

**UNIVERSIDADE FEDERAL DA BAHIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA**

**ESTUDOS GENÔMICO-MOLECULARES PARA CARACTERÍSTICAS
MORFOMÉTRICAS E DE VARIABILIDADE GÊNICA EM EQUINOS**

MARISA SILVA BASTOS

**SALVADOR – BAHIA
FEVEREIRO – 2022**



**UNIVERSIDADE FEDERAL DA BAHIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA**

**ESTUDOS GENÔMICO-MOLECULARES PARA CARACTERÍSTICAS
MORFOMÉTRICAS E DE VARIABILIDADE GÊNICA EM EQUINOS**

MARISA SILVA BASTOS
Zootecnista

**SALVADOR – BAHIA
FEVEREIRO – 2022**

MARISA SILVA BASTOS

**ESTUDOS GENÔMICO-MOLECULARES PARA
CARACTERÍSTICAS MORFOMÉTRICAS E DE VARIABILIDADE
GÊNICA EM EQUINOS**

Tese apresentada ao Programa de Pós-Graduação em Zootecnia, da Universidade Federal Da Bahia como requisito parcial para obtenção do título de Doutor em Zootecnia.

Área de Concentração: Produção Animal

Orientador: Prof. Dr. Raphael Bermal Costa
Coorientadores: Prof. Dr. Gregório Miguel Ferreira de Camargo
Dra. Iara Del Pilar Solar Diaz

SALVADOR – BA
FEVEREIRO – 2022

Bastos, Marisa Silva.

Estudos genômico-moleculares para características morfométricas e de variabilidade gênica em equinos / Marisa Silva Bastos. - 2022.

62 f.: il.

Orientador: Prof. Dr. Raphael Bermal Costa.

Coorientador: Prof. Dr. Gregório Miguel Ferreira de Camargo.

Tese (doutorado) - Universidade Federal da Bahia, Escola de Medicina Veterinária e Zootecnia, Salvador, 2022.

1. Zootecnia. 2. Cavalos - Criação. 3. Cavalos - Melhoramento genético. 4. Genômica. I. Costa, Raphael Bermal. II. Universidade Federal da Bahia. Escola de Medicina Veterinária e Zootecnia. III. Título.

CDD - 636.1

CDU - 636.1

**ESTUDOS GENÔMICO-MOLECULARES PARA
CARACTERÍSTICAS MORFOMÉTRICAS E DE VARIABILIDADE
GÊNICA EM EQUINOS**

Marisa Silva Bastos

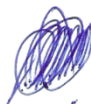
**Tese defendida e aprovada para obtenção do grau de
Doutor em Zootecnia**

Salvador, 23 de fevereiro de 2022

Comissão examinadora:



**Dr. Raphael Bernal Costa
UFBA
Orientador / Presidente**



**Analívia Martins Barbosa
UFBA
Vice-Coordenadora**

“Tudo o que temos de decidir é o que fazer com o tempo que nos é dado.”

Gandalf - J.R.R. Tolkien

Dedico este trabalho aos meus pais, José Gomes Bastos (*in memoriam*) e Maria Aparecida da Silva Bastos, meus maiores exemplos de força, coragem e honestidade. E aos meus irmãos, Simone Pereira Bastos, Lucas Silva Bastos e Gabriel Silva Bastos, com os quais sei que posso contar sempre.

AGRADECIMENTOS

A Deus, por ter me dado saúde e força para superar as dificuldades e realizar esse sonho.

Aos meus pais e aos meus irmãos que com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

A alguém que se tornou muito especial, Lucas Maia, pelo apoio, incentivo, compressão e companheirismo.

Às parceiras de trabalho sem as quais nada disso seria possível, Jackeline Alves e Louise Oliveira.

Aos professores Raphael Bermal e Gregório Camargo, pela compreensão, incentivo e orientação fundamental para o desenvolvimento desta tese.

A Iarita (Dra. Iara Diaz), pela coorientação, ensinamentos e inspiração.

Às Pós-Doutorandas Val (Dra. Valdecy Aparecida) e Fabi (Dra. Fabieli Loise), pela parceria, incentivo e transmissão de conhecimentos.

Aos estagiários de iniciação científica, pela ajuda nas análises.

A todos os professores e colegas do Programa de Pós-Graduação em Zootecnia da UFBA, que foram tão importantes nesse período da minha vida acadêmica, obrigada pelo apoio e convívio.

À Professora Fernanda Godoi (UFRRJ), pela parceria.

A Sabrina Lambert (LBCM – UFBA), Waléria Borges (Hospmev) e Mônica Vieira (LGHM - UFBA), pelo auxílio no desenvolvimento das análises.

Ao Grupo de Estudo em Genética e Melhoramento Animal (GEMA), pelo companheirismo e por todo conhecimento compartilhado.

Ao Grupo de Estudos em Equideocultura do Recôncavo Baiano (GEERB), pelo apoio e incentivo.

Aos integrantes da banca de defesa, pelas valiosas sugestões que certamente irão enriquecer o presente trabalho.

À FAPESB, pela concessão da bolsa.

À Universidade Federal da Bahia e ao Programa de Pós-Graduação em Zootecnia, por ter tornado possível a realização dos meus estudos e deste projeto.

BIOGRAFIA DA AUTORA

MARISA SILVA BASTOS - filha de Maria Aparecida da Silva Bastos e José Gomes Bastos, nasceu em Ruy Barbosa, Estado da Bahia, no dia 02 de junho de 1990. Coursou Zootecnia na Universidade Federal do Recôncavo da Bahia, no período de 2010 a 2015. Obteve o título de Mestra em Zootecnia, área de concentração Produção Animal, pela Universidade Federal de Sergipe em julho de 2017. Em outubro de 2017, iniciou o Doutorado no Programa de Pós-Graduação em Zootecnia da Universidade Federal da Bahia, área de concentração Produção Animal, sob orientação do Professor Dr. Raphael Bernal Costa, do Professor Dr. Gregório Miguel Ferreira de Camargo e da Dra. Iara Del Pilar Solar Diaz. Em 23 de fevereiro de 2022, submeteu-se à banca examinadora para defesa da tese de doutorado.

LISTA DE FIGURAS

CAPÍTULO I

Caracterização do gene *PRDM9* em equinos

Página

Figura 1 - Novos domínios <i>ZF</i> identificados no gene <i>PRDM9</i> em equinos.....	23
Figura 2 – Seguimento identificado após o domínio <i>ZF</i> do gene <i>PRDM9</i> em equinos..	27

CAPÍTULO II

Associação genômica para componentes principais de medidas morfométricas em equinos: identificação de genes relacionados a crescimento ósseo

Figura 1 - Gráfico Manhattan obtido por associação a um índice composto pelo PC1 ponderado pelos EBVs de medidas lineares de equinos da raça Campolina.....	41
---	----

LISTA DE TABELAS

CAPÍTULO I

Caracterização do gene *PRDM9* em equinos

Página

Tabela 1 – Descrição e frequência dos alelos identificados para o gene <i>PRDM9</i> em equinos.....	23
Tabela 2 - Genótipos <i>PRDM9</i> observados em diferentes raças de equinos	25
Tabela 3 - Frequências dos genótipos observados para o gene <i>PRDM9</i> em equinos	27

CAPÍTULO II

Associação genômica para componentes principais de medidas morfométricas em equinos: identificação de genes relacionados a crescimento ósseo

Tabela 1- Estatística descritiva de 15 medidas lineares em equinos da raça Campolina.....	37
Tabela 2 - Identificação das regiões (cromossomos) e respectivos conjuntos de genes que explicaram maior proporção da variância genética aditiva das características analisadas pelo índice HCP1.....	42

SUMÁRIO

Página

1. INTRODUÇÃO GERAL	1
2. REVISÃO DE LITERATURA GERAL	2
2.1. <i>Os Equinos</i>	2
2.2. <i>Importância Da Variabilidade Genética</i>	3
2.3. <i>O Gene PRDM9</i>	4
2.4. <i>Estimação de Parâmetros Genéticos Utilizando Componentes Principais</i>	6
2.5. <i>GWAS em Equinos</i>	8
REFERÊNCIAS	10
CAPÍTULO I.....	18
Caracterização do gene <i>PRDM9</i> em equinos	
Resumo.....	18
Abstract	19
1. Introdução	20
2. Material e Métodos	21
2.1. <i>Amostragem e Extração de DNA</i>	21
2.2. <i>Extração do DNA e PCR</i>	21
2.3. <i>Genotipagem</i>	22
3. Resultados	22
4. Discussão.....	28
5. Conclusão.....	29
Referências	29
CAPÍTULO II	32
Associação genômica para componentes principais de medidas morfométricas em equinos: identificação de genes relacionados a crescimento ósseo	
Resumo.....	32
Abstract	33
1. Introdução	34
2. Material e Métodos.....	36
2.1. <i>Estimação dos parâmetros genéticos</i>	36
2.2. <i>Componentes Principais</i>	37
2.3. <i>Dados genotípicos e controle de qualidade</i>	38
2.4. <i>Análise de associação genômica ampla (GWAS)</i>	38
2.5. <i>Identificação dos genes</i>	39
3. Resultados	40
4. Discussão.....	44
5. Conclusão.....	46
Referências	46
CONSIDERAÇÕES FINAIS	51

1. INTRODUÇÃO GERAL

Os equídeos tem papel relevante no desenvolvimento da civilização, uma vez que, serviram ao homem para diversas funções como fonte de alimento, força de trabalho, matéria prima e transporte.

Estudos paleontológicos indicam que o ancestral dos equídeos modernos teve origem no continente americano e se espalhou pelo mundo a milhares de anos (MacFadden, 2013; Cerdeño, 2017; Heintzman et al., 2017). Estudos como esses permitiram traçar a evolução dos equídeos desde animais com cerca de 30cm e presença de quatro dedos até os animais com a conformação moderna.

Desde a domesticação, os equídeos passaram por intensa seleção fenotípica o que possibilitou o surgimento de diversas raças com diferentes aptidões. No entanto, com os crescentes avanços tecnológicos houve a substituição dos equídeos principalmente nas funções de transporte e tração, causando diminuição do tamanho populacional de diversas raças. Essa redução no tamanho populacional intensifica a ocorrência de acasalamentos endogâmicos, o que pode causar consequências negativas para sua produção e conservação devido ao declínio da variabilidade genética (De Andrade Salles et al., 2013; Todd et al., 2018; Dini et al., 2020).

A manutenção da variabilidade genética é premissa básica para que haja eficiência na implementação de programas de melhoramento genético animal. Assim, é imprescindível a utilização de ferramentas que possibilitem o monitoramento da diversidade genética dentro e entre as populações.

O *crossing over* é um processo de recombinação no qual há a quebra em determinados pontos da sequência de DNA e o rearranjo dos fragmentos durante a meiose, promovendo assim aumento da variabilidade genética. A proteína PRDM9, através da estrutura denominada *zinc fingers (ZF)*, se liga em determinados locais do DNA incitando a quebra e a recombinação da sequência. Nos equinos, o domínio *ZF* é formado por um arranjo sequencial de 28 aminoácidos, e uma única alteração na sequência determina uma nova variante com diferentes pontos de ligação no DNA (Steiner & Ryder, 2013).

Por meio do estudo da codificação da proteína PRDM9 pode ser possível determinar padrões de dissimilaridade genética entre grupos e, quando necessário, sugerir métodos para corrigir problemas ligados a baixa variabilidade genética.

O objetivo do melhoramento genético é a melhoria das características de interesse, e esse resultado é obtido nas progênes em relação aos pais, geração após geração mantendo adicionalmente a homogeneidade fenotípica, de acordo com o sistema de produção.

Na espécie equina, assim como em outras espécies de interesse zootécnico, a homogeneidade fenotípica diz respeito, entre outras características, ao padrão de conformação racial. Na espécie, as características morfométricas são muito utilizadas para critérios de registro. Além disso, a morfometria assume grande importância pois está diretamente relacionada a dinâmica de locomoção do animal (Holmström e Back, 2013).

A raça Campolina registra a medida de 15 características corporais (morfométricas). A mensuração é uma técnica relativamente simples, que possibilita a obtenção de um grande volume de informações, permitindo estimar os parâmetros genéticos que podem auxiliar os criadores na seleção dos animais (Costa et al., 2016; Souza et al., 2019). Entretanto, essas características são altamente correlacionadas (Junqueira et al., 2021), o que as fazem candidatas a serem estudadas em análises de componentes principais (PCA). Esse método estatístico reduz o volume de dados através do descarte de variáveis que apresentam baixa variabilidade ou redundantes (Meira et al., 2013; Solar Diaz et al., 2021).

Adicionalmente, a realização de GWAS associada a um índice composto por PCA pode possibilitar a descoberta de regiões relacionadas a um conjunto de características relevantes para a espécie equina. Além disso, estudos deste tipo propiciam melhor conhecimento sobre a expressão biológica de características quando avaliadas pela proporção conjunta da variância genética aditiva (Mudadu et al., 2016; Vargas et al., 2018; Zhou et al., 2019).

2. REVISÃO DE LITERATURA GERAL

2.1. Os Equinos

O cavalo como conhecemos atualmente evoluiu de animais com hábitos e estrutura corporal muito diferentes das observadas nos indivíduos modernos. Isto graças a um processo evolutivo de aproximadamente 55 milhões de anos *Caballus* (cavalo doméstico) (Heintzman et al., 2017).

Ao se espalhar em diferentes habitats os equinos passaram por processo de seleção e desenvolveram adaptações e características que permitiram sua sobrevivência e o surgimento de diferentes raças - grupos de indivíduos de uma mesma espécie que compartilham características semelhantes que são transmitidas aos seus descendentes e que os tornam diferentes de outros grupos (Heintzman et al., 2017).

Apesar de apresentarem diferenças morfológicas, as raças pertencentes à espécie *Equus caballus*, ou simplesmente equinos, compartilham o mesmo cariótipo ($2n=64$), portanto o cruzamento entre elas gera descendentes viáveis e férteis.

Pouco se sabe a respeito dos mecanismos evolutivos envolvidos na formação das espécies e os processos que as resguardam como organismos distintos. É sabido que existem algumas barreiras reprodutivas que mantêm a integridade das espécies, uma delas é a esterilidade híbrida, uma barreira reprodutiva pós-zigótica gerada por incompatibilidades genéticas em populações divergentes (Islam et al., 2013; Steiner & Ryder, 2013).

Esta incompatibilidade genética pode estar relacionada ao gene *PRDM9* uma vez que este gene codifica uma histona metiltransferase que está diretamente envolvida na gametogênese (Oliver, et al. 2009; Mukaj et al., 2020; Forejt et al., 2021). A literatura ressalta ainda que outro fator importante ligado ao gene *PRDM9* é o seu potencial de aumentar a variabilidade genética nas populações (Schwartz et al., 2014; Grey et al., 2018).

