



UNIVERSIDADE FEDERAL DE JATAÍ (UFJ)
UNIDADE ACADÊMICA ESPECIAL DE CIÊNCIAS AGRÁRIAS (CIAGRA)
PROGRAMA DE PÓS GRADUAÇÃO EM BIOCIÊNCIA ANIMAL

MILENA APARECIDA FERREIRA CAMPOS

**Clusterização por similaridade genética em uma população
bovina da raça Nelore**

JATAÍ-GO

2022



UNIVERSIDADE FEDERAL DE GOIÁS
UNIDADE ACADÊMICA ESPECIAL DE CIÊNCIAS AGRÁRIAS

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFMG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFMG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFMG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data. O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFMG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

☒ Dissertação ☐ Tese

2. Nome completo do autor: MILENA APARECIDA FERREIRA CAMPOS

3. Título do trabalho: CLUSTERIZAÇÃO POR SIMILARIDADE GENÉTICA EM UMA POPULAÇÃO BOVINA DA RAÇA NELORE

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concordo com a liberação total do documento ☐ SIM ☒ NÃO

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(a) autor(a) e ao(a) orientador(a);
- b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por Edgar Alain Collao Seenz, Orientador, em 21/03/2022, às 13:30, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por MILENA APARECIDA FERREIRA CAMPOS, Discente, em 21/03/2022, às 14:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site

https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_documento=2773189

informando o código verificador 2773189 e o código CRC ACAA972A.

MILENA APARECIDA FERREIRA CAMPOS

**Clusterização por similaridade genética em uma população
bovina da raça Nelore**

Dissertação apresentada ao Programa de Pós-Graduação em
Biociência Animal, da Universidade Federal de Jataí (UFJ),
como requisito para a obtenção do título de Mestre em Biociência Animal.

Área de concentração: Saúde e Produção Animal.

Linha de Pesquisa: Nutrição, Produção e Reprodução Animal

Orientador: Professor Doutor Edgar Alain Collao Saez

Co orientadora: Professora Doutora Adriana Luize Bocchi

JATAÍ-GO

2022

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFJ.

Campos, Milena Aparecida Ferreira
Clusterização por similaridade genética em uma população bovina da raça Nelore / Milena Aparecida Ferreira Campos. - 2022.
58 f.

Orientador: Prof. Edgar Alain Collao Saez; co-orientador Prof. Adriana Luize Bocchi.

Dissertação (Mestrado) - Universidade Federal de Jataí, Unidade Acadêmica Especial de Ciências Agrárias, Jataí, Programa de Pós-Graduação em Biociência Animal, Jataí, 2022.

Bibliografia. Apêndice.

Inclui siglas, abreviaturas, gráfico, tabelas, lista de figuras, lista de tabelas.

1. fixação de características. 2. linhagens. 3. variabilidade genética.
I. Saez, Edgar Alain Collao , orient. II. Título.

CDU 635



UNIVERSIDADE FEDERAL DE GOIÁS

UNIDADE ACADÊMICA ESPECIAL DE CIÊNCIAS AGRÁRIAS

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº 002/2022-PPGBA da sessão de Defesa de Dissertação de **MILENA APARECIDA FERREIRA CAMPOS**, que confere o título de Mestre no Programa de Pós-Graduação em **BIOCIÊNCIA ANIMAL**, área de concentração em Saúde e Produção Animal.

Ao vigésimo quinto dia, do mês de fevereiro do ano de dois mil e vinte e dois, a partir das 09:00 horas, realizou-se por meio de videoconferência a sessão pública de Defesa de Dissertação intitulada **"SIMILARIDADE GENÉTICA EM UMA POPULAÇÃO DE BOVINOS DA RAÇA NELORE"**. Os trabalhos foram instalados pelo Orientador, Professor Doutor Edgar Alain Collao Saenz (UACIAGRA/UF) - Universidade Federal de Jataí/GO) com a participação dos demais membros da Banca Examinadora: Professor Doutor Tiago do Prado Palm (IF GOIANO - Rio Verde) - membro titular externo e Professor Doutor Raphael Bernal Costa (Universidade Federal da Bahia - Salvador/BA) - membro titular externo. Durante a arguição os membros da banca fizeram sugestão de alteração do título do trabalho. A Banca Examinadora se reuniu em sessão secreta, a fim de concluir o julgamento da Dissertação, tendo sido a candidata **aprovada** pelos seus membros. Proclamados os resultados pelo Professor Doutor Edgar Alain Collao Saenz, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, ao vigésimo quinto dia, do mês de fevereiro do ano de dois mil e vinte e dois.

