

**INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA  
GOIANO - CAMPUS RIO VERDE  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO**

**ESTRUTURA POPULACIONAL E ASSINATURAS DE SELEÇÃO AO  
LONGO DE GERAÇÕES NO COMPOSTO MONTANA TROPICAL®**

Discente: Camila Alves dos Santos

Orientador: Prof. Dr. Tiago do Prado Paim

Coorientadores: Prof. Dr. José Bento Sterman Ferraz

Prof. Dr. Adriano Carvalho Costa

Rio Verde - GO

Setembro – 2022

**ESTRUTURA POPULACIONAL E ASSINATURAS DE SELEÇÃO  
AO LONGO DE GERAÇÕES NO COMPOSTO MONTANA  
TROPICAL®**

Discente: Camila Alves dos Santos

Orientador: Prof. Dr. Tiago do Prado Paim

Coorientadores: Prof. Dr. José Bento Sterman Ferraz

Prof. Dr. Adriano Carvalho Costa

Dissertação apresentada, como parte das exigências para obtenção do título de MESTRE EM ZOOTECNIA, no programa de Pós-graduação em zootecnia do Instituto Federal de Educação, Ciência e Tecnologia Goiano – Campus Rio Verde  
Área de concentração Zootecnia

Rio Verde - GO

Setembro– 2022

Sistema desenvolvido pelo ICMC/USP  
Dados Internacionais de Catalogação na Publicação (CIP)  
**Sistema Integrado de Bibliotecas - Instituto Federal Goiano**

SSA237 Santos, Camila  
e       ESTRUTURA POPULACIONAL E ASSINATURAS DE SELEÇÃO  
AO LONGO DE GERAÇÕES NO COMPOSTO MONTANA TROPICAL® /  
Camila Santos; orientador Tiago do Prado Paim; co-  
orientador Adriano Carvalho Costa. -- Rio Verde,  
2022.  
59 p.

Dissertação (Mestrado em Programa de Pós Graduação  
em Zootecnia ) -- Instituto Federal Goiano, Campus  
Rio Verde, 2022.

1. : Adaptabilidade. 2. Pecuária. 3. SNP. I. do  
Prado Paim, Tiago, orient. II. Carvalho Costa,  
Adriano , co-orient. III. Título.

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO

Com base no disposto na Lei Federal nº 9.610, de 19 de fevereiro de 1998, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano a disponibilizar gratuitamente o documento em formato digital no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

### IDENTIFICAÇÃO DA PRODUÇÃO TÉCNICO-CIENTÍFICA

- |  |   |
|--|---|
| <input type="checkbox"/> Tese (doutorado)                  | <input type="checkbox"/> Artigo científico              |
| <input checked="" type="checkbox"/> Dissertação (mestrado) | <input type="checkbox"/> Capítulo de livro              |
| <input type="checkbox"/> Monografia (especialização)       | <input type="checkbox"/> Livro                          |
| <input type="checkbox"/> TCC (graduação)                   | <input type="checkbox"/> Trabalho apresentado em evento |

Produto técnico e educacional - Tipo:

Nome completo do autor:  
Camila Alves dos Santos

Matrícula:  
2020102310240015

Título do trabalho:  
ESTRUTURA POPULACIONAL E ASSINATURAS DE SELEÇÃO AO LONGO DE GERAÇÕES NO COMPOSTO MONTANA TROPICAL®

### RESTRIÇÕES DE ACESSO AO DOCUMENTO

Documento confidencial:  Não  Sim, justifique:

Informe a data que poderá ser disponibilizado no RIIF Goiano: 29 / 11 / 2022

O documento está sujeito a registro de patente?  Sim  Não

O documento pode vir a ser publicado como livro?  Sim  Não

### DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O(a) referido(a) autor(a) declara:

- Que o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- Que obteve autorização de quaisquer materiais incluídos no documento do qual não detém os direitos de autoria, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- Que cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Rio Verde

29 / 11 / 2022

Local

Data

*Camila Alves dos Santos*

Assinatura do autor e/ou detentor dos direitos autorais

Ciente e de acordo:

*Tiago do Prado Paím*

Assinatura do(a) orientador(a)



SERVIÇO PÚBLICO FEDERAL  
MINISTÉRIO DA EDUCAÇÃO  
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA  
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA GOIANO

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA  
GOIANO - CAMPUS RIO VERDE  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA**

Estrutura Populacional e Assinatura de Seleção ao Longo de  
Gerações no Composto Montana Tropical

Autora: Camila Alves dos Santos  
Orientador: Tiago do Prado Paim

TITULAÇÃO: Mestre em Zootecnia - Área de Concentração em Zootecnia/Recursos  
Pesqueiros.

APROVADA em 30 de setembro de 2022.

Dr. Luiz Fernando Brito  
Avaliador externo  
Purdue University

Dr. Adriano Carvalho Costa  
Avaliador interno  
IF Goiano/RV

Dr. José Bento Sterman Ferraz  
Avaliador externo  
FZEA/USP

Dr. Karen Martins Leão  
Avaliador interno  
IF Goiano/RV

Dr. Tiago do Prado Paim

Presidente da banca  
IF Goiano/RV

Documento assinado eletronicamente por:

- Luiz Fernando Brito, Luiz Fernando Brito - Professor Avaliador de Banca - Instituto Federal Goiano - Campus Rio Verde (10651417000500), em 01/10/2022 22:46:19.
- José Bento Sterman Ferraz, José Bento Sterman Ferraz - Professor Avaliador de Banca - Instituto Federal Goiano - Campus Rio Verde (10651417000500), em 01/10/2022 17:22:11.
- Adriano Carvalho Costa, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 30/09/2022 17:54:23.
- Karen Martins Leao, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 30/09/2022 12:04:25.
- Tiago do Prado Paim, MEDICO VETERINARIO, em 30/09/2022 11:29:33.

Este documento foi emitido pelo SUAP em 17/08/2022. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifgoiano.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 416617  
Código de Autenticação: f3cc4f7366



INSTITUTO FEDERAL GOIANO  
Campus Rio Verde  
Rodovia Sul Goiana, Km 01, Zona Rural, None, None, RIO VERDE / GO, CEP 75901-970  
(64) 3620-5600

## AGRADECIMENTOS

Agradeço primeiramente a Deus, pelo dom da vida e a Nossa Senhora, por me guiarem na minha jornada aqui na terra e me trazerem conforto espiritual diante as dificuldades que enfrentei desde o primeiro dia em que sai de casa em busca dos meus sonhos.

Aos meus pais Suelton e Vanúzia, por me apoiarem nas decisões que tomei e ficarem felizes com cada pequena conquista da minha história, e por serem meu refúgio nos momentos de cansaço.

Ao meu irmão João Batista, por sempre compartilhar comigo minhas alegrias por me ajudar a ver o lado bom das dificuldades e por sempre me agradar com seus abraços seus causos e boas conversas ao som de Almir Sater.

Ao meu padrinho Lázaro, tia Gilmeire, meus primos Lorryne e Lucas, por me acolherem nos primeiros meses em que cheguei em Rio Verde e por serem tão disponíveis e cuidadosos comigo.

A toda minha família, em especial aos meus avós João Batista, Hilda, Laurentino e Adelia, pelas palavras de incentivo e coragem para seguir com meus sonhos. E, também minha madrinha Ildimar, por ser inspiração de coragem e atitude.

Agradeço a minha amiga irmã Raíza Loiana, que sempre esteve ao meu lado sonhando comigo e me apoiando nas empreitadas que a vida apresentou.

Ao meu orientador Tiago Paim, por todos os ensinamentos, incentivos e confiança, obrigada por me ajudar a ter clareza e vislumbrar oportunidades que muitas vezes eu não acreditei estarem ao meu alcance. Espero que muitos alunos tenham a satisfação, assim como eu, de trabalhar com você e se tornarem pessoas melhores.

Ao professor José Bento Ferraz, Elisangela Chicarone, Rafael Espigolan e Grupo de Melhoramento Animal e Biotecnologia da Faculdade de Zootecnia e Engenharia de Alimentos, pela disponibilidade em atender as minhas dúvidas e meus pedidos de dados para este trabalho.

Ao IF Goiano e ao corpo docente do PPGZ, pelos ensinamentos.

Ao CNPq, pela bolsa concedida.



“Viver é um rasgar-se e remendar-se”

Guimarães Rosa

## **BIOGRAFIA DA AUTORA**

Camila Alves dos Santos nascida em 26 de abril de 1997, foi criada na fazenda em meio a sua maior paixão, cresceu decidida a formar-se Médica Veterinária, inspirada na lida diária de seu pai produtor de leite na cidade de Piranhas, no interior de Goiás. Em 2013 mudou de cidade para fazer ensino médio no Colégio Estadual Ariston Gomes da Silva e técnico em agropecuária no IF Goiano- Campus Iporá, onde teve seu primeiro contato com pesquisa como PIBIC Jr. Em 2020 obteve seu diploma de Bacharel em Medicina Veterinária na Universidade Federal de Goiás, e viveu 5 anos envolvida com pesquisas, monitorias, estágios, grupos de estudo e extensão em reprodução e melhoramento genético animal. Após a Graduação no segundo ano de pandemia do COVID-19, mudou-se para Rio Verde – GO para concluir seu mestrado em Zootecnia no Instituto Federal Goiano e trabalhar em um projeto de protótipo tecnológico em parceria com a Universidade Federal do Tocantins.

## ÍNDICE GERAL

<b>Resumo .....</b>	<b>10</b>
<b>Abstract .....</b>	<b>11</b>
<b>CAPÍTULO I- CONSIDERAÇÕES INICIAIS .....</b>	<b>12</b>
1 INTRODUÇÃO GERAL.....	12
2 REVISÃO DE LITERATURA.....	13
2.1 Genômica no melhoramento genético animal.....	13
2.2 Uso de Cruzamentos e Raças Compostas em bovinos .....	14
2.3 Gerações em melhoramento genético animal .....	15
2.4 Estrutura populacional .....	16
2.5 Endogamia .....	17
2.6 Assinaturas de seleção .....	18
2.7 Genômica de Paisagem.....	20
2.8 Ancestralidade local.....	21
REFERÊNCIAS.....	23
OBJETIVOS .....	30
<b>CAPÍTULO II - ESTRUTURA POPULACIONAL E ASSINATURA DE SELEÇÃO AO LONGO DE GERAÇÕES NO COMPOSTO MONTANA TROPICAL ®.....</b>	<b>31</b>
RESUMO.....	31
ABSTRACT.....	32
INTRODUÇÃO .....	33
MATERIAL E MÉTODOS .....	34
RESULTADOS .....	37
DISCUSSÃO .....	51
CONCLUSÃO .....	54
REFERÊNCIAS.....	55

## ÍNDICE DE FIGURAS

- Figura 1-** Dois primeiros componentes da análise de componentes principais gerados com dados genômicos de 3215 animais Montana. Generations: corresponde ao número de gerações equivalentes calculados com o pedigree..... 37
- Figura 2 -** Proporção dos grupos Raciais NABC do pedigree ao longo de gerações equivalentes no Composto Montrana Tropical ®..... 39
- Figura 3 –** Valores observados no pedigree versus valores preditos pela regressão a partir dos componentes Principais 1, 2, 3 e 4. .... 40
- Figura 4. -** Estrutura populacional do Composto Tropical Montana® inferida pelo software ADMIXTURE. Cada animal é representado por uma única barra vertical dividida em k cores, em que k é o número de agrupamentos ancestrais assumidos, que é representado graficamente para k = 2–4. C1= Azul, C2= Vermelho, C3= laranja e C4=verde. .... 41
- Figura 5-** Proporção dos clusters do ADMIXTURE ao longo de gerações equivalentes no Composto Montrana Tropical ®..... 42
- Figura 6-** Regressão valores de Pedigree estimados pelos clusters C1, C2, C3 e C4 ..... 43
- Figura 7 -** Resultado dos dois primeiros componentes principais da análise de componentes principais utilizando os dados de composição racial do pedigree (N A B C), 4 grupos da análise de admixture (C1 C2 C3 C4) e os quatro primeiros componentes principais da análise de PCA (PC1, PC2, PC3, PC4)..... 44
- Figura 8-** Correlação entre coeficiente de endogamia gnômico ( $F_{ROH}$ ) e coeficiente de endogamia do pedigree ( $F_{ped}$ ). .... 45
- Figura 9-** Correlação encontrada entre o coeficiente de endogamia genômico e gerações equivalentes ..... 46
- Figura 10-** Endogamia genômica baseada em corridas de homozigose ( $F_{ROH}$ ) por classes de geração equivalente e por classes de comprimento ROH. Os resultados da comparação do teste  $t$  são mostrados no topo (ns: não significativo; \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ ; \*\*\*\* $p < 0,0001$ )..... 47

## ÍNDICE DE FIGURAS

**Tabela 1**- Corridas de Homozigosidade (ROH) estatística por animal .....45

**Tabela 2** - Regiões de homozigose observadas em animais Composto Montana Tropical® e identificação dos genes subjacentes aos QTL em cada região .....49

## ÍNDICE DE ABREVIações

A	Raças taurinas adaptadas ( <i>Bos taurus taurus</i> )
B	Raças Britânicas ( <i>Bos taurus taurus</i> )
C	Raças da Europa continental ( <i>Bos taurus taurus</i> )
F	Inbreeding coefficient
IBD	<i>Identical by descent</i> (Idêntico por descendência)
LD	Linkage disequilibrium (Desequilíbrio de Ligação)
MAF	Minor allele frequency (frequência de alelos menores)
N	Raças Zebuínas ( <i>Bos taurus indicus</i> )
PCA	<i>Principal components analysis</i> (Análise de Componentes Principais)
QTL	<i>Quantitative trait loci</i> (Loci de características quantitativas)
ROH	<i>Runs of homozygosity</i> (corridas de homozigidade)
SNP	<i>Single nucleotide polymorphism</i> (Polimorfismo de Nucleotídeo Único)

## ESTRUTURA POPULACIONAL E ASSINATURAS DE SELEÇÃO AO LONGO DE GERAÇÕES NO COMPOSTO MONTANA TROPICAL®

### Resumo

O composto Tropical Montana® é população composta por quatro tipos Biológicos ou grupos raciais diferentes: 1) N = raças Zebuínas (*Bos taurus indicus*), 2) A = raças taurinas adaptadas (*Bos taurus taurus*), 3) B = raças Britânicas (*Bos taurus taurus*), e 4) C = raças da Europa continental (*Bos taurus taurus*). Foram utilizados dados de pedigree e genótipos (51,962 SNP) de 3,215 animais do Composto Montana Tropical® para prever a composição dos grupos biológicos do pedigree com dados genômicos e como essa composição se relaciona com as gerações dentro da raça composta; (2) caracterizar a distribuição de corridas de homozigosidade (ROH) em relação as gerações dentro da raça composta; (3) estimar e comparar o coeficiente de endogamia (F) com base nas informações de marcadores genômicos ( $F_{ROH}$ ) e pedigree ( $F_{PED}$ ); (4) identificar regiões genômicas de ilhas de ROH e (5) identificar características relacionadas em cada região genômica significativa sob seleção. A estrutura populacional genômica do Montana Tropical® tem pouca relação com a composição racial com base no sistema NABC, portanto o uso do NABC no programa de melhoramento genético não representa adequadamente a estrutura da população. Os níveis de endogamia dentro da raça estão baixos, mas demonstram evolução ao longo das gerações sendo importante atentar para esse aspecto, mesmo se tratando de raça que explora cruzamentos de diversos tipos biológicos. As detecções de assinaturas de seleção com QTLs relacionados a qualidade de carne demonstram que o objetivo de produzir animal de raça composta voltado para a produção de carne a pasto.

**Palavras-chave:** Adaptabilidade, Pecuária, SNP.

## Abstract

The Montana Tropical® Composite population was developed based on crossing of four different biological types or racial groups: 1) N = Zebu breeds (*Bos taurus indicus*), 2) A = adapted taurine breeds (*Bos taurus taurus*), 3) B = British breeds (*Bos taurus taurus*), and 4) C = races from continental Europe (*Bos taurus taurus*). Pedigree and genotypes (51,962 SNP) from 3,215 animals from the Montana Tropical® Composite were used to predict the composition of biological groups in the pedigree with genomic data and how this composition relates to each generation within the composite population; (2) characterize a distribution of runs of homozygosity (ROH) in relation to new information within the run; (3) estimate and compare the inbreeding coefficient (F) based on the markers (FROH) and pedigree (FPED); (4) identify genomic regions of the islands of the ROH, and (5) identify quantitative trait loci located in the genomic regions under selection. The population structure of the Montana Tropical Composite is lowly related to the racial composition based on the NABC system. The levels and endogamy within the levels, an evolution over the generations is being important to this same aspect when it comes to a breed that explores the crosses of different biological types. Selection signatures identified overlap with QTLs related to meat quality indicating that the production objective to produce a composite breed animal past the meat of outstanding meat for meat quality is being achieved.

**Keywords:** Adaptability, Livestock, SNP.



## CAPÍTULO I- CONSIDERAÇÕES INICIAIS

### 1 INTRODUÇÃO GERAL

Após o processo de domesticação, os animais mudaram notavelmente por causa da intensa seleção antropogênica (Joost *et al.*, 2011). As raças de bovinos atuais são resultado de anos de seleção, adaptação aos diferentes ambientes e efeitos demográficos como migração, deriva genética e efeito fundador, os quais contribuem para os padrões atuais de diversidade genética (Bruford *et al.*, 2003). No futuro, recursos genéticos com maior eficiência alimentar e ainda capacidade para se adaptar a mudanças climáticas (novas condições agroecológicas e novas doenças) serão necessários (Kantanen *et al.*, 2015).