2.2. Importância Da Variabilidade Genética

Os organismos apresentam diferentes composições em suas sequências do DNA, essas diferenças também definidas como variabilidade genética, podem causar variações no fenótipo dos animais. A manutenção da variabilidade genética é de extrema importância principalmente do ponto de vista evolutivo, pois é a ocorrência de diferentes alelos, e suas conseqüentes variações na expressão fenotípica, que possibilitam a adaptação dos animais as mudanças ambientais (Ortiz-Barrientos et al., 2016; Hecht et al., 2020).

A manutenção da variabilidade genética é premissa básica para a obtenção de ganho genético com a implementação de programas de melhoramento genético animal, e a redução da variabilidade pode levar a prejuízos produtivos e reprodutivos.

Os principais fatores que causam redução na variabilidade genética são os pequenos números de indivíduos aptos à reprodução que compõem as populações e os acasalamento

consanguíneos ou endogamia. Esse tipo de acasalamento foi utilizado para formação da maioria das raças equinas com a finalidade de fixar características de interesse, sem se preocupar com ocorrência de homozigose de alelos deletérios (Al Abri et al., 2017; Todd et al., 2018; Dini et al., 2020).

Estudos relatam preocupação com relação a diversidade genética e as taxas de endogamia em Equídeos. Marchiori et al. (2019), ao analisarem a raça Quarto de Milha dividida em duas linhagens (corrida e trabalho) observaram que a linhagem corrida apresentou pequeno tamanho efetivo de população indicando a necessidade do monitoramento da variabilidade genética e controle da endogamia na raça. Para Assis et al. (2018), houve aumento do coeficiente de endogamia entre os anos de 1962 a 2000 na raça Mangalarga Marchador, e isto pode estar associado ao acasalamento entre animais aparentados para a fixação de características desejáveis na raça. Alves et al. (2021) relataram baixa diversidade genética em populações das raças Lavradeiro e Piquira. Os autores sugerem que os acasalamentos sejam dirigidos para aumentar a variabilidade genética.

É fundamental avaliar a variabilidade genética e a estrutura populacional, só então é possível a implementação de programas de seleção eficazes e com melhor direcionamento dos acasalamentos (Hosseini-Zadeh, 2016).

Em termos genéticos, o *crossing-over* é uma forma de recombinação gênica que aumenta a variabilidade genética, isto por que ao promover a troca de partes dos cromossomos homólogos durante a meiose ocorre o rearranjo dos genes presentes. Assim, os descendentes podem receber cromossomos com alelos de genes diferentes dos existentes nos pais. No entanto, a ocorrência de *crossing-over* ao longo dos cromossomos não é regular, mas concentrada em fragmentos relativamente pequenos do DNA, denominados *hotspots* de recombinação (Mézard et al., 2015).

Segundo a literatura a PRDM9, uma proteína da família PRDM atuante na meiose, é responsável por determinar os *hotposts* em mamíferos, apresentando assim potencial de aumentar a variabilidade genética nas espécies (Hochwagen e Marais, 2010; Myers et al. 2010).

2.3. O Gene PRDM9

O grupo de proteínas PRDM é caracterizado por se apresentar de distintas formas moleculares, diferindo basicamente pela quantidade de domínios de zinco (ZF) presentes

em sua estrutura. Essas proteínas desempenham um papel essencial na tradução de sinais que controlam a proliferação e diferenciação celular (Di Zazzo et al., 2013).

O *PRDM9* foi identificado como o primeiro gene de especiação conferindo isolamento reprodutivo em vertebrados (Myers et al. 2010). A proteína codificada por esse gene contém os domínios KRAB e SET, que são altamente conservados entre metazoários, seguido por um conjunto de domínios de zinco, do inglês *zinc fingers (ZF)*, formados por aminoácidos, e uma única alteração de aminoácidos nos domínios *ZF* pode formar uma nova variante de *PRDM9* (Di Zazzo et al., 2013).

O domínio KRAB, geralmente encontrado em proteínas de ligação, nas proteínas do grupo PRDM está presente apenas na PRDM7 e na PRDM9, e seu truncamento por outras proteínas leva à perda da função da proteína PRDM9 e à alteração da prófase meiótica e da gametogênese (Imai et al., 2017). O domínio SET, por sua vez, possui atividade da trimetiltransferase catalisando o H3K4me3. O H3K4me3 induz a quebra do DNA (Hochwagen e Marais, 2010; Wu et al., 2013).

Os domínios *ZF*, são estruturas funcionais que apresentam entre 28 e 30 resíduos de aminoácidos e sua estabilidade estrutural é determinada pela ligação de pelo menos um átomo de zinco. Nas proteínas do grupo PRDM, essa estrutura pode se apresentar em diferentes arranjos modulares, caracterizando as variações nas proteínas desse grupo. Na PRDM9, cada *ZF* se liga a três bases nitrogenadas da sequência do DNA atraindo as moléculas que promovem a recombinação (Baker et al., 2015).

Segundo Gonen et al. (2017), modificações nos domínios *ZF* no gene *PRDM9* podem aumentar o ganho genético em espécies de produção, já que parálogos desse gene foram identificados, o que aumenta a variabilidade na localização e número de *hotspots* de recombinação. Estudos constataam rápida evolução e seleção do gene *PRDM9* em vários organismos incluindo humanos, ratos, bovinos e peixes (Oliver et al., 2009; Ponting, 2011; Ahlawat et al., 2017). A recombinação homóloga contribui para o enriquecimento da diversidade genética, a exclusão de alelos deletérios e a adequada segregação de cromátides durante a meiose (Coop e Przeworski, 2007).

O envolvimento do *PRDM9* na recombinação genética atraiu muita atenção e fez desse gene um foco de intensos estudos nos últimos anos. Vários alelos de *PRDM9* foram descritos em camundongos (Buard et al., 2014; Kono et al., 2014), chimpanzés (Auton et al., 2012; Groeneveld et al., 2012), e humanos (Berg et al., 2010; Myers et al., 2010).

Ahlawat et al. (2016) ressaltam a notável diversidade encontrada nos arranjos de domínios de *ZF* em pequenos ruminantes, e observaram a ocorrência de 36 alelos

diferentes deste gene para caprinos e 26 alelos para ovinos, sugerem ainda que o *PRDM9* provavelmente ativa *hotspots* de recombinação diferentes nas espécies analisadas, apresentando variabilidade substancial, não apenas entre espécies, mas também dentro das espécies. Em estudo semelhante analisando o gene *PRDM9* em bovinos, bubalinos, mithun e iaques, Ahlawat et al. (2017) encontraram vasta diversidade de arranjos de ZF que resultaram em 41 alelos do gene *PRDM9* exclusivos nessas espécies, os autores destacam ainda a importância deste gene como promotor de aumento na variabilidade genética. Os trabalhos citados demonstraram os alelos mais comuns para cada espécie e quais as raças que apresentam a menor variabilidade para este *locus*, permitindo aliar as informações obtidas a estratégias de melhoramento e conservação dos diferentes grupos genéticos.

Steiner e Ryder (2013) analisaram o gene *PRDM9* em espécies do gênero *Equus*, e descreveram uma sequência de oito domínios ZF na região terminal do gene (DKKDIKB), sendo este o único alelo descrito para a espécie equina. Assim, modificações nas sequências desses domínios de zinco, que resultam na formação de alelos distintos, produzem diferentes *hotspot* de recombinação ao longo da sequência do DNA e aumentam a variabilidade genética da população.

A variabilidade genética é de suma importância para os programas de melhoramento genético, pois sua ocorrência é considerada premissa básica para a obtenção de efetivo ganho genético através da seleção

2.4. Estimação de Parâmetros Genéticos Utilizando Componentes Principais

A correta e acurada estimação dos parâmetros genéticos é fundamental para o sucesso de um programa de melhoramento genético. O objetivo então é predizer os valores de forma com que esses estejam os mais próximos possíveis dos verdadeiros valores presentes na população, e para isso se utilizam as evidências. As evidências nesse caso são os fenótipos, ou melhor, o desempenho dos animais.

Usualmente, as tradicionais avaliações genéticas se utilizam de informações do desempenho do próprio indivíduo, dos seus parentes e de pedigree para estimar o mérito genético dos animais e nesse contexto a dependência sobre o fenótipo faz com que tais medidas devam ser corretamente obtidas. Em animais que são selecionados através de avaliações visuais, como os equídeos, as mensurações ganham maior importância, uma vez que a subjetividade aliada à seleção visual pode levar a viés e a estimativas irreais.

Se a seleção através da avaliação visual for aliada a seleção pelo mérito genético dos animais, os ganhos serão maiores e mais acurados. Assim, avaliar e estimar tais mensurações e métodos de avaliações é de suma importância para o progresso genético dos equídeos.

A técnica de mensuração das proporções corporais dos equídeos é um dos métodos mais antigos utilizados para a seleção dos animais. Atualmente mais de 15 características corporais são mensuradas nos equídeos e utilizadas pelas associações de criadores para definir o padrão racial, e para estudos relacionados a morfologia e aptidão de cada linhagem. Essa é uma técnica relativamente simples pois os marcos de referência para mensurações são determinados por locais de fácil observação, geralmente próximos a extremidades ósseas (Santiago et al., 2013).

As mensurações corporais permitem calcular as proporções entre as partes do corpo do cavalo e assim determinam a sua funcionalidade. Além disso, possibilita analisar a morfologia corporal, característica essencial para compreender a dinâmica da locomoção em equídeos. Permite ainda estimar parâmetros genéticos para melhor direcionamento da seleção dos animais (Barbosa, 1993; Costa et al., 1997; Pinto et al., 2005).

Algumas características como as morfométricas, que estão relacionadas a conformação estrutural do corpo dos equídeos sofrem menos influência do ambiente e por tanto, apesar de possuírem alta herdabilidade, acabam apresentando baixa amplitude de variação individual e alta correlação genética, tornando a utilização de informações de muitas características redundante (Boligon et al., 2016; Bramante et al., 2016; Novotná et al., 2017).

Nesse sentido, a utilização de métodos estatísticos como a análise de componentes principais (PCA), uma técnica de análise multivariada, facilita a interpretação de fenômenos ao elaborar variáveis alternativas a partir das variâncias e covariâncias, através de combinações lineares entre as variáveis originais.

A PCA é uma técnica que foi desenvolvida em 1901 por Karl Pearson, aperfeiçoada por Hotelling em 1933 e 1936, que busca transformar um conjunto de variáveis em outro conjunto reduzido e de mesma dimensão. No entanto, esta técnica só passou a ser amplamente utilizada a partir do advento da computação (Hongyu et al., 2016).

A PCA é o método de análise multivariada que cria combinações lineares com máxima variância, formando componentes através de combinações lineares de variáveis previamente selecionadas. Esses componentes seguem em uma ordem decrescente de variância, em que o primeiro componente representa a máxima variância, o segundo

componente apresenta a segunda maior variância e assim sucessivamente (Price et al., 2010; He et al., 2011). No entanto, os primeiros componentes devem apresentar o máximo de informações em termos de variação total.

Barbosa (1993) observou por meio da PCA que sete componentes eram suficientes para explicar 80% da variação total ao analisar 12 características morfométricas em cavalos Mangalarga Marchador. Ainda segundo o mesmo autor, a utilização de análise de componentes principais associada a análises genéticas possibilitam melhor compreensão das variações existentes e permitem a definição de índices de seleção ajustados. Santos (2006) utilizou a técnica de PCA para avaliar medidas lineares de machos e fêmeas campeões da raça Campolina e observou que foram necessários 5 e 4 componentes principais respectivamente, para explicar 80% da variação total.

Segundo Pinto et al. (2005), a PCA assume grande relevância para estudos sobre melhoramento e diversidade genética pois através dessa técnica é possível converter um conjunto de informações como altura e largura, em outro conjunto análogo passível de análises mais profundas e acuradas.

Adicionalmente, é possível aplicar a PCA em matrizes de covariância de valores genéticos (EBVs) de cada característica, assim os componentes obtidos permitem observar contrastes entre as características (Vargas et al., 2018). Os EBVs podem ser utilizados como pseudo-fenótipos estimados para animais genotipados que não possuem informações de fenótipos, tomando como base informações de desempenho dos pais, irmãos e progênes, levando em consideração a correção para fatores como ambiente, manejo, estação de nascimento, entre outros (Morota et al., 2014).

2.5. GWAS em Equinos

Quando comparada a outras espécies de interesse zootécnico, as pesquisas sobre genômica aplicada a espécie equina ainda têm muito a desenvolver. Os estudos iniciais abordando fermentas moleculares na espécie foram realizados com a raça Puro Sangue Inglês (PSI) e geraram informações que auxiliam na seleção de animais com desempenho superior em práticas esportivas (Pereira et al., 2015).

A realização desse tipo de estudo em raças de equinos utilizadas em práticas esportivas se dá principalmente devido ao alto valor agregado aos animais. No entanto, no Brasil mesmo para animais desta categoria os estudos são escassos. Na prática o que ocorre é a seleção dos animais baseada em seu fenótipo, premiações, desempenho em provas e genealogia (Regatieri e Mota, 2012).

De modo geral, a implementação de programas de melhoramento genético eficientes para a espécie equina é complexa devido principalmente à baixa herdabilidade das características de desempenho e os valores de correlação genética entre as características mais desejadas, como velocidade e performance (h^2 de 0,28 e 0,06 respectivamente; e correlação de 0,87) (Ricard e Touvais, 2007). No entanto, ferramentas genômicas surgem como alternativa viável para o efetivo melhoramento genético mesmo diante dos obstáculos apresentados.

Uma das vantagens da utilização de infamações genômicas na seleção de cavalos é que essas informações podem estar disponíveis antes de o animal atingir a maturidade sexual, independente de fatores como sexo, auxiliando de maneira mais eficaz nas decisões voltadas ao melhoramento genético. Ademais, conhecer previamente a capacidade genética do animal em resposta ao treinamento auxiliam os criadores na tomada de decisões com relação aos investimentos que serão empregados aos cavalos, principalmente para animais utilizados em atividades esportivas (Stock et al., 2016).

Os estudos de associação genômica ampla (GWAS) promovem o mapeamento genético por meio de associação entre *locus* e características fenotípicas com o objetivo de encontrar efeitos dos genes sobre o valor genético dos animais. Nesse sentido, estudos relatam avanços expressivos em análises de genes candidatos relacionados a características de interesse econômico em equinos como desempenho (Hill et al., 2010), pelagem (Corbin et al., 2020; Grilz-Seger et al 2020), fertilidade (Giesecke et al., 2009), resistência a doença (Dunuwille et al., 2020; Rosa et al., 2020), coordenação motora (Andersson et al., 2012; De Oliveira Bussiman et al., 2020; Ricard e Duluard, 2021), comportamento e temperamento (Momozawa et al., 2006), entretanto a utilização da genômica como ferramenta no melhoramento genético da espécies ainda está distante da realidade quando comparada a algumas espécies de exploração zootécnica como, aves, suínos e bovinos (Wolc et al., 2011; Jenko et al., 2017; Yang et al., 2020).

