Science**, v. 5, p. 4-43, 2005.
- BOLORMAA, S.; HAYES, B. J.; SAVIN, K.; HAWKEN, R.; BARENDSE, W.; ARTHUR, P. F.; HERD, R. M.; GODDARD, M. E. Genome-wide association studies for feedlot and growth traits in cattle. **Journal of Animal Science**, v. 89, p. 1684-1697, 2011.

- BROOKS, S. A., GABRESKI, N.; MILLER, D.; BRISBIN, A.; BROWN, H. E.; STREETER, C.; MEZEY, J.; COOK, D.; ANTCZAK, D. F. Whole-genome SNP association in the horse: Identification of a deletion in Myosin Va responsible for Lavender Foal Syndrome. **PLoS Genet.**, v. 6(4):e1000909, 2010.
- CAETANO, A. R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v. 38, n. esp., p. 64-71, 2009.
- CARJAVAL, A. B. **Estrutura populacional e estimação de parâmetros genéticos para características associadas ao desempenho reprodutivo em bovinos da raça Caracu**. 2017. 61 f. Dissertação (mestrado) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2017.
- CHOWDHARY, B. P.; RAUDSEPP, T. The Horse Derby: racing from map to whole genome sequence. **Chromosome Research**, v. 16, p. 109-127, 2008.
- COOK, D.; GALLAGHER, P.; BAILEY, E. Genetics of swayback in American Saddlebred horses. **Animal Genetics**, v. 41, supl. 2, p. 64-71, 2010.
- CORRÊA, M. J. M.; MOTA, M. D. S. Genetic evaluation of performance traits in Brazilian Quarter Horse. **Journal of Applied Genetics**, v. 48, p. 145-151, 2007.
- DU, F. X.; CLUTTER, A. C.; LOHUIS, M. M. Characterizing linkage disequilibrium in pig populations. **Int. J. Biol. Sci.**, v. 3, p. 166-178, 2007.
- DUPUIS, M. C.; ZHANG, Z.; DRUET, T.; DENOIX, J. M.; CHARLIER, C.; LEKEUX, P.; GEORGES, M. Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. **Mammalian Genome**, v. 22, n. 9-10, p. 613-20, 2011.
- EBERTH, J.; SWERCZAK, T.; BAILEY, E. Investigation of Dwarfism Among Miniature Horses using the Illumina Horse SNP50 Bead Chip. **Journal of Equine Veterinary Science**, v. 29, n. 5, p. 315, 2009.
- ELLERSIECK, M. R.; LOCK, W. E.; VOGT, D. W.; AIPPERSPACH, R. Genetic evaluation of cutting scores in horses. **Equine Veterinary Science**, v. 5, p. 287-289, 1985.
- EUCLIDES FILHO, K.; SILVA, L. O. C.; ALVES, R. G. O.; FIGUEIREDO, G. R. Tendência genética na raça Gir. **Pesquisa Agropecuária Brasileira**, v. 35, p. 787-791, 2000.
- EVANS, J. W. Horses: a guide to selection, care and enjoyment. Freeman and Company: New York, 4 ed., 1996. 797 p.

- FALCONER, D. S. Introdução à genética quantitativa. Trad. de Martinho de Almeida e Silva e José Carlos Silva. Viçosa, UFV, Impr. Univ., 1987. 279 p.
- FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: dominant markers and null alleles. **Molecular Ecology Notes**, v. 7, p. 574–578, 2007.
- FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. **Genetics**, v. 164, p. 1567–1587, 2003.
- FAO. Food and Agriculture Organization of United Nations. Disponível em: <<http://www.fao.org/home/en/>>. Acesso em 14 de nov. 2014.
- FARIA, R. A. da S.. **Estrutura populacional e parâmetros genéticos da característica classe de tempo em corridas de equinos da Raça Quarto de Milha**. 2016. 69 f. Dissertação (Mestrado) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal, 2016.
- FARNIR, F.; COPPIETERS, W.; ARRANZ, J. J.; BERZI, P.; CAMBISANO, N.; GRISART, B.; KARIM, L.; MARCQ, F.; MOREAU, L.; MNI, M.; NEZER, C.; SIMON, P.; VANMANSHOVEN, P.; WAGENAAR, D.; GEORGES, M. Extensive genome-wide linkage disequilibrium in cattle. **Genome Research**, v. 10, p. 220-227, 2000.
- FERREIRA, M. E.; GRATTAPAGLIA, D. Introdução ao uso de marcadores moleculares em análise genética. Embrapa-Cenargen: Brasília, 3 ed., 1998. 220 p.
- FONSECA, M. G.; FERRAZ, G. C.; LAGE, J.; PEREIRA, G. L.; CURI, R. A. A genome-wide association study reveals differences in the genetic mechanism of control of the two gait patterns of the Brazilian Mangalarga Marchador breed. **Journal of Equine Veterinary Science**, v. 53, p. 64-67, 2017.
- FREEMAN, D. W. Physical conditioning of horses. Stillwater, Okla: Oklahoma State University, Division of Agricultural Sciences and Natural Resources. 2013.
- GIBBS, R. A.; TAYLOR, J. F.; VAN TASSELL, C. P. Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. **Science**, v. 324, p. 528-532, 2009.
- GLIPHA. Global Livestock Production and Health Atlas. Disponível em: <[http://www.fao.org/ag/againfo/home/en/news\\_archive/AGA\\_in\\_action/glipha.html](http://www.fao.org/ag/againfo/home/en/news_archive/AGA_in_action/glipha.html)>. Acesso em: 11 set. 2017.
- GODDARD, M. E.; HAYES, B. J. Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nat. Rev. Genet.**, v. 10, p. 381-391, 2009.

- GU, S.; PAKSTIS, A. J.; LI, H.; SPEED, W. C.; KIDD, J. R.; KIDD, K. K. Significant variation in haplotype block structure but conservation in tagSNP patterns among global populations. **Eur J Hum Genet**, v. 15, p. 302-312, 2007.
- GUIMARÃES, P. E. M.; COSTA, M. C. R. SNPs: Sutis diferenças de um código. **Biotecnologia Ciência e Desenvolvimento**, v. 26, p. 24-27, 2002.
- GURGUL, A.; EMIK, E.; PAWLINA, K.; SZMATOLA, T.; JASIELCZUK, I.; BUGNOPONIEWIERSKA, M. The application of genome-wide SNP genotyping methods in studies on livestock genomes. **Journal of Applied Genetics**, v. 55, n. 2, p. 197-208, 2014.
- HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J.; GODDARD, M. E. Invited review: Genomic selection in dairy cattle: progress and challenges. **J Dairy Sci**, v. 92, n. 2, p. 433-443, 2009.
- HAYES, B. J.; VISSCHER, P. M.; MCPARTLAN, H. C.; E. GODDARD, M. E. Novel Multilocus Measure of Linkage Disequilibrium to Estimate Past Effective Population Size. **Genome Research**, v. 13, p. 635-643, 2003.
- HESTAND, M. S.; KALBFLEISCH, T. S.; COLEMAN, S. J.; ZENG, Z.; LIU, J.; ORLANDO, L.; MACLEOD, J. N. Annotation of the Protein Coding Regions of the Equine Genome. **PLoS ONE**, 10(6):e0124375, 2015.
- HILL, E. W.; MCGIVNEY, B. A.; GU, J.; WHISTON, R.; MACHUGH, D. E. A genome-wide SNP-association study confirms a sequence variant (g.66493737C>T) in the equine myostatin (MSTN) gene as the most powerful predictor of optimum racing distance for Thoroughbred race horses. **BMC Genomics**, v. 11, n. 552, p. 1-10, 2010.
- HILL, W. Estimation of effective population size from data on linkage disequilibrium. **Genetical Research**, v. 38, n. 3, p. 209-216, 1981.
- HILL, W. G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **Theoretical and Applied Genetics**, v. 38, p. 226-231, 1968.
- HINTZ, R. L. Genetics performance in the horse. **Journal of Animal Science**, v. 51, p. 582-594, 1980.
- HRISTOV, P. I.; TEOFANOVA, D. R.; NEOV, B. S.; ZAGORCHEV, L. I.; RADOSLAVOV, G. A. Population structure of two native bulgarian cattle breeds with regard to CSN3 and CSN1S1 gene polymorphism. **Bulgarian Journal of Veterinary Medicine**, v. 17, p. 18-24, 2014.

- IBGE. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Sistema IBGE de Recuperação Automática – SIDRA. Disponível em: <<http://www.sidra.ibge.gov.br/>>. Acesso em: 7 nov. 2017.
- ISO-TOURU, T.; TAPIO, M.; VILKKI, J.; KISELEVA, T.; AMMOSOV, I.; IVANOVA, Z.; POPOV, R.; OZEROV, M.; KANTANEN, J. Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe. **Anim Genet**, v. 47, p. 647–657, 2016.
- JCS. Especial Quarto de Milha de Corrida. Jockey Club de Sorocaba, Sorocaba, SP, 2002.
- JOMBART, T. Adegnet: a R package for the multivariate analysis of genetic markers. **Bioinformatics**, v. 24, p. 1403–1405, 2008.
- JOMBART, T.; DEVILLARD, S.; BALLOUX, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. **Bmc Genet**, v. 11, p. 94-109, 2010.
- JOMBART, T.; PONTIER, D.; DUFOUR, A. B. Genetic markers in the playground of multivariate analysis. **Heredity**, 102(4), p. 330–341, 2009.
- KHATKAR, M. S.; COLLINS, A.; CAVANAGH, J. A. L.; HAWKEN, R. J.; HOBBS, M.; ZENGER, K. R.; BARRIS, W.; McCLINTOCK, A. E.; THOMSON, P. C.; NICHOLAS, F. W.; RAADSMA, H. W. A first-generation metric linkage disequilibrium map of bovine chromosome 6. **Genetics**, v. 174, p. 79-85, 2006.
- KIJAS, J. W.; TOWNLEY, D.; DALRYMPLE, B. P.; HEATON, M. P.; MADDOX, J. F.; MCGRATH, A.; WILSON, P.; INGERSOLL, R. G.; MCCULLOCH, R.; MCWILLIAM, S.; TANG, D.; MCEWAN, J.; COCKETT, N. V.; ODDY, H.; NICHOLAS, F. W.; RAADSMA, H. A. Genome Wide Survey of SNP Variation Reveals the Genetic Structure of Sheep Breeds. **PLoSOne**, 4(3):e4668, 2009.
- KWOK, P. Y.; GU, Z. Single nucleotide polymorphism libraries: why and how are we building them? **Mol Med Today**, v. 5, p. 538-543, 1999.
- LEWONTIN, R. C. The interaction of selection and linkage. I. General consideration sheterotic models. **Genetics**, v. 49, p. 49-67, 1964.
- LIMA, B. N. O. **Variabilidade na expressão da reatividade de potros brasileiro de hipismo não associada à polimorfismo do gene DBH**. 2017. 38 f. Dissertação (Mestrado em Genética) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal, 2017.

- LIMA, R. A. S.; SHIROTA, R.; BARROS, G. S. C. Estudo do Complexo do Agronegócio Cavalo no Brasil. (Relatório Final), CEPEA – Centro de Estudos Avançados em Economia Aplicada, Piracicaba: ESALQ/USP, SP, Brasil, 2006. 250 p.
- MANSO FILHO, H. C.; COTHRAN, E. G.; JURAS, R.; GOMES FILHO, M. A.; SILVA, N. M. V da.; SILVA, G. B. da.; FERREIRA, L. M. C.; ABREU, J. M. G.; MANSO, H. E. C. da C. C. Alelo DMRT3 mutante em equinos de marcha batida e picada das raças Campolina e Mangalarga Marchador. **Ciência Veterinária nos Trópicos**, v. 18, n. 1, p. 6-11, 2015.
- MAPA. Ministério da Agricultura, Pecuária e Abastecimento. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em: 18 ago. 2017.
- McCUE, M.; MICKELSON, J. **Genomic Tools and Resources: Development and Applications of an Equine SNP Genotyping Array**. In: Equine Genomics (Ed. B. P. Chowdhary), Blackwell Publishing Ltd., Oxford, UK, 2013.
- MCKAY, S. D.; SCHNABEL, R. D.; MURDOCH, B. M.; MATUKUMALLI, L. K.; AERTS, J.; COPPIETERS, W.; CREWS, D.; DIAS NETO, E.; GILL, C. A.; GAO, C.; MANNEN, H.; STOTHARD, P.; WANG, Z.; VAN TASSELL, C. P.; WILLIAMS, J. L.; TAYLOR, J. F.; MOORE, S. S. Whole genome linkage disequilibrium maps in cattle. **BMC Genetics**, v. 8, p. 74, 2007.
- MCKAY, S. D.; SCHNABEL, R. D.; MURDOCH, B. M.; MATUKUMALLI, L. K.; AERTS, J.; COPPIETERS, W.; CREWS, D.; DIAS NETO, E.; GILL, C. A.; GAO, C.; MANNEN, H.; WANG, Z.; VAN TASSELL, C. P.; WILLIAMS, J. L.; TAYLOR, J. F.; MOORE, S. S. An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. **BMC Genetics**, v. 9, p. 37, 2008.
- MCRAE, A. F.; MCEWAN, J. C.; DODDS, K. G.; WILSON, T.; CRAWFORD, A. M.; SLATE, J. Linkage disequilibrium in domestic sheep. **Genetics**, v. 160, p. 1113-1122, 2002.
- MEIRA, C. T.; CURI, R. A.; FARAH, M. M.; OLIVEIRA, H. N.; BELTRAN, N. A. R.; SILVA, J. A. II V.; MOTA, M. D. S. Prospection of genomic regions divergently selected in racing line of Quarter Horses in relation to cutting line. **Animal Journal**, v. 8, ed. 11, p. 1754-1764, 2014a.
- MEIRA, C. T.; FARAH, M. M.; FORTES, M. R. S.; MOORE, S. S.; PEREIRA, G. L.; SILVA, J. A. II V.; MOTA, M. D. S.; CURI, R. A. A Genome-wide association study for morphometric traits in Quarter Horse. **Journal of Equine Veterinary Science**, v. 34, ed. 8, p. 102-1031, 2014b.
- MEIRA, C. T.; FORTES, M. R. S.; FARAH, M. M.; PORTO-NETO, L. R.; CURI, R. A.; MOORE, S. S.; PEREIRA, G. L.; SILVA, J. A. II V. A genome-wide association study



for height at withers in racing quarter horse. **Proc. Assoc. Advmt. Anim. Breed. Genet.**, v. 20, p. 420-23, 2013.