A adaptação ao ambiente é definida como a capacidade de o animal sobreviver, crescer e se reproduzir na presença de estressores endêmicos do ambiente. As implicações econômicas devido ao desequilíbrio adaptativo incluem perdas de produção que afetam tanto os produtos em quantidade e qualidade, mortalidades (principalmente de doenças transmitidas por parasitas) e custos de tratamento (Burrow, 2012).

O cruzamento entre os animais de origem zebuína e os de origem europeia é uma alternativa para otimizar os efeitos de heterose e as características complementares das raças, combinando adaptabilidade ao ambiente e desempenho produtivo. Seguindo este raciocínio, o desenvolvimento de raças a partir dos cruzamentos resulta em bovino composto que, utilizando de diferentes proporções de raças fundadoras, pode se adaptar a diferentes estratégias e ambientes produtivos. Essa estratégia aproveita a heterozigose e variabilidade genética em locais ligados a características quantitativas de importância econômica, resultando em maior produtividade do animal (Marson *et al.*, 2005).

Raças compostas e animais cruzados também podem fornecer informações interessantes sobre como os efeitos de seleção para determinado ambiente atuam sobre o perfil genômico. Além do que quanto maior for o número de informações sobre a estrutura genética das populações, presença de genes de interesse e o desempenho em características produtivas, maior é o conhecimento sobre o seu potencial de utilização. Na maioria dos casos, as características de interesse produtivo e adaptação local são características quantitativas poligênicas e a identificação dos locais que governam a variação de tais características é uma tarefa desafiadora (Savolainen *et al.*, 2013). Algumas ferramentas podem ser usadas para detectar a base genômica destas características tais como Mapeamento de Loci Quantitativo (QTL), Mapeamento de associação através de análise de desequilíbrio de Ligação (LD), Análise de Componentes Principais (PCA) e Corridas de homozigose (ROH) (Somenzi *et al.*, 2020).

## 2 REVISÃO DE LITERATURA

O Brasil possui o segundo maior rebanho bovino do mundo e é o maior produtor e exportador de carne bovina, que respondeu por 17,3% das exportações agrícolas do país em 2018 (OCDE, 2018). A sustentabilidade econômica e ambiental da pecuária brasileira é de grande importância para garantir o fornecimento de proteína animal e segurança alimentar mundial, pois, espera-se grande aumento na demanda de alimentos nas próximas décadas (McManus *et al.*, 2016)

Os avanços atuais na agricultura brasileira podem ser impactados diretamente pelas mudanças climáticas e pelos efeitos biofísicos resultantes (Zilli *et al.*, 2020). Sob condições ambientais adversas, a produtividade animal é afetada, resultando em perdas econômicas para as indústrias pecuárias (Nardone *et al.*, 2010). É importante que sejam feitos esforços para compreender as respostas adaptativas do gado doméstico e garantir a produção e viabilidade econômica frente as demandas ambientais, um passo para isso é utilizar ferramentas de melhoramento genético e conhecer recursos genéticos responsivos aos sistemas de produção do país (Lobato *et al.*, 2014, Rojas-Downing *et al.*, 2017)

Nesse sentido, o presente capítulo pretende revisar tópicos conceituais e tecnologias importantes na identificação de características populacionais, que melhoram a compreensão, gerenciamento genético e tomada de decisões do manejo de bovinos de raça composta.

### 2.1 Genômica no melhoramento genético animal

A capacidade de genotipar milhares de marcadores moleculares em milhares de animais, denominado de estudos genômicos, provocou revolução no melhoramento genético mundial. A implementação de ferramentas genômicas está implicitamente associada ao descobrimento dos polimorfismos de base única (SNP) e o desenvolvimento de ferramentas de genotipagem desse tipo de marcador em larga escala e com relativo baixo custo (Hayes *et al.*, 2009).

Os SNPs (Single Nucleotide Polymorphisms) são a forma mais comum de polimorfismo entre indivíduos surgindo aproximadamente a cada 200 pares de bases nos bovinos (Williams, 2005). Um marcador SNP é uma alteração de base única em uma sequência de DNA (Vignal, 2002). Os métodos para identificação de novas variações de nucleotídeo único são o sequenciamento direto de regiões de interesse, permitindo a identificação de SNPs através de alinhamento e comparação de sequências usando bancos de dados públicos (Pariset *et al.*, 2009). Devido a sua abundância e recente disponibilidade de tecnologias de análise de alto rendimento, os SNP são os marcadores mais adequados para estudos genômicos (Pariset, 2012).

Os painéis contendo os marcadores SNP podem ter diversas densidades, como exemplo, de baixa densidade, o Illumina Bovine\_3k Beadchip que possui cerca de 2900 SNP e, de alta densidade, o Illumina BovineHD Beadchip com 777.000 SNP (Purfield *et al.*, 2012). A densidade necessária para análises genômicas em uma espécie depende do comprimento médio dos blocos cromossômicos, e nível de desequilíbrio de ligação (LD). Dentro desses blocos um ou alguns marcadores podem atuar como proxies para prever todos ou a maioria dos haplótipos presentes no bloco inteiro (Matukumally *et al.*, 2009).

Existem muitas plataformas comerciais disponíveis, que foram otimizadas para alta eficiência, robustez e geração de dados, alcançando excelentes custos-benefícios, como por exemplo, o chip SNP 777k para gado (Bovine HD Genotipagem Beadchip, Illumina, INC., San Diego, CA, EUA) e o chip SNP 600k para ovinos (SNP Bed Chip Ovine, Infinium HD, Illumina, Inc., San Diego, CA, EUA) (Paim *et al.*, 2018). A plataforma de densidade de SNP mais comum para bovinos é de 54.000 SNP (Illumina Bovinechip SNP50) (Purfield *et al.*, 2012).

## **2.2 Uso de Cruzamentos e Raças Compostas em bovinos**

Nenhuma raça se destaca em todas as características que são importantes para produção de carne. Diversas raças são necessárias para explorar a heterose e complementaridade por meio de cruzamentos e combinar o potencial genético para diversos mercados, recursos alimentares e climas (Wheeler *et al.*, 2005). A hibridização é uma forma de explorar o vigor híbrido (heterose) que ocorre quando o híbrido tem maior capacidade adaptativa e produtiva do que a média dos seus progenitores, devido a novas interações gênicas e redução dos alelos recessivos acumulados (Edmands, 2007).

Um dos desafios de implementação do uso de animais cruzados na produção animal é a reposição de matrizes no rebanho. O uso de raças compostas segue a ideia de explorar a heterose e conseguir manter uma população relativamente estável, em que se pode utilizar as fêmeas para reposição do rebanho. Portanto, após a formação da raça composta com a contribuição de 2 ou mais raças, ocorre a seleção dentro da raça (Gregory & Cundiff, 1980).

Entre as raças compostas utilizadas no Brasil podendo citar alguns exemplos como a raça Canchim criada em 1960 a partir do cruzamento entre animais *Bos taurus indicus* (zebu) e *Bos taurus taurus* (charolês). A composição final da raça apresenta 5/8 de charolês e 3/8 de zebu sendo a raça nelore a mais utilizada para compor a porção zebuína (Mokry *et al.*, 2014). A raça Brangus também um exemplo de raça composta com adequada adaptação aos trópicos, esta raça foi desenvolvida em 1949, com objetivo de combinar a produtividade do gado Angus

com 62,5% da composição racial e adaptação ambiental do zebu (Brahama ou nelore) com 37,5% na composição final da raça (Koger, 1980).

O Composto Tropical Montana® é uma população que foi desenvolvida em 1994, após estudos conduzidos por pesquisadores e produtores no Brasil em parceria com o Centro de pesquisa em animais de corte dos Estados Unidos no Clay Center, Departamento de Agricultura dos Estados Unidos (USDA). Buscando explorar a heterose e a capacidade de trabalho dos touros na monta a campo. Esta população composta foi desenvolvida cruzando animais de quatro tipos Biológicos ou grupos raciais diferentes: 1) N = raças Zebuínas (*Bos taurus indicus*), 2) A = raças taurinas adaptadas (*Bos taurus taurus*), 3) B = raças Britânicas (*Bos taurus taurus*), e 4) C = raças da Europa continental (*Bos taurus taurus*) (Grigoletto et al., 2020). O programa de melhoramento da população Montana tropical utiliza o sistema do tipo “aberto” que permite a introdução de novas raças dos diferentes tipos biológicos (FERRAZ *et al.*, 1999).

A composição de um animal da população Tropical Montana® é representada por múltiplos de 1/16 na sequência NABC. Desta forma, um animal Montana 444 se refere ao animal com a composição de ¼ do grupo biológico N, ¼ do grupo biológico A, ¼ do grupo biológico B e ¼ do grupo biológico C.

O gado adaptado tropicalmente é conhecido por sua capacidade de tolerar estresse por calor e resistência a doenças, mantendo a capacidade de produzir leite e carne (Huson *et al.*, 2014). As raças Bonsmara, Belmont Red, Romosinuano, Senepol e Caracu são alguns exemplos de Taurinos adaptados presentes no pedigree do Composto Montana Tropical®. Este grupo também é composto por animais da Família Sanga (África) que são animais adaptados aos trópicos (Makina *et al.*, 2014). O Caracu, por exemplo, é oriundo de animais *Bos taurus taurus* que originalmente foram trazidos para o Brasil no século XVI e que ao longo de sucessivas gerações e acasalamentos aleatórios foram considerados adaptados, pois conseguiram sobreviver aos desafios do clima tropical (Campos *et al.*, 2019).

As diferenças entre os grupos biológicos podem ser utilizadas para adequar o genótipo dos animais ao ambiente e aumentar a eficiência produtiva dos rebanhos, atendendo as demandas dos consumidores por produtos de alta qualidade e respeito ao meio ambiente (Silva *et al.*, 2007).

### **2.3 Gerações em melhoramento genético animal**

A hibridação pode ser definida como o cruzamento de indivíduos de populações distintas, este fenômeno acontece naturalmente em populações selvagens e pode ocorrer por ação antropogênica como no caso do bovino (VonHoldt *et al.*, 2018). O híbrido de primeira

geração (F1) resulta do cruzamento de duas espécies parentais puras, um híbrido de segunda geração (F2) resulta do cruzamento de dois híbridos (F1) e, assim por diante (Jackiw *et al.*, 2016). No entanto, em bovinos a separação de gerações de cruzamento não é tão clara, porque ocorre sobreposição de gerações, ou seja, o F2 é cruzado com o F3 e assim por diante, e o F1 pode ser cruzado com outra linhagem ou raça pura.

O termo geração refere à transição entre progenitores e a prole, sendo uma geração definida como todos os indivíduos que estão no mesmo estágio de descendência de um ancestral comum. Em bovinos, usualmente estes parâmetros são calculados pelas informações de pedigree, podendo obter as seguintes informações número de gerações completas traçadas que se refere à geração mais distante em que todos os ancestrais sejam conhecidos, ou seja, que separa a progênie da mais distante geração. O número máximo de gerações rastreadas que é o número de gerações que separam o indivíduo de seu ancestral mais remoto. O número equivalente de gerações completas, por sua vez, é obtido pelo somatório dos termos  $(1/2)^n$  de todos os ancestrais conhecidos, em que n é o número de gerações que separa o indivíduo de cada ancestral conhecido (Gutiérrez & Goyache, 2005)

O intervalo de gerações é dado pela média da idade dos pais ao nascimento de sua progênie. Podendo ser computado intervalos de gerações para Touro-filho, Touro-filha, Vaca-Filho, Vaca filha (Rendel e Robertson, 1951). Sendo este parâmetro fator importante, pois afeta a taxa de ganho genético das populações, especialmente na espécie bovina que possui intervalos de gerações relativamente longos. Biologicamente o menor intervalo de geração possível é a soma da idade de maturidade sexual e a duração da gestação (Scheffers & Weigel, 2012).

## 2.4 Estrutura populacional

O conhecimento da estrutura populacional e composição racial é importante para gerir as informações da raça dentro do programa de melhoramento genético e tomar decisões mais assertivas e acuradas na conservação da diversidade, seleção de características de interesse e avaliar adaptabilidade de raças mestiças a determinado ambiente de produção (Reist-Marti *et al.*, 2003, Kuehn *et al.*, 2011).

Os valores de composição do pedigree são utilizados juntamente com informações de desempenho para avaliação genética. Essas análises são baseadas em metodologia de modelo misto, em particular a abordagem estatística de modelo animal usando os melhores métodos de predição linear não viesada para obter valores genéticos estimados (EBV) para características economicamente importantes (Montaldo *et al.*, 2012).

Agora que os dados de chip SNP estão disponíveis, esses EBVs de pedigree podem ser complementados com informações de dados genômicos, gerando EBVs genômicos (GEBVs). O GEBV pode ser calculado como função linear dos genótipos SNP ponderados pelo efeito aparente de cada genótipo na característica (Wray *et al.*, 2019). Alguns trabalhos indicaram que a previsão do mérito genômico mestiço pode ser melhorada calculando valores genômicos diretos de acordo com os efeitos ponderados do SNP de cada uma das raças contribuintes (VanRaden *et al.*, 2015).

A estimativa de composição racial com base no pedigree é um método convencionalmente utilizado e acessível, no entanto está sujeito a erros de anotação, dados incompletos ou ausentes prejudicando a confiabilidade da informação, além do que a composição racial por pedigree é apenas uma estimativa probabilística dos efeitos de meiose que pode ter boa aproximação para o rebanho, mas alto grau de erro para o indivíduo (Ron *et al.*, 1996). Os avanços nas tecnologias genômicas permitiram inovações e implementações nos métodos de estimação da composição racial dos animais domésticos, e representa a possibilidade de aumentar a precisão das estimativas dos valores genéticos de raça pura ou mestiça em raças compostas ou populações mestiças (Manel *et al.*, 2005, Montaldo *et al.*, 2012, Huang *et al.*, 2014).

Métodos que utilizaram dados de genoma completo foram utilizados em espécies como bovinos, suínos e ovelhas para predição e estimação da composição geral e a proporção racial genômica com coeficientes de regressão satisfatórios (Dodds *et al.*, 2014, Funkhouser *et al.*, 2017, Gobena *et al.*, 2018). Informações genômicas têm o potencial de fornecer mais estimativas precisas da composição da raça e distância genômica entre as raças em relação ao convencionalmente registrado em dados de pedigree (Kelleher *et al.*, 2017). Portanto estudo de dados genômicos da diversidade e estrutura populacional podem melhorar a compreensão dos recursos genéticos reais da população (Tolone *et al.*, 2012).

## **2.5 Endogamia**

Dois indivíduos são geneticamente relacionados quando possuem ancestrais comuns. A partir disto, o grau de endogamia dependerá da relação de parentesco existente entre seus descendentes. Desta forma, a endogamia corresponde ao acasalamento entre indivíduos mais aparentados que a média da população em que estão inseridos (Burrow, 1993).

Altos níveis de endogamia podem promover decréscimos no número de indivíduos heterozigotos, redução da variabilidade genética, aumento da chance de aparecimento de alelos recessivos com efeitos deletérios (Falconer & Mackay, 1996). Este aumento da homozigose é

responsável pela depressão endogâmica que pode ser caracterizada pela redução da resposta a seleção ou redução do desempenho produtivo, quando os animais são provenientes de acasalamentos endogâmicos (Alexander *et al.*, 1961, Reverter *et al.*, 2017) O conhecimento da taxa de endogamia ao longo das gerações associado a programas de melhoramento genético, pode possibilitar melhor orientação de planos de acasalamento em cruzamentos.

A diversidade genética ou nível de autozigosidade é estimada e gerenciada com a ajuda de coeficientes de endogamia (F) com base em dados de pedigree. Esses coeficientes refletem a proporção do genoma que se espera que seja idêntico por descendência (IBD) em um indivíduo e entre dois indivíduos, respectivamente (Wright, 1921)

Corridas de homozigose (conhecidas como runs of homozygosity – ROHs) são segmentos homozigotos contínuos herdados de ancestrais comuns (Ferencakovic *et al.*, 2011, Purfield *et al.*, 2012). As formações das ROHs podem ser influenciadas por endogamia, deriva genética, gargalo populacional, bem como seleção natural e artificial, ou seja, seu estudo pode caracterizar o histórico da estrutura populacional (Curick *et al.*, 2014).

As variações no tamanho e na distribuição das ROHs observadas pelo genoma podem ser utilizadas para inferir sobre eventos de consanguinidade em uma população específica (Howrigan *et al.* 2011). Raças e linhagens especializadas de animais de produção passaram e continuam sendo submetidas à intensa seleção de pacotes de alelos com efeito positivo sobre características produtivas, reprodutivas ou de padrão racial e, devido a isso, é comum encontrar altas taxas de endogamia e, conseqüentemente, ROHs distribuídas em abundância pelo genoma e presentes em altas frequências nas populações estudadas (Zavarez *et al.*, 2015; Zhang *et al.*, 2015).

De acordo com o tamanho da ROH é possível calcular o coeficiente de endogamia genômico através da FROH que é estimada a partir do somatório de corridas em homozigose (ROHs), separada em diferentes tamanhos mínimos de corridas e pode ser definida como a proporção do genoma autossômico coberto por SNPs que é composta por ROHs. Permitindo classificar a endogamia sendo segmentos longos relacionados com endogamia recente e os curtos com endogamia mais antiga (McQuillan *et al.* 2008).

## **2.6 Assinaturas de seleção**

As assinaturas de seleção delimitam regiões do genoma que são, ou foram, funcionalmente importantes, e, portanto, estiveram sob seleção natural ou artificial (Zinovieva *et al.*, 2020). A seleção normalmente deixa assinaturas no genoma, que geralmente são caracterizadas por alta diferenciação genética entre as raças e ou forte redução na diversidade

genética em regiões associadas a características sob intensa pressão de seleção. O estudo dessas assinaturas vem sendo utilizado para associação com variedade de características incluindo tolerância térmica, resposta imune, funções reprodutivas, ingestão de alimento e metabolismo (Onzima, et al 2018).