Desta forma, estudos que envolvam a aplicação e desenvolvimento de ferramentas aliadas à informação moleculares, que também utilizem de maneira eficiente as informações fenotípicas, se torna de suma importância para obter ganhos genéticos expressivos na equideocultura.

REFERÊNCIAS

- Ahlawat, S., De, S., Sharma, P., Sharma, R., Arora, R., Kataria, R.S., ... e Singh, R.K. 2017. Evolutionary dynamics of meiotic recombination hotspots regulator PRDM9 in bovids. *Molecular genetics and genomics*, 292(1), 117-131.
- Ahlawat, S., Sharma, P., Sharma, R., Arora, R., Verma, N.K., Brahma, B., ... e De, S. 2016. Evidence of positive selection and concerted evolution in the rapidly evolving PRDM 9 zinc finger domain in goats and sheep. *Animal genetics*, 47(6), 740-751.
- Al Abri, M.A., König von Borstel, U., Strecker, V., e Brooks, S.A. 2017. Application of genomic estimation methods of inbreeding and population structure in an Arabian Horse Herd. *Journal of Heredity*, 108(4), 361-368.
- Alves, J.S., da Silva Anjos, M., Bastos, M.S., de Oliveira, L.S.M., Oliveira, I.P.P., Pinto, L.F.B., ... e de Camargo, G.M.F. 2021. Variability analyses of the maternal lineage of horses and donkeys. *Gene*, 769, 145231.
- Andersson, L.S., Larhammar, M., Memic, F., Wootz, H., Schwochow, D., Rubin, C.J., ... e Kullander, K. 2012. Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. *Nature*, 488(7413), 642-646.
- Assis, C.M., Baena, M.M., De Fátima Bretanha, R., e Rocha, S.L.C.M. Endogamia e Relação Genética em Equinos da Raça Mangalarga Marchador No Brasil. 28º Congresso Brasileiro de Zootecnia – Zootec. Goiânia – GO, 2018. Anais: <http://www.adaltech.com.br/anais/zootecnia2018/resumos/trab-0845.pdf>
- Associação Brasileira de Criadores de Cavalos Campolina – ABCCCampolina, 2021. Disponível em:<<http://www.campolina.org.br/sobre.php>>. Acesso em: 14 jan. 2022.
- Auton, A., Fledel-Alon, A., Pfeifer, S., Venn, O., Ségurel, L., Street, T., ... e McVean, G. 2012. A fine-scale chimpanzee genetic map from population sequencing. *science*, 336(6078), 193-198.
- Baker, C.L., Petkova, P., Walker, M., Flachs, P., Mihola, O., Trachtulec, Z., ... e Paigen, K. 2015. Multimer formation explains allelic suppression of PRDM9 recombination hotspots. *PLoS genetics*, 11(9), e1005512.
- Barbosa, C.G. Estudo morfométrico na raça Mangalarga Marchador. Uma abordagem multivariada. Belo Horizonte: Universidade Federal de Minas Gerais, 1993. 76p. Dissertação (Mestrado em Zootecnia) - Universidade Federal de Minas Gerais, 1993.

- Berg, I.L., Neumann, R., Lam, K.W.G., Sarbajna, S., Odenthal-Hesse, L., May, C.A., e Jeffreys, A.J. 2010. PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nature genetics*, 42(10), 859-863.
- Binns, M., Swineburne, J., e Breen, M. 1999. *Molecular genetics of the horse. The genetics of the horse*, 109-118.
- Boligon, A.A., Vicente, I.S., Vaz, R.Z., Campos, G.S., Souza, F.R.P., Carvalheiro, R., e Albuquerque, L.G. 2016. Principal component analysis of breeding values for growth and reproductive traits and genetic association with adult size in beef cattle. *Journal of animal science*, 94(12), 5014-5022.
- Bramante, G., Cito, A.M., Ciani, E., Pieragostini, E., 2016. Multi-trait animal model estimation of genetic parameters for morphometric measurements in the Murgese horse breed. *Livestock Science*, 191, 139-142.
- Buard, J., Rivals, E., Dunoyer de Segonzac, D., Garres, C., Caminade, P., de Massy, B., e Boursot, P. 2014. Diversity of PRDM9 zinc finger array in wild mice unravels new facets of the evolutionary turnover of this coding minisatellite. *PloS one*, 9(1), e85021.
- Cerdeño, E. 2017. Fossil Horses of South America. *Phylogeny, Systemics and Ecology. Ameghiniana*, 54(4), 478-479.
- Corbin, L.J., Pope, J., Sanson, J., Antczak, D.F., Miller, D., Sadeghi, R., e Brooks, S.A. 2020. An Independent Locus Upstream of ASIP Controls Variation in the Shade of the Bay Coat Colour in Horses. *Genes*, 11(6), 606.
- Costa M.D. *Estudo Genético Quantitativo das Medidas Lineares do Pônei da Raça Brasileira*. Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 1997.
- Costa, M.D., Mendes, L.J., Maruch, S., Ramirez, P.A., Meneses, A.C.A., Martins Neto, T., e Chamone, J. M.A. 2016. Efeito da composição genética nas características de conformação em equinos. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 68(6), 1629-1637.
- De Andrade Salles, P., de Oliveira Sousa, L., Barbosa, L.P., Gomes, V.V.B., de Medeiros, G.R., de Sousa, C. M., e Weller, M. 2013. Analysis of the population of equidae in semiarid region of Paraíba. *Journal of Biotechnology*, 4(3), 269-275.
- De Oliveira Bussiman, F., dos Santos, B.A., Silva, B.D.C.A., Mamani, G.C.M., Grigoletto, L., Pereira, G.L., e de Carvalho Balieiro, J.C. 2020. Genome-wide association study: Understanding the genetic basis of the gait type in Brazilian Mangalarga Marchador horses, a preliminary study. *Livestock Science*, 231, 103867.

- Di Zazzo, E., De Rosa, C., Abbondanza, C., e Moncharmont, B. 2013. PRDM proteins: molecular mechanisms in signal transduction and transcriptional regulation. *Biology*, 2(1), 107-141.
- Dini, P., Bartels, T., Revah, I., Claes, A.N., Stout, T.A., e Daels, P. 2020. A retrospective study on semen quality parameters from four different Dutch horse breeds with different levels of inbreeding. *Theriogenology*. 157, 18-23.
- Dunuwille, W.M., YousefiMashouf, N., Balasuriya, U.B., Pusterla, N., e Bailey, E. 2020. Genome-wide association study for host genetic factors associated with equine herpesvirus type-1 induced myeloencephalopathy. *Equine veterinary journal*, 52(6), 794-798.
- Forejt, J., Jansa, P., e Parvanov, E. 2021. Hybrid sterility genes in mice (*Mus musculus*): a peculiar case of PRDM9 incompatibility. *Trends in Genetics*. <https://doi.org/10.1016/j.tig.2021.06.008>
- Giesecke, K., Hamann, H., Stock, K.F., Woehlke, A., Sieme, H., e Distl, O. 2009. Evaluation of SPATA1-associated markers for stallion fertility. *Animal genetics*, 40(4), 359-365.
- Gonen, S., Battagin, M., Johnston, S.E., Gorjanc, G., e Hickey, J.M. 2017. The potential of shifting recombination hotspots to increase genetic gain in livestock breeding. *Genetics Selection Evolution*, 49(1), 1-12.
- Grilz-Seger, G., Reiter, S., Neuditschko, M., Wallner, B., Rieder, S., Leeb, T., e Druml, T. 2020. A genome-wide association analysis in noriker horses identifies a snp associated with roan coat color. *Journal of equine veterinary science*, 88, 102950.
- Groeneveld, L.F., Atencia, R., Garriga, R.M., e Vigilant, L. 2012. High diversity at PRDM9 in chimpanzees and bonobos. *PloS one*, 7(7), e39064.
- He, H., Zhang, X., Ding, L., Baye, T.M., Kurowski, B.G., e Martin, L.J. 2011. Effect of population stratification analysis on false-positive rates for common and rare variants. *Biomed Medical Central Proceedings*, v.5 (Suppl 9), S116.
- Hecht, L.B., Thompson, P.C., e Rosenthal, B.M. 2020. Assessing the evolutionary persistence of ecological relationships: a review and preview. *Infection, Genetics and Evolution*, 104441.
- Heintzman, P.D., Zazula, G.D., MacPhee, R D., Scott, E., Cahill, J.A. 2017. A new genus of horse from Pleistocene North America. *eLife* 6:e29944.

- Hill, E.W., Gu, J., Eivers, S.S., Fonseca, R.G., McGivney, B.A., Govindarajan, P., ... e MacHugh, D. 2010. A sequence polymorphism in MSTN predicts sprinting ability and racing stamina in thoroughbred horses. *PloS one*, 5(1), e8645.
- Hochwagen, A., e Marais, G.A. 2010. Meiosis: a PRDM9 guide to the hotspots of recombination. *Current Biology*, 20(6), R271-R274.
- Holmström, M., e Back, W. The effects of conformation. In *Equine Locomotion*, 2nd ed.; Back, W., Clayton, H., Eds.; WB Saunders: London, UK, 2013; pp. 229–243.
- Hongyu, K., Sandanielo, V. L. M., e de Oliveira Junior, G. J. 2016. Análise de componentes principais: resumo teórico, aplicação e interpretação. *E&S Engineering and science*, 5(1), 83-90.
- Hossein-Zadeh, N.G. 2016. Analysis of population structure and genetic variability in Iranian buffaloes (*Bubalus bubalis*) using pedigree information. *Animal Production Science*, 56(7), 1130-1135.
- Imai, Y., Baudat, F., Taillepierre, M., Stanzione, M., Toth, A., e de Massy, B. 2017. The PRDM9 KRAB domain is required for meiosis and involved in protein interactions. *Chromosoma*, 126(6), 681-695.
- Islam, FB., Ishishita, S., Uno, Y., Mollah, M.B.R., Srikulnath, K., e Matsuda, Y. 2013. Male hybrid sterility in the mule duck is associated with meiotic arrest in primary spermatocytes. *The Journal of Poultry Science*, 0130011.
- Jenko, J., Wiggans, G.R., Cooper, T.A., Eaglen, S.A.E., Luff, W.D.L., Bichard, M., ... e Woolliams, J.A. 2017. Cow genotyping strategies for genomic selection in a small dairy cattle population. *Journal of dairy science*, 100(1), 439-452.
- Jones, W. 1987. *Genética e Criação de Cavalos*. Editora Roca LTDA, São Paulo, 666p.
- Junqueira, G.S.B., Diaz, I.D.P. S., da Cruz, V.A.R., de Araújo Oliveira, C.A., de Godoi, F.N., de Camargo, G.M.F., e Costa, R.B. 2021. Influence of coat color on genetic parameter estimates in horses. *Journal of Applied Genetics*, 62(2), 297-306.
- Kono, H., Tamura, M., Osada, N., Suzuki, H., Abe, K., Moriwaki, K., ... e Shiroishi, T. 2014. Prdm9 polymorphism unveils mouse evolutionary tracks. *Dna Research*, 21(3), 315-326.
- MacFadden B.J. 2013. Dispersal of Pleistocene Equus (family Equidae) into South America and calibration of GABI 3 based on evidence from Tarija, Bolivia. *PLOS ONE* 8(3):e59277.
- Marchiori, C.M., Pereira, G.L., Maiorano, A.M., Rogatto, G.M., Assoni, A.D., Silva, J.A.I.V., ... e Curi, R.A. 2019. Linkage disequilibrium and population structure

- characterization in the cutting and racing lines of Quarter Horses bred in Brazil. *Livestock Science*, 219, 45-51.
- Meira, C.T., Pereira, I.G., Farah, M.M., Pires, A.V., Garcia, D.A., e Cruz, V.A.R. 2013. Seleção de características morfofuncionais de cavalos da raça Mangalarga Marchador por meio da análise de componentes principais. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 65(6), 1843-1848.
- Mézard, C., Jahns, M.T., e Grelon, M. 2015. Where to cross? New insights into the location of meiotic crossovers. *Trends in Genetics*, 31(7), 393-401.
- Momozawa, Y., Takeuchi, Y., Tozaki, T., Kikusui, T., Hasegawa, T., Raudsepp, T., ... e Mori, Y. 2006. Polymorphism identification, RH mapping, and association analysis with the anxiety trait of the equine serotonin transporter (SLC6A4) gene. *Journal of veterinary medical science*, 68(6), 619-621.
- Morota, G.; Boddhireddy, P.; Vukasinovic, N.; Gianola, D.; e Denise, S. 2014. Kernel-based variance component estimation and whole-genome prediction of precorrected phenotypes and progeny tests for dairy cow health traits. *Frontiers in Genetics*, v.5.
- Mudadu, M.A., Porto-Neto, L.R., Mokry, F.B., Tizioto, P.C., Oliveira, P.S., Tullio, R.R., e Regitano, L.C. 2016. Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *BMC genomics*, 17(1), 1-16.
- Mukaj, A., Piálek, J., Fotopulosova, V., Morgan, A.P., Odenthal-Hesse, L., Parvanov, E.D., e Forejt, J. 2020. PRDM9 intersubspecific interactions in hybrid male sterility of house mouse. *Molecular biology and evolution*, 37(12), 3423-3438.
- Myers, S., Bowden, R., Tumian, A., Bontrop, R.E., Freeman, C., MacFie, T.S., ... e Donnelly, P. 2010. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science*, 327(5967), 876-879.
- Navas, F.J., Jordana, J., León, J.M., Barba, C., e Delgado, J.V. 2017. A model to infer the demographic structure evolution of endangered donkey populations. *Animal*, 11(12), 2129-2138.
- Novotná, A., Svitáková, A., Veselá, Z., Vostrý, L., 2017. Estimation of genetic parameters for linear type traits in the population of sport horses in the Czech Republic. *Livestock Science*, 202, 1-6.
- Oliver, P.L., Goodstadt, L., Bayes, J.J., Birtle, Z., Roach, K.C., Phadnis, N., ... e Ponting, C.P. 2009. Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. *PLoS genetics*, 5(12), e1000753.