MEIRA, C. T.; FORTES, M. R. S.; FARAH, M. M.; PORTO NETO, L. R.; KELLY, M.; MOORE, S. S.; PEREIRA, G. L.; CHARDULO, L. A. L.; CURI, R. A. Speed Index in the Racing Quarter Horses: A Genome wide association study. **Journal of Equine Veterinary Science**, v. 34, ed. 11-12, p. 1263-1268, 2014c.

MEUWISSEN, T. H. E.; GODDARD, M. E. Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. **Genetics**, v. 155, p. 421- 430, 2000.

MOREIRA, H. L.; CANOVA, E. B.; MUNARI, D. P.; BEZERRA, L. A. F., LÔBO, R. B.; PAZ, C. C. P. de. Parâmetros genéticos para período de gestação e características de crescimento pré e pós desmame em bovinos Nelore. **Boletim de Indústria Animal**, v. 72, n. 2, p. 130-135, 2015.

MUIR, W. M.; WONGB, G. K.; ZHANG, Y.; WANGC, J.; GROENEND, M. A. M.; CROOIJMANS, R. P. M. A.; MEGENSD, H. J.; ZHANG, H.; OKIMOTO, R.; VEREIJENG, A.; JUNGIERUSG, A.; ALBERSG, G. A. A.; LAWLEYH, C. T.; DELANVI, M. E.; MACEACHERNE, S.; CHENG, H. H. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, p. 17312-17317, 2008.

NAKAYA, H. L.; AMARAL, P. P.; LOURO, R.; LOPES, A.; FACHEL, A. A.; MOREIRA, Y. B.; EL-JUNDI, T. A.; SILVA, A. M.; REIS, E. M.; VERJOVSKI-ALMEIDA, S. Comment reviews reports deposited research refereed research interactions information Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. **Genome Biology**, v. 8(R43), p.1-25, 2007.

NCHA. National Cutting Horse Association. Disponível em: <<http://www.nchacutting.com>>  
Acesso em: 13 set. 2017

ODANI, M.; NARITA, A.; WATANABE, T.; YOKOUCHI, K.; SUGIMOTO, Y.; FUJITA, T.; OGUNI, T.; MATSUMOTO, M.; SASAKI, Y. Genome-wide linkage disequilibrium in two Japanese beef cattle breeds. **Animal Genetics**, v. 37, p. 139-144, 2006.

PEREIRA, G. L.; CHUD, T.; BERNARDES, P.; VENTURINI, G.; CHARDULO, L. A.; CURI, R. Genotype Imputation and Accuracy Evaluation in Racing Quarter Horses Genotyped Using Different Commercial SNP Panels. **Journal of Equine Veterinary Science**, v. 58, p. 89-96, 2017.

- PEREIRA, G. L. **Identificação de regiões cromossômicas, genes e polimorfismos de DNA associados ao desempenho de equinos de corrida da raça quarto de milha.** 2017. 98 f. Tese (Doutorado melhoramento animal corrente em Genética e) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2017.
- PEREIRA, G. L.; MATTEIS, R.; REGITANO, L. C. A.; CHARDULO, L. A. L.; CURI, R. A. MSTN, CKM and DMRT3 gene variants in different lines of Quarter Horses. **Journal of Equine Veterinary Science**, v. 39, p. 33-37, 2016a.
- PEREIRA, G. L.; MATTEIS, R.; REGITANO, L. C. A.; MEIRA, C. T.; SILVA, J. A. II V.; CHARDULO, L. A. L.; CURI, R. A. Comparison of sequence variants in the PDK4 and COX4I2 genes between racing and cutting lines of quarter horses and associations with the speed index. **Journal of Equine Veterinary Science**, v. 39, p. 1-6, 2016b.
- PETERSEN, J. L.; MICKELSON, J. R.; RENDAHL, A. K.; VALBERG, S. J.; ANDERSSON, L. S.; AXELSSON, J.; BAILEY, E.; BANNASCH, D.; BINNS, M. M.; BORGES, A. S.; BRAMA, P.; MACHADO, A. da C.; CAPOMACCIO, S.; CAPPELLI, K.; COTHRAN, E. G.; DISTL, O.; FOX-CLIPSHAM, L.; GRAVES, K. T.; GUÉRIN, G.; HAASE, B.; HASEGAWA, T.; HEMMANN, K.; HILL, E. W.; LEEB, T.; LINDGREN, G.; LOHI, H.; LOPES, M. S.; MCGIVNEY, B. A.; MIKKO, S.; ORR, N.; PENEDO, M. C. T.; PIERCY, R. J.; RAEKALLIO, M.; RIEDER, S.; RØED, K. H.; SWINBURNE, J.; TOZAKI, T.; VAUDIN, M.; WADE, C. M.; MCCUE, M. E. Genome-Wide Analysis Reveals Selection for Important Traits in Domestic Horse Breeds. **PLoS Genetics**, v. 9, n. 1, p. 1-17, 2013.
- PIRES, M. P.; FARAH, M. M.; CARREÑO, L. O. D.; UTSUNOMIYA, A. T. H.; ONO, R. K.; BERTIPAGLIA, T. S.; FONSECA, R. Estimates of genetic parameters for growth traits in Suffolk sheep in Brazil. **Arg. Bras. Med. Vet. Zootec.**, Belo Horizonte, v. 67, n. 4, p. 1119-1124, 2015.
- PORTO-NETO, L.; KIJAS, J.; REVERTER, A. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. **Genetics Selection Evolution**, v. 46, n. 2, 2014.
- PRADO, R. S. A.; MOTA, M. D. S. Correlações genéticas entre deslocamento e conformação em equinos Mangalarga. **Archivos de Zootecnia**, v. 57, p. 165-169, 2008.
- PRICE, A. L.; PATTERSON, N. J.; PLENGE, R. M.; WEINBLATT, M. E.; SHADICK, N. A.; REICH, D. Principal components analysis corrects for stratification in genome-wide association studies. **Nature Genetics**, v. 38, p. 904–909, 2006.

- QANBARI, S.; SIMIANER, H. Mapping signatures of positive selection in the genome of livestock. **Livestock Science**, v. 166, p. 133-143, 2014.
- REGATIERI, I. C.; PEREIRA, G. L.; TEIXEIRA NETO, A. R.; FERRAZ, G. C.; CURI, R. A.; QUEIROZ-NETO, A. Polymorphisms in MCT1, CD147, PDK4 and DMRT3 genes in Arabian and Quarter Horses. **Journal of Equine Veterinary Science**, v. 48, p. 161-165, 2016.
- SANTOS, L. M. **Morfologia e genética do cavalo Campolina**. 2006. 48 f. Dissertação (Mestrado em Genética) – Universidade Federal de Minas Gerais, Belo Horizonte, 2006.
- SARGOLZAEI, M.; SCHENKEL, F. S.; JANSEN, G. B.; SCHAEFFER, L. R. Extent of linkage disequilibrium in Holstein cattle in North America. **Journal of Dairy Science**, v. 91, p. 2106-2117, 2008.
- SCHRÖDER, W.; KLOSTERMANN, A.; STOCK, K. F.; DISTL, O. A genome-wide association study for quantitative trait loci of show-jumping in Hanoverian Warm blood horses. **Animal Genetics**, v. 43, n. 4, p. 392-400, 2011.
- SOUZA, G. A. de.; CARVALHO, M. R. de O.; MARTINS, E. R.; GUEDES, R. N. C.; OLIVEIRA, L. O. de. Diversidade genética estimada com marcadores ISSR em populações brasileiras de *Zabrotes subfasciatus*. **Pesquisa Agropecuária Brasileira**, v. 43, n. 7, p. 843-49, 2008.
- STEFANI, G. **Associação entre longevidade e características de tipo, produção de leite e saúde do úbere de vacas da raça Holandesa**. 2016. 36 f. Dissertação (Mestrado) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal, 2016.
- SVED, J. A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. **Theor. Popul. Biol.**, v. 2, p. 125-141, 1971.
- TEYSSÈDRE, S.; DUPUIS, M. C.; GUÉRIN, G.; SCHIBLER, L.; DENOIX, J. M.; ELSÉN, J. M.; RICARD, A. Genome-wide association studies for osteochondrosis in French Trotters. **Journal Animal Science**, v. 90, p. 45-53, 2012.
- VOZZI, P. A.; MARCONDES, C. R.; MAGNABOSCO, C. de U.; BEZERRA, L. A. F.; LÔBO, R. B. Structure and genetic variability in Nelore (*Bos indicus*) cattle by pedigree analysis. **Genetics and Molecular Biology**, v. 29, p. 482-485, 2006.
- WADE, C. M.; GIULOTTO, E.; SIGURDSSON, S.; ZOLI, M.; GNERRE, S.; IMSLAND, F.; LEAR, T. L.; ADELSON, D. L.; BAILEY, E.; BELLONE, R. R.; BLOCKER, H.; DISTL, O.; EDGAR, R. C.; GARBER, M.; LEEB, T.; MAUCELI, E.; MACLEOD, J. N.; PENEDO,

M. C. T.; RAISON, M.; SHARPE, T.; VOGEL, J.; ANDERSSON, L.; ANTCZAK, D. F.; BIAGI, T.; BINNS, M. M.; CHOWDHARY, B. P.; COLEMAN, S. J.; DELLA VALLE, G.; FRYC, S.; GUERIN, G.; HASEGAWA, T.; HILL, E. W.; JURKA, J.; KIIALAINEN, A.; LINDGREN, G.; LIU, J.; MAGNANI, E.; MICKELSON, J. R.; MURRAY, J.; NERGADZE, S. G.; ONOFRIO, R.; PEDRONI, S.; PIRAS, M. F.; RAUDSEPP, T.; ROCCHI, M.; ROED, K. H.; RYDER, O. A.; SEARLE, S.; SKOWN, L.; SWINBURNE, J. E.; SYVANEN, A. C.; TOZAKI, T.; VALBERG, S. J.; VAUDIN, M.; WHITE, J. R.; ZODY, M. C.; Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team; LANDER, E. S.; LINDBLAD-TOH, K. Genome sequence, comparative analysis, and population genetics of the domestic horse. **Science**, v. 326, p. 865-867, 2009.

WALLIN, L.; STRANDBERG, E.; PHILIPSSON, J. Genetic correlations between field test results of Swedish Warmblood Riding Horses as 4-year-olds and life time performance results in dressage and show jumping. **Livestock Production Science**, v. 82, p. 61-71, 2003.

YOUNG, A.; BOYLE, T.; BROWN, T. The population genetic consequences of habitat fragmentation for plants. **Trends Ecol. Evol.**, v. 11, p. 413-418, 1996.

ZENGER, K. R.; KHATKAR, M. S.; CAVANAGH J. A. L.; HAWKEN, R. J.; RAADSMA, H. W. Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. **Animal Genetics**, v. 38, p. 7-14, 2006.

ZHU, M.; ZHAO, S. Candidate Gene Identification Approach: Progress and Challenges. **International Journal of Biological Sciences**, v. 3, n. 7, p. 420-427, 2007.

## **CAPÍTULO 2 – Caracterização genômica de equinos das linhagens de trabalho e de corrida da raça Quarto de Milha**

**RESUMO** – Na raça Quarto de Milha de equinos (*Equus caballus*), diferentes objetivos de seleção levaram à formação de linhagens com distintas habilidades, entre as quais a de trabalho e a de corrida. Estes grupos também diferem em relação às características morfológicas, fisiológicas e em relação à consanguinidade. O objetivo deste estudo foi caracterizar, por meio da genotipagem de SNP em larga escala, o desequilíbrio de ligação (LD), calculado por  $r^2$ , nas linhagens de trabalho e de corrida de cavalos Quarto de Milha criados no Brasil. Também foi investigado o tamanho efetivo ( $N_e$ ) das duas populações, bem como as suas estruturas e relações. Foram utilizados 428 equinos Quarto de Milha, de ambos os sexos, registrados na associação Brasileira de criadores da raça (ABQM), sendo 68 da linhagem de trabalho e 360 da de corrida. Sessenta e oito animais da linhagem de trabalho e 120 da de corrida foram genotipados utilizando arranjo de 54.602 SNP (54K). As demais amostras da linhagem de corrida ( $n = 240$ ) foram genotipadas com arranjo de 65.157 SNP (65K). O  $r^2$  genômico médio entre pares de marcadores foi 0,22 para a linhagem de trabalho e 0,27 para a de corrida. Na linhagem de trabalho o  $r^2$  foi inferior a 0,2 entre 100 e 150 Kb, enquanto na de corrida foram encontrados valores de  $r^2$  menores que 0,2 entre 300 e 350 Kb. A estimativa dos valores de  $N_e$  foi de 60 e 50 animais na última geração para as linhagens de trabalho e de corrida, respectivamente. A maior extensão do LD e o menor  $N_e$  na linhagem de corrida podem ser explicados pela sua estrutura populacional mais fechada em decorrência do maior rigor no registro genealógico dos animais, além da grande influência do Puro-Sangue Inglês na sua formação. O estudo da estrutura da população indicou clara distinção entre as duas linhagens da raça. Na linhagem de corrida, contudo, foram evidenciadas subestruturas populacionais decorrentes da formação de famílias que descendem de diferentes e importantes reprodutores.