Os estudos de assinatura de seleção vêm se tornando cada vez mais populares porque oferecem estratégia complementar para estudos de associação ampla do genoma (GWAS) sobre variantes de mapeamento que afetam características de interesse, ajudando a vincular fenótipos a função do gene (Utsunomiya *et al.*, 2015). Os resultados provenientes deste tipo de estudo podem ser úteis para conhecer a variabilidade genética da raça e auxiliar no desenvolvimento dos programas de melhoramento genético (Zinovieva *et al.*, 2020).

A assinatura de seleção pode ser detectada utilizando diferentes abordagens e métodos como alta frequência de alelos derivados que pegaram carona na população, a duração e estrutura de haplótipos, medidos por homozigosidade, e por diferenciação genética entre as populações medidas por  $F_{ST}$  ou estatísticas relacionadas (Fariello *et al.*, 2013).

A análise de hapFLK consiste em analisar através da estrutura hierárquica da população amostrada as diferenças nas frequências de haplótipos detectando, assim regiões sob seleção (Fariello *et al.*, 2013). Para estimar a estrutura hierárquica das populações calcula-se a distância de Reynolds e é convertida em matriz de parentesco com scripts fornecidos pela página da Web do hapFLK (<https://forge-dga.jouy.inra.fr/projects/hapflk>). O uso de uma combinação de informações de haplótipos e da estrutura hierárquica de populações resulta em maior robustez para detecção de assinatura de seleção (Manunza *et al.*, 2016).

A metodologia de XP-EHH é baseada na detecção de assinaturas de seleção entre populações, comparando seus perfis de EHH. O método de EHH explora a decadência de homozigosidade do haplótipo como uma função de distância genética de um SNP focal. Ou seja, o XP-EHH detecta assinaturas de seleção nas quais o alelo selecionado se aproximou ou atingiu a fixação em uma população. Este método usa comparação entre populações de comprimentos de haplótipos para controlar a variação local das taxas de recombinação. Essa comparação entre populações é complicada pelo fato de que os comprimentos dos haplótipos também dependem da história da população, como gargalos e expansões. O Teste de XP-EHH normaliza para diferenças em todo o genoma no comprimento do haplótipo entre as populações (Sabeti *et al.*, 2006).

A frequência que os marcadores SNP aparecem dentro de uma corrida de homozigose (ROH) na população pode indicar uma assinatura de seleção dentro da população. Assim, é possível identificar os efeitos de pressões seletivas em diferentes lapsos temporais e o uso intenso de alguns reprodutores na população, uma vez que a ROH determina a existência de um



ancestral comum no pedigree do animal. Assim, é possível observar as consequências da seleção ao longo do desenvolvimento da raça para características de importância econômica e/ou características relacionadas a adaptação ao ambiente (Peripolli *et al.*, 2018).

## 2.7 Genômica de Paisagem

A adaptação local implica na existência de alelos vantajosos conferindo a população em seu habitat nativo maior aptidão do que qualquer outra população alóctone vivendo no mesmo habitat (Kawecki & Ebert, 2004). A genômica de paisagem correlaciona padrões de variação genética com variáveis geográficas para investigar como as características geográficas e ambientais afetam a estrutura genética das populações, um campo que combina marcadores moleculares, genética e ambiente (Pariset *et al.*, 2012).

Genômica de paisagem é uma das abordagens usadas para detectar assinaturas de adaptação local e se tornaram cada vez mais populares, principalmente pela diminuição do custo de sequenciamento, mas também pela recente disponibilidade de conjuntos de dados ambientais de alta resolução (Osipova *et al.*, 2019). Nos animais de produção, o estudo da adaptação de diferentes raças ao meio ambiente é de crucial importância para apoiar sistemas de produção baseados em raças adaptadas reduzindo o impacto ambiental, assim a genômica de paisagem oferece ferramenta para identificar os genótipos adaptados a determinado ambiente (Joost *et al.*, 2010).

Na genômica de paisagem, a informação genética é incorporada dentro de um contexto geográfico. No gado consiste simplesmente em atribuir valores de latitude e longitude, topografia, altitude, clima, informações epidemiológicas, ecológicas, cobertura de solo, estradas, demografia, umidade relativa do ar, precipitação e dados econômicos a qualquer amostra de DNA colhida em animais amostrados (Joost *et al.*, 2011).

Para analisar a interação de variáveis ambientais e padrões genéticos, a genômica de paisagem requer alto número de marcadores moleculares para fornecer alta resolução e uma precisão de localização geográfica. O principal sistema de marcadores utilizado são as tecnologias de SNPS (Pariset *et al.*, 2012). E, conseqüentemente, a genômica de paisagem é altamente dependente do sistema de informações geográficas (SIG).

Um exemplo de ferramenta para este estudo é o SAM (Método de análise espacial) que é uma abordagem que oferece a possibilidade de identificar loci, provavelmente sob seleção natural. O SAM analisa a associação entre as frequências alélicas nos marcadores moleculares e dados de várias variáveis ambientais. Para esse fim, ele usa um ou mais variáveis ambientais que descrevem o local de amostragem e uma matriz de marcadores moleculares. Usando uma regressão logística esse método associa a frequência de marcadores moleculares com parâmetros em cada site e destaca os possíveis marcadores vinculados a regiões genômicas envolvidas na adaptação (Joost *et al.*, 2008).

## 2.8 Ancestralidade local

Por causa da grande resolução alcançada hoje em dia com as plataformas de genotipagem, é possível utilizar ferramentas de computação capazes de traçar a origem dos haplótipos dentro de uma população de animais de raças compostas. Essa compreensão da ancestralidade de cada região genômica é conhecida atualmente como ancestralidade local (Chen *et al.*, 2020).

Raças compostas são frequentemente manejadas como única população, isso significa que a seleção natural e seleção artificial são aplicadas usando informações fenotípicas ou diferença esperada na progênie (DEP para peso em diferentes idades, rendimento de carcaça etc) disponíveis, essa seleção poderá causar alterações genômicas na população (Goszczyński *et al.*, 2018).

Essas regiões genômicas sob seleção em ambientes desafiadores podem conter grande desvio de ancestralidade ao longo das gerações na direção de um dos fundadores e podem indicar forte pressão de seleção ou deriva genética (Gautier & Naves, 2011).

A obtenção dessa informação nos abre novas formas para compreender como os alelos são mantidos na população após cruzamento e pode-se analisar se as vantagens que foram buscadas na criação do composto (produção de carne e adaptabilidade) estão sendo mantidas ao longo das gerações. E, tomar decisões de como manejar a hibridização definindo melhores práticas de seleção para as raças compostas, visando produção, saúde, bem-estar e sustentabilidade (Paim *et al.*, 2020).

O software ADMIXTURE pode ser utilizado para estudar a ancestralidade local de raças compostas. Ele estima o coeficiente de ancestralidade como parâmetros de modelo estatístico, adotando o modelo de verossimilhança. Estima simultaneamente as frequências alélicas juntamente com as proporções de ascendência. Capaz de executar testes de clusters

supervisionados para avaliar a ancestralidade de híbridos simulados e reais (Alexander *et al.*, 2009).

Outra maneira eficiente de identificar informações importantes sobre ancestralidade a partir de dados densos de genótipos é com o software Chromo Painter. O Chromo Painter encontra haplótipos nos dados de sequência e cada indivíduo é pintado como uma combinação de todas as outras sequências (Lawson *et al.*, 2012).

## REFERÊNCIAS

- Alexander, G. I., & Bogart, r. 1961. Effect of inbreeding and selection on performance characteristics of beef cattle. *Journal of Animal Science*, 20(4), 702-707. <https://doi.org/10.2527/jas1961.204702x>
- Alexander D.H., Novembre J., Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19: 1655–64. <http://doi10.1101/gr.094052.109>
- Bruford, M.W., Bradley, D.G., & Luikart, G. 2003. Dna markers reveal the complexity of livestock domestication. *Nature*4:900-910. <https://doi.org/10.1038/nrg1203>
- Burrow, H. M. The effects of inbreeding in beef cattle. *Animal Breeding Abstracts*. 1993. p. 737-751.
- Burrow, H. M. 2012 Importance of adaptation and genotype x environment interactions in tropical beef breeding systems. *Animal: an international journal of animal bioscience*, v. 6, n. 5, p. 729. <https://doi.org/10.1017/S175173111200002X>
- Campos, B. M., Carmo, A. S., Egito, A. A., Mariante, A. S., Albuquerque, M. S. M., Gouveia, J. J. S., ... & Carneiro, P. L. S. 2017. Genetic diversity, population structure, and correlations between locally adapted zebu and taurine breeds in Brazil using SNP markers. *Tropical animal health and production*,49(8), 1677-1684. <https://doi.org/10.1007/s11250-017-1376-7>
- Caye, K., Jumentier, B., Lepeule, J., & François, O. 2019. LFMM 2: fast and accurate inference of gene-environment associations in genome-wide studies. *Molecular biology and evolution* 36: 852-860. <https://doi.org/10.1093/molbev/msz008>
- Chen, Q., Zhan, J., Shen, J., Qu, K., Hanif, Q., Liu, J., ... & Lei, C. (2020). Whole-genome resequencing reveals diversity, global and local ancestry proportions in yunling cattle. *Journal of Animal Breeding and Genetics*,137(6), 641-650. <https://doi.org/10.1111/jbg.12479>
- Curik, I., Ferenčaković, M., & Sölkner, J. 2014. Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livestock Science*,166, 26-34. <https://doi.org/10.1016/j.livsci.2014.05.034>
- Dodds, K. G., Auvray, B., Newman, S. A. N., & McEwan, J. C. (2014). Genomic breed prediction in New Zealand sheep. *BMC genetics*, 15(1), 1-15. <https://doi.org/10.1186/s12863-014-0092-9>
- Edmands S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology* 16:463–475 <https://doi.org/10.1111/j.1365-294X.2006.03148.x>
- Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*, Ed 4. Longmans Green, Harlow, Essex, UK.
- Fariello, M. I., Boitard, S., Naya, H., Sancristobal, M., & Servin, B. 2013. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics*,193(3), 929-941. <https://doi.org/10.1534/genetics.112.147231>

- Ferencakovic, M., Hamzic, E., Gredler, B., Curik, I., & Sölkner, J. 2011. Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. *Agriculturae Conspectus Scientificus*, 76(4), 325-329.
- Funkhouser, S. A., Bates, R. O., Ernst, C. W., Newcom, D., & Steibel, J. P. 2017. Estimation of genome-wide and locus-specific breed composition in pigs. *Translational animal science*, 1(1), 36-44. <https://doi.org/10.2527/tas2016.0003>
- Gautier, M., & Naves, M. 2011. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Molecular ecology*, 0(15), 3128-3143. <https://doi.org/10.1111/j.1365-294X.2011.05163.x>
- Gobena, M., Elzo, M. A., & Mateescu, R. G. 2018. Population structure and genomic breed composition in an angus–brahman crossbred cattle population. *Frontiers in genetics*, 9, 90. <https://doi.org/10.3389/fgene.2018.0009>
- Goszczynski, D. E., Corbi-Botto, C. M., Durand, H. M., Rogberg-Muñoz, A., Munilla, S., Peral-Garcia, P., ... & Giovambattista, G. (2018). Evidence of positive selection towards Zebuine haplotypes in the BoLA region of Brangus cattle. *animal*, 12(2), 215-223.
- Gutiérrez, J. P., & Goyache, F. 2005. A note on ENDOG: a computer program for analysing pedigree information. *Journal of Animal Breeding and genetics*, 122(3), 172-176. <https://doi.org/10.1111/j.1439-0388.2005.00512.x>
- Gregory, K. E., & Cundiff, L. V. 1980. Crossbreeding in beef cattle: evaluation of systems. *Journal of animal science*, 51(5), 1224-1242. <https://doi.org/10.2527/jas1980.5151224x>
- Grigoletto, L., Ferraz, J., Oliveira, H. R., Eler, J. P., Bussiman, F. O., Abreu Silva, B. C., ... & Brito, L. F. 2020. Genetic Architecture of Carcass and Meat Quality Traits in Montana Tropical® Composite Beef Cattle. *Frontiers in genetics*, 11, 123. <https://doi.org/10.3389/fgene.2020.00123>
- Hayes, B.J., Chamberlain, A.J., Maceachern, S., Savin, K., McPartlan, H., MacLeod, I., Sethuraman, L. and Goddard, M.E., 2009. A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. *Animal genetics*, 40(2), pp.176-184. <https://doi.org/10.1111/j.1365-2052.2008.01815.x>
- Howrigan, Daniel P, Matthew A Simonson, and Matthew C Keller. 2011. Detecting Autozygosity through Runs of Homozygosity: A Comparison of Three Autozygosity Detection Algorithms. *BMC genomics* 12(1): 460. <https://doi.org/10.1186/1471-2164-12-460>.
- Huson, H. J., Kim, E. S., Godfrey, R. W., Olson, T. A., McClure, M. C., Chase, C. C., ... & Sonstegard, T. S. (2014). Genome-wide association study and ancestral origins of the slick-hair coat in tropically adapted cattle. *Frontiers in Genetics*, 5, 101. <https://doi.org/10.3389/fgene.2014.00101>
- Jackiw, R.N., Mandil, G. & Hager, H.A. 2015. A framework to guide the conservation of species hybrids based on ethical and ecological considerations. *Conserv. Biol* 29: 1040-1051. <https://doi.org/10.1111/cobi.1252>
- Joost, S., Coll, L., Baret, P.V., Garcia, J.F., Boettcher P.J., Tixier-Boichard, M., Ajmone-Marsan, P., & the Globaldiv Consortium. 2010. Integrating geo-referenced multiscale and

- multidisciplinary data for the management of biodiversity in livestock genetic resources. *Animal Genetics* 41: 47–63. <https://doi.org/10.1111/j.1365-2052.2010.02037.x>
- Joost, S., Colli, L., Bonin, A., Biebach, I., Allendorf, F. W., Hoffmann, I., ... & Globaldiv Consortium. 2011. Promoting collaboration between livestock and wildlife conservation genetics communities. *Conservation genetics resources* 3:785-788. <https://doi.org/10.1007/s12686-011-9457-9>.
- Joost, S., Kalbermatten, M., & Bonin, A. 2008. Spatial Analysis Method (SAM): a software tool combining molecular and environmental data to identify candidate loci for selection. *Molecular Ecology Resources* 8:957–960. <https://doi.org/10.1111/j.1755-0998.2008.02162.x>
- Kantanen, J., Løvendahl, P., Strandberg, E., Eythorsdottir, E., LI, M. H., Kettunen-PRÆBEL, A., ... & Meuwissen, T. 2015. Utilization of farm animal genetic resources in a changing agro-ecological environment in the Nordic countries. *Frontiers in genetics*. 6:52. <https://doi.org/10.3389/fgene.2015.00052>
- Kawecki, T. J., & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecology letters*, 7(12), 1225-1241. <https://doi.org/10.1111/j.1461-0248.2004.00684.x>
- Kelleher, M. M., Berry, D. P., Kearney, J. F., McParland, S., Buckley, F., & Purfield, D. C. 2017. Inference of population structure of purebred dairy and beef cattle using high-density genotype data. *Animal*, 11(1), 15-23. <https://doi.org/10.1017/S1751731116001099>
- Koger M. 1980. Sistemas de cruzamento eficazes utilizando zebuínos. *Journal of Animal Science* 50.1215–20. <https://doi.org/10.2527/jas1980.5061215x>
- Kuehn, L. A., Keele, J. W., Bennett, G. L., McDanel, T. G., Smith, T. P. L., Snelling, W. M., ... & Thallman, R. M. (2011). Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. *Journal of animal science*, 89(6), 1742-1750. <https://doi.org/10.2527/jas.2010-3530>
- Lawson D.J., Hellenthal G., Myers S., Falush D., Zhang F. 2012 Inference of population structure using dense haplotype data G. P. Copenhaver. *PLoS Genetics* 8. 10.1371 / journal.pgen.1002453. <https://doi.org/10.1371/journal.pgen.1002453>
- Lobato, J. F. P., Freitas, A., Devincenzi, T., Cardoso, L. L., Tarouco, J. U., Vieira, R. M., ... & Castro, I. 2014. Brazilian beef produced on pastures: Sustainable and healthy. *Meat science*, 98(3), 336-345. <https://doi.org/10.1016/j.meatsci.2014.06.022>
- Maiorano, A. M., Lourenco, D. L., Tsuruta, S., Ospina, A. M. T., Stafuzza, N. B., Masuda, Y., ... & Silva, J. A. I. D. V. 2018. Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PLoS One*, 13(8), e0200694. <https://doi.org/10.1371/journal.pone.0200694>
- Makina, S. O., Muchadeyi, F. C., Van Marle-Köster, E., Macneil, M. D., & Maiwashe, A. 2014. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. *Frontiers in genetics* 5:333. <https://doi.org/10.3389/fgene.2014.00333>