- Ortiz-Barrientos, D., Engelstädter, J., e Rieseberg, L.H. 2016. Recombination rate evolution and the origin of species. *Trends in ecology & evolution*, 31(3), 226-236.
- Pereira, G.L., Regatieri, I.C., Ferraz, G.D.C., Neto, A.D.Q., e Curi, R.A. 2015. Perspectivas do uso de marcadores moleculares no melhoramento genético de equinos de corrida da raça Quarto de Milha. *Veterinária e Zootecnia*, 22(3), 347-369.
- Pinto, L.F.B., Almeida, F.Q.D., Quirino, C.R., Azevedo, P.C.N.D., Cabral, G.C., e Corassa, A. 2005. Análise multivariada das medidas morfométricas de potros da raça Mangalarga Marchador: análise de componentes principais. *Revista Brasileira de Zootecnia*, 34(2), 589-599.
- Ponting, C.P. 2011. What are the genomic drivers of the rapid evolution of PRDM9? *Trends in Genetics*, 27(5), 165-171.
- Price, A. L., Zaitlen, N. A., Reich, D. e Patterson, N. 2010. New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics*, v.11, p.459-463.
- Regatieri, I.C., e Mota, M. 2012. Horse Breeding Program: Biochemical Aspects. Melhoramento genético de equinos: Aspectos bioquímicos. *Ars Veterinaria*, 28(4), 227-233.
- Ricard, A., e Duluard, A. 2021. Genomic analysis of gaits and racing performance of the French trotter. *Journal of Animal Breeding and Genetics*, 138(2), 204-222.
- Ricard, A., e Touvais, M. 2007. Genetic parameters of performance traits in horse endurance races. *Livestock Science*, 110(1-2), 118-125.
- Rosa, L.P., Mallicote, M.F., Long, M.T., e Brooks, S.A. 2020. Metabogenomics reveals four candidate regions involved in the pathophysiology of Equine Metabolic Syndrome. *Molecular and Cellular Probes*, 53, 101620.
- Santiago, J.M., de Rezende, A.S.C., Fonseca, M.G., Abrantes, R.G.P., Lage, J., e Lana, Â.M.Q. 2013. Comparação entre as medidas morfométricas do rebanho atual de machos Mangalarga Marchador e dos campeões da raça. *Boletim de Indústria Animal*, 70(1), 46-52.
- Santos, L. M. Morfologia e genética do cavalo Campolina. Belo Horizonte: Universidade Federal de Minas Gerais, 2006. Dissertação (Mestrado em Genética) Universidade Federal de Minas Gerais, 2006.
- Schwartz, J.J., Roach, D.J., Thomas, J.H., e Shendure, J. 2014. Primate evolution of the recombination regulator PRDM9. *Nature communications*, 5(1), 1-7.

- Solar Diaz, I.D.P., Strauss Borges Junqueira, G., Aparecida Rocha Cruz, V., Albano Araújo de Oliveira, C., Nunes de Oliveira, H., Miguel Ferreira de Camargo, G., e Bernal Costa, R. 2021. Principal components for morphometric traits in Campolina horses. *Journal of Animal Breeding and Genetics*, 138(2), 179-187.
- Souza, F.A.C.D., Fernandes, T.J., Cunha, F.O., Ribeiro, R.A., Muniz, F.R., Meirelles, S.L.C., e Moura, R.S. 2019. Características morfométricas de equinos da raça Mangalarga Marchador determinadas por modelos não lineares. *Pesquisa Agropecuária Brasileira*, 54.
- Steiner, C.C., e Ryder, O.A. 2013. Characterization of Prdm9 in equids and sterility in mules. *PloS one*, 8(4), e61746.
- Stock, K. F., Jönsson, L., Ricard, A., e Mark, T. 2016. Genomic applications in horse breeding. *Animal Frontiers*, 6(1), 45-52.
- Todd, E.T., Ho, S.Y., Thomson, P.C., Ang, R.A., Velie, B.D., e Hamilton, N.A. 2018. Founder-specific inbreeding depression affects racing performance in Thoroughbred horses. *Scientific reports*, 8(1), 1-10.
- Vargas, G., Schenkel, F.S., Brito, L.F., Neves, H.H.R., Munari, D.P., Boligon, A., e Carvalheiro, R. 2018. Unravelling biological biotypes for growth, visual score and reproductive traits in Nelore cattle via principal component analysis. *Livestock Science* 207:37-43.
- Warmuth, V., Eriksson, A., Bower, M.A., Barker, G., Barrett, E., Hanks, B.K., ... e Manica, A. 2012. Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proceedings of the National Academy of Sciences*, 109(21), 8202-8206.
- Wolc, A., Stricker, C., Arango, J., Settar, P., Fulton, J.E., O'Sullivan, N.P., ... e Lamont, S.J. 2011. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. *Genetics Selection Evolution*, 43(1), 5.
- Wu, H., Mathioudakis, N., Diagouraga, B., Dong, A., Dombrowski, L., Baudat, F., ... e Kadlec, J. 2013. Molecular basis for the regulation of the H3K4 methyltransferase activity of PRDM9. *Cell reports*, 5(1), 13-20.
- Yang, A Q., Chen, B., Ran, M.L., Yang, G.M., e Zeng, C. 2020. The application of genomic selection in pig cross breeding. *Yi Chuan= Hereditas*, 42(2), 145-152.

Zhou, J., Liu, L., Chen, C.J., Zhang, M., Lu, X., Zhang, Z., e Shi, Y. 2019. Genome-wide association study of milk and reproductive traits in dual-purpose Xinjiang Brown cattle. *BMC genomics*, 20(1), 1-11.

CAPÍTULO I

Caracterização do gene *PRDM9* em equinos

Resumo

A baixa variabilidade genética pode prejudicar diversos aspectos produtivos e reprodutivos em equídeos. Processos de recombinação gênica, como o *crossing-over*, podem aumentar a variabilidade genética ao promover trocas entre partes dos cromossomos homólogos durante a meiose. Entre as proteínas envolvidas no *crossing-over* destacamos o papel da *PRDM9* que determina os pontos de quebra da sequência de DNA (*hotspots*) através de uma estrutura chave denominada domínio de dedos de zinco, do inglês *zinc fingers* (*ZF*). Modificações nas sequências desses domínios *ZF* promovem formação de diferentes *hotspot* de recombinação ao longo do DNA. Assim, esse estudo teve como objetivo descrever a diversidade de arranjos de domínios de zinco e distribuição de alelos do gene *PRDM9* em 26 raças de equinos. Para obtenção do DNA foram utilizadas amostras de pelos de 95 animais. A amplificação do gene de interesse foi realizada utilizando primers específicos e os produtos da amplificação foram qualificados, quantificados e então sequenciados. Para a análise e identificação dos polimorfismos, as sequências obtidas foram visualizadas com os programas CodonCode Aligner. Foram encontrados 14 novos domínios *ZF*, 19 alelos e 19 genótipos. O alelo 1 assim como o seu genótipo homocigoto (1|1) apresentaram maiores frequências (0,889 e 0,775, respectivamente). Os resultados permitem a condução de manejo reprodutivo de animais cujo genótipo do *PRDM9* seja menos frequente em populações com pouca variabilidade genética. Promovem-se novas recombinações e auxilia-se no aumento de sua variabilidade. Adiciona-se a isso o fato de o gene ter papel importante no processo evolutivo, podendo essas informações contribuir para o melhor entendimento desse mecanismo.

Palavras-chave: domínio *ZF*, genótipo, PCR, recombinação, variabilidade genética

Characterization of the *PRDM9* gene in horses

Abstract

The low genetic variability can cause several productive and reproductive prejudices in horses. Genetic recombination processes, such as crossing-over, can increase genetic variability by promoting changes between parts of homologous chromosomes during meiosis. Among the proteins involved in crossing-over, we highlight the role of *PRDM9*, which determines DNA sequence hotspots through a key structure called zinc fingers (*ZF*). Modifications in the sequences of these *ZF* promote the recombination of different hotspots along the DNA. Thus, this study aimed to describe the diversity of *ZF* arrangements and allele distribution of the *PRDM9* gene in 26 horse breeds. To obtain the DNA, hair samples from 95 animals were used. Amplification of the gene of interest was performed using specific primers and the amplification products were qualified, quantified and then sequenced. For the analysis and identification of polymorphisms, the sequences obtained were visualized using CodonCode Aligner programs. A total of 14 new *ZF* domains, 19 alleles and 19 genotypes were found. Allele 1 as well as its homozygous genotype (1|1) showed higher frequencies (0.889 and 0.775, respectively). The results allow better matings, it permits to choose animals whose *PRDM9* genotype is less frequent in populations with little genetic variability. It promotes new recombinations and helps to increase variability. Moreover, the gene plays an important role in the evolutionary process, and this information may contribute to a better understanding of this mechanism.

Keywords: genotype, genetic variability, PCR, recombination, *ZF* domain

1. Introdução

Os equídeos domésticos passaram por intensa seleção fenotípica o que proporcionou o surgimento de diversas raças com diferentes aptidões. A substituição de equinos das funções de transporte e tração pode ter causado diminuição de tamanho populacional em algumas raças. O reduzido número de representantes de um determinado grupo genético pode levar a acasalamentos endogâmicos que geram consequências para sua produção e conservação.

Algumas raças apresentam taxas de endogamia e variabilidade genética preocupantes (Gonçalves et al., 2011; Santana et al., 2017; Marchiori, 2019; Assis et al., 2018; Alves et al., 2021). É imprescindível a atenção na manutenção da variabilidade genética existente, pois ela está diretamente relacionada a respostas mais eficientes a mudanças ambientais, além de melhores índices produtivos e reprodutivos (Ortiz-Barrientos et al., 2016; Hecht et al., 2020).

A variabilidade genética se dá pela combinação e quantidade de alelos intra e inter-loci. Processos de recombinação gênica, como o *crossing-over*, aumentam a variabilidade genética, pois possibilitam a troca de partes dos cromossomos homólogos durante a meiose. Os descendentes podem receber cromossomos com alelos de genes diferentes em combinação distinta da que ocorre nos pais. Esse processo de recombinação ocorre em determinadas regiões do DNA, denominadas *hotspots* de recombinação (Mézard et al., 2015; Durbano, 2015).

Em mamíferos a proteína PRDM9 está envolvida na determinação dos *hotspots* (Myers et al., 2009; Baudat et al., 2010). Essa proteína é composta por três diferentes domínios: o KRAB, o SET e os domínios de zinco (*ZF*), sendo o último que se liga ao DNA e determina o *hotspot* de recombinação (Hohenauer e Moore 2012; Di Zazzo et al., 2013; Wu et al. 2013). Os domínios *ZF* são estruturas chaves da proteína PRDM9, sendo cada domínio composto por 28 aminoácidos. Cada *ZF* sequencial liga-se a um trinucleotídeo na molécula de DNA alvo e, portanto, influencia a localização do *hotspot* (Neale, 2010). Assim, uma única alteração de aminoácidos nos domínios *ZF* pode formar uma nova variante de PRDM9 com novas especificações de ligação ao DNA, mostrando que esse gene e suas variações alélicas podem auxiliar no aumento/manutenção da variabilidade genética por promover recombinações em regiões cromossômicas distintas.

Ahlawat et al. (2016) e Ahlawat et al. (2017) caracterizaram os alelos do *PRDM9* em ruminantes domésticos. Mostraram que há uma vasta diversidade nos arranjos *ZF*, resultando em 36 alelos diferentes em caprinos e 26 alelos em ovinos e 41 alelos em

bovídeos (bovinos, bubalinos, mithuns e iaques). Os autores ressaltam que as modificações não ocorrem apenas entre espécies, mas também dentro das espécies. Os estudos demonstraram quais alelos são mais comuns em cada espécie e raça. As informações geradas podem ser aliadas a outras estratégias de melhoramento e conservação animal de diferentes grupos genéticos. Raças com problemas de perda de variabilidade, podem usar reprodutores cujos genótipos do *PRDM9* são compostos por alelos menos frequentes na população e assim promover novas recombinações gênicas que auxiliam no aumento da variabilidade.

Em equinos, o gene *PRDM9* foi descrito por Steiner e Ryder (2013). O gene em equinos possui oito domínios *ZF* (DAKKDIKB). Todavia, não há relatos sobre a caracterização da variação alélica desse gene nessas espécies. Esse conhecimento pode gerar uma ferramenta auxiliar interessante na definição de estratégias de acasalamento de populações pequenas e/ou endogâmicas.

Assim, o presente estudo teve como objetivo verificar a diversidade de arranjos de domínios *ZF* e caracterizar os alelos do gene *PRDM9* em vinte e seis raças de equinos, a fim de obter informações auxiliares para produção e conservação das espécies.

2. Material e Métodos

2.1. Amostragem e Extração de DNA

A coleta de material biológico das amostras utilizadas foi aprovada pela Comissão de Ética no Uso de Animais da Escola de Medicina Veterinária e Zootecnia da Universidade Federal da Bahia sob numeração: 08/2019 CEUA.

Foram coletadas amostras de pelos com bulbos capilares preservados da cauda de 95 equinos, de 26 raças (Appaloosa (1), Árabe (10), Brasileiro de Hipismo (1), Campolina (8), Crioulo (4), Gypsy (2), Hanoveriano (1), Holsteiner (2), KWPN (4), Lavradeiro (8), Lusitano (5), Mangalarga Machador (6), Mangalarga (7), Mini Horse (5), Nordestino (3), Paint Horse (1), Passo Peruano (1), Percheron (2), Piquira (3), Pônei Brasileiro (6), Puro Sangue Inglês (6), Quarto de Milha (5), Sela Belga (1), Sela Francesa (1), Sela Sueca (1), e Zongersheide (1)).

2.2. Extração do DNA e PCR

O DNA foi extraído a partir do bulbo capilar utilizando o kit de extração de DNA NucleoSpin® Tissue, de acordo com recomendações do fabricante.

Os primers utilizados nas reações de amplificação do gene *PRDM9* (F: GAGGCTTCAATGACAGGGCAAGTCTTAT; R: CGTTGGAGCTGGAGTATGGAGT) foram descritos por Steiner e Ryder (2013).

Para as reações em cadeia da polimerase (PCR) foram realizadas em um volume de 20 µl usando eppendorf graduado para termociclador. Cada reação incluía entre 100-120ng de DNA, 10,45 µl de Taq Mix (com MgCl₂ 1,5 mM, trifosfatos de desoxinucleósidos 10 mM), e 0,4 µl IM de cada primer. As condições do ciclo de PCR foram 95°C por 6 min, seguido por 34 ciclos de desnaturação a 94°C por 1 min, anelamento a 57,5°C por 1 min e ao final extensão a 72°C por 10 min.

Os produtos de PCR foram eletroforados juntamente com marcadores de peso molecular de DNA em gel de agarose 1,2% e visualizado por coloração com Siber Gold com incidência de luz UV. Os fragmentos amplificados foram purificados com solução de polietilenoglicol (PEG) a 20% (Sambrook et al. 1989) e sequenciados usando um kit de sequenciamento BigDye v3.1 (Applied Biosystems, Foster City, CA, EUA) e sequenciador de DNA ABI 3130 (Applied Biosystems) de acordo com o manual do fabricante.

2.3. Genotipagem

O alinhamento de bases foi realizado para todas as sequências de *PRDM9* de equinos tendo como base a sequência gerada por Steiner e Ryder (2013). Para a análise e identificação dos polimorfismos, as sequências obtidas foram analisadas e visualizadas com o programa CodonCode Aligner disponível no site (<http://www.codoncode.com/aligner/download.htm>). As sequências foram aparadas depois de marcar o início e fim da região de codificação os domínios *ZF*. Isso foi feito para remover ambiguidades sem alterar o quadro de leitura, de modo que incluía apenas o domínio *ZF* do *PRDM9* para análise posterior.