TÍTULO SUGERIDO PELA BANCA

CLUSTERIZAÇÃO POR SIMILARIDADE GENÉTICA EM UMA POPULAÇÃO BOVINA DA RAÇA NELORE



Documento assinado eletronicamente por **Edgar Alain Collao Saenz**, Orientador, em 25/02/2022, às 10:44, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Raphael Bernal Costa**, Usuário Externo, em 25/02/2022, às 10:53, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Tiago do Prado Palm**, Usuário Externo, em 25/02/2022, às 10:53, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufg.br/sei/controlador-externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador 2692443 e o código CRC 807EF81B.

file:///C:/Users/milena/OneDrive/Documents/MESTRADO BIOCENCIA ANIMAL-UFU/DISSERTAÇÃO/DISSERTAÇÃO FINALIZAÇÃO DOCS/Ata... 1/2

FOLHA DE MENÇÃO

Os Programas de Pós-Graduação stricto sensu, ora em funcionamento na Universidade Federal de Jataí (UFJ), em virtude de procedimentos técnicos relacionados à CAPES, continuam provisoriamente vinculados à Universidade Federal de Goiás (UFG), no entanto, todos os elementos pré-textuais do trabalho apresentado estão identificados como Universidade Federal de Jataí, em função da migração da BDTD ter ocorrido a partir de 16 de agosto de 2021, e pelo fato das pesquisas e produções estarem sendo realizadas na UFJ.

todos os colegas, professores e funcionários envolvidos do programa de Biociência Animal, que mesmo no meio das adversidades causadas pela pandemia de COVID-19 se desdobraram para dar continuidade no seu trabalho, dedico.

AGRADECIMENTOS

Agradeço a minha família e amigos pelo apoio e sustentação em todos os desafios enfrentados.

Aos meus orientadores Adriana Luize Bocchi e Edgar Alain Collao Saenz, pela amizade, pelo incentivo, paciência e por acreditarem sempre nas nossas capacidades.

Ao programa de Biociência Animal da Universidade Federal de Jataí.

À agência financiadora CAPES pela bolsa, possibilitando a continuidade do trabalho.

RESUMO

A importância da pecuária de corte para o Brasil é reflexo de intensos esforços de diversas áreas de conhecimento, e principalmente do melhoramento genético animal, que vem selecionando animais produtivos e eficientes, gerando lucratividade e eficiência ao setor. Ferramentas e estratégias de seleção dos animais têm sido desenvolvidas e aprimoradas para que o progresso e ganho genético da população seja notável. Fatores como a quantidade e a intensidade de uso dos animais selecionados podem afetar o progresso genético da população, porque a resposta a seleção poderá ser amplificada com a redução do número de reprodutores. Para objetivos de melhoramento, deve-se levar em consideração o coeficiente de endogamia existente no grupo selecionado, não permitindo que o uso excessivo de determinado touro comprometa desempenhos futuros em características de baixa herdabilidade, causados por depressão endogâmica. Dentro da produção animal eficiente, se busca animais superiores em mais de um atributo, para isso mais de uma característica precisa ser analisada e selecionada. A utilização desse número restrito de animais, porém de forma planejada, leva à alteração nas frequências genéticas e consequentemente maior proporção na população de fenótipos de interesse podendo resultar na formação de linhagens, com maior especialização em conjuntos específicos de características. Estas características devem ser identificadas e utilizadas como fonte de informação para seleção e acasalamentos que promovam a complementariedade genética de progênie. Ferramentas de análise estatística para verificação de diversidade genética, características fixadas e formação de linhagens por meio de agrupamentos são uma técnica adicional que pode ser utilizada para a escolha dos pais das próximas gerações.

Palavras-chave: fixação de características, linhagens, variabilidade genética.