- Manel, S., Joost, S., Epperson, B. K., Holderegger, R., Storfer, A., Rosenberg, M. S., ... & Fortin, M. J. (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, 19(17), 3760-3772. <https://doi.org/10.1111/j.1365-294X.2010.04717.x>
- Manunza, A., Cardoso, T. F., Noce, A., Martínez, A., Pons, A., Bermejo, L. A., ... & Amills, M. 2016. Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. *Scientific reports*, 6(1), 1-10. doi: 10.1038 / srep27296 (2016).
- Marson, E. P., Ferraz, J. B. S., Meirelles, F. V., Balieiro, J. C. D. C., Eler, J. P., Figuerido, L. G. G., & Mourão, G. B. 2005. Genetic characterization of European-Zebu composite bovine using RFLP markers. *Genet. Mol. Res* 4: 496-505.
- Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., ... & Van Tassell, C. P. 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PloS one* 4. <https://doi.org/10.1371/journal.pone.0005350>
- McQuillan, Ruth et al. 2008. Runs of Homozygosity in European Populations. *American journal of human genetics* 83(3): 359–72. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2556426&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1016/j.ajhg.2008.08.007>
- MCmanus, C., Barcellos, J. O. J., Formenton, B. K., Hermuche, P. M., Carvalho JR, O. A. D., Guimarães, R., ... & Neto, J. B. 2016. Dynamics of cattle production in Brazil. *PloS one*, 11(1), e0147138. <https://doi.org/10.1371/journal.pone.0147138>
- Mokry, F. B., Buzanskas, M. E., de Alvarenga Mudadu, M., do Amaral Grossi, D., Higa, R. H., Ventura, R. V., ... & de Almeida Regitano, L. C. 2014. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC genomics*, 15(7), 1-9. <https://doi.org/10.1186/1471-2164-15-S7-S6>
- Montaldo, H. H., Casas, E., Sterman Ferraz, J. B., Vega-Murillo, V. E., & Román-Ponce, S. I. 2012. Opportunities and challenges from the use of genomic selection for beef cattle breeding in Latin America. doi:10.2527/af.2011-0029
- Nardone, A., Ronchi, B., Lacetera, N., Ranieri, M. S., & Bernabucci, U. 2010. Effects of climate changes on animal production and sustainability of livestock systems. *Livestock Science*, 130(1-3), 57-69. <https://doi.org/10.1016/j.livsci.2010.02.011>
- Osipova, L., Okello, M. M., Njumbi, S. J., Ngene, S., Western, D., Hayward, M. W., & Balkenhol, N. 2019. Using step-selection functions to model landscape connectivity for African elephants: accounting for variability across individuals and seasons. *Animal conservation* 22: 35-48. <https://doi.org/10.1111/acv.12432>
- Onzima, RB, Upadhyay, MR, Doekes, HP, Brito, L., Bosse, M., Kanis, E., ... & Crooijmans, RP 2018. Caracterização de todo o genoma de assinaturas de seleção e execuções de homozigidade em raças de cabras de Uganda. *Fronteiras em genética*, 9, 318. <https://doi.org/10.3389/fgene.2018.00318>
- Paim, T. D. P., Ianella, P., PAIVA, S. R., CAETANO, A. R., & PIMENTEL, C. M. M. 2018. Detection and evaluation of selection signatures in sheep. *Pesquisa Agropecuária Brasileira*, 53:527-539. <https://doi.org/10.1111/age.12907>

- Paim, T. D. P., Hay, E. H. A., Wilson, C., Thomas, M. G., Kuehn, L. A., Paiva, S. R., ... & Blackburn, H. 2020. Genomic breed composition of selection signatures in Brangus beef cattle. *Frontiers in genetics*, 11, 710. <https://doi.org/10.1590/s0100-204x2018000500001>
- Pariset, L., Joost, S., Gargani, M., & Valentini, A. 2012. Landscape genomics in livestock (No. BOOK\_CHAP, p. 360)
- Pariset, L., Joost, S., Marsan, P. A., & Valentini, A. 2009. Econogene Consortium (EC): Landscape genomics and biased Fst approaches reveal single nucleotide polymorphisms under selection in goat breeds of North-East Mediterranean. *BMC Genet*, 10(7). <https://doi.org/10.1186/1471-2156-10-7>
- Peripolli, E., Stafuzza, N. B., Munari, D. P., Lima, A. L. F., Irgang, R., Machado, M. A., ... & da Silva, M. V. G. B. 2018. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. *BMC genomics*, 19(1), 1-13. <https://doi.org/10.1186/s12864-017-4365-3>
- Peripolli, E., Stafuzza, N. B., Amorim, S. T., de Lemos, M. V. A., Grigoletto, L., Kluska, S., ... & Baldi, F. 2020. Genome-wide scan for runs of homozygosity in the composite Montana Tropical® beef cattle. *Journal of Animal Breeding and Genetics*, 137(2), 155-165. <https://doi.org/10.1111/jbg.12428>
- Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. 2012. Runs of homozygosity and population history in cattle. *BMC genetics* 13:70. <https://doi.org/10.1186/1471-2156-13-70>
- Rendel, J. M., & Robertson, A. 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *Journal of Genetics*, 50(1), 1-8.
- Reist-Marti, S. B., Simianer, H., Gibson, J., Hanotte, O., & Rege, J. E. O. 2003. Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. *Conservation Biology*, 17(5), 1299-1311. <https://doi.org/10.1046/j.1523-1739.2003.01587.x>
- Rojas-Downing, M. M., Nejadhashemi, A. P., Harrigan, T., & Woznicki, S. A. 2017. Climate change and livestock: Impacts, adaptation, and mitigation. *Climate Risk Management*, 16, 145-163. <https://doi.org/10.1016/j.crm.2017.02.001>
- Savolainen, O., Lascoux, M., & Merilä, J. 2013. Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14 (11), 807-820. <https://doi.org/10.1038/nrg3522>
- Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., ... & Lander, E. S. 2006. Positive natural selection in the human lineage. *science*, 312(5780), 1614-1620. DOI: 10.1126/science.1124309
- Schepers, J. M., & Weigel, K. A. 2012. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. *Animal Frontiers*, 2(1), 4-9. <https://doi.org/10.2527/af.2011-0032>
- Silva, A. M. D., Alencar, M. M. D., Regitano, L. C. D. A., Oliveira, M. C. D. S., & Barioni Júnior, W. 2007 Artificial infestation of *Boophilus microplus* in beef cattle heifers of four



genetic groups. *Genetics and Molecular Biology*, 30(4), 1150-1155. <https://doi.org/10.1590/S1415-47572007000600020>

Somenzi, E; Ajmone-Marsan, P; Barbato, Mario 2020 Identification of Ancestry Informative Marker (AIM) Panels to Assess Hybridisation between Feral and Domestic Sheep. *Animals*: 10:582; <https://doi.org/10.3390/ani10040582>

Tolone, M, Mastrangelo, S, Rosa, AJM and Portolano, B. 2012. Genetic diversity and population structure of Sicilian sheep breeds using microsatellite markers. *Small Ruminant Research* 102, 18–25. <https://doi.org/10.1016/j.smallrumres.2011.09.010>

Utsunomiya, Y. T., Pérez O'brien, A. M., Sonstegard, T. S., Sölkner, J., & Garcia, J. F. 2015. Genomic data as the “hitchhiker's guide” to cattle adaptation: tracking the milestones of past selection in the bovine genome. *Frontiers in Genetics*, 6, 36. <https://doi.org/10.3389/fgene.2015.00036>

VanRaden, P. M., & Cooper, T. A. 2015. Genomic evaluations and breed composition for crossbred US dairy cattle. *Interbull bulletin*, (49).

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics selection evolution*, 34(3), 275-305. DOI: 10.1051/gse:2002009

VonHoldt, B. M., Brzeski, K. E., Wilcove, D. S., & Rutledge, L. Y. 2018. Redefining the role of admixture and genomics in species conservation. *Conservation Letters*, 11(2), e12371. <https://doi.org/10.1111/conl.12371>

Wheeler, T. L., Cundiff, L. V., Shackelford, S. D., & Koohmaraie, M. 2005. Characterization of biological types of cattle (Cycle VII): Carcass, yield, and longissimus palatability traits. *Journal of animal science*, 83(1), 196-207. <https://doi.org/10.2527/2005.831196x>

Williams, J. L. 2005. The use of marker-assisted selection in animal breeding and biotechnology. *Revue Scientifique et Technique-Office International des Epizooties*, 24(1), 379. <http://dx.doi.org/10.20506/rst.24.1.1571>

Wray, N. R., Kemper, K. E., Hayes, B. J., Goddard, M. E., & Visscher, P. M. (2019). Complex trait prediction from genome data: contrasting EBV in livestock to PRS in humans: genomic prediction. *Genetics*, 211(4), 1131-1141.

Wright, S. 1921. Systems of mating. *Genetics* 6: 111– 178.

Zavarez, Ludmilla B. et al. 2015. Assessment of Autozygosity in Nellore Cows (*Bos Indicus*) through High-Density SNP Genotypes. *Frontiers in Genetics* 6(January): 1–8. <http://journal.frontiersin.org/journal/10.3389/fgene.2015.00005/abstract>.

Zilli, M., Scarabello, M., Soterroni, A. C., Valin, H., Mosnier, A., Leclere, D., ... & Ramos, F. M. 2020. The impact of climate change on Brazil's agriculture. *Science of the Total Environment*, 740, 139384. <https://doi.org/10.1016/j.scitotenv.2020.139384>

Zinovieva, N. A., Dotsev, A. V., Sermyagin, A. A., Deniskova, T. E., Abdelmanova, A. S., Kharzinova, V. R., ... & Brem, G. 2020. Selection signatures in two oldest Russian native cattle

breeds revealed using high-density single nucleotide polymorphism analysis. *Plos one*, 15(11), e0242200. <https://doi.org/10.1371/journal.pone.0242200>

## **OBJETIVOS**

### ***Objetivos Gerais***

- Compreender a formação genômica e as consequências do processo seletivo na população composta Montana Tropical®.

### ***Objetivos específicos***

- Caracterizar a estrutura da população do composto Montana Tropical®;
- Identificar diferenças em assinaturas de seleção e composição de raças fundadoras no composto Montana Tropical® ao longo de gerações;
- Detectar assinaturas de seleção no composto Montana Tropical® e identificar genes presentes nas regiões de assinatura.

## CAPÍTULO II - ESTRUTURA POPULACIONAL E ASSINATURA DE SELEÇÃO AO LONGO DE GERAÇÕES NO COMPOSTO MONTANA TROPICAL®

### RESUMO

O conhecimento da estrutura populacional e composição racial é importante para gerir as informações da raça dentro do programa de melhoramento genético. O Composto Montana Tropical® foi desenvolvido cruzando animais de quatro tipos biológicos ou grupos raciais diferentes. Ao usar os dados genômicos é possível realizar diversos estudos de estrutura de população e assinaturas de seleção. Foram utilizados dados de pedigree e genótipos (51962 SNP) de 3215 animais do Composto Tropical Montana® para prever a composição dos grupos biológicos do pedigree com dados genômicos e como essa composição se relaciona com as gerações dentro da raça composta; (2) caracterizar a distribuição de corridas de homozigidade (ROH) em relação as gerações dentro da raça composta; (3) estimar e comparar o coeficiente de endogamia (F) com base nas informações de marcadores genômicos ( $F_{ROH}$ ) e pedigree ( $F_{PED}$ ); (4) identificar regiões genômicas de ilhas de ROH e (5) identificar características relacionadas em cada região genômica significativa sob seleção. A estrutura de população (PCA e ADMIXTURE) não está relacionada com a composição racial do pedigree e possui forte relação com o número de gerações dentro da raça. Ocorreu aumento na  $F_{ROH}$  à medida que o número de gerações equivalentes aumentou. As estimativas de  $F_{ROH}$  na população de Montana também demonstram que a partir da geração equivalente 5 o número de animais com  $F_{ROH} = 0$  é reduzido. Acredita-se que utilizar os resultados de PCA no programa de melhoramento genético da raça para correção da estrutura de população seja uma alternativa interessante, em substituição a correção pelas proporções NABC baseada no pedigree. A população estudada apresentou assinaturas de seleção em regiões dos cromossomos 3, 4, 6, 20 e 26. A frequência de observação dos marcadores em homozigose nessas regiões aumentou conforme as gerações dentro da raça, demonstrando ser realmente um processo de seleção que está ocorrendo após a formação da raça composta. Grande parte dessas regiões de assinaturas de seleção apresentam QTLs de qualidade de carcaça, demonstrando que os objetivos da criação da raça em desenvolver um animal a pasto com qualidade de carne, estão sendo atingidos em nível genômico ao longo das gerações da raça.

**Palavras-chave:** Adaptação; pedigree; Genômica.

## ABSTRACT

Knowledge of population structure and racial composition is important for managing breed information within the breeding program. Composite Montana Tropical® was developed by crossing animals of four different biological types or racial groups. By using the genomic data, it is possible to carry out several studies of population structure and selection signatures. Pedigree and genotype data (51962 SNP) from 3215 animals from the Composite Tropical Montana® were used to predict the composition of biological groups in the pedigree with genomic data and how this composition relates to generations within the composite breed; (2) characterize the distribution of races of homozygosity (ROH) in relation to generations within the composite breed; (3) estimate and compare the inbreeding coefficient (F) based on information from genomic markers (FROH) and pedigree (FPED) ; (4) identify genomic regions of islands of ROH and (5) identify related traits in each significant genomic region under selection. The population structure (PCA and ADMIXTURE) is not related to the racial composition of the pedigree and has a strong relationship with the number of generations within the breed. There was an increase in FROH as the number of equivalent generations increased. Estimates of FROH in the Montana population also show that from the equivalent generation 5 onwards the number of animals with FROH= 0 is reduced. It is believed that using PCA results in the breed genetic improvement program to correct the population structure is an interesting alternative to replace the correction by NABC proportions based on pedigree. The population studied showed selection signatures in regions of chromosomes 3, 4, 6, 20 and 26. The frequency of observation of markers in homozygosity in these regions increased with the generations within the breed, demonstrating that it is really a selection process that is occurring after the formation of the composite race. Most of these selection signature regions have carcass quality QTLs, demonstrating that the breed's breeding goal of developing a pasture animal with meat quality, are being achieved at the genomic level over the generations of the breed.

**Key-words:** Adaptation, pedigree, genomics

## INTRODUÇÃO

O conhecimento da estrutura populacional e composição racial é importante para gerir as informações da raça dentro do programa de melhoramento genético e tomar decisões mais assertivas e acuradas na conservação da diversidade, seleção de características de interesse e avaliar adaptabilidade de raças compostas a determinado ambiente de produção (Reist-Marti *et al.*, 2003, Kuehn *et al.*, 2011).

A estimativa de composição racial com base no pedigree é um método convencionalmente utilizado e acessível, no entanto está sujeito a erros de anotação, dados incompletos ou ausentes prejudicando a confiabilidade da informação (Ron *et al.*, 1996). Os avanços nas tecnologias genômicas permitiram inovações e implementações nos métodos de estimação da composição racial dos animais domésticos (Huang *et al.*, 2014). Métodos que utilizaram dados de genoma completo foram utilizados satisfatoriamente em espécies como bovinos, suínos e ovelhas para predição e estimação da composição geral e a proporção racial genômica com coeficientes de regressão (Dodds *et al.*, 2014, Funkhouser *et al.*, 2017, Gobena *et al.*, 2018).

Ao usar os dados genômicos a estimativa da composição racial é menos sujeita a erros de entrada de informação e ao efeito de desvios de amostragem mendeliana, durante a gametogênese que afeta as estimativas com base no pedigree (Kuehn *et al.*, 2011). O uso desta tecnologia na predição da composição racial apresenta como desvantagem o maior custo para obtenção da informação (Gurgul *et al.*, 2014). Esse impacto do custo pode ser mitigado pela diversidade de informações que podem ser obtidas ao utilizar os mesmos dados genômicos, tais como teste de parentesco, estimar valores genéticos genômicos e rastrear doenças (Gobena *et al.*, 2018).

Outra informação que pode ser obtida com dados genotípicos é a identificação das corridas de homozigose (conhecidas como runs of homozygosity – ROHs) que são segmentos homozigotos contínuos herdados de ancestrais comuns (Ferençakovic *et al.*, 2011, Purfield *et al.*, 2012). As formações das ROHs podem ser influenciadas por endogamia, deriva genética, gargalo populacional, bem como seleção natural e artificial, ou seja, seu estudo pode caracterizar o histórico da estrutura populacional (Curik *et al.*, 2014).

De acordo com o tamanho das ROH é possível calcular o coeficiente de endogamia genômico e classificar a endogamia, sendo segmentos longos relacionados com endogamia recente e os curtos com endogamia mais antiga (McQuillan *et al.* 2008). Essas regiões dependendo da frequência que aparecem na população podem ser consideradas assinaturas de

seleção e permitem identificar características de importância produtiva que estão sob seleção ao longo do desenvolvimento da raça (Peripolli *et al.*, 2018).

O Composto Montana Tropical® foi desenvolvido cruzando animais de quatro tipos biológicos ou grupos raciais diferentes 1) Raças Zebuínas (N), 2) Raças taurinas adaptadas (A), 3) Raças Britânicas (B), e 4) raças da Europa continental (C) (Grigoletto *et al.*, 2020). As diferenças entre os grupos biológicos podem ser utilizadas para adequar o genótipo dos animais ao ambiente e aumentar a eficiência produtiva, permitindo melhores desempenhos em ambientes que seriam mais desafiadores para raças puras (Silva *et al.*, 2007).

Essa raça composta já foi objeto de estudo em trabalhos anteriores e demonstrou possuir traços de assinatura de seleção recente (Peripolli *et al.*, 2020). Utilizando a raça composta Montana Tropical®, buscando com este trabalho (1) pela primeira vez prever a composição dos grupos biológicos do pedigree com dados genômicos e como essa composição se relaciona com as gerações dentro da raça composta; (2) caracterizar a distribuição de corridas de homoziguidade (ROH) em relação as gerações dentro da raça composta; (3) estimar e comparar o coeficiente de endogamia (F) com base nas informações de marcadores genômicos ( $F_{ROH}$ ) e pedigree ( $F_{PED}$ ); (4) identificar regiões genômicas de ilhas de ROH e (5) identificar características relacionadas em cada região genômica significativa sob seleção.