**Palavras-chave:** *BeadChip*, cavalo, marcadores moleculares, polimorfismos de DNA, tamanho efetivo

## 1. INTRODUÇÃO

O cavalo (*Equus caballus*) Quarto de Milha da linhagem de trabalho, destina-se às provas de caráter funcional, explorando habilidades como agilidade e obediência, características consideradas de grande importância no manejo do gado a campo (MEIRA et al., 2014b). A linhagem de corrida explora a aptidão dos animais quanto à velocidade em pistas retas e de curta distância. Equinos desta linhagem da raça Quarto de Milha têm melhor desempenho em corridas de curtas distâncias do que qualquer outra raça (ABQM, 2016).

Caracterizados por sua versatilidade, equinos Quarto de Milha podem realizar variados tipos de exercícios físicos, indo do predominantemente aeróbio ao predominantemente anaeróbio. Alguns exercícios de arena, como apartação e rédeas, modalidades executadas pela linhagem de trabalho, intercalam turnos curtos de exercício anaeróbio com maiores períodos de atividade aeróbia. Já animais da linhagem de corrida apresentam metabolismo muscular predominantemente anaeróbio ao longo de suas corridas de curta distância (FREEMAN, 2013).

Em geral, o cavalo de corrida da raça Quarto de Milha é mais alto que o de trabalho devido à influência de animais Puro-Sangue Inglês (PSI), utilizados para melhorar a sua velocidade (MEIRA et al., 2013; PETERSEN et al., 2014). Entretanto, a seleção direcionada a diferentes objetivos (trabalho ou corrida) também promoveu mudanças significativas nas características corporais desses equinos (MEIRA et al., 2013). Nessa mesma pesquisa foi verificado maior coeficiente de endogamia médio nos animais de corrida quando comparados aos de trabalho, sugerindo a ocorrência de maior prática de acasalamentos endogâmicos nos primeiros.

As metodologias de genotipagem de SNP em larga escala possibilitam a realização de inúmeros estudos, como a identificação de regiões genômicas modificadas pela seleção e a identificação de regiões cromossômicas, genes e polimorfismos de nucleotídeo único (*Single Nucleotide Polymorphisms* – SNP) que contribuem para a variação de características de interesse nas principais raças de equinos criadas no mundo (PEREIRA et al., 2018). Muitos destes estudos requerem o conhecimento ou acesso ao desequilíbrio de ligação (*linkage disequilibrium* – LD) existente entre os marcadores moleculares em dada população da espécie de interesse.

O desequilíbrio de ligação é definido como uma associação ou ligação originada da proximidade física existente entre dois locos e é influenciado por eventos como mutação, deriva, seleção e recombinação. A quantificação do nível de LD é importante, visto que a sua extensão pode auxiliar nas estimativas do tamanho efetivo da população atual e anterior, principalmente quando os registros de *pedigree* não estão disponíveis. Além da relação com o tamanho efetivo da população (WANG et al., 2013), a extensão do LD e a variabilidade haplotípica podem revelar informações sobre o histórico populacional e desenvolvimento de raças (WANG et al., 2013), melhor compreensão da arquitetura genômica e da estrutura histórica da população (HAYES et al., 2009).

Dada a importância dos cavalos da raça Quarto de Milha para o agronegócio nacional e considerando as diferentes pressões de seleção exercidas, as diferentes práticas de acasalamento, os diferentes graus de introdução de animais de outras raças e as diferentes aptidões existentes entre as linhagens de trabalho e de corrida, o objetivo desta pesquisa foi caracterizar, por meio da genotipagem de SNP em larga escala, o desequilíbrio de ligação e o tamanho efetivo das duas populações, bem como as suas estruturas e relações.

## **2. MATERIAL E MÉTODOS**

### **2.1. Animais**

Foram utilizados 428 equinos Quarto de Milha, de ambos os sexos, registrados na Associação Brasileira de Criadores da raça (ABQM), sendo 68 da linhagem de trabalho e 360 da de corrida. Os animais da linhagem de trabalho (25 machos e 43 fêmeas) eram filhos de 44 garanhões e 65 éguas, resultando em média de 1,5 progênie/garanhão e 1,04 progênie/égua. Estes animais destinavam-se a provas de rédeas e apartação, pertenciam a 45 criadores e encontravam-se alojados em três propriedades no interior do estado de São Paulo, Brasil. Por sua vez, os animais da linhagem de corrida (78 machos e 282 fêmeas) eram filhos de 83 garanhões e 249 éguas, resultando em média de 4,3 progênies/garanhão e 1,4 progênie/égua. Estes animais pertenciam a 156 criadores e encontravam-se alojados em 25 propriedades no interior do estado de São Paulo, Brasil e no Jockey Club de Sorocaba, São Paulo, Brasil.

A amostra total foi estabelecida de forma a representar, da melhor maneira possível, a diversidade da população das duas linhagens da raça no Brasil. Foram coletados animais com grande número de filhos e descendentes de indivíduos influentes na formação da raça. A amostragem de irmãos completos foi evitada em ambas as linhagens.

## **2.2. Coleta de sangue e extração do DNA**

As colheitas de amostras de sangue foram realizadas nos anos de 2011, 2013 e 2015. Amostras de 5 mL de sangue total de cada animal foram colhidas por venipunção da jugular esquerda, na região do pescoço, utilizando-se tubos com vácuo, contendo 7,5 mg de EDTA. Todos os procedimentos envolvendo a manipulação de animais foram realizados segundo as normas brasileiras de bem-estar animal (Protocolo nº 110/2016-CEUA expedido pela Comissão de Ética no Uso de Animais – CEUA, da Faculdade de Medicina Veterinária e Zootecnia da Unesp, Botucatu, São Paulo, Brasil).

O processo de extração do DNA a partir de alíquotas de sangue foi realizado nas estruturas do Laboratório de Genética e Melhoramento Animal (LGMA) do Departamento de Melhoramento e Nutrição Animal da FMVZ, Unesp, Botucatu/SP, utilizando o *Illustra Blood GenomicPrep Mini Spin Kit* (GE Healthcare, EUA), de acordo com as instruções do fabricante. Ao final do procedimento de extração, o DNA foi quantificado e analisado quanto à pureza (relações 260/280 e 260/230) em espectrofotômetro de pequenos volumes do modelo *NanoDrop* (Thermo Scientific, EUA) e analisado quanto a sua integridade em gel de agarose a 0,8%. Após a diluição para a concentração de trabalho necessária à genotipagem dos SNP (aproximadamente 50 ng/uL), as amostras de DNA foram estocadas em mini-tubos de 0,2 mL e armazenadas a -20°C.

## **2.3. Genotipagem e controle de qualidade**

Cento e oitenta e oito dos 428 animais utilizados neste estudo, 68 da linhagem de trabalho e 120 da de corrida foram genotipados no ano de 2011, utilizando o *Equine SNP50 BeadChip* (Illumina, Inc., EUA), que dispõe, em sua matriz, de 54.602 SNP (54K). As demais amostras de DNA da linhagem de corrida (n = 240) foram genotipadas no ano de 2015, com o *Equine SNP70 BeadChip* (Illumina, Inc., EUA),



que possui 65.157 SNP (65K). Ambos os chips foram escaneados utilizando o sistema *HiScan* (Illumina, Inc., EUA).

Os controles de qualidade das informações de genotipagem, para indivíduos e SNP, foram realizados por meio do pacote *snpStats* (CLAYTON, 2015) do programa *R* (R Core Team, 2016). Em relação aos indivíduos foi utilizado o *call rate*, em que animais com valores abaixo de 0,90 foram excluídos do conjunto amostral. Em relação aos SNP, foram excluídos os localizados no cromossomo X, os com *call frequency* inferior a 0,9 e os com *p-value* inferior a  $1 \times 10^{-5}$  para o teste de equilíbrio de *Hardy-Weinberg* (*Hardy-Weinberg Equilibrium* – HWE). Na linhagem de trabalho, foram excluídos os SNP com frequências de alelos menores (*minor allele frequency* – MAF) inferiores a 0,05. A princípio, não se utilizou a MAF como critério para exclusão de SNP na linhagem de corrida uma vez que as diferentes faixas de MAF foram utilizadas no processo de verificação da eficiência de imputação de genótipos em duas vias. Os números de marcadores de ambos os arranjos (54K e 65K) que permaneceram após controle de qualidade estão apresentados na Tabela 1.

**Tabela 1** – Número de marcadores dos chips *Equine SNP50 Beadchip* e *Equine SNP70 Beadchip* e número de SNP excluídos por cada um dos critérios utilizados no controle de qualidade das linhagens de trabalho e de corrida da raça Quarto de Milha.

Critérios de exclusão de amostras	Trabalho SNP50 (n)	Corrida SNP50 (n)	Corrida SNP70 (n)
<i>Call rate</i> < 0,9	7	4	9
Amostras restantes	61	116	231
Critérios de exclusão de SNP	Trabalho SNP50 (n)	Corrida SNP50 (n)	Corrida SNP70 (n)
SNP genotipados	52.602	54.602	65.157
<i>Callfrequency</i> < 0,9	432	148	504
Localizados no cromossomo X	539	3.223	3.411
HWE $p < 1 \times 10^{-5}$	147	4.519	4.638
MAF < 0,05	8367	-	-
SNP restantes	43.117	46.712	56.604

\*HWE: Equilíbrio de Hardy-Weinberg (*Hardy-Weinberg Equilibrium*)

## 2.4. Imputação de genótipos na linhagem de corrida da raça Quarto de Milha

O procedimento de imputação de genótipos nos animais da linhagem de corrida da raça Quarto de Milha foi realizado por meio do programa FImpute v.2.2 (SARGOLZAEI et al., 2014).

Nos animais genotipados com o arranjo de 54K foram imputados os SNP exclusivos do arranjo de 65K, e nos indivíduos genotipados com o arranjo de 65K foram imputados os SNP exclusivos do arranjo de 54K. As informações utilizadas para a realização da imputação de genótipos em duas vias (i.e. do 54K para o de 65K e do 65K para o de 54K), como descrita por Pereira et al. (2017), foram relativas ao número de indivíduos e SNP obtidos após o controle de qualidade descrito no item 2.3. A acurácia de imputação foi determinada utilizando a taxa de concordância (*concordance rate* – CR), à qual equivale à proporção de genótipos imputados corretamente, e o  $r^2$  alélico, que é determinado pelo quadrado da correlação entre a contagem de alelos dos genótipos imputados ( $0x(AA)+1x(AB)+2x(BB)$ ) e a contagem de alelos dos genótipos originais. Os valores de acurácia foram obtidos tanto para SNP (médias por intervalos de MAFs e cromossomos) quanto para amostras.

Após o processo de imputação, o número de SNP informativos na linhagem de corrida da raça Quarto de Milha, considerando a MAF de 0,05, foi 55.114. Desta forma, para as análises genéticas subsequentes de desequilíbrio de ligação e tamanho efetivo realizadas nas linhagens de trabalho ( $n = 61$ ) e de corrida ( $n = 347$ ) da raça Quarto de Milha foram utilizados 43.117 e 55.114 SNP, respectivamente.

## 2.5. Desequilíbrio de ligação

O desequilíbrio de ligação (LD) nas linhagens de trabalho e de corrida da raça Quarto de Milha foi avaliado por meio da estatística  $r^2$  (HILL e ROBERTSON, 1968) para todos os pares de SNP em cada cromossomo utilizando o programa estatístico PLINK 1.9 (PURCELL et al., 2007). O LD médio foi calculado para cada cromossomo e para o genoma como um todo. Também foi calculado o LD médio para intervalos de diferentes distâncias físicas entre os marcadores. Assim, os genótipos para todos os pares de SNP com menos de 5 Mb de distância foram avaliados para estimar o LD em intervalos de 25 a 100 Kb, 100 a 200 Kb, 200 a 500 Kb, 500 a 1000 Kb, e 1000 a 5000 Kb.

O gráfico do decaimento da média de desequilíbrio de ligação em função da distância física entre os marcadores foi gerado para as linhagens de trabalho e de corrida da raça Quarto de Milha com a utilização do programa R (R Core Team, 2016).

## 2.6. Estimação do tamanho efetivo das populações

A relação entre o LD ( $r^2$ ) e o tamanho efetivo da população ( $N_e$ ) é expressa pela equação:  $r^2 = \frac{1}{(4N_e c + 1)}$ , em que  $c$  é a distância genética entre dois SNP expressa em centimorgans (cM) (SVED, 1971).