,

ABSTRACT

CLUSTERING BY GENETIC SIMILARITY IN A NELORE CATTLE POPULATION

The importance of beef cattle in Brazil is a reflection of the intense efforts of several areas of concentration, and mainly of animal genetic improvement, which has been selecting productive and efficient animals, generating profitability and longevity to the sector. Tools and strategies for the selection of animals have been developed and improved so that the progress and genetic gain of the population is notable. Factors such as the number and intensity of use of selected animals can affect the genetic progress of the population, because the response to selection can be amplified with the reduction of the number of reproducers. The inbreeding coefficient existing in the selected group must be taken into account, not allowing the excessive use of a particular bull to compromise future performances in traits of low heritability, caused by inbreeding depression. Within efficient animal production, superior animals are sought in more than one attribute, for which more than one characteristic needs to be analyzed and selected. The use of this restricted number of animals, however in a planned way, leads to the establishment of certain characteristics and the formation of lines, with greater specialization in specific sets of characteristics, which must be identified and used as a source of information for selection and mating that promote the genetic complementarity of progeny. Statistical analysis tools to verify genetic diversity, fixed traits and lineage formation through clusters is an additional technique that can be used to choose the parents of the next generations.

Keywords: trait fixation, lines, genetic variability.

SUMÁRIO

CAPÍTULO 1- CONSIDERAÇÕES GERAIS	16
1 INTRODUÇÃO GERAL	16
2 REVISÃO DE LITERATURA.....	17
2.1. Seleção.....	17
2.2. Métodos de seleção	17
2.3. Tipos de seleção	18
2.3.1. Seleção direta e indireta	18
2.3.2. Seleção em tandem	19
2.3.3 Níveis independentes de descarte (rejeição)	19
2.3.4. Índices de seleção	19
2.4. Critérios de seleção	20
2.5. Variabilidade Genética	21
2.5.1. Variabilidade genética e Identificação de linhagens especializadas.....	21
2.6. Programas de Melhoria de Gado de Corte.....	24
2.7. Agrupamentos como forma de seleção	26
2.7.1. Análise de Componentes Principais.....	26
2.7.2. Clusterização	27
REFERÊNCIAS	30
CAPÍTULO 2- SIMILARIDADE GENÉTICA EM POPULAÇÃO BOVINA DA RAÇA NELORE	34
RESUMO	34
ABSTRACT	35
GENETIC SIMILARITY IN NELORE CATTLE POPULATION.....	35
1 Introdução.....	36
2 Material e Métodos	37
3 Resultados.....	39

3.2 Análise de componentes principais.....	40
3.3 Clusterização.....	43
3.4 Representatividade dos reprodutores por grupo	46
4 Discussão	46
5 Conclusões.....	51
6-Agradecimentos	52
REFERÊNCIAS	53
APÊNDICE- Tabela de correlação de Pearson entre as características estudadas e análise descritiva dos agrupamentos.....	56

LISTA DE FIGURAS

FIGURA 1. Correlações de Pearson entre características estudadas.....	39
FIGURA 2. Porcentagem de variação explicada pelos componentes principais.....	40
FIGURA 3. Porcentagem de variação explicada pelos 5 primeiros componentes principais...	41
FIGURA 4. Contribuição das variáveis para as dimensões utilizadas.....	41
FIGURA 5. Correlação das variáveis com os componentes principais escolhidos.....	42
FIGURA 6. Dendrograma demonstrando clusterização e escolha de 4 clusters.....	44
FIGURA 7. Indivíduos agrupados de acordo com o cluster a que pertencem, representando a primeira e segunda dimensão da análise de HCPC.....	44
FIGURA 8. Representação do valor médio de DEP dos animais que compõe os clusters.....	45

LISTA DE TABELAS

Tabela 1. Análise descritiva das diferenças esperadas na progênie da população analisada...	38
Tabela 2. Proporção da variância explicada pelas variáveis em estudo (E_{ta}^2).....	43
Tabela 3. Médias e porcentagens de representação dos dez touros com maior número de filhos em cada cluster, assim como o coeficiente médio de endogamia de cada cluster.....	46
Tabela 4. Comparação entre médias dos animais agrupados no CLUSTER 2 com a média nacional dos animais avaliados pelo programa de melhoramento genético Nelore Brasil da ANCP.....	49