## **MATERIAL E MÉTODOS**

### ***Animais***

Foram utilizados dados de pedigree e genótipos imputados (51962 SNP) de 3,215 animais do composto Tropical Montana®. As amostras foram provenientes de fazendas do Brasil (Cento-Oeste, Sul e Sudeste) e do Uruguai e fornecidos pelo Grupo de Melhoramento Animal e Biotecnologia da Faculdade de Zootecnia e Engenharia de Alimentos (Pirassununga, São Paulo, Brasil).

O controle de qualidade dos genótipos foi realizado no software Plink versão 1.09 (<https://www.cog-genomics.org/plink/1.9/>) (Purcell *et al.*, 2007). Após o controle de qualidade com filtragem de 0,5 para GC score (Gen Call), call rate amostra (>90%), call rate SNP (>90%), MAF (>1%) e poda por LD (--indep-pairwise 50 5 0.5) o conjunto retido para análise de dados foi de 40,634 SNP em 29 cromossomos autossômicos. Para as análises de ROH o filtro de MAF e a poda por LD não foram executadas devido a sua capacidade de afetar a detecção de ROH (Meyermens *et al.*, 2020), sendo o conjunto final de SNP de 51962 SNP dos 29 cromossomos autossômicos.

### ***Avaliação do Pedigree***

O arquivo de pedigree foi avaliado utilizando o pacote optiSel (Wellmann, 2017) no software R 3.4.2 (R Core Team, 2017). Os animais que apresentaram composição racial 1 para qualquer grupo biológico na escala N:A:B:C foram considerados ancestrais.

O número de gerações equivalente foi calculado através da equação  $g = \sum(1/2)^n$  em que  $g$  é o número de gerações equivalentes e  $n$  o número de gerações que separa o indivíduo de cada ancestral conhecido Welsh et al (2010).

### ***PCA***

O software Plink 1.9 foi utilizado para executar a Análise de Componentes Principais (PCA) baseada na matriz de relacionamento padronizada de variância para verificar estratificação e distância genética entre os indivíduos da população (Purcell *et al.*, 2007). As análises foram realizadas com até 10 componentes principais. Os autovetores individuais foram plotados ao longo dos diferentes eixos de um gráfico. A variância total explicada pelos componentes principais também foi estimada.

### ***ADMIXTURE***

Análise de cluster com base nos dados de genotipagem foi executada usando o método não supervisionado de máxima verossimilhança implementado no programa ADMIXTURE versão 1.3.0 (Alexander *et al.*, 2009). Valores K (parâmetro que descreve o número de subpopulações que compõem a população total em um conjunto de dados fornecido como entrada) de 2, 3 e 4 foram usados para o conjunto de dados para estabelecer os padrões de agrupamento na população (Ahmad *et al.*, 2020). Ao tentar encontrar o valor de  $k$  ideal da população, foi utilizado o K de 2 a 20 mas não foi encontrado um ponto que o CV fosse adequado para determiná-lo, pois o decaimento foi contínuo sem um ponto o mesmo aumentasse.

As saídas de  $K=2$  a 4 foram plotadas utilizando os pacotes ggplot2 (Wickham, 2016), forcats (Wickham, 2021), ggthemes (Arnold, 2021) e patchwork (Pedersen, 2020) no software R e os dados organizados por classes de gerações equivalentes ( $<2.5$ ,  $2.5 \leq \text{equiGen} < 4.5$  e  $>4.5$ ).

As análises de correlação, regressão e gráficos entre resultados de PCA, Gerações equivalentes, ADMIXTURE e composição NABC do pedigree foram executados utilizando os pacotes do software R Hmisc (Frank & Harrell, 2022), MASS (Venables & Ripley, 2002), ggplot2 (Wickham, 2016) e corrplot (Wei & Simko, 2019) respectivamente.

### ***Corridas de Homozigose***



As ilhas de ROH foram identificadas utilizando o software Plink 1.9 a partir da função --homozyg obedecendo os seguintes critérios (Saravanan *et al.*, 2020):

- I. uma janela deslizante de 50 SNPs ao longo do genoma (--homozyg-windows-snp);
- II. uma razão de sobreposição de janela homocigótica de 0.05(--homozyg-windows-threshold);
- III. um comprimento mínimo para considerar um segmento homocigoto de 1000 kb (--homozyg-kb);
- IV. um gap máximo entre SNPs homocigotos consecutivos de 1000 kb (--homozyg-gap);
- V. a densidade de um SNP por 100 kb (--homozyg-density);
- VI. requisito de conter pelo menos 20 SNPs (--homozyg-snp);
- VII. exigência de não permitir genótipos ausentes e não permitir heterocigotos (Caivio-Nasner *et al.*, 2021).

A estimativa de endogamia genômica foi calculada da seguinte forma:

$$FROH = \sum \frac{LROH}{LAUTO}$$

em que FROH é a medida de homocigosidade por animal, LROH é a soma da ROH por animal e LAUTO o comprimento total (2715853792) do genoma com referência ARS 1.2.

A incidência de ROH comum foi transformada para a frequência de classe de geração, dividindo-se pelo número de animais de cada classe de geração na análise. Testes de normalidade foram realizados e o limiar de frequência que define o 1% superior das observações para cada classe de geração equivalente foi determinado. As regiões homocigotas acima do limiar de frequência de cada classe do pedigree de geração equivalente (5% para todas as gerações, 4.3% para equiGen<2.5, 4.9% para 2.5<equiGen<4.5 e 12% para equiGen>4.5) foram consideradas como regiões selecionadas.

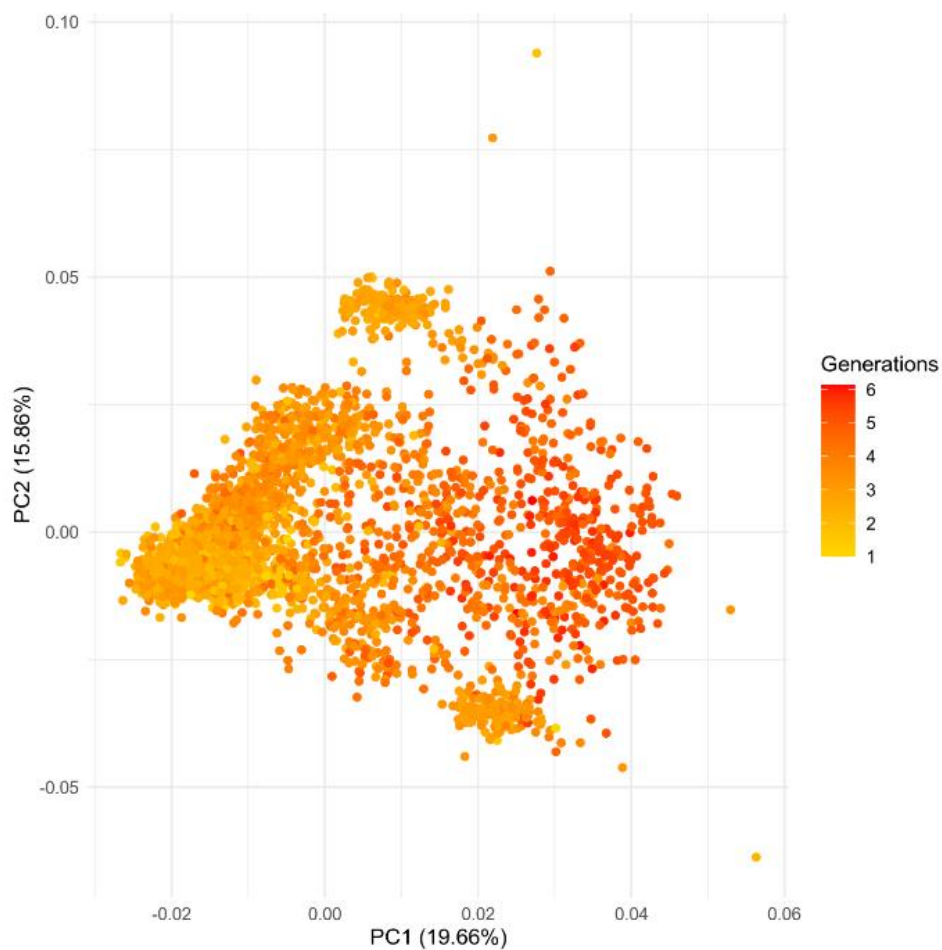
De acordo com o comprimento da ROH, é possível estimar o número de gerações traçadas até o ancestral comum, que gera a homocigosidade naquela região. As ROH foram classificadas em 4 categorias (1 = mais de 6 gerações, 2 = entre 2.5 e 4.5, 3 = entre 4.5 e 2.5 e 4 = menos de 2.5 gerações) utilizando a equação proposta por Curik *et al.* (2014):  $E(L IBD-H / gcA) = 100 / (2 gcA)$ , sendo  $E(L IBD-H / gcA)$  é o comprimento esperado de um haplótipo idêntico por descendente (IBD) (em centiMorgans – cM), e  $gcA$  é o número de gerações do ancestral comum. A conversão da métrica de taxa de recombinação para distância física (de cM para Mb) foi realizada utilizando a média dos resultados de Arias *et al.* (2009) e Weng *et al.* (2014). Com base no Curik *et al.* (2014), por exemplo, um ROH maior que 15.6 Mb provavelmente se originou de um ancestral comum há menos de 2.5 gerações.

Os genes em cada região selecionada foram pesquisados no Genome Data Viewer (NCBI platform) usando ARS-UCD1.3 (GCF\_002263795.2). Para identificar características relacionadas a genes localizados em cada região genômica significativa foi realizada busca no banco de dados do Animal Genome QTL Database (<https://www.animalgenome.org/cgi-bin/QTLdb/EC/index>).

## RESULTADOS

### *Estrutura da população*

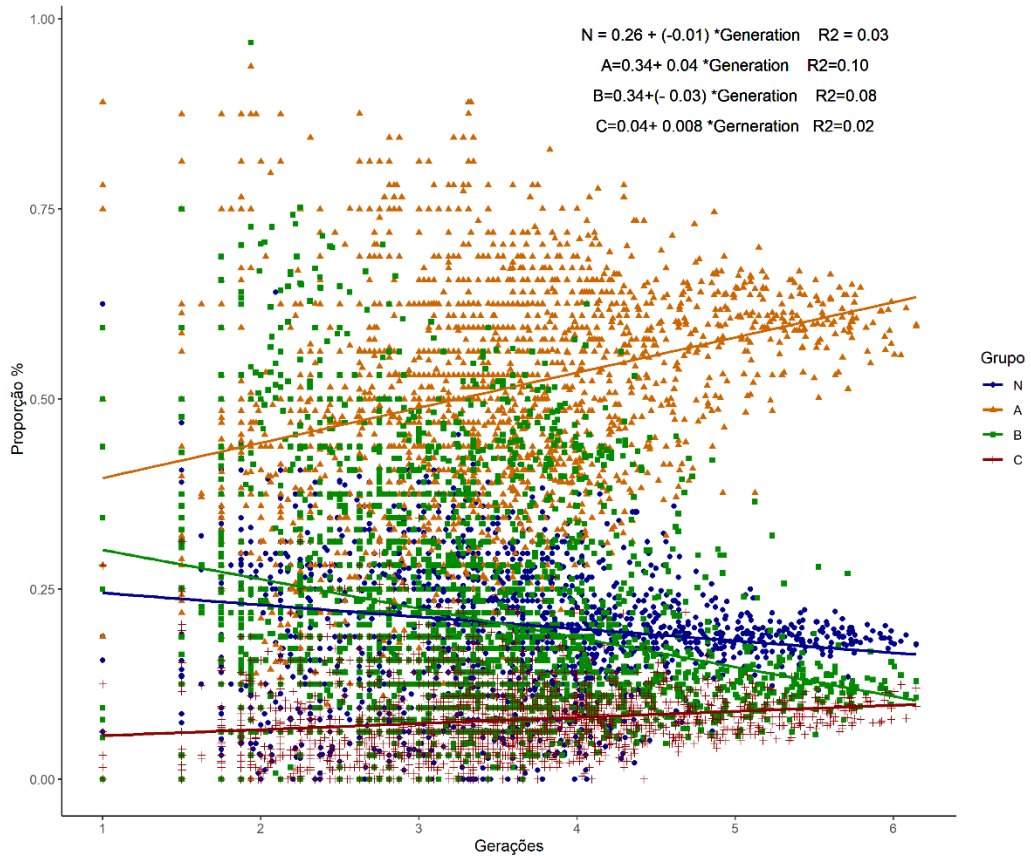
De acordo com a PCA, o primeiro componente principal (PC1) explicou 19.66% da proporção da variância entre os animais do Composto Tropical Montana®, sendo a estrutura hierárquica relacionada com a geração equivalente estimada pelo pedigree (Figura 1). O PC 2 explicou 15.86% da variação entre os animais.



**Figura 1-** Dois primeiros componentes da análise de componentes principais gerados com dados genômicos de 3215 animais Montana. Generations: corresponde ao número de gerações equivalentes calculados com o pedigree.

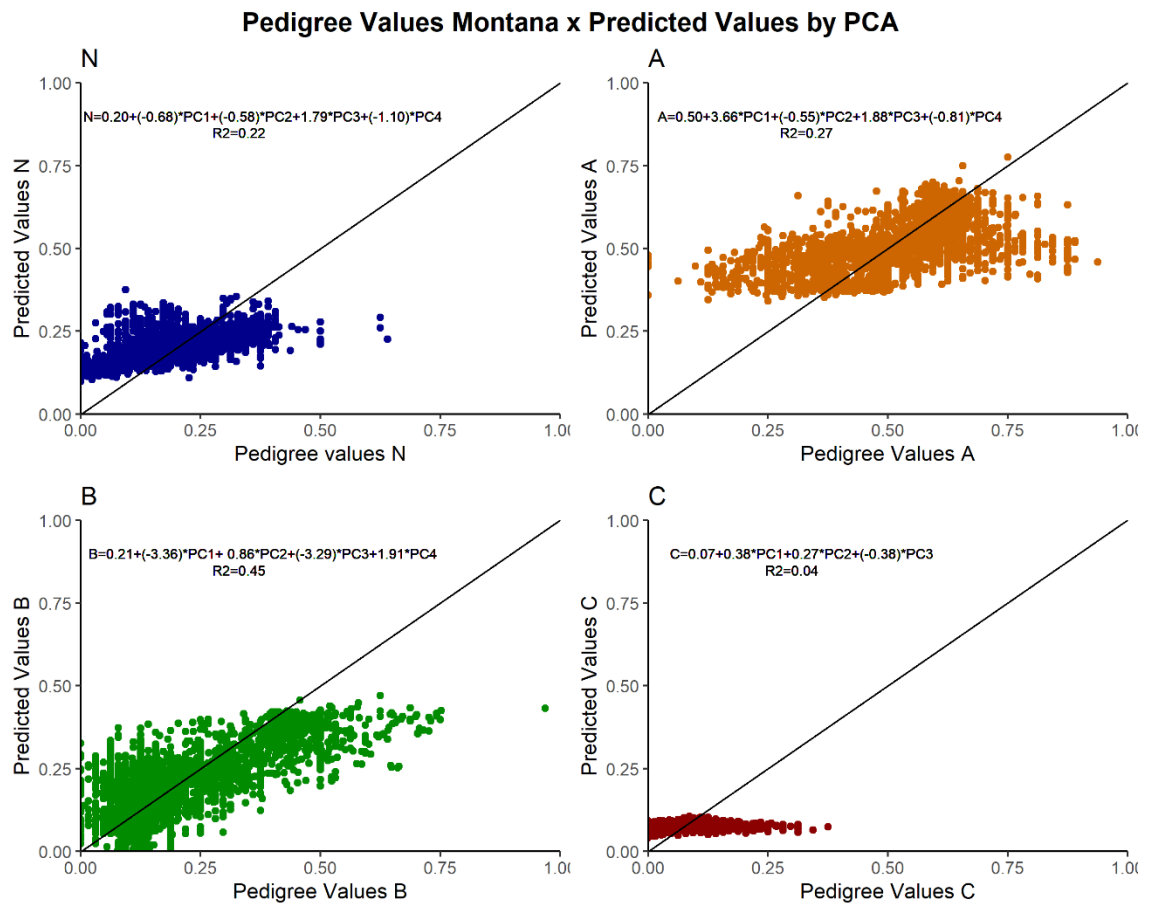
O resultado da correlação entre PC e proporção racial dos animais derivada do pedigree demonstrou correlação positiva para o PC1 e o grupo racial A ( $r=0.45$ ;  $p<0.001$ ) e negativa com grupo racial B ( $r=-0.44$   $p<0.001$ ). O PC3 apresentou correlação positiva com o grupo racial N ( $r=0.37$ ;  $p<0.001$ ) e negativa com o grupo racial B ( $r=-0.43$ ;  $p<0.001$ ). O PC 4 apresentou correlação negativa com o componente N do pedigree ( $r=-0.23$ ;  $p<0.001$ ) e positiva com o componente racial B ( $r=0.25$ ;  $p<0.001$ ).

As análises de correlação revelaram correlações positivas significativas entre os componentes biológicos do pedigree A ( $r=0.33$ ,  $p<0.001$ ) e C ( $r=0.16$ ,  $p<0.001$ ) com a gerações equivalentes estimadas na população em estudo. Correlações negativas e significativas foram observadas entre o componentes biológico N ( $r=-0.19$ ,  $p<0.001$ ) e B ( $r=-0.29$ ,  $p<0.001$ ) e as gerações equivalentes estimadas pelo pedigree. A análise regressão confirmou a tendência de aumento de 4% na composição de taurino adaptado (A) no pedigree dos animais a cada geração e decréscimo de 3% na proporção de raças britânicas (B) no pedigree para cada geração dentro da raça composta (Figura 2). Observou-se importante diferença de padrão de composição racial no pedigree dos animais acima de 4,5 gerações, sendo que menor número de gerações apresenta grande variabilidade nas proporções raciais e após o ponto de 4,5 gerações equivalentes a composição no sistema NABC apresenta menor variabilidade (Figura 2).