Os tamanhos efetivos das populações das linhagens de trabalho e de corrida da raça Quarto de Milha foram estimados com base nos cromossomos autossomos e na distância física entre SNP (<http://www.ensembl.org>). As distâncias genéticas foram convertidas assumindo que 1cM equivale a 1Mb. Os tamanhos efetivos populacionais foram estimados utilizando a fórmula:  $NeT = \left(\frac{1}{4c}\right) \left(\frac{1}{r^2 - 1}\right)$ , em que  $NeT$  é o tamanho efetivo populacional há  $T$  gerações, e  $T = \frac{1}{2c}$  (HAYES et al., 2003; GODDARD et al., 2006). Pares de marcadores com  $c < 10^{-6}$  (~100bp), isto é,  $T > 500.000$ , não foram utilizados uma vez que a aproximação é válida apenas para  $c$  muito maior que a taxa de mutação (~ $10^{-8}$  por loco por geração). As estimativas pontuais finais da  $NeT$  foram obtidas por meio da média da estimativa de valores  $N_e$  dentro  $T$ . Uma vez que  $r^2$  é parâmetro limitado e as estimativas pontuais de  $N_e$  são infinitas em  $r^2 = 0$  e zero em  $r^2 = 1$ , apenas valores de  $0,01 < r^2 < 0,99$  foram utilizados para estimar  $N_e$ . Este cálculo foi realizado por meio do programa R (R Core Team, 2016).

## 2.7. Estrutura populacional

Com o intuito de agrupar animais geneticamente semelhantes e de avaliar subdivisões na raça Quarto de Milha criada no Brasil a partir de dados genômicos, a estrutura da população foi inicialmente observada por meio da análise de componentes principais (*principal component analysis* – PCA). As linhagens de trabalho e de corrida da raça Quarto de Milha foram consideradas de forma conjunta ( $n = 408$ ), em que foram mantidos somente os SNP comuns entre ambas (33.944). Para a visualização geral da estrutura genética básica da população realizada por

análise de componentes principais por meio do programa PLINK 1.9 (PURCELL, 2007), os resultados foram analisados por meio de gráficos gerados pelo pacote ggplot2 do programa R.

A análise discriminante dos componentes principais (*discriminant analysis of principal component* – DAPC) foi implementada por meio do pacote adegenet (JOMBART, 2008) do programa R. O número de  $K$  grupos e número ótimo de componentes principais (PCs) a serem retidos na análise foram também previamente definidos por meio desse pacote. Para a determinação do número de grupos ( $K$ ) dentro da população foi utilizado o algoritmo de agrupamento *K-means* (JOMBART; DEVILLARD; BALLOUX, 2010). Este algoritmo utiliza os mesmos passos a serem dados na DAPC, em que, de forma prévia, é feita a transformação dos dados por meio de PCA (principalmente para reduzir os números de variáveis) com posterior análise discriminante, com a ressalva que, neste caso, é considerado determinado número de grupos a cada rodada. O algoritmo de agrupamento *K-means* foi implementado para  $K = 2$  a 20 no conjunto total de dados, incluindo as linhagens de trabalho e corrida da raça Quarto de Milha. Para isso foram considerados o número de PCs que explicassem em torno de 70 a 80% da variância acumulada. Para identificar o número ótimo de grupos genéticos foi analisado o critério de informação bayesiana (*bayesian information criterion* – BIC).

A DAPC baseia-se na transformação de dados por meio de PCA, como um passo anterior e necessário à análise discriminante (*discriminant analysis* – DA) propriamente dita. Nesse sentido, é preciso manter número de PCs suficientes para garantir o poder de discriminação sem, contudo, ter um número excessivo de dimensões, o que poderia levar a um *overfit* das funções discriminantes. Para escolha do número ideal de PCs a ser retido na análise, foi calculado o  $\alpha$ -score por meio do pacote adegenet (JOMBART, 2008) do programa R. O procedimento é baseado no cálculo do  $\alpha$ -score, em que se mede a diferença entre a proporção da re-atribuição bem-sucedida de indivíduos aos grupos (discriminação observada) e os valores obtidos usando grupos aleatórios (discriminação aleatória). Para diminuir o uso intensivo do processamento, o método utiliza interpolação, reduzindo o número de  $\alpha$ -score estimados na faixa de PCs escolhida, a qual foi estipulada como de 1 a 50.

Após a inferência do número de grupos pelo método *K-means*, foi possível analisar subestruturas familiares importantes dentro da população. A classificação de indivíduos com base no grupo possibilitou, por análise simples do *pedigree*,

estabelecer os principais reprodutores dentro destes, possibilitando assim aumentar a compreensão da dinâmica e estrutura populacional de equinos da raça Quarto de Milha criados no Brasil. Para isso foram analisados dentro de cada grupo os indivíduos mais frequentes (pai, avô paterno e avô materno).

### 3. RESULTADOS E DISCUSSÃO

#### 3.1. Desequilíbrio de ligação

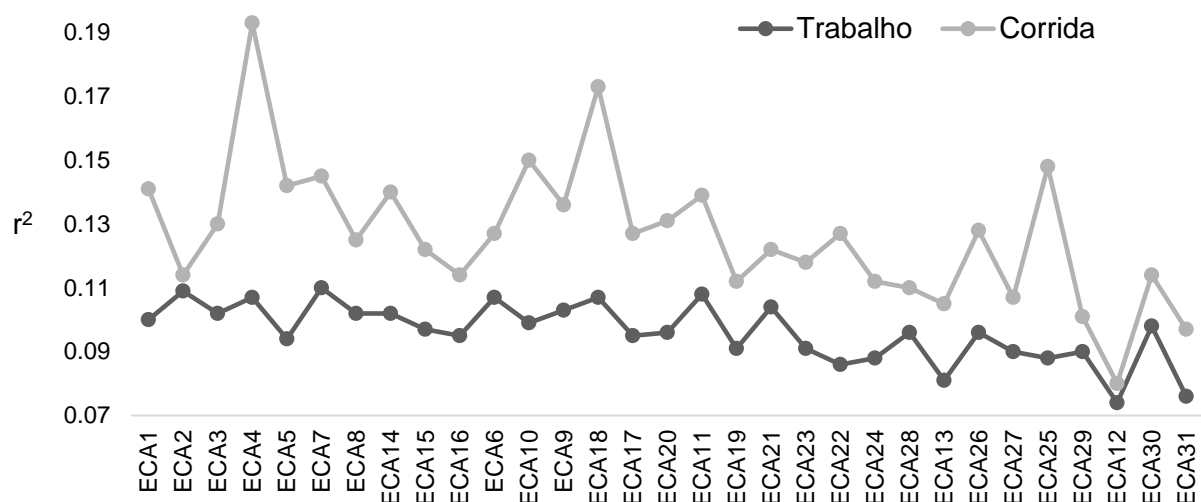
Na Tabela 2 estão descritos os tamanhos dos 31 cromossomos autossomos de equinos, os números de SNP após a realização do processo de imputação de genótipos e os valores médios de  $r^2$  médio por cromossomo nas linhagens de trabalho e de corrida da raça Quarto de Milha. Observa-se na Tabela 2 e na Figura 1 que a linhagem de trabalho apresentou o maior valor médio de  $r^2$  (0,110) no ECA7 e o menor valor no ECA12 ( $r^2 = 0,074$ ). Na linhagem de corrida,  $r^2$  foi maior no ECA4 ( $r^2 = 0,193$ ) e menor no ECA12 ( $r^2 = 0,080$ ). Mostra-se assim, como esperado, grande diversidade de valores de LD entre os diferentes cromossomos nas duas linhagens. Uma possível explicação é que a grande variabilidade nas taxas de recombinação nos cromossomos autossomos, além de outros fatores, leva à considerável diversidade no padrão de LD em distintas regiões genômicas (ARIAS et al., 2009).

Em parte dos estudos realizados em diversas espécies, observou-se que as diferenças nas taxas de recombinação estão associadas ao tamanho do cromossomo, o que foi observado de forma moderada na presente pesquisa ( $r = 0,612$  na linhagem de trabalho e  $r = 0,514$  na linhagem de corrida). A observação de que o maior tamanho do cromossomo se relaciona com o aumento da taxa média de recombinação foi descrita em estudos com bovinos (ARIAS et al., 2009), ratos (SHIFMAN, 2006), e humanos (KONG, 2002). Contudo, em trabalho conduzido por Bohmanova, Sargolzaei e Schenkel (2010), utilizando população de bovinos *Bos taurus* da raça *Holstein* genotipados com o *Bovine SNP50 BeadChip*, não foi encontrada relação entre o tamanho dos cromossomos e a extensão do LD.

**Tabela 2** – Tamanho dos cromossomos em megabases (Mb), número de SNP após a imputação e valores médios de  $r^2$ , por cromossomo, na amostra estudada de equinos das linhagens de trabalho e de corrida da raça Quarto de Milha.

ECA	Tamanho Chr (Mb)	SNP (n)		Valor de $r^2$	
		Trabalho	Corrida	Trabalho	Corrida
1	185,84	3123	4526	0,100	0,141
2	120,86	1958	3111	0,109	0,114
3	119,48	1985	2915	0,102	0,130
4	108,57	1863	2649	0,107	0,193
5	99,68	1654	2357	0,094	0,142
6	84,72	1318	2132	0,107	0,127
7	98,54	1608	2381	0,110	0,145
8	94,06	1545	2329	0,102	0,125
9	83,56	1415	2106	0,103	0,136
10	83,98	1326	2020	0,099	0,150
11	61,31	996	1540	0,108	0,139
12	33,09	464	735	0,074	0,080
13	42,58	652	1010	0,081	0,105
14	93,9	1591	2296	0,102	0,140
15	91,57	1560	2372	0,097	0,122
16	87,37	1522	2245	0,095	0,114
17	80,76	1364	1955	0,095	0,127
18	82,53	1338	1811	0,107	0,173
19	59,98	1039	1566	0,091	0,112
20	64,17	1067	1597	0,096	0,131
21	57,72	959	1438	0,104	0,122
22	49,95	868	1203	0,086	0,127
23	55,73	879	1308	0,091	0,118
24	46,75	788	1194	0,088	0,112
25	39,54	661	938	0,088	0,148
26	41,87	682	1089	0,096	0,128
27	39,96	665	937	0,090	0,107
28	46,18	765	1173	0,096	0,110
29	33,67	560	759	0,090	0,101
30	30,06	496	762	0,098	0,114
31	24,98	452	645	0,076	0,097

\*ECA: autossomo *Equus caballus*; Mb= megabases



**Figura 1** – Valores médios de  $r^2$  por cromossomos (ordenados por tamanho e de forma decrescente), na amostra estudada de equinos das linhagens de trabalho e de corrida da raça Quarto de Milha.

Nas Tabelas 3 e 4 encontram-se descritos os valores médios de  $r^2$  para os diferentes intervalos de distância nas linhagens de trabalho e de corrida da raça Quarto de Milha, respectivamente. As maiores médias de  $r^2$  na distância de 0-25 Kb na linhagem de trabalho foram 0,479, 0,467 e 0,463, nos cromossomos ECA7, ECA18 e ECA 6, respectivamente. Os menores valores, 0,050; 0,052; 0,059, foram encontrados nos cromossomos ECA12, ECA31, ECA13, respectivamente, no intervalo de 1000 a 5000 Kb. Os maiores valores médios de  $r^2$  no intervalo de 0-25 Kb na linhagem de corrida foram encontrados nos cromossomos ECA4, ECA25 e ECA18, com valores de 0,560, 0,543 e 0,542, respectivamente, e os menores valores de  $r^2$  foram encontrados no intervalo de 1000-5000 Kb nos cromossomos ECA 12 ( $r^2=0,053$ ), ECA31 ( $r^2=0,062$ ) e ECA29 ( $r^2=0,069$ ) (Tabela 4). Observa-se que, para ambas as linhagens estudadas, como esperado, os maiores valores de  $r^2$  foram encontrados nos intervalos de 0-25 Kb e os menores valores na distância de 1000-5000 Kb. Esses resultados são similares aos encontrados em bovinos da raça Nelore, que apresentaram maiores  $r^2$  nas distâncias de 30 Kb, 50 Kb, 100 Kb com valores de 0,20, 0,16 e 0,1 respectivamente (ESPIGOLAN et al., 2013), evidenciando que o LD decai com aumento da distância independentemente da espécie analisada.

**Tabela 3** – Valores médios de  $r^2$  nos diferentes intervalos de distância física entre SNP em equinos da linhagem de trabalho da raça Quarto de Milha.