LISTA DE ABREVIATURAS

PIB	produto interno bruto
ABIEC	Associação brasileira das indústrias exportadoras de carnes
EMBRAPA	Empresa brasileira de pesquisa agropecuária
IPF	Integração pecuária floresta
ILPF	Integração lavoura pecuaria floresta
DEP	Diferença esperada na progenie
SNP	Single nucleotide polymorphism
NeS	Nelore Seleção
NeT	Nelore tradicional
NeC	Nelore Controle
LARRL	Livestock and Range Research Laboratory
DP	Desvio padrão
Var	Variância;
Min	Mínimo;
Max	Máximo;
MP120	DEP materna 120 dias;
MP210	DEP materna 210 dias;
DAOL	DEP área de olho de lombo;
DACAB	DEP acabamento;
DMAR	DEP marmoreio;
DCAR	DEP consumo alimentar residual;
DIPM	DEP idade precocidade em machos;
D3P	DEP probabilidade de parto precoce;
DP210	DEP peso aos 210 dias;
DP365	DEP peso aos 365 dias;
DP450	DEP peso aos 450 dias;
DIMS	DEP ingestão de matéria seca;

CAPÍTULO 1- CONSIDERAÇÕES GERAIS

1 INTRODUÇÃO GERAL

A pecuária de corte representou 10% do total do PIB brasileiro no ano de 2020, sendo que de 1990 a 2020 a produtividade elevou de 1,63@/ha/ano para 4,2@/ha/ano, um aumento de 159% de produtividade no setor, tornando o Brasil o maior exportador de carne bovina do mundo (ABIEC, 2021). Os rumos da produção pecuária são determinados pelas exigências do mercado consumidor, fatores como a qualidade da carne, preocupações como esse animal foi criado e o impacto da atividade no meio ambiente, são determinantes na hora de comercializar o produto com valor agregado.

Para conseguir atender as demandas básicas e específicas de cada mercado para o qual o Brasil exporta, o trabalho integrado entre técnicos, professores, pesquisadores e produtores de diversas áreas relacionadas à pecuária é direcionado para uma produção eficiente e com menor impacto ambiental. A Associação Brasileira das Indústrias Exportadoras de Carne (ABIEC) estimou que devido a intensificação da produção por meio de pesquisas e uso de tecnologias em setores como nutrição, bem estar animal, reprodução, sanidade e em melhoramento genético, 253,8 milhões de hectares deixaram de ser desmatados entre o período de 1990 e 2020, ressaltando que produtividade e conservação de recursos naturais caminham juntas em sinergia.

Como os principais sistemas de produção brasileiros são por meio de uso de pastagens, o Brasil possui uma vantagem competitiva em relação ao mercado internacional, podendo se valer de iniciativas e programas de certificação que autenticuem a sustentabilidade da atividade. Um exemplo é a carne carbono neutro, conceito desenvolvido pela EMBRAPA, que fornece um selo às propriedades que produzem animais em sistemas de integração pecuária-floresta (IPF) ou lavoura-pecuária-floresta (ILPF), permitindo o sequestro de carbono da atmosfera pelas árvores, compensando assim a emissão de metano pelos bovinos (ALVES et al., 2015).

Nossos sistemas de produção, em sua maioria já sustentáveis, possuem potencial de crescimento, devido ao contínuo aumento de produtividade por área e intensos esforços realizados por grupos de pesquisas e programas de melhoramento genético animal para a utilização de características que visem a identificação de animais com melhor aproveitamento de alimentos. A seleção desses animais não causa nenhum prejuízo com relação às características de crescimento, são animais que possuem menores exigências de manutenção do que outros menos eficientes, menor ingestão de alimentos e permite a mitigação de emissão de gases que tenham impacto no efeito estufa (DELVEAUX et al., 2020; MENEZES et al., 2020).

Todas essas iniciativas além de proporcionar uma atividade sustentável, permite rentabilidade ao setor e um produto final de qualidade ao consumidor.