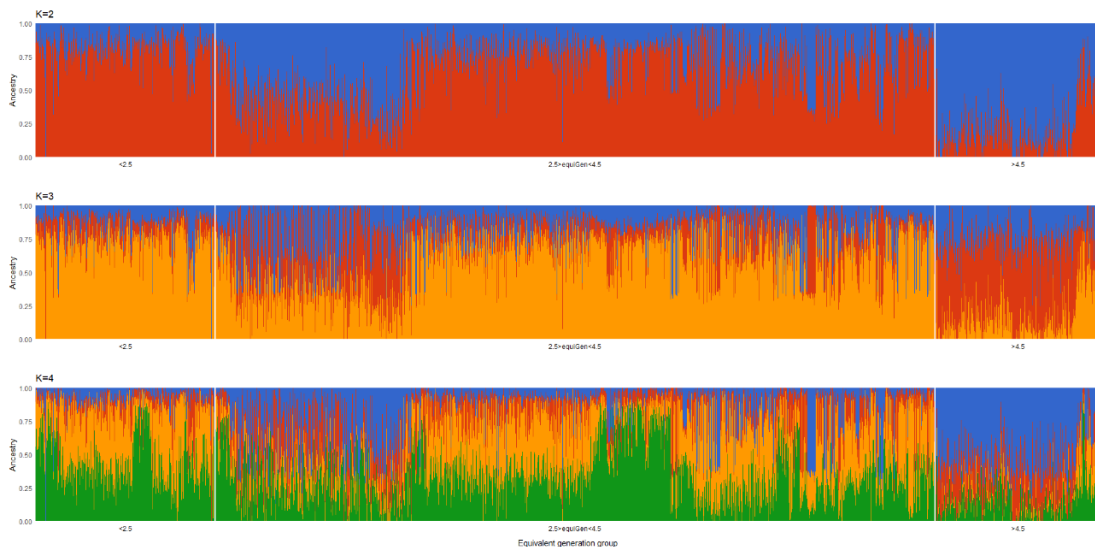
**Figura 2** - Proporção dos grupos Raciais NABC do pedigree ao longo de gerações equivalentes no Composto Montrana Tropical ®.

Foi executada uma análise de regressão para determinar o valor da proporção do grupo biológico do animal através dos valores de PCA. As equações de regressão com os valores dos PCs foram capazes de prever os componentes do NABC no pedigree, no entanto estas apresentaram coeficientes de regressão ( $R^2$ ) baixos não sendo adequados para empregar o modelo de predição do componente biológico do pedigree dos animais (Figura 3). Indicando que a estrutura de componentes principais representa uma estrutura da população diferente da composição racial do pedigree.



**Figura 3** – Valores observados no pedigree versus valores preditos pela regressão a partir dos componentes Principais 1, 2, 3 e 4.

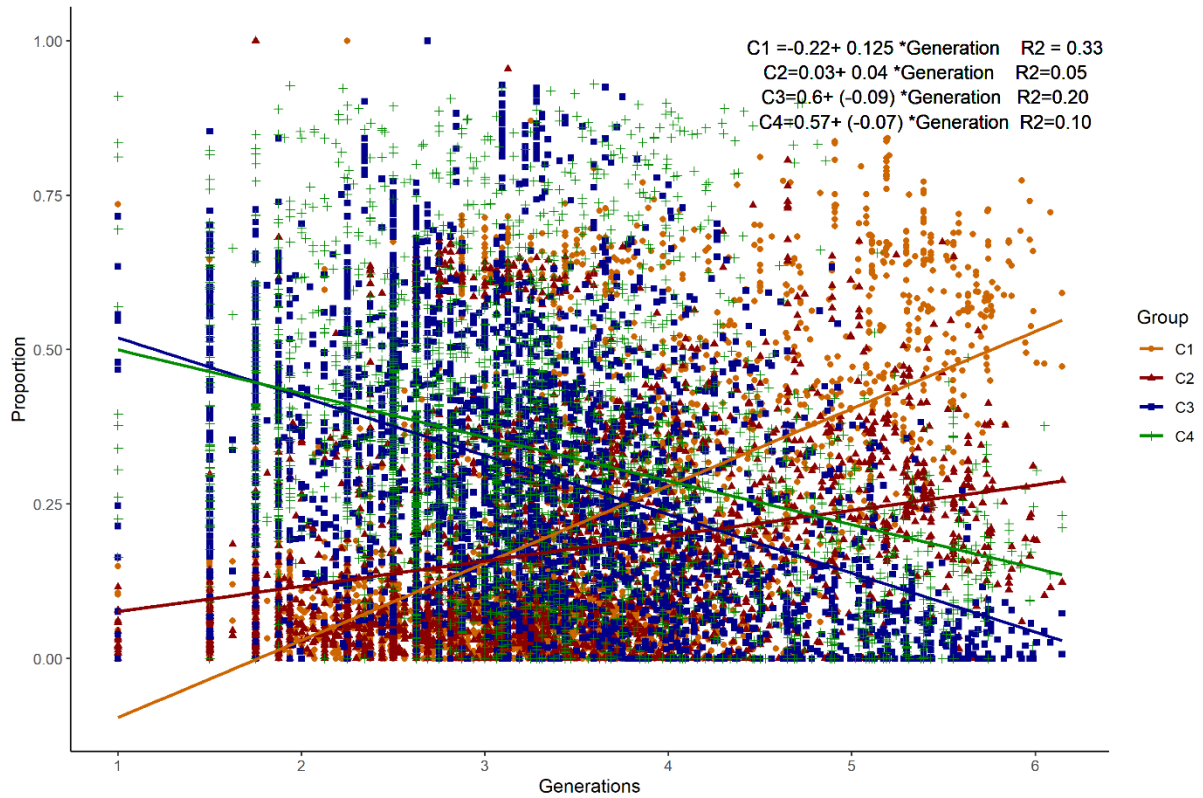
A análise no software ADMIXTURE demonstrou no valor de  $K=2$  que os animais com geração equivalente  $<2.5$  apresentam composição uniforme dos clusters e existindo no grupo de animais  $2.5 < \text{equiGen} < 4.5$  um agrupamento mais evidente dentro dessa classe com proporção menor do C2 comparado aos outros animais, e nas gerações equivalentes  $>4.5$  também pode ser observado um grupo de animais mais diferenciados com maiores proporções do C2. No  $K=3$  o padrão de agrupamento dentro das classes de gerações se manteve semelhante ao  $K=2$  com maior presença do C2 nos animais com geração maior que 4,5. Já os resultados de  $K=4$ , separa claramente um cluster (C1) para as gerações maiores que 4,5 e revelam grande variação de composição nas gerações entre 2,5e 4,5 (Figura 4).



**Figura 4.** - Estrutura populacional do Composto Tropical Montana® inferida pelo software ADMIXTURE. Cada animal é representado por uma única barra vertical dividida em k cores, em que k é o número de agrupamentos ancestrais assumidos, que é representado graficamente para  $k = 2-4$ . C1= Azul, C2= Vermelho, C3= laranja e C4=verde.

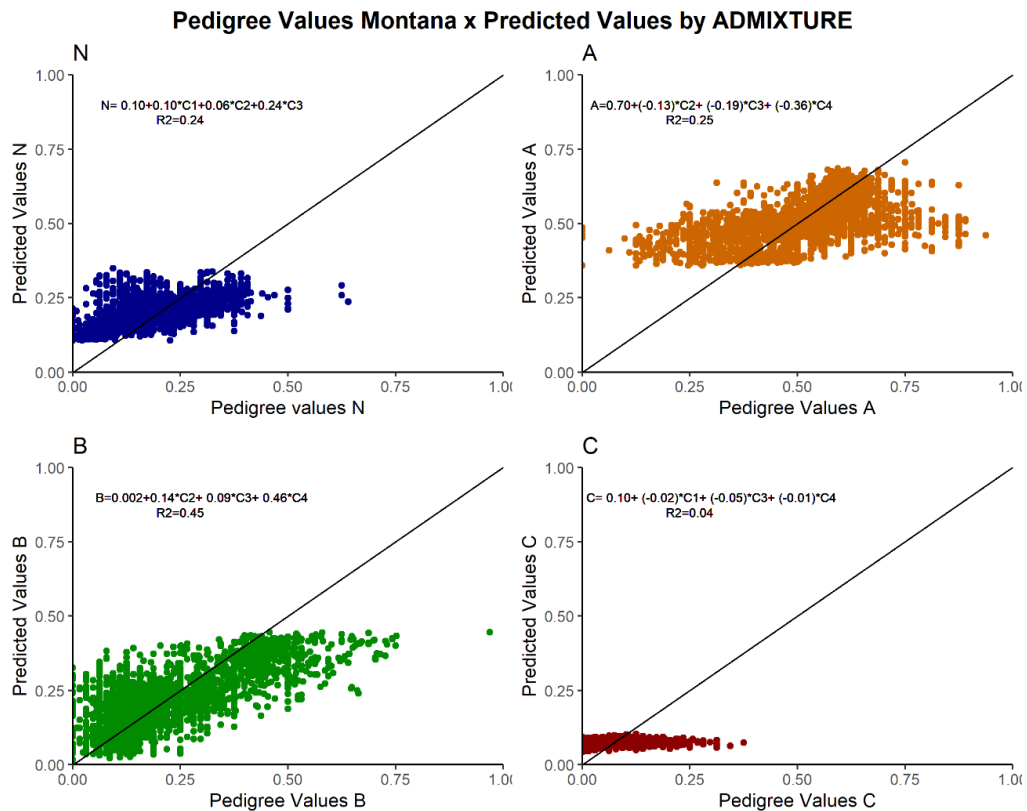
Os clusters apresentaram os seguintes resultados de correlação com os componentes do Pedigree, C1 apresentou correlação significativa e positiva com A ( $r=0.42$ ,  $p<0.001$ ) e negativa significativa com B ( $r=-0.45$ ,  $p<0.001$ ) o C3 apresentou correlação positiva e significativa com o componente N ( $r=0.43$ ,  $p<0.001$ ) e o C4 apresentou correlação positiva com B ( $r=0.65$ ,  $p<0.001$ ) e negativa significativa com A ( $r= -0.44$ ,  $p<0.001$ ).

A análise de entre os clusters do ADMIXTURE revelaram correlação positiva entre o C1( $r=0.58$ ,  $p<0.001$ ) e C2( $r=0.23$ ,  $p<0.001$ ) com as gerações equivalentes. E, os Clusters C3( $r=-0.45$ ,  $p<0.001$ ) e C4( $r=-0.42$ ,  $p<0.001$ ) apresentaram correlação negativa com as gerações equivalentes calculadas pelo pedigree. A análise regressão demonstrou a tendência de aumento de 12% de C1, 4% de C2 e redução de 9% de C3 e redução de 0.7% de C4 a cada geração. Não é possível afirmar que o C1 representa o grupo biológico A, pois o seu intercepto é negativo (-0.22) indicando que esse cluster surge dentro do Montana Tropical. O C2 aparentemente também não representa o grupo biológico A, pois os seus valores de intercepto (0.03) não condizem com o uso de A na estratégia de formação da raça. Já a representação de C3 pelo Nelore e C4 pelo Britânico parecem ser condizentes com o uso destes grupos biológicos ao longo das gerações (Figura 5).



**Figura 5-** Proporção dos clusters do ADMIXTURE ao longo de gerações equivalentes no Composto Montrana Tropical ®.

Foi executada uma análise de regressão para determinar o valor da proporção do grupo biológico do animal através dos valores ADMIXTURE (Figura 6). As equações de regressão com os valores dos PCs foram capazes de prever os componentes do NABC no pedigree, porém apresentaram coeficientes de regressão ( $r^2$ ) baixos não sendo adequados para empregar o modelo de predição do componente racial do pedigree dos animais.



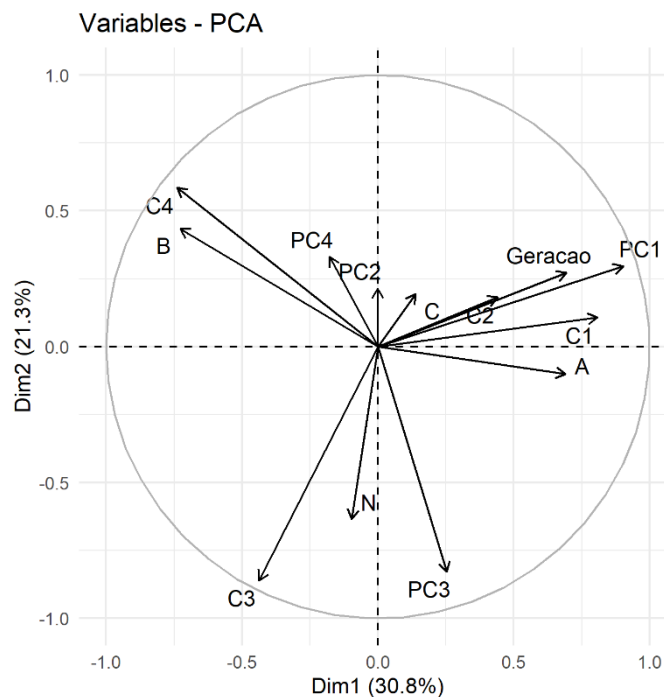
**Figura 6-** Regressão valores de Pedigree estimados pelos clusters C1, C2, C3 e C4.

O resultado dos dois métodos de estimar a composição NABC do pedigree foram semelhantes, ou seja, apesar de estatisticamente significantes com coeficiente de regressão baixo. O Componente B foi o que obteve a melhor estimativa tanto pelo PCA, quanto pelo ADMIXTURE ( $r^2=0.45$ ). Os coeficientes de regressão dos grupos N (PCA  $r^2= 0.22$  ADMIXTURE  $R^2= 0.24$ ) e A (PCA  $r^2= 0.27$  ADMIXTURE  $r^2= 0.25$ ) foram semelhantes. O componente C apresentou o mesmo  $r^2 = 0.04$  na equação com PCA e ADMIXTURE sendo o valor mais baixo dentre os grupos raciais. Esses resultados demonstram que a estrutura genômica do Composto Montana Tropical® é diferente da estrutura demonstrada pelo pedigree.

A análise de componentes principais envolvendo os resultados dos 4 primeiros componentes principais dos genótipos, os 4 clusters do ADMIXTURE e os componentes NABC do pedigree (Figura 7) demonstraram que os componentes de variação do PC1 dos genótipos, geração equivalente, C1 e C2 e o componente racial A apresentam direcionamento semelhante. Esse resultado condiz com as correlações descrições apresentadas da estrutura de populações, sugerindo que o C1 e o C2 juntos podem estar representando o comportamento de uso do grupo racial A, e o surgimento de um novo cluster dentro do Montana Tropical®, mas não podendo individualizá-los entre os clusters. C3, PC3 e N no pedigree também estão no mesmo sentido. C4 e B também no mesmo sentido. Portanto, os clusters do ADMIXTURE



indicam estar associados da seguinte forma: C1 com A e número de gerações dentro da raça, C2 com número de gerações dentro da raça, C3 com N e C4 com B.



**Figura 7** - Resultado dos dois primeiros componentes principais da análise de componentes principais utilizando os dados de composição racial do pedigree (N A B C), 4 grupos da análise de admixture (C1 C2 C3 C4) e os quatro primeiros componentes principais da análise de PCA (PC1, PC2, PC3, PC4).

### *Corridas de Homozigose*

Dos 3215 animais avaliados 3140 apresentaram algum segmento de ROH no genoma, totalizando 2349 blocos de haplótipos. A ROH média dos animais Montana foi de  $7,48 \pm 0,08$  Mb variando de 0 a 39 (Tabela 1). A soma da média de todos os seguimentos de ROH por animal foi de 61,21 resultando em comprimento médio de 8,18 MB.

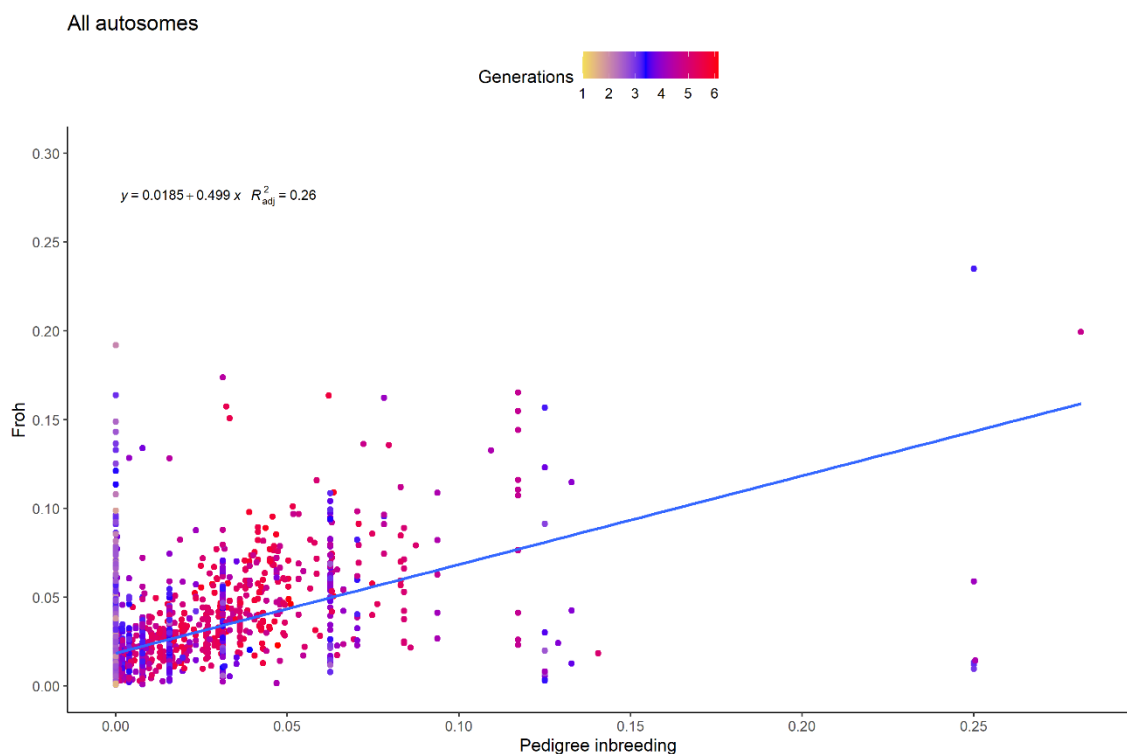
**Tabela 1-** Corridas de Homozigosidade (ROH) estatística por animal.