ECA	0- 25 Kb	25- 100 Kb	100- 200 Kb	200- 500 Kb	500- 1000 Kb	1000- 5000 Kb
1	0,427±0,375	0,280±0,310	0,228±0,275	0,167±0,223	0,133±0,187	0,074±0,115
2	0,431±0,382	0,299±0,325	0,252±0,293	0,185±0,241	0,146±0,202	0,080±0,125
3	0,423±0,380	0,300±0,324	0,247±0,291	0,180±0,240	0,136±0,198	0,073±0,121
4	0,439±0,384	0,299±0,323	0,249±0,291	0,183±0,243	0,142±0,203	0,078±0,127
5	0,404±0,379	0,271±0,316	0,212±0,276	0,156±0,225	0,121±0,186	0,069±0,115
6	0,463±0,386	0,311±0,326	0,253±0,293	0,184±0,241	0,141±0,200	0,078±0,124
7	0,479±0,392	0,319±0,341	0,262±0,307	0,191±0,257	0,147±0,213	0,079±0,130
8	0,455±0,383	0,319±0,334	0,251±0,294	0,183±0,247	0,139±0,206	0,072±0,123
9	0,438±0,378	0,319±0,336	0,251±0,298	0,177±0,244	0,137±0,203	0,074±0,122
10	0,432±0,388	0,293±0,330	0,233±0,288	0,168±0,234	0,128±0,193	0,072±0,117
11	0,405±0,373	0,289±0,324	0,239±0,293	0,183±0,249	0,147±0,216	0,079±0,131
12	0,371±0,359	0,247±0,291	0,185±0,246	0,128±0,190	0,095±0,152	0,050±0,089
13	0,325±0,352	0,237±0,293	0,190±0,254	0,139±0,207	0,107±0,168	0,059±0,099
14	0,444±0,390	0,304±0,338	0,238±0,294	0,173±0,238	0,135±0,199	0,074±0,121
15	0,439±0,378	0,289±0,315	0,222±0,275	0,164±0,226	0,128±0,188	0,072±0,117
16	0,408±0,378	0,285±0,321	0,230±0,283	0,168±0,230	0,127±0,189	0,068±0,113
17	0,462±0,383	0,296±0,331	0,230±0,288	0,164±0,231	0,122±0,187	0,068±0,113
18	0,467±0,389	0,317±0,336	0,254±0,300	0,187±0,245	0,145±0,204	0,076±0,124
19	0,376±0,369	0,265±0,311	0,215±0,274	0,161±0,226	0,122±0,185	0,066±0,111
20	0,436±0,377	0,298±0,329	0,236±0,291	0,164±0,231	0,127±0,191	0,068±0,114
21	0,439±0,375	0,299±0,325	0,238±0,284	0,175±0,230	0,139±0,191	0,077±0,118
22	0,377±0,364	0,276±0,312	0,220±0,273	0,154±0,217	0,116±0,177	0,060±0,103
23	0,430±0,376	0,275±0,314	0,217±0,271	0,157±0,219	0,119±0,178	0,065±0,108
24	0,355±0,360	0,231±0,285	0,186±0,247	0,138±0,199	0,115±0,167	0,067±0,105
25	0,418±0,388	0,273±0,312	0,216±0,276	0,159±0,224	0,117±0,181	0,062±0,107
26	0,449±0,389	0,291±0,325	0,235±0,288	0,173±0,236	0,132±0,193	0,067±0,112
27	0,406±0,366	0,269±0,306	0,215±0,268	0,154±0,215	0,120±0,176	0,063±0,106
28	0,415±0,370	0,287±0,316	0,229±0,276	0,164±0,224	0,128±0,183	0,070±0,111
29	0,376±0,358	0,249±0,296	0,197±0,260	0,153±0,215	0,122±0,180	0,065±0,109
30	0,417±0,379	0,278±0,314	0,219±0,267	0,162±0,215	0,129±0,182	0,071±0,112
31	0,358±0,341	0,239±0,280	0,185±0,241	0,133±0,192	0,104±0,158	0,052±0,091

\*ECA: autossomo *Equus caballus*; Kb= kilobases



**Tabela 4** – Valores médios de  $r^2$  nos diferentes intervalos de distância física entre SNP em equinos da linhagem de corrida da raça Quarto de Milha.

ECA	0- 25 Kb	25- 100 Kb	100- 200 Kb	200- 500 Kb	500- 1000 Kb	1000- 5000 Kb
1	0,479±0,380	0,351±0,334	0,303±0,306	0,242±0,266	0,199±0,233	0,107±0,154
2	0,445±0,377	0,319±0,328	0,272±0,298	0,210±0,252	0,165±0,215	0,082±0,133
3	0,464±0,386	0,331±0,335	0,285±0,307	0,219±0,263	0,176±0,226	0,099±0,152
4	0,560±0,385	0,434±0,357	0,385±0,342	0,318±0,315	0,264±0,285	0,150±0,206
5	0,473±0,393	0,333±0,340	0,286±0,311	0,228±0,271	0,189±0,242	0,111±0,179
6	0,477±0,382	0,351±0,334	0,299±0,306	0,231±0,264	0,182±0,226	0,092±0,144
7	0,521±0,393	0,365±0,349	0,316±0,318	0,247±0,276	0,201±0,243	0,108±0,164
8	0,500±0,389	0,359±0,347	0,297±0,314	0,223±0,267	0,177±0,229	0,091±0,144
9	0,474±0,383	0,349±0,346	0,294±0,316	0,227±0,272	0,185±0,238	0,104±0,164
10	0,504±0,388	0,367±0,348	0,308±0,316	0,245±0,276	0,204±0,247	0,116±0,182
11	0,455±0,373	0,343±0,334	0,298±0,312	0,240±0,280	0,197±0,252	0,104±0,170
12	0,380±0,355	0,253±0,287	0,211±0,255	0,151±0,205	0,117±0,169	0,053±0,099
13	0,390±0,369	0,292±0,314	0,243±0,282	0,187±0,238	0,149±0,202	0,075±0,127
14	0,495±0,389	0,359±0,346	0,302±0,316	0,238±0,274	0,196±0,241	0,104±0,157
15	0,448±0,372	0,320±0,322	0,269±0,293	0,212±0,252	0,173±0,219	0,090±0,138
16	0,449±0,388	0,322±0,332	0,268±0,299	0,208±0,253	0,165±0,216	0,081±0,132
17	0,511±0,383	0,360±0,340	0,298±0,309	0,229±0,263	0,179±0,224	0,092±0,140
18	0,542±0,389	0,408±0,361	0,350±0,338	0,286±0,301	0,242±0,269	0,132±0,186
19	0,411±0,374	0,315±0,325	0,269±0,299	0,211±0,255	0,165±0,217	0,078±0,131
20	0,481±0,381	0,343±0,339	0,295±0,311	0,230±0,267	0,188±0,233	0,096±0,149
21	0,465±0,389	0,336±0,337	0,283±0,306	0,217±0,259	0,177±0,225	0,088±0,139
22	0,449±0,373	0,361±0,334	0,300±0,305	0,232±0,264	0,185±0,227	0,091±0,143
23	0,460±0,385	0,327±0,336	0,272±0,305	0,211±0,260	0,168±0,222	0,084±0,139
24	0,412±0,372	0,289±0,315	0,248±0,285	0,199±0,244	0,161±0,210	0,082±0,130
25	0,543±0,398	0,386±0,345	0,336±0,318	0,263±0,278	0,215±0,243	0,106±0,158
26	0,503±0,390	0,344±0,336	0,296±0,306	0,232±0,262	0,189±0,228	0,090±0,144
27	0,465±0,380	0,310±0,318	0,257±0,287	0,192±0,240	0,153±0,202	0,076±0,125
28	0,427±0,367	0,310±0,318	0,258±0,286	0,196±0,240	0,154±0,204	0,080±0,130
29	0,397±0,363	0,278±0,305	0,237±0,277	0,185±0,235	0,148±0,202	0,069±0,120
30	0,470±0,386	0,326±0,328	0,270±0,293	0,210±0,247	0,171±0,214	0,078±0,130
31	0,459±0,365	0,307±0,318	0,253±0,283	0,191±0,232	0,148±0,196	0,062±0,115

\*ECA: autossomo *Equus caballus*; Kb= kilobases

O  $r^2$  genômico (média geral de todos os  $r^2$  calculados) entre pares de marcadores foi 0,22 e 0,27 para as linhagens de trabalho e de corrida da raça Quarto de Milha, respectivamente. A fim de avaliar o decaimento da curva de LD em função da distância física entre marcadores ao longo do genoma, os pares de SNP foram analisados em intervalos. Os valores médios de  $r^2$  por intervalo considerando todo o genoma estão descritos na Tabela 5 para a linhagem de trabalho e para a de corrida. O LD diminuiu, em ambas as linhagens, na medida que a distância física entre os

marcadores aumentou, sendo que o declínio mais acentuado da curva foi observado nos primeiros 200 Kb, decaindo mais lentamente com o aumento da distância. A curva do decaimento do LD foi menos pronunciada na linhagem de corrida, como mostra a Figura 2. Esse padrão de declínio inicial acentuado seguido de menor intensidade ao longo de maiores distâncias foi similar ao encontrado por Corbin, Blott e Swinburne (2010), em equinos da raça Puro-Sangue Inglês.

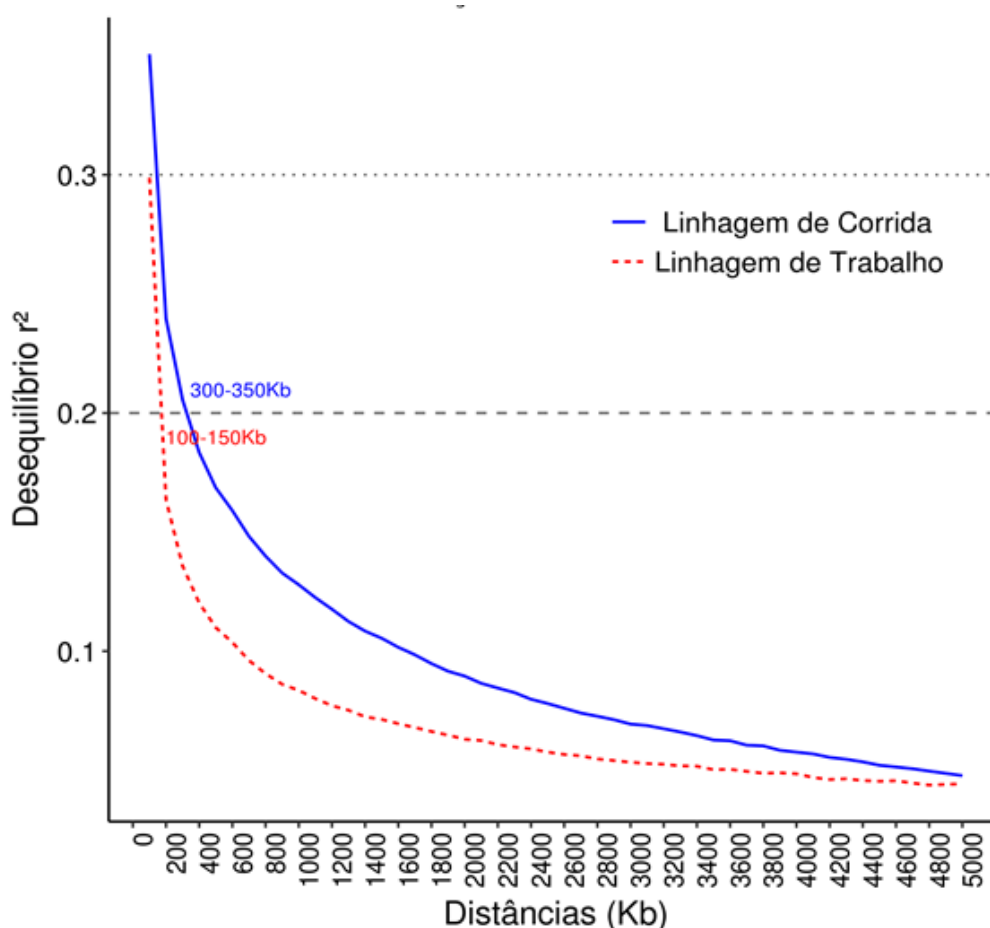
**Tabela 5** – Valores médios de  $r^2$ , desvio padrão e mediana nos diferentes intervalos de distância física entre SNP nas linhagens de trabalho e de corrida de equinos Quarto de Milha.

	Distância	Média $r^2$	Desvio Padrão	Mediana
Trabalho	0-25 kb	0,427±0,00241	0,380	0,283
	25-100 kb	0,290±0,00119	0,322	0,147
	100-200 kb	0,233±0,00078	0,284	0,110
	200-500 kb	0,170±0,00042	0,232	0,076
	500-1000 kb	0,131±0,00025	0,193	0,057
	1000-5000 kb	0,072±0,00007	0,117	0,031
Corrida	0-25 kb	0,475±0,00196	0,384	0,363
	25-100 kb	0,344±0,00091	0,338	0,205
	100-200 kb	0,292±0,00050	0,309	0,163
	200-500 kb	0,229±0,00032	0,267	0,117
	500-1000 kb	0,185±0,00020	0,233	0,090
	1000-5000 kb	0,097±0,00006	0,153	0,039

Observa-se na Figura 2 que o LD decai a curtas distâncias em ambas as linhagens, sendo que a partir de distâncias superiores a 100 Kb o  $r^2$  médio foi inferior a 0,3. Na linhagem de trabalho o  $r^2$  foi inferior a 0,2 entre 100 e 150 Kb, enquanto na de corrida foram encontrados valores de  $r^2$  menores que 0,2 entre 300 e 350 Kb. Em estudo conduzido por McCue e Mickelson (2013), a extensão do LD foi caracterizada em 14 raças equinas. Em geral, os valores calculados de  $r^2$  caíram abaixo de 0,2 à distâncias entre 100 e 150 Kb. Notáveis exceções foram o Puro-Sangue Inglês, o *Standardbred* e o *French Trotter*, nos quais os valores de  $r^2$  permaneceram acima de 0,2 a distâncias superiores a 400 Kb. Entre as raças mais difundidas analisadas nesse estudo, a Quarto de Milha se destacou por seus valores de LD relativamente baixos quando comparados à outras raças equinas, com decaimento de valores de  $r^2$  abaixo de 0,2 entre 50 e 100 Kb. (McCUE; MICKELSON, 2013).