O trabalho do melhoramento genético em bovinos de corte, permite que as outras tecnologias tenham maior eficiência na produção, pois permite uma elevada expressão genética dos animais em desempenho, resistência a doenças e adaptabilidade.

2 REVISÃO DE LITERATURA

2.1. Seleção

A seleção é uma das principais forças que alteram as frequências gênicas de uma população, de acordo com a intensidade estabelecida, um número reduzido de animais são eleitos como pais da próxima geração, fazendo com que se aumente a frequência de genes desejáveis (ELER, 2017a). Essa escolha pode ser feita com base em fenótipos ou valores genéticos desses animais, a depender do objetivo de seleção (PEREIRA, 2008). Os objetivos podem ser determinados pelo mercado em que se está inserido ou pretende atender (cliente final, ou venda de reprodutores), ou visando melhora de algum problema específico da propriedade (maior resistência a parasitas e doenças, adaptabilidade ao clima, etc.).

2.2. Métodos de seleção

A confiança nos valores genéticos depende da herdabilidade e da fonte e qualidade da informações sobre os indivíduos avaliados. Segundo ELER, 2017b as metodologias de seleção são aplicadas de diversas formas a depender das particularidades das características envolvidas, do sistema de criação e das fontes de informações disponíveis, podendo ser: seleção com base em informações fenotípicas, avaliações genéticas ou genômicas, mais recentemente (ROSA et al., 2013). As avaliações genéticas e genômicas geram DEPs (Diferenças Esperadas na Progenie) que servem como estimativas do que será passado às próximas gerações, essas DEPs podem ser calculadas utilizando uma variedade de fonte de informação em conjunto ou isoladas (PEREIRA, 2008).

Algumas avaliações podem ser calculadas utilizando apenas a informação de pedigree, que se refere aos dados de filiação ou genealogia desse indivíduo, incluído os desempenhos desses pais durante a vida produtiva (ELER, 2017b). A acurácia de estimativas feitas apenas com a informação de pedigree é baixa, porque existem fatores, incluindo a segregação cromossômica, que pode não refletir na progênie o desempenho dos pais (ROSA et al., 2013).

Os dados produtivos coletados do próprio indivíduo podem ser utilizados, porém esses dados isolados, ou somados com os dados de pedigree ainda apresentam baixa acurácia (PEREIRA, 2008). Uma forma de driblar essa acurácia baixa, seria o uso de um teste empregado e bem conhecido, principalmente na pecuária leiteira que é o teste de progênie, onde o material genético de determinado indivíduo, é distribuído em diferentes propriedades para que seja avaliado o desempenho de suas progênies, resultando em seu valor genético mais acurado. A principal limitação é um aumento no intervalo de gerações e custos elevados (HAGAN; CUE, 2020).

As informações provenientes de desempenhos de irmãos, tios, tias, primos e primas também servem como fonte de informação para cálculo do valor genético do animal, mesmo que não sejam irmãos completos (ROSA et al., 2013), existem modelos matemáticos que conseguem fazer estimativas, garantindo que mesmo jovem, esse animal tenha um mérito genético, que apesar da acurácia baixa, não impede o uso do animal, mas implicando em uma intensidade de uso menor (ELER, 2017b). A elaboração de avaliações genéticas a partir de uma única fonte de informação pode se demonstrar inviável caso a característica não possua uma herdabilidade alta e não possa ser medida em ambos os sexos (PEREIRA, 2008). Metodologias têm sido frequentemente desenvolvidas para utilizar múltiplas fontes de informação no desenvolvimento de predições genéticas para que sejam mais acuradas e tenham grandes reflexos no progresso genético da população.

A seleção de animais utilizando múltiplas fontes de informações se tornou possível a partir do desenvolvimento de modelos matemáticos e da evolução computacional, permitindo que informações de ancestrais, de progênies, desempenho do próprio indivíduo, de colaterais, pudessem ser incluídos ao modelo de forma conjunta, possibilitando a estimação de valores genéticos mais confiáveis (SWALVE, 2000).