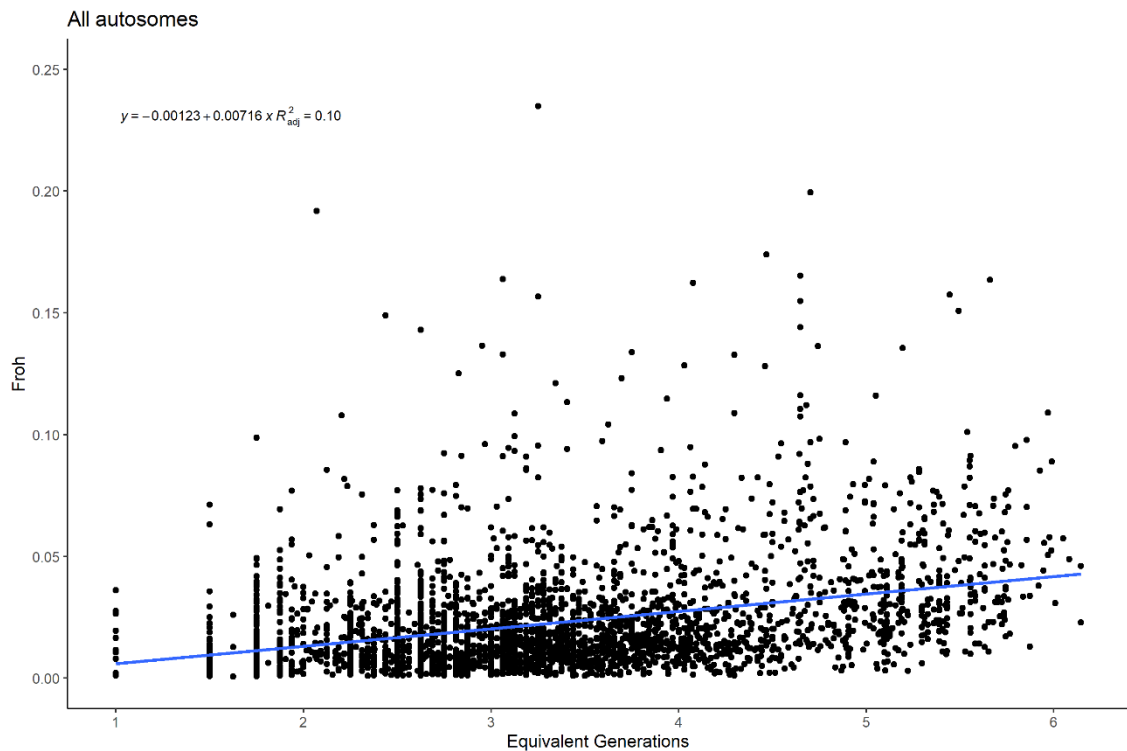
Total Sample	Média	SE	SD	Min	Max
<b>SROH</b>	61,21	1,10	62,89	0	845,53
<b>NROH</b>	7,48	0,08	4,97	0	39
<b>LROH</b>	8,18	0,05	7,81	1,01	109,11
<b>FROH</b>	0,02	0,0004	0,02	0	0,31

Erro padrão (*SE*), média (Mb), desvio padrão (*SD*), min (Mb), max (Mb), SROH (comprimento médio do genoma coberto por ROH), NROH (Número médio de ROH), LROH (comprimento médio do ROH em Mb) e FROH (coeficiente de endogamia).

O coeficiente de endogamia genômico (FROH) apresentou correlação positiva com os valores de endogamia calculados pelo pedigree (Figura 7). O FROH global do Composto Montana Tropical® calculado a partir de diferentes categorias de comprimento foi de 2% e a endogamia do pedigree 0.8%.

Ocorreu aumento na FROH à medida que o número de gerações equivalentes aumentou conforme observado na figura 8. As estimativas de FROH na população de Montana também demonstram que a partir da geração equivalente 5 o número de animais com FROH= 0 é reduzido.

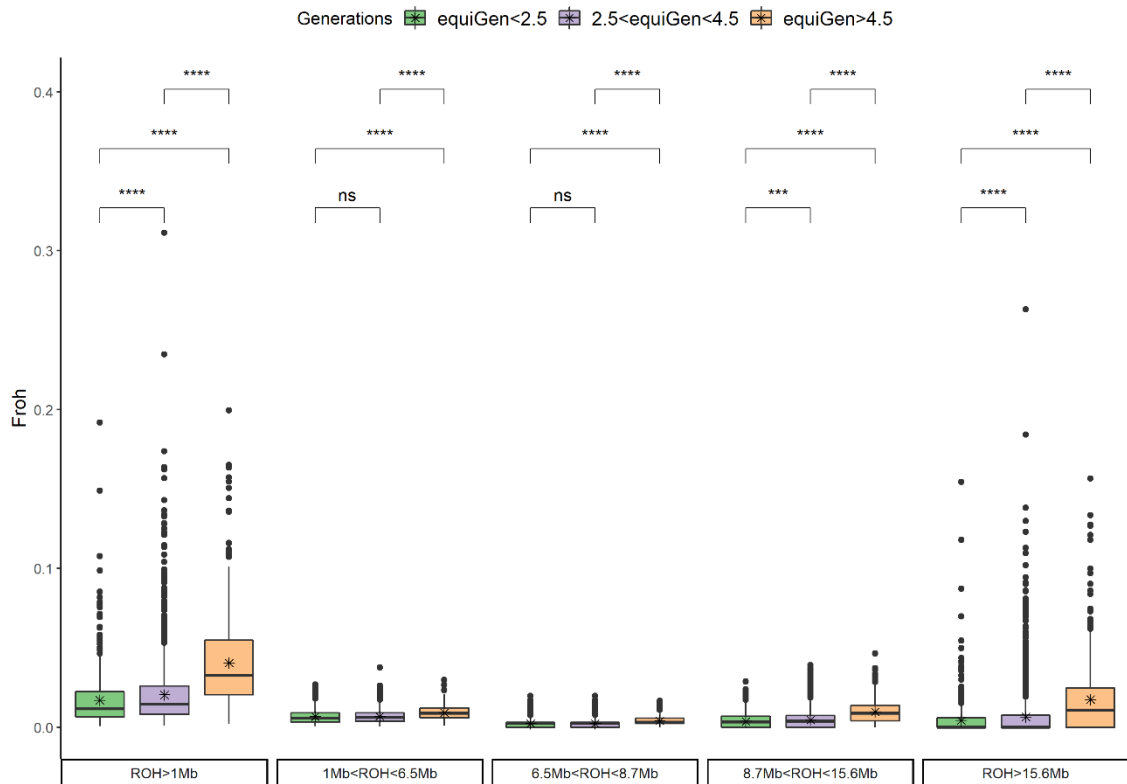
**Figura 8-** Correlação entre coeficiente de endogamia gnômico (Froh) e coeficiente de endogamia do pedigree (Fped).



**Figura 9-** Correlação encontrada entre o coeficiente de endogamia genômico e gerações equivalentes.

Dos animais estudados 95,89% apresentaram alguma sequência de ROH de tamanho  $1\text{Mb} < \text{ROH} < 6.5\text{Mb}$ , 33,81% de tamanho  $1\text{Mb} < \text{ROH} < 6.5\text{Mb}$ , 34% de tamanho  $8.7\text{Mb} < \text{ROH} < 15.7\text{Mb}$ . As maiores sequências de  $\text{ROH} > 15.6\text{ Mb}$  (com ancestral comum nas 2,5 gerações anteriores) foram identificadas 39,45% dos animais do Composto Montana Tropical® indicando surgimento de endogamia recente (Figura 8).

As gerações equivalentes  $> 4.5$  apresentaram valores significativamente maiores ( $p < 0,0001$ ) de coeficiente endogâmico (FROH) que todas as gerações especialmente nas corridas de homozigose  $> 15.6\text{ Mb}$ .

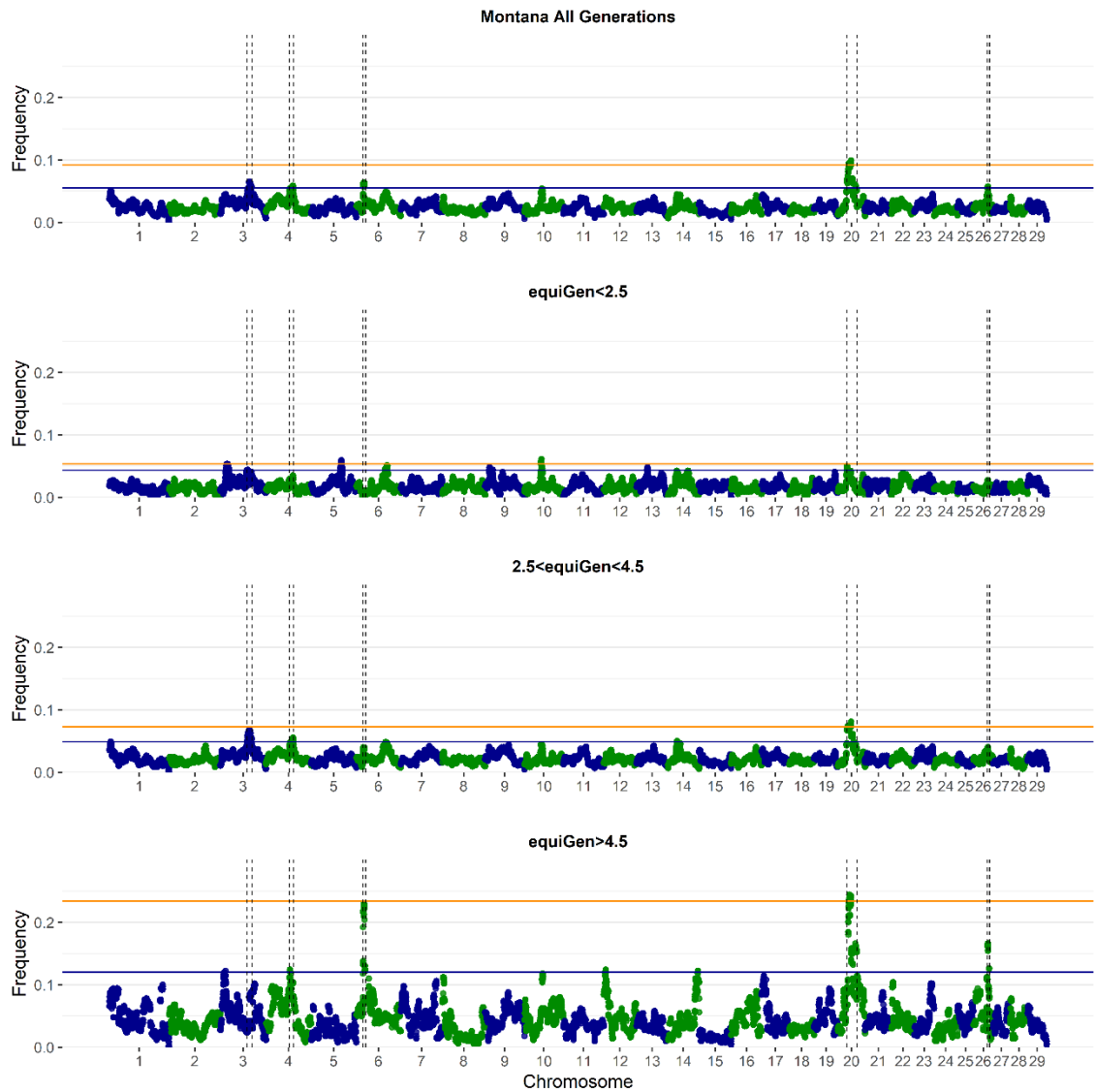


**Figura 10-** Endogamia genômica baseada em corridas de homozigose ( $F_{ROH}$ ) por classes de geração equivalente e por classes de comprimento ROH. Os resultados da comparação do teste  $t$  são mostrados no topo (ns: não significativo; \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ ; \*\*\*\* $p < 0,0001$ ).

Em cada classe de geração equivalente, foram localizadas 5 regiões de maior frequência em homozigose sobrepostas (Figura 11), BTA 3 (posição: 70241129-83472550), BTA 4 (posição: 62316125-73792670), BTA 6 (posições: 16674075- 24813978), BTA 20 (posição: 23007509-51003225) e BTA 26 (posição: 36413577- 42757206).

O BTA3 não foi identificado como assinatura de seleção nas gerações maiores que 4,5. Portanto, essa região pode ser relacionada com uma pressão de seleção que vem sendo aplicada nos últimos 20 anos (aproximadamente 4 gerações) e que não era utilizada anteriormente.

As ilhas de ROH encontradas em BTA4, BTA 6, BTA 20 e BTA 26 apresentaram a sua frequência em homozigose crescendo conforme o aumento do número de gerações equivalentes. Assim, isso indica que realmente essas regiões estão sendo selecionadas dentro da raça composta.



**Figura 11** - A distribuição de ROH entre autossomos nas três classes de gerações equivalentes. O eixo x representa a coordenada genômica e o eixo y exibe a frequência de sobreposição de ROH entre os indivíduos.

Nas regiões identificadas como ilhas de ROH, foram encontrados 1272 genes (Tabela 2) e alguns QTLs já relatados na literatura relacionados com características como qualidade de carcaça, resistência a doenças e produção e qualidade do leite em resumo na tabela 2.

**Tabela 2** - Regiões de homozigose observadas em animais Composto Montana Tropical® e identificação dos genes subjacentes aos QTL em cada região.

Chr	Start regions (Mb)	End regions (Mb)	n Genes	Genes	QTL traits
3	70.24	83.47	1069	<a href="https://drive.google.com/drive/folders/13dEJt0kNbsZi5zJ5_6HcnzH6fPdT4dK?usp=sharing">https://drive.google.com/drive/folders/13dEJt0kNbsZi5zJ5_6HcnzH6fPdT4dK?usp=sharing</a>	Força de cisalhamento, teor de ácido caproico no leite, Quantidade de tecido conjuntivo, pontuação de marmoreio.
4	62.31	73.79		ROH	Teor de ácido mirístico do leite, porcentagem de gordura do leite, pontuação de marmoreio, força de cisalhamento, doença da linha branca, quantidade de tecido conjuntivo, dermatite digital, teor de acetona do leite.
6	16.67	24.81	38	RPL34, LEF1, HADH, CYP2U1, SGMS2, TRNAC-GCA, TRNAG-CCC, PAPSS1, DKK2, GIMD1, AIMP1, TBCK, NPNT, GSTCD, INTS12, ARHGEF38, PPA2, TET2, CXXC4, TACR3, CENPE, BDH2, SLC9B2, NHEDC1, CISD2, UBE2D3, MANBA, NFKB1, SLC39A8, BANK1, PPP3CA, EMCN, DDIT4L, H2AFZ, DNAJB14, LAMTOR3, TRNAS-GGA, DAPP1	Pontuação de marmoreio, Rendimento longe do leite, Porcentagem de proteína do leite, Rendimento da proteína do leite, Força de cisalhamento, Quantidade de tecido conjuntivo, Rendimento do leite.
20	23	51	14	ANKRD55, IL6ST, MIR2359, IL31RA, DDX4, SLC38A9, PLPP1, MTREX, DHX29, CCNO, MCIDAS, CDC20B, MIR449C, MIR449B, MIR449A, GPX8, GZMA, GZMK, ESM1, SNX18, HSPB3, ARL15,	Produção de leite, porcentagem de gordura do leite, porcentagem de proteína do leite.