3. Resultados

Após realizar o controle de qualidade e o alinhamento com a sequência referência foram identificados 14 novos domínios *ZF* ainda não relatados na literatura para equinos nesse gene. Tomando como referência os domínios descritos por Steiner e Ryder (2013), os novos domínios são caracterizados por variações nos aminoácidos presentes em determinadas posições (9, 10, 13 e 16) da sequência, como descritos na Figura 1.

Figura 1 - Novos domínios *ZF* identificados no gene *PRDM9* em equinos

		9	10	13	16	
A	C R E C E Q G F	T	R	K S S	L I A	H Q R T H T G E K P Y V
R	S	E
S	S	V	. . .	R
T	S	V	. . .	T
U	S	E	. . N
V	S	E
X	Q	. . N
Y	Q
Z	H
A'	S	E	. . .	R
B'	H	. . T
C' R
D'	A	Q
F'	S	Q T
G'	S	V	. . R

Cada domínio *ZF* é codificado por letra em negrito e a variação de aminoácidos é indicada nas posições 9, 10, 13 e 16.

Os alelos do gene *PRDM9* são caracterizadas pela combinação dos oito *ZF* presentes na região terminal do gene. Para os equinos foram identificados 19 alelos, descritos de 1 a 19. O alelo de maior frequência é o alelo 1 (0,889). Todos os outros alelos possuem frequências muito inferiores e muitos com uma única ocorrência no espaço amostral (Tabela 1).

Tabela 1 – Descrição e frequência dos alelos identificados para o gene *PRDM9* em equinos

(*Continua*)

Alelos	Domínio <i>ZF</i>	N*	Frequência
1	DAKKDIKB	169	0,889
2	DAKIVIKB	1	0,005
3	DAZZA'XDU	1	0,005
4	DAKIVIMB	1	0,005
5	DTZZSZB'U	2	0,012
6	DAKIIIMC'	1	0,005
7	DD'ISIIMB	1	0,005
8	DAKIIIMB	1	0,005
9	RTIIIMB	1	0,005
10	DF'DIIKA	1	0,005
11	DRDDNYDV	2	0,012

Tabela 1 – Descrição e frequência dos alelos identificados para o gene *PRDM9* em equinos

(Continuação)

Alelos	Domínio ZF	N	Frequência
12	DTDDTYDV	2	0,012
13	DAKIIIMA	1	0,005
14	DAIIVIMD	1	0,005
15	DIDYTDDA'	1	0,005
16	DYDDIIG'D	1	0,005
17	DYDDTYIC	1	0,005
18	DYDDTYDC	1	0,005
19	DRDDTYDV	1	0,005
Total			1,000

N, número de alelos observados

A Tabela 2 mostra a distribuição dos genótipos para cada raça de equinos. Observe que o genótipo de maior frequência é o homozigoto 1|1.

Tabela 2 - Genótipos *PRDM9* observados em diferentes raças de equinos
(*Continua*)

Raças	N	1 1	1 2	1 3	1 4	1 5	1 6	1 7	1 8	1 9	1 10	1 11	1 12	1 13	1 14	1 15	1 16	1 17	1 18	1 19
Árabe	10	6	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Campolina	8	7	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-
Lavradeiro	8	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Puro Sangue Inglês	6	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mangalarga Pônei	7	6	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
Brasileiro	6	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mangalarga Marchador	6	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Lusitano	5	3	-	-	-	-	-	-	-	-	-	1	1	-	-	-	-	-	-	-
Quarto de Milha	5	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1	1
Mini Horse	5	1	-	-	-	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-
Crioulo	4	3	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-
KWPN	4	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Piquira	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Nordestino	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Gypsy	2	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Holsteiner	2	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-
Percheron	2	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Brasileiro de Hipismo	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Appaloosa	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Hanoveriano	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabela 2 - Genótipos *PRDM9* observados em diferentes raças de equinos
(*Continuação*)

Raças	N	1 1	1 2	1 3	1 4	1 5	1 6	1 7	1 8	1 9	1 10	1 11	1 12	1 13	1 14	1 15	1 16	1 17	1 18	1 19
Paint Horse	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Passo Peruano	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sela Belga	1	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
Sela Francesa	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sela Sueca	1	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
Zongersheide	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

N, número de amostras

Todos os equinos estudados apresentam o alelo 1 em homozigose (1|1) ou em heterozigose com os demais alelos. Deste modo, o genótipo 1|1 apresentou frequência (0,775), representando todos os homozigotos. Enquanto aproximadamente 22% dos indivíduos apresentaram genótipo heterozigoto com outros alelos (Tabela 3).

Tabela 3 - Frequências dos genótipos observados para o gene *PRDM9* em equinos

<i>PRDM9</i>	
Genótipos	Frequência
1 1	0,775
1 2	0,011
1 3	0,011
1 4	0,011
1 5	0,020
1 6	0,011
1 7	0,011
1 8	0,011
1 9	0,011
1 10	0,011
1 11	0,020
1 12	0,020
1 13	0,011
1 14	0,011
1 15	0,011
1 16	0,011
1 17	0,011
1 18	0,011
1 19	0,011
Total	1,000

Ao final dos oito ZF do gene *PRDM9*, há um seguimento que se assemelha ao ZF, mas é composto por 26 aminoácidos ao invés de 28. Cerca de 3% dos animais estudados apresentam variação específica na posição 13 onde ocorre a substituição de uma asparagina (N) por uma histidina (H) (Figura 2).

Figura 2 – Seguimento identificado após o domínio ZF do gene *PRDM9* em equinos

A	C R E C E Q G F S V K S	¹³ N	L I R H Q R T H T G E K L
B	H

4. Discussão

O presente estudo analisou amostras de 95 equinos. Entretanto, tentou-se amplificar outras 175 amostras sem êxito. Isso pode ter acontecido por diversos motivos: sequências diferentes na região de pareamento dos primers, repetição de domínios/sequências ao longo do fragmento amplificado ou até mesmo variações no número de domínios *ZF* no gene *PRDM9* como ocorre em outras espécies (Ahlawat et al 2016 e Ahlawat et al 2017). Sugerem-se a utilização de outras técnicas em estudos futuros para garantir o sucesso de amplificação da região genômica.

A estrutura chave da proteína *PRDM9* é a sequência de domínios *ZF*, pois é ela que ativa os pontos de quebra do material genético. Steiner e Ryder (2013) analisaram o gene *PRDM9* em espécies do gênero *Equus* e descreveram 17 domínios *ZF*. Um domínio *ZF* é uma estrutura composta por 28 aminoácidos. O presente estudo encontrou mais 14 domínios de *ZF* (Figura 1). No gene *PRDM9*, os alelos são dados pela quantidade e sequência dos domínios de *ZF*. Em equinos, foi descrita uma sequência de oito domínios *ZF* (DKKDIKB) (Steiner e Ryder, 2013), sendo esse o único alelo descrito para a espécie até então.

Encontraram-se aqui domínios *ZF* não relatos para o gene *PRDM9* e combinações de *ZF* diferentes. A combinação dos domínios de *ZF* geraram 19 alelos e 19 genótipos (Tabelas 1, 2 e 3). Isso indica clara potencialidade de uso de reprodutores com alelos e/ou genótipos menos frequentes em populações pequenas e/ou endogâmicas como ferramenta adicional para melhoria da variabilidade genética.

Todavia, a adoção dessa estratégia pode ser um pouco dificultada pela alta frequência do alelo 1 e do genótipo homozigoto formado por ele (Tabela 3), mas pode ser feito. Por exemplo, há estudos que indicam populações de equinos da raça Árabe sob depressão endogâmica (Sierszchulski et al., 2005). Nossos resultados mostram existência de outros genótipos do *PRDM9* que não os mais frequentes na raça (Tabela 2). Isso indica a possibilidade de reprodução desses indivíduos de genótipos diferentes, aliada a outras estratégias, como facilitador do aumento da variabilidade genética.

Marchiori et al. (2019), ao estudarem populações da raça Quarto de Milha destacaram a necessidade de monitoramento da variabilidade genética e endogamia na raça, devido ao tamanho efetivo populacional ser pequeno. Em nosso estudo, observamos que a raça apresentou diversos genótipos para o gene *PRDM9* (Tabela 2). Nesse caso, também é possível sugerir a utilização de reprodutores com genótipos diferentes em acasalamentos afim de aumentar/manter a variabilidade genética.

Ahlawat et al. (2016) e Ahlawat et al. (2017) reforçam a ideia de participação do gene *PRDM9* no processo de evolução, especiação e infertilidade de híbridos. Tendo as espécies de equídeos grande importância nos temas citados (Steiner e Ryder, 2013), os presentes resultados apresentam informações relevantes para estudos comparativos e de geração de conhecimento biológico básico e evolutivo, além das aplicações na produção animal.

5. Conclusão

A partir da análise de vinte e seis raças de equinos foi possível caracterizar a estrutura terminal do gene *PRDM9*, denominada domínios *ZF*. Foi possível ainda identificar novos domínios *ZF*, bem como novos alelos e genótipos e suas distribuições em equinos. As variações do gene podem ter potencialidade de contribuir para o aumento da variabilidade genética de populações e agregar informações junto a estudos de evolução, especiação e hibridação.

Referências

- Ahlawat, S., De, S., Sharma, P., Sharma, R., Arora, R., Kataria, R.S., e Singh, R.K. 2017. Evolutionary dynamics of meiotic recombination *hotspots* regulator *PRDM9* in bovids. *Molecular Genetics e Genomics*, 292(1), 117-131.
- Ahlawat, S., Sharma, P., Sharma, R., Arora, R., Verma, N.K., Brahma, B., e De, S. 2016. Evidence of positive selection and concerted evolution in the rapidly evolving *PRDM9* zinc finger domain in goats and sheep. *Animal genetics*, 47(6), 740-751.
- Alves, J.S., da Silva Anjos, M., Bastos, M.S., de Oliveira, L.S.M., Oliveira, I.P.P., Pinto, L.F.B., ... e de Camargo, G.M.F. 2021. Variability analyses of the maternal lineage of horses and donkeys. *Gene*, 769, 145231.
- Assis, C.M., Baena, M.M., De Fátima Bretanha, R., Rocha, S.L.C.M. 2018. Endogamia e Relação Genética em Equinos da Raça Mangalarga Marchador No Brasil. 28º Congresso Brasileiro de Zootecnia, Goiânia – GO. Anais... 2018.
- Baker, C.L., Petkova, P., Walker, M., Flachs, P., Mihola, O., Trachtulec, Z., Paigen, K. 2015. Multimer formation explains allelic suppression of *PRDM9* recombination *hotspots*. *PLoS genetics*, 11(9), e1005512.

- Baudat F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., De Massy, B. 2010. *PRDM9* is a major determinant of meiotic recombination *hotspots* in humans and mice. *Science* 327, 836–40.
- Carneiro, G.F., Lucena, J.E.C., e De Oliveira Barros, L. 2018. The current situation and trend of the donkey industry in South America. *Journal of equine veterinary science*, 65, 106-110.
- Cothran, E.G., MacCluer, J.W., Weitkamp, L.R., Guttormsen, S.A. 1986. Genetic variability, inbreeding, and reproductive performance in standardbred horses. *Zoo Biology*, 5(2), 191-201.
- Di Zazzo, E., De Rosa, C., Abbondanza, C., Moncharmont, B. 2013. PRDM proteins: molecular mechanisms in signal transduction and transcriptional regulation. *Biology*, 2(1), 107-141.
- Durbano, J.P.D.M. 2015. As pesquisas de Barbara McClintock sobre o crossing-over em Zeamays: 1925-1932. *Filosofia e História da Biologia*, 10(1), 49-65.
- Falconer, D.S. 1989. *Introduction to quantitative genetics* 3.ed. New York: Logman Scientific and Technical. 438p.
- Gonçalves, R.W., Costa, M.D., Da Costa, M.R., Silva, E.D.S.P., Ribeiro Filha, A.D.M. 2011. Efeito da endogamia sobre características reprodutivas em um rebanho da raça Mangalarga Marchador. *Revista Brasileira de Saúde e Produção Animal*, 12(3).
- Gonen, S., Battagin, M., Johnston, S.E., Gorjanc, G., Hickey, J.M. 2017. The potential of shifting recombination *hotspots* to increase genetic gain in livestock breeding. *Genetics Selection Evolution*, 49(1), 55.
- Hecht, L.B., Thompson, P.C., e Rosenthal, B.M. 2020. Assessing the evolutionary persistence of ecological relationships: a review and preview. *Infection, Genetics and Evolution*, 104441.
- Hohenauer, T., e Moore, A.W. 2012. The PRDM family: expanding roles in stem cells and development. *Development*, 139(13), 2267-2282.
- Marchiori, C.M., Pereira, G.L., Maiorano, A.M., Rogatto, G.M., Assoni, A.D., Silva, J.A.I.V., ... e Curi, R.A. 2019. Linkage disequilibrium and population structure characterization in the cutting and racing lines of Quarter Horses bred in Brazil. *Livestock Science*, 219, 45-51.
- Mézard, C., Jahns, M.T., e Grelon, M. 2015. Where to cross? New insights into the location of meiotic crossovers. *Trends in Genetics*, 31(7), 393-401.

- Myers, S., Bowden, R., Tumian, A., Bontrop, R.E., Freeman, C., MacFie, T.S., Donnelly, P. 2009. Drive against *hotspot* motifs in primates implicates the *PRDM9* gene in meiotic recombination. *Science*, e1166202.
- Neale, M.J. 2010. *PRDM9* points the zinc finger at meiotic recombination *hotspots*. *Genome biology*, 11 (2), 104.
- Ortiz-Barrientos, D., Engelstädter, J., e Rieseberg, L.H. 2016. Recombination rate evolution and the origin of species. *Trends in ecology & evolution*, 31(3), 226-236.
- Queiroz, S.A.D., Albuquerque, L.G.D., Lanzoni, N.A. 2000. Efeito da endogamia sobre características de crescimento de bovinos da raça Gir no Brasil. *Revista Brasileira de Zootecnia*, 29, 1014-1019.
- Santana, B.F., Mamani, G., Junior, G., Laitano, L., Molento, M., Ventura, R., Ferraz, J.B. 2017. Caracterização de segmentos de homozigose em uma população de cavalos puro sangue inglês. *Ciência and Tecnologia Fatec-JB*, 9.
- Sierszchulski, J., Helak, M., Wolc, A., Szwaczkowski, T., e Schlote, W. 2005. Inbreeding rate and its effect on three body conformation traits in Arab mares. *Animal Science Papers and Reports*, 23(1), 51-59.
- Steiner, C.C., e Ryder, O.A. 2013. Characterization of *PRDM9* in equids and sterility in mules. *PloSone*, 8(4), e61746.
- Wu, H., Mathioudakis, N., Diagouraga, B., Dong, A., Dombrovski, L., Baudat, F., Kadlec, J. 2013. Molecular basis for the regulation of the H3K4 methyltransferase activity of *PRDM9*. *Cell reports*, 5(1), 13-20.