2.3. Tipos de seleção

2.3.1. Seleção direta e indireta

A seleção direta se refere à possibilidade de mensurar diretamente a característica que se pretende selecionar (PEREIRA, 2008). Na indireta se executam medições em outras características que sejam indicativas ou possuam alta correlação com aquela que se pretende selecionar, e é realizada quando existe dificuldade na mensuração, ou por ser uma característica expressa em fase da vida mais tardia no animal.

2.3.2. Seleção em tandem

São estipulados critérios de seleção, entretanto são selecionados um por vez por gerações até atingir o objetivo (ELER, 2017a). Esse tipo de seleção tem antecedentes na pecuária de corte, quando características de peso passaram a ser prioridade nos sistemas de seleção, isso porque o preço de mercado pago ao produtor depende quase que exclusivamente do que foi mensurado na balança. Como resultado, a seleção individual dessa característica levou a animais cada vez mais pesados, porém com pouca precocidade de carcaça e reprodutiva, e um considerável aumento de partos distócitos causados por bezerros muito grandes, causando um impacto negativo na rentabilidade da atividade pecuária.

Assim a evolução da escolha de critérios de seleção partiu para um meio combinado, onde os animais são selecionados na mesma geração levando em consideração mais de uma característica, a depender das necessidades do rebanho.

2.3.3 Níveis independentes de descarte (rejeição)

Nesse tipo de seleção um grupo de características são eleitas com valores de corte, caso o animal não atinja o valor mínimo em alguma das características ele é descartado. Mesmo que apresente valores superiores na maioria das características, e em apenas uma tenha valor inferior ao estipulado para a seleção, não será utilizado (PEREIRA, 2008). Permite uma intensidade maior de seleção e consequentemente, progresso genético, porém a escolha das características e os níveis de descarte precisam levar em consideração o objetivo de seleção do criador e o mercado consumidor.

2.3.4. Índices de seleção

Os índices de seleção empíricos utilizados dentro das fazendas, como nos programas de melhoramento genético animal, permitem classificar animais para características que podem ser combinadas e ponderadas levando em consideração a importância das características, herdabilidades, correlações e valor econômico (TWOMEY et al., 2020). Os animais com valor de índice elevado, podem ser considerados superiores para aquelas características e selecionados como reprodutores. Um grande desafio no uso dos índices é a determinação do fator de ponderação, mas esse contratempo é facilmente compensado pela vantagem da grande eficiência na escolha dos animais.

2.4. Critérios de seleção

A definição dos critérios de seleção deve ser direcionada para se atingir o objetivo, pois a partir deles obtemos medidas, direta ou indiretamente, que permitirá comparar os animais disponíveis na população e definir quais serão considerados para a reprodução (ELER, 2017b). Eles devem ser preferencialmente utilizados a partir de avaliações genéticas, porque apenas valores fenotípicos, sem considerar interações de ambiente, formas de criação ou dados de pedigree, podem levar a grandes equívocos com relação à quais animais realmente tem potencial para passar aos seus descendentes genes desejáveis (ARAÚJO NETO et al., 2018).

A quantidade de critérios incorporados aos programas de melhoramento genético aumentam com estudos de fenotipagem dos animais. Isso permite que o produtor disponha de um vasto catálogo de informações, que deve ser adaptado a cada realidade. Visto que é inviável e o progresso genético pode ser reduzido quando se seleciona para uma quantidade grande de critérios, deve-se definir e escolher de modo equilibrado as características que precisam ser fixadas no rebanho de modo prioritário.

É importante avaliar sempre a resposta à seleção das características, para acompanhamento do progresso genético do rebanho e definir estratégias futuras (ELER, 2017a). Isso porque se trabalha com estimativas, que dependendo da acurácia, podem alterar de acordo com a inclusão de mais dados, fazendo com que o ganho genético seja dependente da continuidade da seleção. O desempenho dos animais também depende de fatores não aditivos (dominância e epistasia) e interação genótipo-ambiente, que exige o contínuo acompanhamento para a utilização das ferramentas do melhoramento de forma eficiente (KUMAR et al., 2016).