Destaca-se que, em estudos como os realizados por McCue e Mickelson (2013), cavalos da linhagem de trabalho da raça Quarto de Milha são mais amplamente utilizados por serem mais numerosos e de fácil acesso se comparados aos da linhagem de corrida. Dessa forma, no presente estudo, valores semelhantes aos descritos por esses autores foram encontrados somente na linhagem de trabalho.

Uma possível explicação para o maior LD observado na linhagem de corrida da raça Quarto de Milha advém da grande influência do Puro-Sangue Inglês ao longo da sua formação. Também pode ser destacado que, além da influência do Puro-Sangue Inglês, a sua estrutura populacional mais fechada em relação à linhagem de trabalho, em decorrência do maior rigor no registro genealógico dos animais, exerce influência sobre a extensão média do LD observado nos animais de corrida.



**Figura 2** – Decaimento da média do desequilíbrio de ligação (LD), analisado por meio de  $r^2$ , nas linhagens de corrida e de trabalho da raça Quarto de Milha.

Em grande parte dos estudos, realizados em diversas espécies, que avaliaram a média do LD ao longo das diferentes distâncias entre marcadores (Kb), observou-se resultados similares aos apresentados no presente estudo, ou seja, médias de  $r^2$  maiores para distâncias menores (AI; HUAN; REN, 2013; MCKAY et al., 2007; MEADOWS; CHAN; KIJAS, 2008). A comparação de níveis de LD obtidos em estudos distintos é difícil, pois ocorre variação no tamanho da amostra, diferenças no método de mensuração, diferença nos tipos e densidade de marcadores utilizados, bem como na história evolutiva da população analisada (PRITCHARD e PRZEWORSKI, 2001). Em geral, esse padrão de decaimento da curva de desequilíbrio de ligação ao longo de maiores distâncias segue resultado esperado, uma vez que com o aumento na distância física entre marcadores do mesmo cromossomo, maiores serão as possibilidades de ocorrerem eventos de recombinação genética e, portanto, a quebra da ligação entre marcadores (DU; CLUTTER; LOHUIS, 2007).

A maior parte dos estudos genômicos já realizados com equinos da raça Quarto de Milha não consideraram as diferentes linhagens separadamente, ou seja, trabalharam a raça como um todo. Porém, as diferenças da extensão de LD entre as linhagens de trabalho e de corrida descritas neste estudo, fornecem informações estratégicas para futuras pesquisas genômicas na raça Quarto de Milha, como, por exemplo, os estudos de associação ampla. Em geral, o LD é variável no genoma e depende da espécie e população sob estudo (ORAGUZIE et al., 2007). Assim, a quantificação do LD é de grande utilidade em estudos evolutivos e possibilita determinar a resolução de mapeamento e a densidade de marcadores necessários na identificação de QTL (NORDBORG et al., 2002; WEI; JACKSON.; MCINTYRE, 2006; ZHU et al., 2008).

O sucesso do mapeamento das regiões em LD depende de um balanço apropriado entre a extensão deste desequilíbrio e a densidade de marcadores (SARGOLZAEI et al., 2008). Portanto, a densidade dos marcadores deve ser suficientemente alta para garantir que todos os QTL estejam em LD com um marcador ou haplótipo de marcadores. Quando o LD estende-se a maiores distâncias no genoma, como no caso da linhagem de corrida da raça Quarto de Milha, será requerido menor quantidade de marcadores. Porém, se o LD estende-se a menores distâncias no genoma, como no caso da linhagem de trabalho da raça Quarto de Milha, uma elevada quantidade de marcadores será necessária (YU; BUCKLER, 2006; ORAGUZIE et al., 2007; ZHU et al., 2008). Assim, os mapas de LD ou de

haplótipos são ferramentas fundamentais para explorar a base genética de características economicamente importantes (MEIRA et al., 2014a).

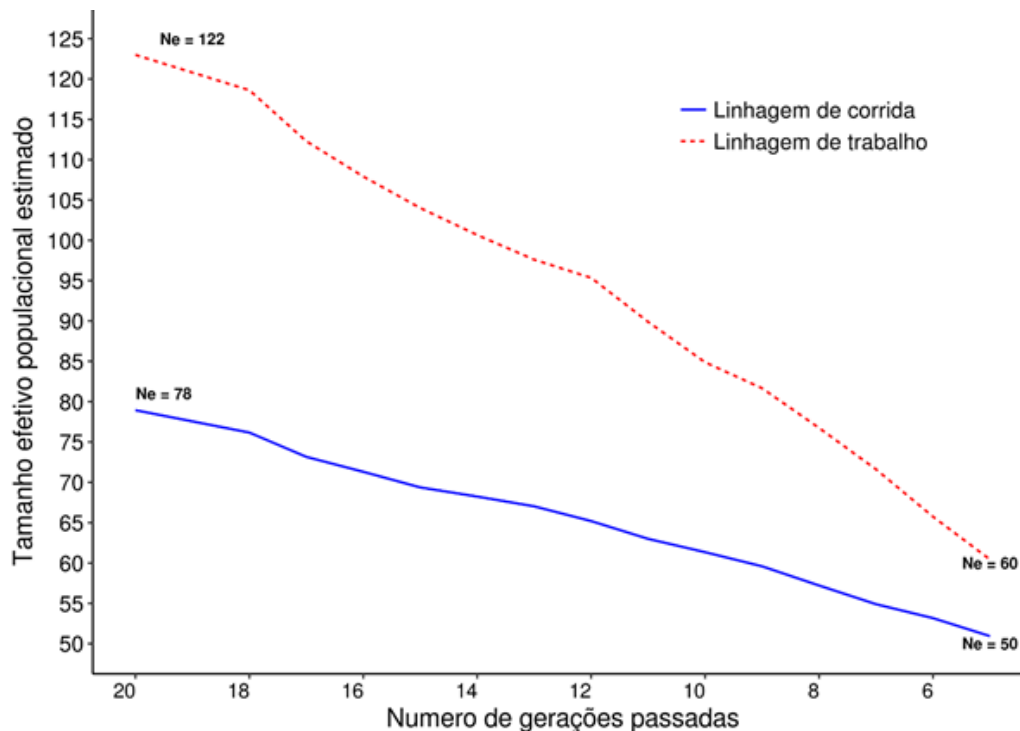
### 3.2. Estimação do tamanho efetivo das populações

O tamanho efetivo populacional ( $N_e$ ), calculado com base em informações de marcadores moleculares, estabelece o número de haplótipos distintos presentes na população, sendo que em populações pequenas este número é limitado. Na Figura 3 observa-se que o  $N_e$  inicial na linhagem de corrida da raça Quarto de Milha foi sensivelmente menor em relação à linhagem de trabalho. A estimativa do tamanho efetivo apresentou em ambos os grupos redução nas últimas 20 gerações, com um máximo de aproximadamente 60 e 50 animais na última geração para as linhagens de trabalho e de corrida, respectivamente. As diferenças observadas entre as linhagens estudadas são em parte reflexo das diferentes práticas utilizadas nos sistemas de acasalamentos. De acordo com Meira et al. (2013) o coeficiente de endogamia médio encontrado para populações de animais de trabalho e de corrida da raça Quarto de Milha criados no Brasil foi de 0,01% e 0,36%, respectivamente. Esses resultados indicaram a maior incidência da prática de acasalamentos endogâmicos na linhagem de corrida em relação à de trabalho.

Ao analisar a curva do valor do  $N_e$  ao longo das gerações, observou-se que a população de animais da linhagem de corrida apresenta menor queda de  $N_e$  quando comparada à linhagem de trabalho. Nota-se que o  $N_e$  há 20 gerações para a linhagem de trabalho foi de 122 animais e para a de corrida 78. Estes resultados apontam a necessidade de incremento na diversidade genética, principalmente, na linhagem de corrida. Com isso, ressalta-se, mais uma vez, a importância do monitoramento da variabilidade genética e o controle da endogamia de populações submetidas à processos de seleção.

De acordo com Faria (2016), que realizou estudo com informações de *pedigree* de 47.681 equinos da raça Quarto de Milha, nascidos no período de 2000 a 2008, o valor calculado de  $N_e$  foi 114. Embora haja alguma discrepância entre os resultados encontrados no presente trabalho e os de Faria (2016), possivelmente em função das metodologias empregadas (informações genômicas x *pedigree*) e da abrangência amostral, ambos os achados mostram  $N_e$  reduzido para cavalos da raça Quarto de Milha criados no Brasil, indicando a necessidade da introdução de novos reprodutores,

não parentes, para aumento da variabilidade genética na raça. De fato, na literatura, trabalhos com informações de *pedigree* de equinos mostraram, em geral, valores de  $N_e$  maiores do que os realizados com dados moleculares. Dias et al. (2000) estudaram o *pedigree* da raça Brasileiro de Hipismo e encontraram valor de  $N_e$  igual a 253, indicando grande variabilidade genética na raça.



**Figura 3** – Tamanho efetivo populacional estimado ( $N_e$ ) em função do número de gerações passadas em equinos da raça Quarto de Milha.

Em estudo conduzido na Coreia do Sul, utilizando equinos Puro-Sangue Inglês e informação de haplótipos, foi encontrado decréscimo gradual do tamanho efetivo populacional até 10 gerações atrás, seguido por leve aumento até a geração atual, com valor de  $N_e$  próximo de 78. Os autores sugeriram que o ligeiro aumento do  $N_e$  até a geração atual é resultado da importação de cavalos PSI pelo país (LEE; WOO LEE; KIM, 2014). Os animais PSI da Coreia apresentaram valor menor de tamanho efetivo populacional quando comparados aos relatados em outros estudos da raça. Corbin, Blott e Swinburne (2010), observaram, por meio de estimativas de LD, em equinos da raça PSI da Irlanda, um valor de  $N_e$  próximo a 90 há 20 gerações, seguido de aumento até o momento presente, com  $N_e$  próximo de 180 na última geração.

Cervantes et al. (2011), com base em informações de *pedigree* observaram em PSIs da Espanha tamanho efetivo populacional maior ( $N_e = 322$ ).

Uma possível explicação para o maior tamanho efetivo observado na raça PSI, em comparação à raça Quarto de Milha, são as diferentes formas de manejo reprodutivo praticadas em ambas. Diferentemente da raça Quarto de Milha, que se utiliza de técnicas de inseminação artificial, e mais recentemente e de forma menos frequente da transferência de embrião e fertilização *in vitro*, animais PSI são sempre acasalados naturalmente por imposição da sua associação de criadores.

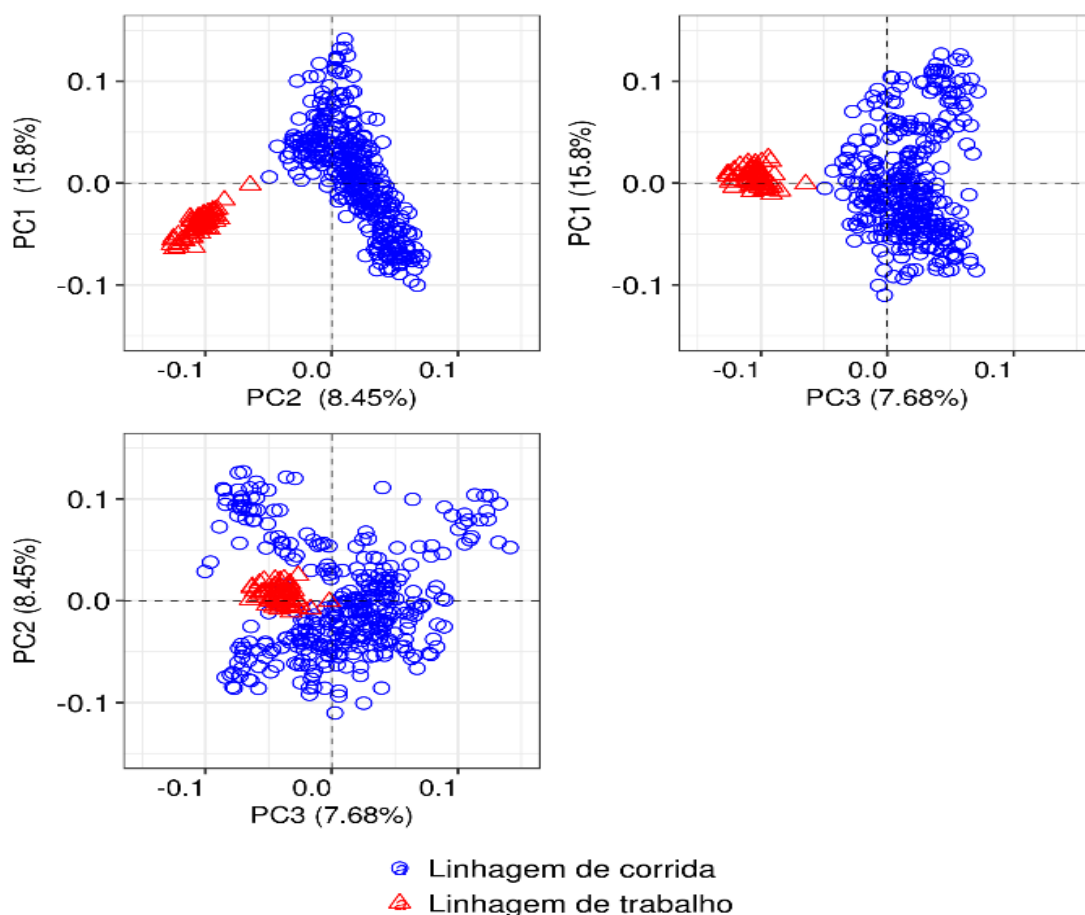
A queda do tamanho efetivo mostrada em ambas as linhagens da raça Quarto de Milha analisadas no presente trabalho sugere perda de variabilidade genética, em que genótipos são impedidos de participar na geração seguinte (ARAKI; TACHIDA, 1997), resultado, provavelmente, de possível seleção intensiva oriunda do uso de pequeno número de reprodutores ao longo das gerações. Estudos com declínio do tamanho efetivo também foram encontrados em bovinos analisados com informações moleculares, em que se observou  $N_e$  igual a 73 nas últimas cinco gerações na raça Simental (ZHU et al., 2013) e  $N_e$  igual a 39 nas últimas duas gerações da raça Gir (SILVA et al., 2010).