				<p> NDUFS4, FST,  MOCS2, ITGA2,  ITGA1, PELO , ISL1,  PARP8, EMB, HCN1,  MRPS30, FGF10,  NNT, PAIP1,  C20H5orf34,  TMEM267, CCL28 ,  HMGCS1, TRH ,  NIM1K, ZNF131,  SELENOP,  CCDC152, GHR ,  FBXO4,  C20H5orf51, OXCT1,  TRNAY-GUA,  PLCXD3, C6,  MROH2B, C7,  CARD6, RPL37,  PRKAA1, TTC33,  PTGER4, TRNAE-  CUC, DAB2, C9,  FYB1, RICTOR,  OSMR, LIFR,  EGFLAM, GDNF,  WDR70, NUP155,  MIR2360, CPLANE1,  NIPBL, SLC1A3,  RANBP3L, NADK2,  SKP2, LMBRD2,  UGT3A2, CAPSL,  IL7R, SPEF2, PRLR,  AGXT2, DNAJC21,  BRIX1, RAD1,  TTC23L, RAI14,  C1QTNF3, AMACR,  SLC45A2, RXFP3,  ADAMTS12,  TRNAT-UGU, TARS,  NPR3, SUB1,  TRNAG-CCC, ZFR ,  MTMR12, TRNAW-  CCA, GOLPH3,  PDZD2, C20H5orf22,  DROSHA , CDH6,  TRNAC-GCA, CDH9  , TRNAC-ACA,  TRNAS-GGA,  TRNAF-AAA,  CDH10 </p>	
--	--	--	--	---	--

26	36.41	42.75	50	GFRA1, CCDC172, PNLIPRP3, PNLIP, PNLIPRP2, C26H10orf82, HSPA12A, MIR2395, ENO4, SHTN1, VAX1, KCNK18, SLC18A2, PDZD8, EMX2, TRNAC-AC, RAB11FIP2, FAM204A, PRLHR, CACUL1, TRNAG- CCC, NANOS1, EIF3A, FAM45A, SFXN4, PRDX3, GRK5, TRNAC- GCA, RGS10, TIAL1, BAG3, INPP5F, MCMBP, SEC23IP, PLPP4, WDR11, FGFR2, ATE1, NSMCE4A, TACC2, MIR2396, BTBD16, PLEKHA1, HTRA1, DMBT1, SPADH2, SPADH1, C26H10orf120, FAM24A, C26H10orf88	Teor de ácido tetracosanoico do leite, Quantidade de tecido conjuntivo, suscetibilidade a doenças respiratórias bovinas, pontuação de marmoreio, quantidade de tecido conjuntivo.
----	-------	-------	----	--	--

## DISCUSSÃO

A estrutura populacional do Composto Montana Tropical® demonstrou nas análises de componentes principais distanciamento genômico dos animais de gerações mais avançadas. Os resultados da análise de admixture demonstraram que os animais de geração equivalente maior que 4,5 apresentaram perfil de composição genômico mais homogêneo quando comparados com as gerações iniciais.

Após o desenvolvimento de uma raça composta, no mínimo 5 gerações são necessárias para formar um novo perfil genômico em um composto de duas raças (Paim *et al.*, 2020, Hay *et al.*, 2022). A estrutura genômica da raça Montana é muito mais complexa pelo uso de múltiplas raças na sua formação, e mesmo assim, observa-se maior homogeneidade e a formação de um cluster específico da raça próxima a quinta geração.

No composto Montana Tropical® constatou-se que as equações de predição utilizando resultados de PCA e ADMIXTURE não foram capazes de predizer os componentes do pedigree



da população, demonstrando que existe diferença na estrutura genética do pedigree e estrutura genômica no composto. A baixa capacidade de predizer a composição racial do pedigree no Montana Tropical® e discrepância encontrada entre a composição genômica e de pedigree em outros trabalhos com composto de duas raças, pode ser devido o pedigree estar incompleto ou incorreto, ou por efeito da amostragem mendeliana durante a recombinação na meiose (Kuehn *et al.*, 2011, Gobena *et al.*, 2018).

O componente taurino adaptado apresenta proporção muito maior nas gerações mais avançadas no Montana, representado no pedigree e pelo C1 e C2 nas análises de ADMIXTURE que podem estar explicando a composição genômica desse grupo biológico. E, os componentes Nelore e Britânico apresentam redução na proporção do pedigree e na composição genômica do C3 e C4. O que está de acordo com os resultados descritos por Grigoletto *et al.*, 2020, e demonstram troca do uso do componente zebuino nelore pelo componente adaptado, explicado pela preferência por características de qualidade de carne que é proporcionado pelo uso de taurinos Adaptados.

Para indivíduos cujos ancestrais se originam em diferentes populações, sua composição genética exibe múltiplas ancestralidades associadas a múltiplos agrupamentos genéticos ou populações diferentes que podem ser descritos por modelos de mistura (Pritchard *et al.*, 2000, Alexander *et al.*, 2009) Isso foi provado com êxito na identificação a composição racial em cruzamento de raças compostas com número de fundadores com 2 e 3 raças, e as ferramentas capazes de estimar a diferenciação dos clusters conforme a proporção dos fundadores através de regressão (Akanno *et al.*, 2017, Gobena *et al.*, 2018, Hay *et al.*, 2022), no entanto o Montana Tropical® apresenta histórico de uso com grande quantidade de raças na sua formação, representados pelos grupos biológicos NABC, tornando a determinação do grupo racial que o cluster representa mais complexo.

O uso de matrizes de relacionamento baseados em SNPs e estrutura de populações nas estimativas de valores genéticos genômicos possibilita melhorar os resultados das avaliações, pois conseguem capturar as diferenças de amostragem mendeliana e os relacionamentos mais distantes do que o pedigree (Hayes *et al.*, 2009). Desta forma, considerar o uso de informações de estrutura populacional utilizando análises como PCA pode prover melhorias nas estimativas de valores genéticos no Composto Montana Tropical®.

Os resultados dos níveis de homoziguidade encontrados na população do Composto Montana Tropical demonstraram valor de ROH ligeiramente superior aos relatados em outro estudo de ROH em população do composto Montana Tropical® (Peripolli *et al.*, 2020). No entanto os níveis encontrados no estudo atual ainda são considerados baixos (Purfield *et al.*, 2012).

Como esperado a FPED (0.8%) subestimou os resultados de endogamia da população quando comparado com FROH (2%), isso provavelmente se deve as falhas do uso de informações provindas do pedigree, confirmando a importância do uso de dados genômicos para melhor entender o cenário de endogamia real da população (Caivio-Nasner *et al.*, 2021). A FROH é um método Genômico que fornece uma das medidas mais confiáveis do verdadeiro valor de endogamia em populações animais, pois considera o relacionamento passado e recente dos indivíduos sendo boa fonte de informação, e é indicado para cálculo de endogamia em populações de raça compostas (Stella *et al.*, 2010, Reverter *et al.*, 2017).

Ao observar a evolução das gerações notou-se aumento de frequência das regiões de assinatura de seleção nas gerações equivalentes maiores que 4,5 confirmando o relacionamento dessas regiões com a seleção dentro da raça composta. O composto Montana tropical® foi desenvolvido para servir como recurso para atender à necessidade dos consumidores de carne de alta qualidade em condições extensivas de climas tropicais (Grigoletto *et al.*, 2020). As regiões de assinatura de seleção no Montana revelaram que QTLs já relatados na literatura envolvidos com características de qualidade de carne como pontuação de marmoreio e força de cisalhamento estão sendo selecionadas na população.

O índice de seleção do Montana Tropical é composto por 70% de características de crescimento, sendo que 20% para peso ao desmame (Grigoletto *et al.*, 2019). A seleção para o peso ao desmame pode ter consequências não intencionais. Embora a produção de leite de vaca seja fenotipicamente significativamente correlacionada com o peso ao desmame do bezerro em diferentes idades de desmame (MacNeil & Mott, 2006), a correlação genética entre ambas as características pode ser negativa (McHugh *et al.*, 2014). Apesar dessa possibilidade o Montana Tropical apresentou assinaturas de seleção com QTLs para produção de leite.

As regiões de assinatura de seleção possuem grandes quantidades de genes, que podem ser explorados mais profundamente quanto a suas funções fisiológicas e estruturais. Como, por exemplo, o gene GHR encontrado na região de assinatura do cromossomo 20. Este gene possui efeito pleiotrópico que afeta a produção de proteína e quantidade de leite em gado leiteiro e crescimento em gado de corte (Yurchenko, *et al.*, 2018).

## **CONCLUSÃO**

A estrutura populacional genômica do Montana tropical tem pouca relação com a composição racial com base no sistema NABC. Os níveis de endogamia dentro da raça estão baixos, mas demonstram evolução ao longo das gerações sendo importante atentar para esse aspecto, mesmo se tratando de uma raça que explora cruzamentos de diversos tipos biológicos. As detecções de assinaturas de seleção com QTLs relacionadas a qualidade de carne demonstram que o objetivo de produzir animal de raça composta voltado para a produção de carne a pasto, destacando para qualidade de carne que está sendo atingida.

## REFERÊNCIAS

- Ahmad, S. F., Panigrahi, M., Chhotaray, S., Pal, D., Parida, S., Bhushan, B., ... & Singh, R. K. 2020. Revelation of genomic breed composition in a crossbred cattle of India with the help of Bovine50K BeadChip.Genomics,112(2), 1531-1535. <https://doi.org/10.1016/j.ygeno.2019.08.025>
- Akanno, E. C., Chen, L., Abo-Ismael, M. K., Crowley, J. J., Wang, Z., Li, C., ... & Plastow, G. 2017. Genomic prediction of breed composition and heterosis effects in Angus, Charolais, and Hereford crosses using 50K genotypes. *Canadian Journal of Animal Science*, 97(3), 431-438. <https://doi.org/10.1139/cjas-2016-0124>
- Alexander D.H., Novembre J., Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19: 1655–64. <http://doi10.1101/gr.094052.109>
- Arias, J. A., Keehan, M., Fisher, P., Coppieters, W., & Spelman, R. (2009). A high density linkage map of the bovine genome. *BMC genetics*, 10(1), 1-12. <https://doi.org/10.1186/1471-2156-10-18>
- Arnold J. B. 2021. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R package version 4.2.4.
- Caivio-Nasner, S., López-Herrera, A., González-Herrera, L. G., & Rincón, J. C. 2021. Diversity analysis, runs of homozygosity and genomic inbreeding reveal recent selection in Blanco Orejinegro cattle. *Journal of Animal Breeding and Genetics*, 138(5), 613-627.
- Curik, I., Ferenčaković, M., & Sölkner, J. 2014. Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livestock Science*,166, 26-34. <https://doi.org/10.1016/j.livsci.2014.05.034>
- Dodds, K. G., Auvray, B., Newman, S. A. N., & McEwan, J. C. (2014). Genomic breed prediction in New Zealand sheep. *BMC genetics*, 15(1), 1-15. <https://doi.org/10.1186/s12863-014-0092-9>
- Frank & Harrell Jr. 2022. Hmisc: Harrell Miscellaneous. R package version 4.7-0
- Ferencakovic, M., Hamzic, E., Gredler, B., Curik, I., & Sölkner, J. 2011. Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. *Agriculturae Conspectus Scientificus*,76(4), 325-329.
- Funkhouser, S. A., Bates, R. O., Ernst, C. W., Newcom, D., & Steibel, J. P. 2017. Estimation of genome-wide and locus-specific breed composition in pigs. *Translational animal science*,1(1), 36-44. <https://doi.org/10.2527/tas2016.0003>
- Gobena, M., Elzo, M. A., & Mateescu, R. G. 2018. Population structure and genomic breed composition in an angus–brahman crossbred cattle population. *Frontiers in genetics*, 9, 90. <https://doi.org/10.3389/fgene.2018.00094>

- Grigoletto, L., Brito, L. F., Mattos, E. C., Eler, J. P., Bussiman, F. O., Silva, B. D. C. A., ... & Ferraz, J. B. S. (2019). Genome-wide associations and detection of candidate genes for direct and maternal genetic effects influencing growth traits in the Montana Tropical® Composite population. *Livestock Science*, 229, 64-76. <https://doi.org/10.1016/j.livsci.2019.09.013>
- Grigoletto, L., Ferraz, J., Oliveira, H. R., Eler, J. P., Bussiman, F. O., Abreu Silva, B. C., ... & Brito, L. F. 2020. Genetic Architecture of Carcass and Meat Quality Traits in Montana Tropical® Composite Beef Cattle. *Frontiers in genetics*, 11, 123. <https://doi.org/10.3389/fgene.2020.00123>
- Gurgul, A., Semik, E., Pawlina, K., Szmatola, T., Jasielczuk, I., & Bugno-Poniewierska, M. 2014. The application of genome-wide SNP genotyping methods in studies on livestock genomes. *Journal of applied genetics*, 55(2), 197-208. <https://doi.org/10.1007/s13353-014-0202-4>
- Hay, E. H., Toghiani, S., Roberts, A. J., Paim, T., Kuehn, L. A., & Blackburn, H. D. 2022. Genetic architecture of a composite beef cattle population. *Journal of Animal Science*, 100(9), skac230. <https://doi.org/10.1093/jas/skac230>
- Huang, Y., Bates, R. O., Ernst, C. W., Fix, J. S., & Steibel, J. P. 2014. Estimation of US Yorkshire breed composition using genomic data. *Journal of Animal Science*, 92(4), 1395-1404. <https://doi.org/10.2527/jas.2013-6907>
- Kuehn, L. A., Keele, J. W., Bennett, G. L., McDaneld, T. G., Smith, T. P. L., Snelling, W. M., ... & Thallman, R. M. (2011). Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. *Journal of animal science*, 89(6), 1742-1750. <https://doi.org/10.2527/jas.2010-3530>
- McHugh, N., Cromie, A. R., Evans, R. D., & Berry, D. P. (2014). Validation of national genetic evaluations for maternal beef cattle traits using Irish field data. *Journal of Animal Science*, 92(4), 1423-1432. <https://doi.org/10.2527/jas.2013-6658>
- MacNeil, M. D., & Mott, T. B. 2006. Genetic analysis of gain from birth to weaning, milk production, and udder conformation in Line 1 Hereford cattle. *Journal of animal science*, 84(7), 1639-1645. <https://doi.org/10.2527/jas.2005-697>
- McQuillan, Ruth et al. 2008. Runs of Homozygosity in European Populations. *American journal of human genetics* 83(3): 359-72. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2556426&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1016/j.ajhg.2008.08.007>
- Pedersen, T. L. 2020. patchwork: The Composer of Plots. R package version 1.1.1.
- Peripolli, E., Stafuzza, N. B., Amorim, S. T., de Lemos, M. V. A., Grigoletto, L., Kluska, S., ... & Baldi, F. 2020. Genome-wide scan for runs of homozygosity in the composite Montana Tropical® beef cattle. *Journal of Animal Breeding and Genetics*, 137(2), 155-165. <https://doi.org/10.1111/jbg.12428>
- Peripolli, E., Stafuzza, N. B., Munari, D. P., Lima, A. L. F., Irgang, R., Machado, M. A., ... & da Silva, M. V. G. B. 2018. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. *BMC genomics*, 19(1), 1-13. <https://doi.org/10.1186/s12864-017-4365-3>

- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959. <https://doi.org/10.1093/genetics/155.2.945>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3), 559-575. <https://doi.org/10.1086/519795>
- Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. 2012. Runs of homozygosity and population history in cattle. *BMC genetics* 13:70. <https://doi.org/10.1186/1471-2156-13-70>
- Reist-Marti, S. B., Simianer, H., Gibson, J., Hanotte, O., & Rege, J. E. O. 2003. Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. *Conservation Biology*, 17(5), 1299-1311. <https://doi.org/10.1046/j.1523-1739.2003.01587.x>
- Reverter, A., Porto-Neto, L. R., Fortes, M. R. S., Kasarapu, P., De Cara, M. A. R., Burrow, H. M., & Lehnert, S. A. 2017. Genomic inbreeding depression for climatic adaptation of tropical beef cattle. *Journal of animal science*, 95(9), 3809-3821. <https://doi.org/10.2527/jas.2017.1643>
- Ron, M., Blanc, Y., Band, M., Ezra, E., & Weller, J. I. (1996). Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. *Journal of Dairy Science*, 79(4), 676-681. [https://doi.org/10.3168/jds.S0022-0302\(96\)76413-5](https://doi.org/10.3168/jds.S0022-0302(96)76413-5)
- Saravanan, K. A., Panigrahi, M., Kumar, H., Parida, S., Bhushan, B., Gaur, G. K., ... & Singh, R. K. (2022). Genome-wide assessment of genetic diversity, linkage disequilibrium and haplotype block structure in Tharparkar cattle breed of India. *Animal biotechnology*, 33(2), 297-311. [doi.org/10.1080/10495398.2020.1796696](https://doi.org/10.1080/10495398.2020.1796696)
- Silva, A. M. D., Alencar, M. M. D., Regitano, L. C. D. A., Oliveira, M. C. D. S., & Barioni Júnior, W. 2007 Artificial infestation of *Boophilus microplus* in beef cattle heifers of four genetic groups. *Genetics and Molecular Biology*, 30(4), 1150-1155. <https://doi.org/10.1590/S1415-47572007000600020>
- Stella, A., Ajmone-Marsan, P., Lazzari, B., & Boettcher, P. 2010. Identification of selection signatures in cattle breeds selected for dairy production. *Genetics*, 185(4), 1451-1461. <https://doi.org/10.1534/genetics.110.116111>
- Venables, W. N. & Ripley, B. D. 2002. *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
- Yurchenko, A. A., Daetwyler, H. D., Yudin, N., Schnabel, R. D., Vander Jagt, C. J., Soloshenko, V., ... & Larkin, D. M. (2018). Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Scientific reports*, 8(1), 1-16. <https://doi.org/10.1038/s41598-018-31304-w>
- Wei, T. and Simko, V. 2021. R package 'corrplot': Visualization of a Correlation Matrix
- Wellmann R. (2017) optiSel: optimum contribution selection and population genetics.

Weng, Z. Q., Saatchi, M., Schnabel, R. D., Taylor, J. F., & Garrick, D. J. (2014). Recombination locations and rates in beef cattle assessed from parent-offspring pairs. *Genetics Selection Evolution*, 46(1), 1-14. <https://doi.org/10.1186/1297-9686-46-34>

Welsh, C. S., Stewart, T. S., Schwab, C., & Blackburn, H. D. (2010). Pedigree analysis of 5 swine breeds in the United States and the implications for genetic conservation. *Journal of Animal Science*, 88(5), 1610-1618. <https://doi.org/10.2527/jas.2009-2537>

Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York

Wickham H. 2021. *forcats: Tools for Working with Categorical Variables (Factors)*. R package version 0.5.1.