CAPÍTULO II

Associação genômica para componentes principais de medidas morfométricas em equinos: identificação de genes relacionados a crescimento ósseo

Resumo

A mensuração de características corporais de equinos é importante para a determinação do padrão racial, e constitui um dos principais critérios de seleção na espécie. A elaboração de um índice (HCP1) que incorpore componentes principais ponderados pelos valores genéticos aditivos permite explorar as relações mais relevantes usando um número reduzido de variáveis que explicam a maior variação de dados. A análise de associação genômica ampla (GWAS) relacionada ao HCP1 é uma abordagem relativamente nova e possibilita a descoberta de regiões que podem estar relacionadas a um conjunto de características. Esse estudo teve como objetivo realizar GWAS utilizando como variável explanatória o HCP1 para 15 medidas lineares, com o intuito de identificar regiões genômicas associadas e esclarecer os mecanismos biológicos ligados a esse componente em equinos da raça Campolina. Para GWAS, a metodologia “*single-step – GBLUP, wssGBLUP*” foi aplicada sobre o HCP1. Foram identificadas 8 principais janelas que capturaram a maior proporção (95,89%) da variância genética aditiva explicada pelos marcadores. Nas principais janelas foram identificados genes (*SPRY2, COL9A2, MIR30C, HEYL, BMP8B, LTBP1, FAM98A* e *CRIM1*) relacionados ao desenvolvimento ósseo, constituindo potenciais candidatos posicionais para as medidas lineares avaliadas nesse estudo. Os resultados encontrados são fortes indícios da relação entre o HCP1 e características de crescimento e conformação corporal em equinos da raça Campolina, demonstrando assim ser uma eficiente alternativa para reduzir as 15 características usualmente mensuradas. Os genes encontrados têm papel biológico condizente com a expressão da característica. Os resultados permitem estudos futuros de mapeamento-fino para desenvolvimento de marcadores para seleção.

Palavras-chave: conformação, Campolina, GWAS, medidas lineares, morfometria, SNP

Genomic association for principal components of morphometric measurements in horses: identification of genes related to bone growth

Abstract

The measurement of body characteristics of horses is important for the determination of the breed pattern, and constitutes one of the main selection criteria in the species. The development of an index (HCP1) that incorporates principal components weighted by additive breeding values allows exploring the most relevant relationships using a reduced number of variables that explain the greater variation in data. Genomic-wide association analysis (GWAS) related to HCP1 is a relatively new approach and makes it possible to discover regions that may be related to a set of traits. This study aimed to perform GWAS using HCP1 as an explanatory variable for 15 linear measurements, in order to identify associated genomic regions and clarify the biological mechanisms linked to this component in Campolina horses. For GWAS, the “single-step – GBLUP, wssGBLUP” methodology was applied on the HCP1. Eight main windows were identified that captured the highest proportion (95.89%) of the additive genetic variance explained by the markers. Genes (*SPRY2*, *COL9A2*, *MIR30C*, *HEYL*, *BMP8B*, *LTBP1*, *FAM98A* and *CRIMI*) related to bone development were identified in the main windows, constituting potential positional candidates for the linear measurements evaluated in this study. The results found are strong indications of the relationship between HCP1 and growth characteristics and body conformation in Campolina horses, thus demonstrating to be an efficient alternative to reduce the 15 characteristics usually measured. The genes found have a biological role consistent with the expression of the trait. The results allow future fine-mapping studies to develop markers for selection.

Keywords: conformation, Campolina, GWAS, linear measurements, morphometry, SNP

1. Introdução

A raça de equinos Campolina é originária do estado de Minas Gerais e atualmente está difundida por todo território nacional. Na formação da raça, a seleção fenotípica, influenciada principalmente por fatores mercadológicos, enfatizaram principalmente características que elevaram o porte do animal (com enfoque para aumento da altura da cernelha). Assim, ficaram estabelecidas as principais características da raça Campolina, que os classificam como equinos de grande porte, ágeis e que apresentam andamento cômodo, caracterizado pela marcha, e ideais para trabalho, sela e lazer (ABCCCampolina, 2021).

As características físicas dos equinos são fatores fundamentais para a determinação do padrão racial, e é a partir de mensurações realizadas em pontos específicos do corpo dos animais que as associações de criadores avaliam a caracterização racial dos equinos para finalidade de registro. Além disso, a conformação é um importante critério de seleção na espécie equina, uma vez que, a morfometria está diretamente relacionada não somente a aparência, mas também a dinâmica de locomoção do animal (Holmström e Back, 2013).

Atualmente mais de 15 características corporais podem ser medidas em equinos da raça Campolina, e são comumente utilizadas para analisar a morfometria dos animais de forma padronizada (Junqueira et al., 2021; De Oliveira Bussiman et al., 2021; De Oliveira Bussiman et al., 2022). Os marcos de referência para mensurações são determinados por locais de fácil observação, geralmente próximos a extremidades ósseas. Essa técnica, relativamente simples, possibilita a obtenção de um grande volume de informações que permitem entre outras coisas, estimar os parâmetros genéticos que podem auxiliar o melhor direcionamento para os criadores na seleção dos animais, promovendo efetivo progresso genético da raça (Lucena et al., 2015; Costa et al., 2016; Souza et al., 2019).

Estudos de associação genômica ampla (GWAS) com características morfométricas em equinos, permitiram a descoberta de informações importantes relacionadas ao desenvolvimento corporal dos animais. Makvandi-Nejad et al. (2012), demonstram que variações genéticas em apenas quatro loci pode explicar grande parte da variação do tamanho do cavalo, mostrando que assim como outros mamíferos domésticos, os equinos carregam um pequeno número de loci relacionados a tamanho, com alelos de grande efeito.

Neste sentido, a altura de cernelha é uma característica diretamente relacionada ao tamanho em equinos, e conseqüentemente ao porte do animal. Em equinos German Warmbloods, foi identificado um único QTL principal no ECA3 para a altura da cernelha

explicando aproximadamente 18% da variação fenotípica (Tetens et al., 2013). Novos estudos vêm sendo desenvolvidos, e já foram identificados novos QTLs potencialmente envolvidos na conformação corporal de cavalos (Gmel et al., 2019).

Segundo a literatura, características relacionadas à conformação estrutural do corpo dos equídeos sofrem menos influência do ambiente, apresentam alta correlação genética, baixa amplitude de variação individual e valores de herdabilidade com magnitude de moderada a alta (Duensing et al., 2014; Bramante et al., 2016; Novotná et al., 2017; Bussiman et al., 2018), portanto a utilização de todas essas informações em estudos, apesar de serem importantes, pode acabar sendo redundante.

Como alternativa, a análise de componentes principais (PCA) vem sendo amplamente utilizada para lidar com grandes volumes de informações, nos estudos de genética quantitativa, pois reduz a dimensão das variáveis originais correlacionadas transformando-as em um conjunto menor de variáveis não correlacionadas e independentes, explicando o máximo da variabilidade com perda mínima de informações (Hair et al., 2009; Boligon et al., 2016). Muitos estudos já demonstraram a eficiência da utilização da PCA para descartar variáveis que apresentaram baixa variabilidade ou foram redundantes, possibilitando análises mais acuradas com relação a conformação corporal de equinos (Pinto et al., 2005; Meira et al., 2013; Rezende et al., 2016).

Solar Diaz et al. (2021), em uma análise genética tradicional (sem genótipos), sintetizaram 15 medidas lineares de equinos da raça Campolina através da análise de componentes principais, resultando em um índice ponderado pelos valores genéticos dos animais com perda mínima de informações. Os autores ressaltaram que este método possibilita que a seleção seja feita com bases mais objetivas e abrangentes.

A realização de GWAS associada a um índice composto por PCA ainda é algo novo em estudos com equinos, e sua utilização possibilita a descoberta de regiões que podem estar relacionadas a um conjunto de características. Estudos desse tipo propiciam melhor conhecimento sobre a expressão biológica de características quando avaliadas pela proporção conjunta da variância genética aditiva (Mudadu et al., 2016; Vargas et al., 2018; Zhou et al., 2019).

Assim, o presente estudo teve como objetivos realizar GWAS utilizando como variável explanatória o índice (HCP1) composto pelo primeiro componente principal ponderado pelos valores genéticos aditivos (EBVs) dos animais para as características consideradas, com o intuito de identificar regiões genômicas associadas ao HCP1 e

esclarecer os mecanismos biológicos ligados a esse componente em equinos da raça Campolina.

2. Material e Métodos

2.1 Estimação dos parâmetros genéticos

Os componentes de variâncias e, conseqüentemente a matriz que compõe a variância e covariância entre as características, necessária para a obtenção dos componentes principais, foram obtidos por Junqueira et al. (2021), a partir de informações de 18.876 animais da raça Campolina provenientes da Associação Brasileira dos Criadores do Cavalo Campolina (ABCCCampolina). Adicionalmente, os componentes principais e a composição dos índices, foram obtidos por Sola Diaz et al. (2021) para o mesmo banco de dados, sem a utilização de informações genômicas (genótipos dos animais).

No total, 15 medidas corporais foram avaliadas: altura da cernelha (AC), altura da garupa (AG), comprimento da garupa (CG), comprimento da cabeça (CCB), comprimento da espadua (CE), comprimento do corpo (CC), largura da cabeça (LC), largura da anca (LA), perímetro do tórax (PT), altura do dorso (AD), altura dos costados (ACO), comprimento do pescoço (CP), largura do peito (LP), comprimento dorso-lombo (CDL) e perímetro da canela (PC). A estatística descritiva das medidas estudadas está descrita na Tabela 1.

Tabela 1- Estatística descritiva de 15 medidas lineares em equinos da raça Campolina

Variável	N	Média	Desvio padrão	Mínimo	Máximo
AC	18.876	1,56	0,054	1,45	1,71
AG	18.876	1,55	0,051	1,43	1,71
CG	18.876	0,52	0,040	0,33	0,66
CCB	18.876	0,61	0,028	0,50	0,70
CE	18.876	0,56	0,035	0,45	0,70
CC	18.876	1,58	0,062	1,42	1,77
LC	18.876	0,21	0,011	0,17	0,25
LA	18.876	0,53	0,028	0,41	0,61
PT	18.876	1,83	0,075	1,60	2,10
AD	18.876	1,48	0,051	1,35	1,63
ACO	18.876	0,65	0,029	0,53	0,75
CP	18.876	0,64	0,036	0,50	0,80
LP	18.876	0,39	0,030	0,30	0,50
CDL	18.876	0,59	0,056	0,44	0,78
PC	18.876	0,19	0,011	0,16	0,23

(AC) altura da cernelha; (AG) altura da garupa; (CG) comprimento da garupa; (CCB) comprimento de cabeça; (CE) comprimento da espadua; (CC) comprimento do corpo; (LC) largura da cabeça; (LA) largura da anca; (PT) perímetro do tórax; (AD) altura do dorso; (ACO) altura dos costados; (CP) comprimento do pescoço; (LP) largura do peito; (CDL) comprimento dorso-lombo; (PC) perímetro da canela; (N) número de observações. Fonte: Junqueira et al. (2021)

A obtenção dos parâmetros genéticos, foram obtidos por Junqueira et al. (2021) através de inferência Bayesiana, em um modelo multicaracterística por amostrador de Gibbs utilizando os softwares desenvolvidos por Misztal et al. (2020). A descrição completa das análises, assim como das pressuposições do modelo, para a estimação dos parâmetros genéticos encontra-se em Junqueira et al. (2021).

2.2 Componentes Principais

O índice obtido (HCP1) foi composto pelo primeiro componente principal (PC1) ponderado pelos valores genéticos aditivos dos animais para todas as características. A estatística, modelo e algoritmo estão detalhados em Solar Diaz et al. (2021).

Para este estudo, utilizou-se o índice proposto por Solar Diaz et al. (2021), com o mesmo banco de dados, como variável dependente sobre os genótipos dos animais. O HCP1 foi obtido calculando os pesos (coeficientes de ponderação padronizados para o valor de cada característica) do PC1 e dividindo o autovetor pela raiz quadrada dos valores próprios dos valores genéticos padronizados (Solar Diaz et al., 2021) obtendo a seguinte equação:

$$\begin{aligned} \text{HCP1} = & -0,466\text{BVWH} - 0,459\text{BVCH} - 0,157\text{BVCL} - 0,164\text{BVHL} - 0,155\text{BVSL} \\ & - 0,451\text{BVBL} - 0,033\text{BVHW} - 0,144\text{BVHiW} - 0,432\text{BVTG} - 0,452\text{BVBH} - \\ & 0,182\text{BVBaH} - 0,198\text{BVNL} - 0,110\text{BVCW} - 0,153\text{BVBL} - 0,03\text{BVCG} \end{aligned}$$

Onde, BV indica o valor genético da característica. O HCP1 foi composto apenas pelo primeiro componente principal por ser este o que explica a maior proporção da variância genética aditiva (97,8%) entre todas as características, obtidas através da matriz de (co) variâncias descrita por Solar Diaz et al. (2021).

2.3 *Dados genotípicos e controle de qualidade*

Um total de 48 equinos da raça Campolina foram genotipados usando o chip Illumina® Equine (65.000 SNPs; Illumina, Inc., San Diego, CA, EUA). O controle de qualidade de genotipagem (QC) filtrou marcadores localizados em regiões autossômicas e não autossômicas mapeadas para a mesma posição, desviado do teste de equilíbrio de Hardy-Weinberg (HWE) ($P < 10^{-5}$), com pontuação GenCall (GC) menor que 0,15, SNP call rate inferior a 0,95 e minor allele frequency (MAF) inferior a 0,02. O número restante de SNPs e animais após o QC foram 41.860 e 42, respectivamente. Para esse procedimento foi utilizado o programa PREGSf90 (Misztal, 2020).

2.4 *Análise de associação genômica ampla (GWAS)*

O método estatístico usado para estimar os efeitos do índice (HCP1) sobre as características, foi o de passo único “single step” ponderado (wssGBLUP) proposto por Wang et al. (2012), como descrito no modelo:

$$y^* = \mu + Z_a a + e$$

Onde y^* é o vetor da variável explicatória ou dependente (HCP1 neste estudo); μ é um vetor da média geral; Z_a é uma matriz de incidência que relaciona animais a variável explicatória; a é o vetor de efeitos genéticos aditivos diretos e e é o vetor dos resíduos aleatórios. Foi assumido que $a \sim N(0, H\sigma_a^2)$ e $e \sim N(0, R\sigma_e^2)$, onde H é a matriz de parentesco genômica, σ_a^2 é variância genética aditiva, R é uma matriz diagonal, cujos elementos são relacionados aos resíduos relacionados a y , e σ_e^2 é a variância do resíduo.