Como mencionado, o  $N_e$  pode ser obtido por meio de diferentes fontes de informações, como dados de *pedigree* ou dados moleculares, sendo que vários fatores podem alterá-lo. Kijas et al. (2012), ao genotiparem 49.034 SNP em animais de 74 raças ovinas, observaram que 75% das raças modernas analisadas mostraram tamanho efetivo populacional acima de 300, maior do que os encontrados na maioria das raças caninas e bovinas. Os referidos autores sugeriram que este valor foi devido a elevada heterogeneidade antes da domesticação, ou seja, a partir de uma ampla base genética, com gargalo genético menor do que o apresentado em outras espécies.

### **3.3. Estrutura populacional**

Por meio da análise de PCA, observou-se, na linhagem de corrida da raça Quarto de Milha, a dispersão dos valores ao longo de três quadrantes do gráfico (Figura 4 - PC1xPC2). Ao analisar os gráficos dos componentes PC1xPC3 e PC2xPC3, observa-se mais facilmente essa dispersão, em que, claramente, nota-se subestruturas dentro da linhagem de corrida. Na literatura, pesquisas realizadas com

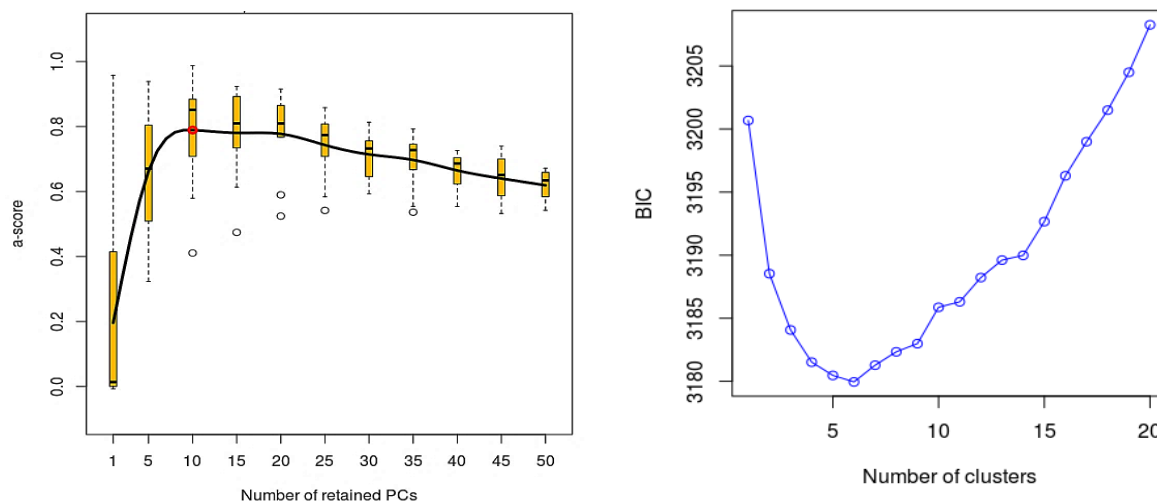
análise de *pedigree*, já haviam levantado a hipótese da existência de subestruturas populacionais dentro da raça Quarto de Milha (TRYON et al., 2009; MEIRA et al., 2013). O método de análise de componentes principais (*principal component analysis* – PCA) têm sido ferramenta eficiente para a detecção da estrutura populacional em estudo com equinos (COSTA et al., 2005), suínos (AI; HUAN; REN, 2013), ovinos (KIJAS et al., 2012), e humanos (PATTERSON; PRICE; REICH, 2006). Da mesma forma que neste estudo, em trabalho com informações de *pedigree* de equinos da raça Mangalarga Marchador, foi evidenciada, a partir de PCA, a existência de subestruturas. Os autores sugeriram a existência de cinco grupos formadores da raça Mangalarga Marchador (COSTA et al., 2005).



**Figura 4** – Representação dos três primeiros componentes principais, PC1 e PC2 (superior à esquerda), PC1 e PC3 (superior à direita) e PC2 e PC3 (inferior à esquerda). Triângulos em vermelho representam animais da linhagem de trabalho e círculos em azul representam animais da linhagem de corrida da raça Quarto de Milha.



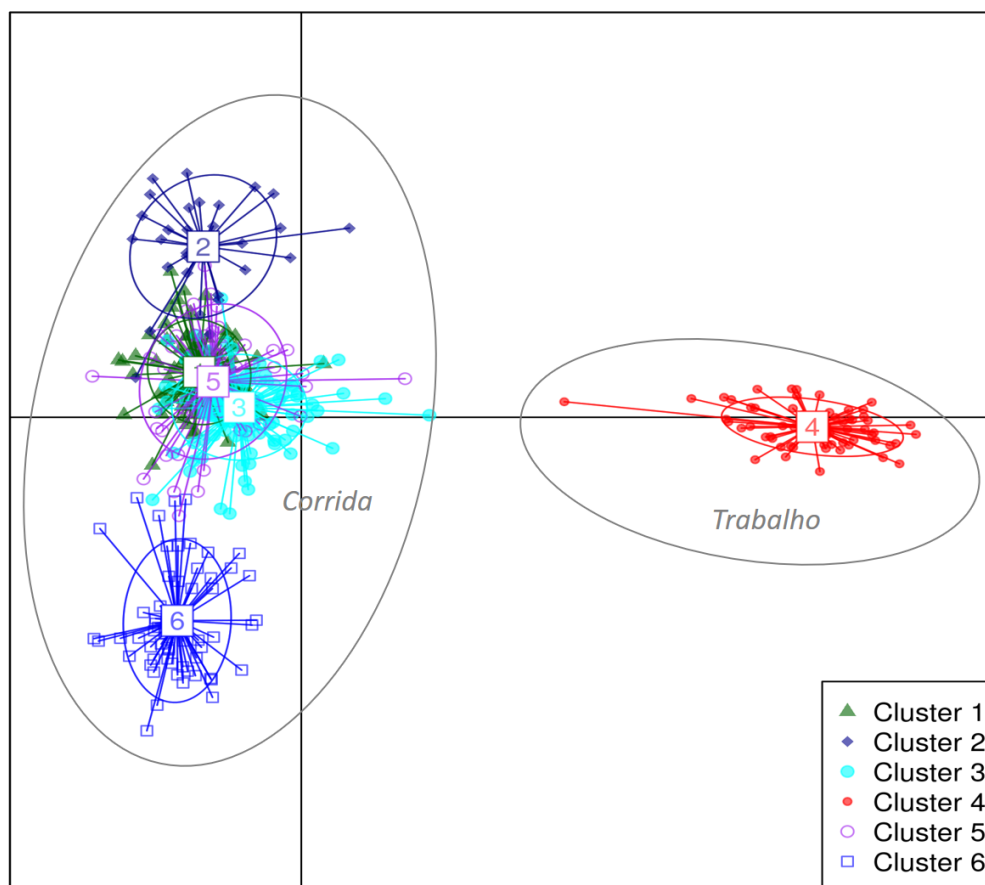
Para o maior entendimento da estrutura populacional da raça Quarto de Milha no Brasil, sobretudo, na linhagem de corrida, a qual apresentou notável estratificação, os dados genômicos foram analisados por meio de análise discriminante dos componentes principais (*discriminant analysis of principal component – DAPC*). O número ideal de grupos ( $K = 6$ ) e de componentes principais (PCs = 10) a serem retidos na DAPC foram escolhidos a partir do menor valor de BIC e maior valor de  $\alpha$ -score, 3135 e 0,8 respectivamente (Figura 5). A partir da definição de grupos ( $K = 6$ ), foram então analisados os *pedigrees* e os principais reprodutores (os mais frequentes) dentro de cada grupo, os quais notou-se tratar de grupos familiares, com um ou poucos indivíduos aparentados como principais reprodutores.



**Figura 5** – Variações do número ótimo de componentes principais a serem retidos (à esquerda) e número ótimo de grupos (à direita) dentro da amostra populacional da raça Quarto de Milha constituída de cavalos das linhagens de trabalho e de corrida.

A partir da análise de DAPC foi possível notar acentuada distância genética entre a linhagem trabalho e grupos pertencentes à linhagem de corrida (Figura 6), o que, portanto, mostra clara distinção entre as duas linhagens da raça Quarto de Milha. Na linhagem de corrida, contudo, foram demonstradas subestruturas decorrentes da formação de famílias que descendem de diferentes e importantes reprodutores. Observou-se, de forma branda, algumas distinções genéticas entre os grupos

centrais. Formados por famílias mais antigas, se considerada a história recente da raça (principais reprodutores das décadas de 1960 e 1970), estes grupos mais próximos representam uma base genética mais homogênea dentro da linhagem de corrida. Isto ocorre, devido, principalmente, à acasalamentos interfamiliares de seus descendentes e semelhança genética dos principais reprodutores entre grupos. Famílias mais recentes com alto grau de parentesco entre si são representadas pelos *clusters* mais afastados (principais reprodutores das décadas 1980 e 1990) e formam grupos familiares explicitamente distintos, o que leva à sensível formação de subestruturas e estratificação dentro da linhagem de corrida, além de evidenciar que os seus reprodutores são também mais distintos do ponto de vista genético, em relação aos dos grupos mais centrais.



**Figura 6** – Gráfico da análise discriminante dos componentes principais a partir de dados genômicos de cavalos de trabalho e corrida da raça Quarto de Milha. Dentro das elipses estão os grupos observados dentro das linhagens de trabalho (à direita) e de corrida (à esquerda). Cores e pontos distintos representam cada grupo (K) obtido previamente.

As informações genômicas, juntamente com as de *pedigree* (principais reprodutores), puderam elucidar as distâncias reais entres famílias e indivíduos, além de possibilitar futuros estudos do desempenho em corridas dentro e entre famílias, bem como servir para correção de subestruturas para estudos genômicos de associação.

Petersen et al. (2014) encontraram na raça Quarto de Milha criada nos Estados Unidos, a partir de análises de dados genômicos, seis grupos com habilidades distintas (trabalho com bovinos, corrida, rédeas, apartação, conformação e lazer), com clara distância entre a linhagem de corrida e o restante da população, a qual formou grupo único e separado. Nessa mesma pesquisa, a análise de componentes principais mostrou que animais de trabalho com bovinos apresentaram alta similaridade com equinos destinados às provas de apartação e rédeas.

Os dados do presente trabalho mostraram ainda que, dentro da linhagem de corrida, os principais reprodutores dos grupos mais centrais, são de origem PSI, desde animais puros até cruzados, o que evidencia a forte influência dessa raça na base genética da população de cavalos Quarto de Milha de corrida do Brasil. Os reprodutores de grupos mais distantes, além da influencia PSI, têm também, por parte materna, importante influência genética de famílias antigas importantes da linhagem de corrida da raça Quarto de milha.

Estudos recentes com equinos da raça Quarto de Milha criados nos Estados Unidos mostraram que a diversidade genética da raça é moderada a alta em comparação a outras raças modernas de cavalos criados em diversos países (PETERSEN et al., 2014), como as raças equinas portuguesas (LUÍS et al., 2007), raças equinas criadas na França (LEROY et al. 2009), e diversas raças equinas da Europa, Ásia e América do Norte, como Andaluz, PSI, e equinos da Mongólia (McCUE et al., 2012). Os autores discorreram que isso é devido a uma população grande e em expansão, diversas ações fundadoras e uma mistura contínua com o PSI. Nesse mesmo estudo, foi confirmada a influência da raça PSI na linhagem de corrida, visto a alta similaridade encontrada no compartilhamento de haplótipos.

#### **4. CONCLUSÃO**

Neste trabalho foram descritas notáveis diferenças genéticas entre as linhagens de trabalho e de corrida de cavalos da raça Quarto de Milha criadas no Brasil. Os resultados evidenciaram o maior desequilíbrio de ligação e o menor tamanho efetivo na linhagem de corrida. Embora diferentes, o tamanho efetivo pode ser considerado relativamente pequeno em ambas, indicando a necessidade do monitoramento da variabilidade genética e controle da endogamia na raça. Ainda em relação aos animais de corrida, foram observadas subestruturas decorrentes da formação de famílias que descendem de diferentes e importantes reprodutores.

## 5. BIBLIOGRAFIA CITADA

- ABQM. Associação brasileira dos criadores de cavalos Quarto de Milha. Disponível em: <<http://www.abqm.com.br/>>. Acesso em: 26 out. 2017.
- AI, H.; HUANG, L.; REN, J. Genetic Diversity, Linkage Disequilibrium and Selection Signatures in Chinese and Western Pigs Revealed by Genome-Wide SNP Markers. **PLoS ONE** 8(2): e56001, 2013.
- ARAKI, H.; TACHIDA, H. Bottleneck effect on evolutionary rate in then early neutral mutation model. **Genetics**, v. 147, p. 907–914, 1997.
- ARIAS, J. A.; KEEHAN, M.; FISHER, P.; COPPIETERS, W., SPELMAN, R. A high density linkage map of the bovine genome. **BMC Genetics**, v. 10, p. 18, 2009.
- BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. Characteristics of linkage disequilibrium in North American Holsteins. **BMC Genomics**, v. 11, p. 421, 2010.
- CERVANTES, I.; GOYACHE, F.; MOLINA, A.; VALERA, M.; GUTIERREZ, J. P. Estimation of effective population size from the rate of coancestry in pedigreed populations. **J. Anim. Breed Genet.**, v.128, n. 1, p. 56-63, 2011.
- CLAYTON, D. snpStats: SnpMatrix and XSnpmatrix classes and methods, R package version 1,24,0, 2015. Disponível em: <<http://bioconductor.org/packages/release/bioc/html/snpStats.html>>.
- CORBIN, L. J.; BLOTT, S. C.; SWINBURNE, J. E. Linkage disequilibrium and historical effective population size in the Thoroughbred horse. **Anim Genet**, v. 41, p. 8- 15, 2010.
- COSTA, M. D.; BERGMANN, J. A. G.; REZENDE, A. S. C.; FONSECA, C. G., FARIA, F. J. C. Estudo da subdivisão genética da raça Mangalarga Marchador. **Arq. Bras. Med. Vet. Zootec.**, v. 57, n. 2, p. 272-280, 2005.
- DIAS, I. M. G.; BERGMANN, J. A. G.; REZENDE, A. S. C; et al. Formação e estrutura populacional do equino Brasileiro de Hipismo. **Arq. Bras. Med. Vet. Zootec.**, v. 52, n. 6, p. 647-654, 2000.
- DU, F-X.; CLUTTER, A. C.; LOHUIS, M. M. Characterizing Linkage Disequilibrium in Pig Populations. **International Journal of Biological Sciences**, v. 7, p. 166-178, 2007.

- ESPIGOLAN, R.; BALDI, F.; BOLIGON, A. A.; SOUZA, F. R. P.; GORDO, D. G. M.; TONUSSI, R. L. L.; CARDOSO, D. F.; OLIVEIRA, H. N.; TONHATI, H.; SARGOLZAEI, M.; SCHENKEL, F. S.; CARVALHEIRO, R.; FERRO, J. A.; ALBUQUERQUE, L. G. Study of whole genome linkage disequilibrium in Nelore cattle. **BMC Genomics**, v. 14, p. 305, 2013.
- FARIA, R. A. da S. **Estrutura populacional e parâmetros genéticos da característica classe de tempo em corridas de equinos da Raça Quarto de Milha**. 2016. 69 f, Dissertação (mestrado) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2016.
- FREEMAN, D. W. Physical conditioning of horses. Stillwater, Okla: Oklahoma State University, Division of Agricultural Sciences and Natural Resources. 2013.
- GODDARD, M. E.; HAYES, B. J.; MCPARTLAN, H.; CHAMBERLAIN, A. J. "Can the same genetic markers be used in multiple breeds?". In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, CD-ROM communication n<sup>o</sup> 22-16, Brasil: Belo Horizonte, MG, p.13-18, 2006.
- HAYES, B. J.; VISSCHER, P. M.; MCPARTLAN, H. C.; GODDARD, M. E. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Research*, v. 13, p. 635–643, 2003.
- HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J.; GODDARD, M. E. Genomic selection in dairy cattle: Progress and challenges. **Journal of Dairy Science**, v. 92, p. 433- 443, 2009.
- HILL, W. G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **Theoretical and Applied Genetics**, v. 38, p. 226-231, 1968.
- JOMBART, T. Adegnet: a R package for the multivariate analysis of genetic markers. **Bioinformatics**, v. 24, p. 1403–1405, 2008.
- JOMBART, T.; DEVILLARD, S.; BALLOUX, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. **Bmc Genet**, v. 11, p. 94-109, 2010.
- KIJAS, J. W.; LENSTRA, J. A.; HAYES, B.; BOITARD, S.; PORTO NETO, L. R.; et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. **PLoS ONE**, v. 10, e1001258, 2012.

- KONG, A.; GUDBJARTSSON, D. F.; SAINZ, J.; JONSDOTTIR, G. M.; GUDJONSSON, S. A.; RICHARDSSON, B.; SIGURDARDOTTIR, S.; BARNARD, J.; HALLBECK, B.; MASSON, G.; SHLIEN, A.; PALSSON, S. T.; FRIGGE, M. L.; THORGEIRSSON, T. E.; GULCHER, J. R.; STEFANSSON, J. R. A high-resolution recombination map of the human genome. **Nat Genet.**, v. 31, p. 241–24, 2002.
- LEE, Y-S.; WOO LEE, J.; KIM, H. Estimating effective population size of thoroughbred horses using linkage disequilibrium and theta ( $4N\mu$ ) value. **Livestock Science**, v. 168, p. 32-37, 2014.
- LEROY, G.; CALLÈDE, L.; VERRIER, E.; MÉRIAUX, J. C.; RICARD, A.; DANCHIN-BURGE, C.; ROGNON, X. Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. **Genet Sel Evol.**, v. 41, n. 5, 2009.
- LUÍS, C.; JURAS, R.; OOM, M. M.; COTHRAN, E. G. Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation, **Anim Genet.**, v. 38, p. 20–27, 2007.
- MCCUE, M. E.; BANNASCH, D. L.; PETERSEN, J. L.; GURR, J.; BAILEY, E.; BINNS, M. M.; et al. A high density SNP array for the domestic horse and extant perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. **PLoS Genet.**, v. 8, p. 1-14, 2012.
- McCUE, M.; MICKELSON, J. Genomic Tools and Resources: Development and Applications of an Equine SNP Genotyping Array. In: CHOWDARY, B.H. (ed.) *Equine Genomics*, Ames: Blackwell Publishing Ltd, p.113–14, 2013.
- MCKAY, S. D.; SCHNABEL, R. D.; MURDOCH, B. M.; MATUKUMALLI, L. K.; AERTS, J.; WOUTER COPPIETERS, W.; CREWS, D.; DIAS NETO, E.; GILL, C. A.; GAO, C.; MANNEN, H.; STOTHARD, P.; WANG, Z.; VAN TASSELL, C.P.; WILLIAMS, J. L.; TAYLOR, J. F.; MOORE, S. S. Whole genome linkage disequilibrium maps in cattle. **BMC Genetics**, v. 8, p. 74, 2007.
- MEADOWS, J. R. S.; CHAN, E. K. F.; KIJAS, J. W. Linkage disequilibrium compared between five populations of domestic sheep. **BMC Genetics**, v. 9, p. 61, 2008.
- MEIRA, C. T.; FORTES, M. R. S.; FARAH, M. M.; PORTO-NETO, L. R.; CURI, R. A.; MOORE, S. S.; et al. A genome-wide association study for height at withers in racing quarter horse. **Association for the Advancement of Animal Breeding and Genetics**, v. 20, p. 420-423, 2013.

- MEIRA, C. T.; CURI, R. A.; FARAH, M. M.; OLIVEIRA, H. N.; BELTRAN, N. A. R.; SILVA, J. A. II V.; MOTA, M. D. S. Prospection of genomic regions divergently selected in racing line of Quarter Horses in relation to cutting line. **Animal**, v. 8, n. 11, p. 1754-1764, 2014a.
- MEIRA, C. T.; FARAH, M. M.; FORTES, M. R. S.; MOORE, S. S.; PEREIRA, G. L.; SILVA, J. A. II V.; MOTA, M. D. S.; CURI, R. A. A Genome-wide association study for morphometric traits in Quarter Horse. *Journal of Equine Veterinary Science*, v. 34, ed. 8, p. 102-1031, 2014b.
- NORDBORG, M.; BOREVITZ, J. O.; BERGELSON, J.; BERRY, C. C.; CHORY, J.; HAGENBLAD, J.; KREITMAN, M.; MALOOF, J. N.; NOYES, T.; OEFNER, P. J.; et al. The extent of linkage disequilibrium in *Arabidopsis thaliana*. **Nat Genet.**, v. 30, n. 2, p.190–193, 2002.
- ORAGUZIE, N. C., RIKKERINK, E. H. A., GARDINER, S. E., SILVA, H. N. de. (Eds.) Association Mapping in Plants. New York: Springer, p. 11-39, 2007.
- PATTERSON, N.; PRICE, A. L.; REICH, D. Population structure and eigenanalysis. **Plos Genetics**, Cambridge, v. 2, p. 2074-2093, 2006.
- PEREIRA, G. L.; CHUD, T.; BERNARDES, P.; VENTURINI, G.; CHARDULO, L. A.; CURI, R. A. Genotype Imputation and Accuracy Evaluation in Racing Quarter Horses Genotyped Using Different Commercial SNP Panels. **Journal of Equine Veterinary Science**, v. 58, p. 89 - 96, 2017.
- PEREIRA, G. L.; SILVA, J. A. II V.; CHARDULO, L. A.; FARIA, R. A. da S.; CURI, R. A. Genomic regions associated with performance in racing line of Quarter Horses. *Livestock Science*, v. online, p. 1, 2018.
- PETERSEN, J. L.; MICKELSON, J. R.; CLEARY, K. D.; MCCUE, M. E. The American Quarter Horse: Population Structure and Relationship to the Thoroughbred. **Journal of Heredity**, v. 105, n. 2, p. 148–162, 2014.
- PRITCHARD, J. K.; PRZEWORSKI, M. Linkage disequilibrium in humans: models and data. **American Journal of Human Genetics**, v. 69, p. 1–14, 2001.
- PURCELL, S.; NEALE, B.; TODD-BROWN, K.; THOMAS, L.; FERREIRA, M. A.; BENDER, D.; MALLER, J.; SKLAR, P.; DE BAKKER, P. I.; DALY, M. J.; SHAM, P. C. PLINK: a tool set for wholegenome association and population-based linkage analyses. **American Journal of Human Genetics**, v. 81, p. 559-575, 2007.



- R CORE TEAM. R: A language and Environment for Statistical Computing. R foundation for Statistical Computing 2016. <<http://www.R-project.org/>>.
- SARGOLZAEI, M.; CHESNAIS, J. P.; SCHENKEL, F. S. A new approach for efficient genotype imputation using information from relatives. **BMC Genomics**, v. 15, p. 478, 2014.
- SARGOLZAEI, M.; SCHENKEL, F. S.; JANSEN, G. B.; SCHAEFFER, L. R. Extent of linkage disequilibrium in Holstein cattle in North America. **Journal of Dairy Science**, v. 91, p. 2106-2117, 2008.
- SHIFMAN, S.; BELL, J. T.; COPLEY, R. R.; TAYLOR, M. S.; WILLIAMS, R. W.; MOTT, R.; et al. A High-Resolution Single Nucleotide Polymorphism Genetic Map of the Mouse Genome. **PLoS Biol**, 4(12): e395, 2006.
- SILVA, C. R.; NEVES, H. H. R.; QUEIROZ, S. A.; et al. Extent of linkage disequilibrium in Brazilian Gyr dairy cattle based on genotypes of AI sires for dense SNP markers. In: 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Alemanha. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, 2010.
- SVED, J. A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. **Theor. Popul. Biol.**, v. 2, p. 125-41, 1971.
- TRYON, R. C.; PENEDO, M. C. T.; MCCUE, M. E.; VALBERG, S. J.; MICKELSON, J. R.; FAMULA, T. R.; WAGNER, M. L.; JACKSON, M.; HAMILTON, M. J.; NOOTEBOOM, S. et al. Evaluation of allele frequencies of inherited disease genes in subgroups of American Quarter Horses. **J Am Vet Med Assoc.**, v. 234, p. 120–125, 2009.
- WANG, L.; SØRENSEN, P.; JANSS, L.; OSTERSEN, T.; EDWARDS, D. Genome wide and local pattern of linkage disequilibrium and persistence of phase for 3 Danish pig breeds. **BMC Genetics**, v. 14, p. 115-126, 2013.
- WEI, X.; JACKSON, P. A.; MCINTYRE, C. L. Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. *Theoretical and Applied Genetics*, New York, v. 114, p. 155-164, 2006.
- YU, J.; BUCKLER, E. S. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, London, v. 17, p. 155-160, 2006.
- ZHU, C.; GORE, M.; BUCKLER IV, E. S.; YU, J. Status and prospects of association mapping in plants. *The Plant Genome*, Madison, v. 1, p. 5-20, 2008.

ZHU, M.; ZHU, B.; WANG, Y. H.; WU, Y.; XU., L.; GUO, L. P.; YUAN, Z. R.; ZHANG, L. P.; GAO, X.; GAO, H. J.; XU, S. Z.; LI, J. Y. Linkage Disequilibrium Estimation of Chinese Beef Simmental Cattle Using High-density SNP Panels. Asian-Australasian. **Journal of Animal Science**, v. 26, p. 772-779, 2013.