O erro padrão relacionado ao índice HCP1 foi obtido através da seguinte equação $SE(\text{HCP1}) = \sqrt{e_{i1}^2 * SE_{j1}^2 + e_{i2}^2 * SE_{j2}^2 + \dots + e_{i15}^2 * SE_{j15}^2}$, onde $SE(\text{HCP1})$ é o erro padrão aproximado do HCP1 para o PC1 do j^{th} animal, e_{i1}^2 é o quadrado do coeficiente

do vetor próprio para o i^{th} PC para a primeira característica, e SE_{j1}^2 é o quadrado do erro padrão do EBV do j^{th} animal para a primeira característica, e_{i2}^2 é o quadrado do coeficiente do vetor próprio para o i^{th} PC para a segunda característica, e SE_{j2}^2 é o quadrado do erro padrão do EBV do j^{th} animal para a segunda característica e assim por diante. O inverso de $SE(HCP1)$ foi usado na diagonal da matriz R, procedimento realizado por Solar Diaz et al. (2021).

Para a matriz H, utilizada na metodologia “single step” o inverso da matriz H, foi definido por (Aguilar et al., 2010) como:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

onde A é a matriz de relacionamento do numerador com base na linhagem para todos os animais; A_{22} é a matriz de relacionamento do numerador baseada na linhagem apenas para animais genotipados; e G é a matriz de relacionamento genômico para animais genotipados. As análises pela metodologia *single-step* (wssGBLUP) foram realizadas usando os programas da família BLUPF90 (Miształ, 2020). As soluções dos efeitos SNP foram obtidas a partir dos valores genéticos usando o algoritmo de Wang et al. (2012). O processo de iteração foi repetido duas vezes para obter a ponderação do efeito da solução dos SNPs. Neste estudo, os resultados de GWAS foram relatados como a proporção de variância genética aditiva explicada por janelas de 100 SNPs adjacentes.

2.5 Identificação dos genes

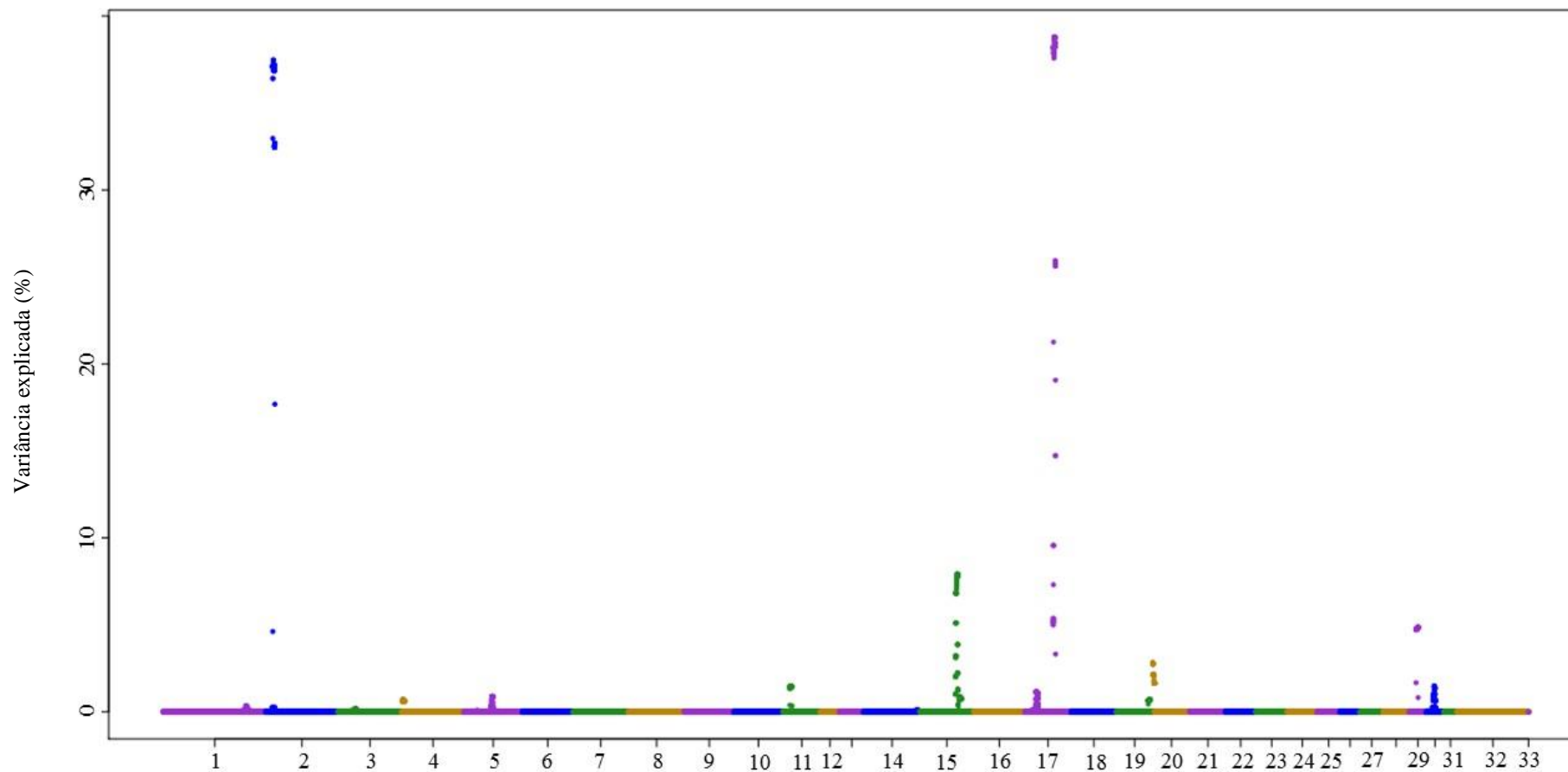
O genoma equino depositado no NCBI foi usado para identificação dos genes, usando o conjunto EquCab3.0 como mapa de referência (<https://www.ncbi.nlm.nih.gov/genome/?term=EQUUS+CABALLUS>). Assim, foi possível identificar segmentos localizados dentro ou próximos a genes que poderiam explicar a variabilidade na expressão do índice estudado.

3. Resultados

Oito janelas genômicas, que explicam que mais 1% de variância genética aditiva, foram identificadas como sendo as mais importantes para as características avaliadas tomadas pelo índice HCP1 (Tabela 2). As janelas selecionadas explicaram juntas 95,89% da variância genética aditiva total. A participação delas pode ser observada graficamente (Figura 1). Para um componente principal formado por características expressão quantitativa, há poucos loci participando.

As principais regiões genômicas associados ao HCP1 (Tabela 2), estão nos ECA17, ECA2 e ECA15, que explicam com variâncias de 38,78%; 37,45% e 7,93%, respectivamente. As demais cinco janelas explicaram juntas 11%. Destacam-se os genes *SPRY2*, *COL9A2*, *MIR30C*, *HEYL*, *BMP8B*, *LTBP1*, *FAM98A* e *CRIMI*, relacionados ao desenvolvimento de tecidos estruturais como ossos e cartilagens e associados a características morfométricas.

Figura 1 - Gráfico Manhattan obtido por associação a um índice composto pelo PC1 ponderado pelos EBVs de medidas lineares de equinos da raça Campolina



O eixo y indica a proporção da variância aditiva explicada pelo janelas de 100 SNPs adjacentes e a identificação dos cromossomos é dada no eixo x.

Tabela 2 - Identificação das regiões (cromossomos) e respectivos conjuntos de genes que explicaram maior proporção da variância genética aditiva das características analisadas pelo índice HCP1

Chr	Posição	Gene	Var
17	51,721,402-55,548,778	<i>NDFIP2, SLITRK1, SPRY2, RBM26, LOC111768697, LOC111768696, LOC111768695, LOC111768694, LOC111768613, LOC111768612, LOC111768611, LOC100055659, LOC100054692, LOC100050152</i>	38,78
2	14,417,590-18,438,851	<i>COL9A2, EDN2, MIR30E, MIR30C, MIR9061, BMP8B, TMEM269, FAM183A, HEYL, TRIT1, MYCL, MFSD2A, CAP1, PPT1, RLF, SMAP2, ZFP69B, ZFP69, ZNF684</i>	37,45
15	61,768,926-64,960,936	<i>LTPB1, RASGRP3, FAM98A, CRIM1, FEZ2, VIT, STRN, LOC111768227, LOC111768127, LOC111768126, LOC111768125, LOC111768124, LOC111768123, LOC106781706, LOC106781705, LOC102149932</i>	7,93
29	17,409,873-21,640,418	<i>VIM, MEIG1, ACBD7, MALRD1, SUV39H2, NMT2, FAM171A1, PTER, CUBN, TRDMT1, HACD1, STAM, TMEM236, MRC1, ARL5B, FAM107B, CDNF, HSPA14, RPP38, ITGA8</i>	4,84
20	358,932-4,149,020	<i>SERPINB1, PSMG4, FOXC1, FOXF2, TUBB2A, SLC22A23, BPHL, RIPK1, NQO2, WRNIP1, GMDS, DUSP22, SERPINB6, SERPINB9, MYLK4, FOXQ1, EXOC2, IRF4, LOC106782295, LOC100050361</i>	2,81
30	13,110,131-17,084,137	<i>SPATA17, RRP15, LYPLAL1, ESRRG, GPATCH2, TGFB2, USH2A, LOC111771332, LOC111771317, LOC111771316, LOC111771312, LOC111771293, LOC111771281, LOC100056341</i>	1,44
11	13,962,473-17,417,986	<i>GHI, ACE, SCN4A, PECAM1, ITGB3, CACNG4, PRR29, EFCAB3, METTL2A, MRC2, MARCHF10, CYB561, MAP3K3, STRADA, CD79B, ICAM2, ERN1, MILR1, POLG2, DDX5</i>	1,44

17	20,198,683-24,080,957	<i>HTR2A, MIR16-2, MIR15A, MIR3613, ARL11, LPAR6, EBPL, LRCH1, NUDT15, RB1, RCBTB2, CYSLTR2, FNDC3A, CDADC1, SETDB2, RCBTB1, KPNA3, SPRYD7, TRIM13, RNASEH2B</i>	1,13
----	-----------------------	--	------

4. Discussão

Encontraram-se oito regiões genômicas que explicam mais de 1% da variância aditiva do primeiro componente principal (HCP1) para quinze características morfométricas em equinos. As oito janelas explicam juntas 95,89% da variância aditiva e o HCP1 97,75%, portanto as janelas explicam ~94% da variação. Poucas janelas explicam muito da variação das características morfométricas em equinos, estando de acordo com estudos prévios (Tetens et al., 2013 Makvandi-Nejad et al. 2012)

Makvandi-Nejad et al. (2012) identificaram quatro loci (ECA3, ECA6, ECA9 e ECA11) que explicam 83% da variação do tamanho corporal de equinos. Em trabalho semelhante Signer-Hasler et al. (2012) encontraram regiões no ECA1, ECA3, ECA6 e ECA9 que influenciam características de conformação, principalmente a altura de cernelha, em cavalos da raça Franches-Montagnes. Meira et al. (2014) identificaram regiões genômicas no ECA2, ECA3, ECA6, ECA7, ECA8, ECA9, ECA21, e ECA26 associados a características morfométricas em cavalos da raça Quarto de Milha.

No presente estudo as principais regiões genômicas foram identificadas no ECA17, ECA2, ECA15, ECA29, ECA20, ECA30 e ECA11, onde foram encontrados alguns genes importantes para o desenvolvimento corporal. Embora nossos resultados tenham mostrado regiões genômicas no ECA2 como encontrado por Meira et al. (2014), e no ECA11 como observado por Makvandi-Nejad et al. (2012), as regiões genômicas não foram as mesmas. Todavia, a arquitetura genética de poucas regiões genômicas, com grande influência na variância, manteve-se. Vale ressaltar que os trabalhos citados realizaram GWAS utilizando como variável dependente as características morfométricas individualmente em outras raças. O presente estudo inova ao utilizar como variável dependente um índice que reúne um conjunto características morfométricas. Essa abordagem nos parece bastante coerente uma vez que, a seleção animal voltada ao melhoramento genético se baseia, mais comumente, em conjuntos de características.

Destacamos as janelas identificadas no ECA17, ECA2 e ECA15 que juntas representam 84% da variância genética aditiva explicada pelos marcadores relacionados ao HCP1. Nessas regiões foram identificados genes que estão relacionados ao desenvolvimento ósseo e características morfométricas.

Na janela do ECA17, que mais explica a aditividade da variável, encontra-se o gene *SPRY2* que participa das fases iniciais da osteogênese, pois a deficiência dessa proteína causa redução na formação óssea prejudicando o crescimento de camundongos (Vesela et al. 2019). Corbi-Botto et al (2019) não identificaram variação no número de cópias

para o ECA17 em equinos, onde essa janela se localiza, sendo, a variante causal, possivelmente, de outra forma.

Na segunda janela que mais explica a característica, localizada no ECA2, encontra-se o gene *COL9A2* que participa da síntese do colágeno associado à cartilagem presente nos ossos (Kemper et al., 2019). Segundo Wypchło et al. (2017), polimorfismos detectados no gene *COL9A2* podem estar associados a distúrbios durante a fase de crescimento dos equinos, causando ossificação endocondral e diminuição de cartilagem. Encontra-se também o RNA-gene *MIR30C* (Mahjoor et al., 2021) e o gene *HEYL* (Khan et al., 2020) expressos na medula óssea. Outro candidato é gene *BMP8B* que codifica uma proteína morfogenética óssea, considerada reguladora de crescimento pois possui atividade indutora e reguladora do processo de reparação e manutenção do tecido ósseo (Wozney et al., 1998). Cao et al. (2013) encontraram SNPs no gene associados a características de crescimento e morfométricas em bovinos.

Na terceira janela, localizada no ECA15, há o gene *LTBP1* associado a medidas morfométricas em ovinos (Wang et al., 2021), *FAM98A* expresso nos osteoclastos (Fujiwara et al., 2016) e uma mutação sem sentido no gene *CRIMI* já foi relatada como causa de nanismo em camundongos (Furuichi et al., 2019).

Estudos que analisaram assinatura de seleção relacionada às características de conformação corporal em equinos identificaram regiões principalmente nos ECA3, ECA9 e ECA11 como potenciais marcas de alterações nas frequências alélicas provocadas pela seleção artificial (Grilz-Seger et al., 2019; Gurgul et al., 2019; Srikanth et al., 2019). Embora nosso estudo também tenha observado influência do ECA11 sobre o índice composto por características de conformação corporal, a posição da janela difere dos estudos citados, reforçando a necessidade da realização de estudos raça-específicos.

Segundo Nanaei et al. (2020), é equivocada a ideia de um único mecanismo evolutivo influenciando os padrões de variação genética em todas as raças de cavalos, isso porque segundo os autores muitas raças passaram por processos de seleção com adaptações para diferentes ambientes e objetivos.

Nesse estudo aplicamos uma abordagem inovadora ao realizarmos o GWAS utilizando um índice calculado com base no PC (Solar Diaz et al., 2021). A importância do GWAS é aumentada, pois são compostos de múltiplas características (Kominakis et al., 2017). Foi possível observar a associação de regiões com genes cuja função está relacionada ao crescimento e desenvolvimento ósseo, sendo completamente condizente com o fenótipo. É importante salientar que a seleção dos equinos de sela é ainda

fenotípica. Ela leva em consideração além de fenótipos de marcha, características raciais muito associadas a características morfométricas. No caso específico do Campolina, há grande enfoque para aumento da altura da cernelha. A identificação de genes candidatos a um componente que leva em consideração a característica alvo junto com outras características morfométricas, pode permitir um aumento de porte de maneira harmônica pela seleção de variantes de efeito grande em mapeamento fino nos genes candidatos aqui indicados.

5. Conclusão

Oito regiões genômicas de grande participação na variância aditiva de índice de componentes principais de medidas morfométricas em equinos foram encontradas. Genes candidatos com ação na formação e crescimento dos ossos foram identificados. Análises de associação genômica ampla para componentes principais permitem identificação de marcadores de grande efeito para várias características e se torna uma eficiente ferramenta como alternativa para a redução de uma quantidade elevada de características mensuradas para o mesmo propósito.

Referências

- Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A., Tsuruta, S., Lawlor, T.J., 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science* 93:743-752.
- Associação Brasileira de Criadores de Cavalos Campolina – ABCCCampolina, 2021. Disponível em: <<http://www.campolina.org.br/sobre.php>>. Acesso em: 14 jul. 2021.
- Boligon, A.A., Vicente, I.S., Vaz, R.Z., Campos, G.S., Souza, F.R.P., Carneiro, R., Albuquerque, L.G., 2016. Principal component analysis of breeding values for growth and reproductive traits and genetic association with adult size in beef cattle. *Journal of animal science*, 94(12), 5014-5022.
- Bramante, G., Cito, A.M., Ciani, E., Pieragostini, E., 2016. Multi-trait animal model estimation of genetic parameters for morphometric measurements in the Murgese horse breed. *Livestock Science*, 191, 139-142.
- Bussiman, F.O., Perez, B.C., Ventura, R.V., Silva, F.F., Peixoto, M.G.C.D., Vizoná, R.G., Mattos, E.C., Ferraz J.B.S., Eler, J.P., Curi, R.A., Balieiro, J.C.C., 2018. Genetic

- analysis of morphological and functional traits in Campolina horses using Bayesian multi-trait model. *Livestock Science*, v.216, p.119-129.
- Cao, X.K., Wang, J., Lan, X.Y., Lei, C.Z., Zhang, C.L., Qi, X.L., Chen, H., 2013. Genetic variants in BMP8B gene are associated with growth traits in Chinese native cattle. *Gene*, 532(1), 115-120.
- Corbi-Botto, C.M., Morales-Durand, H., Zappa, M.E., Sadaba, S.A., Peral-García, P., Giovambattista, G., Díaz, S. 2019. Genomic structural diversity in Criollo Argentino horses: Analysis of copy number variations. *Gene*, 695, 26-31. <https://doi.org/10.1016/j.gene.2018.12.067>
- Costa, M.D., Mendes, L.J., Maruch, S., Ramirez, P.A., Meneses, A.C.A., Martins Neto, T., Chamone, J.M.A., 2016. Efeito da composição genética nas características de conformação em equinos. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 68(6), 1629-1637.
- De Oliveira Bussiman, F., Carvalho, R.S.B., e Silva, F.F., Ventura, R.V., Ferraz, J.B.S., Mattos, E.C., ... e Balieiro, J.C.D.C. 2021. Reduced rank analysis of morphometric and functional traits in Campolina horses. *Journal of Animal Breeding and Genetics*.
- De Oliveira Bussiman, F., e Silva, F.F., Carvalho, R.S.B., Ventura, R.V., Mattos, E.C., Ferraz, J.B.S., ... e de Carvalho Balieiro, J.C. 2022. Confirmatory factor analysis and structural equation models to dissect the relationship between gait and morphology in Campolina horses. *Livestock Science*, 255, 104779.
- Duensing, J., Stock, K.F., Krieter, J. 2014. Implementation and prospects of linear profiling in the warmblood horse. *J. Equine Vet. Sci.* 34, 360–368.
- Fujiwara, T., Ye, S., Castro-Gomes, T., Winchell, C.G., Andrews, N.W., Voth, D.E., Zhao, H. 2016. PLEKHM1/DEF8/RAB7 complex regulates lysosome positioning and bone homeostasis. *JCI insight*, 1(17).
- Furuichi, T., Tsukamoto, M., Saito, M., Sato, Y., Oiji, N., Yagami, K., Tomii, K. 2019. CRIM1C140S mutant mice reveal the importance of cysteine 140 in the internal region 1 of CRIM1 for its physiological functions. *Mammalian Genome*, 30(11), 329-338.
- Gmel, A. I., Druml, T., von Niederhäusern, R., Leeb, T., Neuditschko, M. 2019. Genome-wide association studies based on equine joint angle measurements reveal new QTL affecting the conformation of horses. *Genes*, 10(5), 370.

- Grilz-Seger, G., Druml, T., Neuditschko, M., Mesarič, M., Cotman, M., Brem, G. 2019. Analysis of ROH patterns in the Noriker horse breed reveals signatures of selection for coat color and body size. *Animal genetics*, 50(4), 334-346.
- Gurgul, A., Jasielczuk, I., Semik-Gurgul, E., Pawlina-Tyszko, K., Stefaniuk-Szmukier, M., Szmatoła, T., Bugno-Poniewierska, M. 2019. A genome-wide scan for diversifying selection signatures in selected horse breeds. *PLoS One*, 14(1), e0210751.
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2009. *Multivariate data analysis*. Prentice Hall, Upper Saddle River, p. 761.
- Holmström, M., and Back, W. The effects of conformation. In *Equine Locomotion*, 2nd ed.; Back, W., Clayton, H., Eds.; WB Saunders: London, UK, 2013; pp. 229–243.
- Junqueira, G.S.B., Diaz, I.D.P.S., da Cruz, V.A.R., de Araújo Oliveira, C.A., de Godoi, F.N., de Camargo, G.M.F., Costa, R.B. 2021. Influence of coat color on genetic parameter estimates in horses. *Journal of Applied Genetics*, 62(2), 297-306.
- Kemper, A.M., Drnevich, J., McCue, M.E., McCoy, A.M., 2019. Differential gene expression in articular cartilage and subchondral bone of neonatal and adult horses. *Genes*, 10(10), 745.
- Khan, A.A., Huat, T.J., Al Mutery, A., El-Serafi, A.T., Kacem, H.H., Abdallah, S.H., Jaafar, H. 2020. Significant transcriptomic changes are associated with differentiation of bone marrow-derived mesenchymal stem cells into neural progenitor-like cells in the presence of bFGF and EGF. *Cell & bioscience*, 10(1), 1-18.
- Kominakis, A., Hager-Theodorides, A.L., Zoidis, E., Saridaki, A., Antonakos, G., Tsiamis, G., 2017. Combined GWAS and ‘guilt by association’-based prioritization analysis identifies functional candidate genes for body size in sheep. *Genetics Selection Evolution*, 49(1), 1-16.
- Lucena, J.E.C., de Barros Vianna, S.A., Neto, F.B., Sales Filho; R.L.M., Da Silva Diniz, W.J., 2015. Estudo comparativo das proporções morfométricas entre garanhões e castrados da raça Campolina. *Semina: Ciências Agrárias*, v. 36, n. 1, p. 353-366.
- Mahjoor, M., Afkhami, H., Mollaei, M., Nasr, A., Shahriary, S., Khorrami, S. 2021. MicroRNA-30c delivered by bone marrow-mesenchymal stem cells induced apoptosis and diminished cell invasion in U-251 glioblastoma cell line. *Life Sciences*, 119643.
- Makvandi-Nejad, S., Hoffman, G.E., Allen, J.J., Chu, E., Gu, E., Chandler, A.M., Sutter, N.B. (2012). Four loci explain 83% of size variation in the horse. *PLoS One*, 7(7), e39929.

- Meira, C.T., Farah, M.M., Fortes, M.R., Moore, S.S., Pereira, G.L., Silva, J.A.I.V., Curi, R.A. 2014. A genome-wide association study for morphometric traits in quarter horse. *Journal of Equine Veterinary Science*, 34(8), 1028-1031.
- Meira, C.T., Pereira, I.G., Farah, M.M., Pires, A.V., Garcia, D.A., Cruz, V.A.R., 2013. Seleção de características morfofuncionais de cavalos da raça Mangalarga Marchador por meio da análise de componentes principais. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 65(6), 1843-1848.
- Misztal, I., 2020. BLUPF90 - A Flexible Mixed Model Program in Fortran 90. University of Georgia Retrieved from <http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90.pdf>
- Mudadu, M.A., Porto-Neto, L.R., Mokry, F.B., Tizioto, P.C., Oliveira, P.S., Tullio, R.R., Regitano, L.C., 2016. Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *BMC genomics*, 17(1), 1-16.
- Nanaei, H.A., Esmailizadeh, A., Mehrgardi, A.A., Han, J., Wu, D.D., Li, Y., Zhang, Y.P. 2020. Comparative population genomic analysis uncovers novel genomic footprints and genes associated with small body size in Chinese pony. *BMC genomics*, 21(1), 1-10.
- Novotná, A., Svitáková, A., Veselá, Z., Vostrý, L., 2017. Estimation of genetic parameters for linear type traits in the population of sport horses in the Czech Republic. *Livestock Science*, 202, 1-6.
- Pinto, L.F.B., Almeida, F.Q.D., Quirino, C.R., Azevedo, P.C.N.D., Cabral, G.C., Corassa, A., 2005. Análise multivariada das medidas morfométricas de potros da raça Mangalarga Marchador: análise de componentes principais. *Revista Brasileira de Zootecnia*, 34(2), 589-599.
- Rezende, M.P.G., Sousa, J.C.D., Mota, M.F., Oliveira, N.M., Jardim, R.J.D., 2016. Conformation index of horses of different genetic groups. *Ci. Anim. bras.*, 316-326.
- Solar Diaz, I.D.P., Strauss Borges Junqueira, G., Aparecida Rocha Cruz, V., Albano Araújo de Oliveira, C., Nunes de Oliveira, H., Miguel Ferreira de Camargo, G., e Bernal Costa, R. 2021. Principal components for morphometric traits in Campolina horses. *Journal of Animal Breeding and Genetics*, 138(2), 179-187.
- Souza, F.A.C.D., Fernandes, T.J., Cunha, F.O., Ribeiro, R.A., Muniz, F.R., Meirelles, S.L.C., Moura, R.S., 2019. Características morfométricas de equinos da raça

- Mangalarga Marchador determinadas por modelos não lineares. *Pesquisa Agropecuária Brasileira*, 54.
- Srikanth, K., Kim, N.Y., Park, W., Kim, J.M., Kim, K.D., Lee, K.T., Lim, D. 2019. Comprehensive genome and transcriptome analyses reveal genetic relationship, selection signature, and transcriptome landscape of small-sized Korean native Jeju horse. *Scientific reports*, 9(1), 1-16.
- Tetens, J., Widmann, P., Kühn, C., Thaller, G. 2013. A genome-wide association study indicates LCORL/NCAPG as a candidate locus for withers height in German Warmblood horses. *Animal genetics*, 44(4), 467-471.
- Vargas, G., Schenkel, F.S., Brito, L.F., Neves, H.H.R., Munari, D.P., Boligon, A., Carneiro, R., 2018. Unravelling biological biotypes for growth, visual score and reproductive traits in Nelore cattle via principal component analysis. *Livestock Science* 207:37-43.
- Vesela, B., Svandova, E., Hovorakova, M., Peterkova, R., Kratochvilova, A., Pasovska, M., Matalova, E. 2019. Specification of Sprouty2 functions in osteogenesis in in vivo context. *Organogenesis*, 15(4), 111-119.
- Wang, H., Misztal, I., Aguilar, I., Legarra, A., Muir, W.M., 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research* 94:73-83.
- Wang, S., Yi, X., Wu, M., Zhao, H., Liu, S., Pan, Y., Sun, X. 2021. Detection of key gene InDels in TGF- β pathway and its relationship with growth traits in four sheep breeds. *Animal biotechnology*, 32(2), 194-204.
- Wozney, J.M. 1998. The bone morphogenetic protein family: multifunctional cellular regulators in the embryo and adult. *European Journal of Oral Sciences*, 106(S1), 160-166.
- Wypchło, M., Korwin-Kossakowska, A., Bereznowski, A., Hecold, M., Lewczuk, D., 2017. Polymorphisms of the COL9A2, AOA1 and FRZB Genes in the Horse Genome and their Association with the Occurrence of Osteochondrosis. *Annals of Animal Science*, 17(1), 143-153.
- Zhou, J., Liu, L., Chen, C.J., Zhang, M., Lu, X., Zhang, Z., Shi, Y. 2019. Genome-wide association study of milk and reproductive traits in dual-purpose Xinjiang Brown cattle. *BMC genomics*, 20(1), 1-11.

CONSIDERAÇÕES FINAIS

O presente estudo caracterizou a região terminal do gene *PRDM9*, denominada domínio de dedos de zinco (*ZF*). Assim, foi possível identificar quatorze novos domínios *ZF* ainda não relatados na literatura para equinos nesse gene. Também foram identificados dezenove alelos e dezenove genótipos para o gene *PRDM9* em equinos. Deste modo, o uso de reprodutores com alelos e/ou genótipos menos frequentes em populações pequenas e/ou endogâmicas pode ser uma ferramenta adicional para melhoria da variabilidade genética. Os achados descritos neste trabalho apresentam ainda informações relevantes para estudos comparativos e de geração de conhecimento biológico básico e evolutivo.

Realizou-se ainda um GWAS utilizando como variável explanatória o índice (HCP1) composto pelo primeiro componente principal ponderado pelos valores genéticos aditivos (EBVs) para características morfométricas em equinos da raça Campolina. Foram identificadas oito regiões genômicas que explicam mais de 1% da variância aditiva. Nas principais regiões genômicas associadas ao HCP1 identificaram-se os genes *SPRY2*, *COL9A2*, *MIR30C*, *HEYL*, *BMP8B*, *LTBP1*, *FAM98A* e *CRIM1*, que estão relacionados ao desenvolvimento de tecidos estruturais como ossos e cartilagens e já associados a características morfométricas. O presente estudo inova ao realizar a análises de associação genômica ampla para componentes principais o que permite a identificação de marcadores de grande efeito para várias características. Uma abordagem bastante coerente uma vez que, a seleção animal voltada ao melhoramento genético se baseia, mais comumente, em conjuntos de características.