A quantidade e a intensidade de uso dos animais selecionados afetam o progresso genético da população, pois a resposta à seleção poderá ser amplificada com a redução do número de reprodutores (MALHADO et al., 2009). Deve-se levar em consideração o coeficiente de endogamia existente no grupo selecionado, não permitindo que o uso excessivo de determinado touro comprometa desempenhos futuros em características de baixa herdabilidade, causados por depressão endogâmica (MACNEIL et al., 1989). Dentro da produção animal eficiente, se busca animais superiores em mais de um atributo, para isso mais de uma característica precisa ser analisada e selecionada. Existem grupos de características selecionáveis com o mesmo objetivo, por exemplo, quando na bovinocultura de corte se deseja melhorar a carcaça dos animais, características como área de olho de lombo, acabamento de gordura e marmoreio são todas possíveis de serem medidas e proporcionar ganhos à carcaça e qualidade de carne do animal.

Dentro de rebanhos fechados a predição do valor genético como já dito anteriormente, utiliza além das informações dos animais, informações do pedigree e colaterais, consequentemente se observa a seleção de animais pertencentes à um número restrito de famílias, sendo essas as que possuem maior valor genético na população. A utilização desse número restrito de animais, porém de forma planejada, leva ao aumento na frequência de alelos desejáveis que afetam determinadas características e consequentemente formação de linhagens, com maior especialização em conjuntos determinados de características, que devem ser identificadas e utilizadas como fonte de informação para seleção e acasalamentos que promovam a complementariedade genética nas progênes (GINJA et al., 2019).

2.5. Variabilidade Genética

Populações genéticas de espécies ou raças de bovinos são compostas de subpopulações devido à seleção conduzida de forma independente por diferentes ambientes de criação (STRONEN et al., 2019). Esses grupos de animais podem se tornar mais semelhantes daqueles que estão sofrendo o mesmo processo de seleção e mais distintos daqueles externos a essas condições devido ao parentesco e à endogamia. As diferenças dessas subpopulações podem ser consideráveis o suficiente para afetar a heterose em cruzamentos de animais de diferentes grupos porém de mesmas raças (MACNEIL et al., 1989).

É previsto como limites da seleção, quando a variabilidade populacional é marcadamente reduzida, não sendo possível observar fontes de variação, e quando o limite biológico e fisiológico é atingido e quando não permite a adaptação do animal a diferentes ambientes, tornando-o limitado a poucas situações de criação. Esses limites podem ser pensados e contornados para que o progresso seja mantido por um longo período de tempo. A estrutura populacional deve ser estudada e novas formas de seleção de animais buscadas. O isolamento genético de raças e populações pode diminuir a diversidade em nível molecular e implicando perdas no progresso do melhoramento genético (KUKUČKOVÁ et al., 2018). A perda primária de diversidade genética acontece devido ao uso intenso de um número pequeno de animais, cujo material genético é distribuído por meio das centrais de touros e com os avanços de biotecnologias reprodutivas.

2.5.1. Variabilidade genética e Identificação de linhagens especializadas

Estudos de conservação de raças e linhagens utilizam metodologias de estudo de variação populacional e diversidade genética, pois animais que passaram por processos de seleção e acasalamentos sem direcionamento ocasionaram aumento significativo de

homozigose nas populações, o que causa grande perda de genes adaptativos. Conhecer a variabilidade das características possibilita o entendimento e direcionamento dos trabalhos de conservação de espécies e direcionamento de acasalamentos em animais de produção (KUKUČKOVÁ et al., 2018; MKIZE & ZISHIRI, 2020; SIMČIČ et al., 2015; UPADHYAY et al., 2019).

Os estudos em diversidade genética podem ser estimados por registros de linhagem com base no coeficiente de consanguinidade e parentesco, assim como a proporção do genoma que se espera que seja idêntica por descendência. Em programas de conservação se combinam diferentes fontes de informação visto que a diversidade baseada em linhagem reflete apenas a diversidade como resultado de ancestralidade recente e documentada. Se as informações de pedigree estiverem disponíveis, uma primeira análise dos animais menos relacionados, após sua genotipagem com um chip single nucleotide polymorphism (SNP) de alta densidade, e análise de contribuição ótima com base em informações moleculares permite a identificação dos animais que mais contribuem para a diversidade genética da população. Posteriormente, como a análise do SNP é uma amostragem do genoma total, os animais selecionados podem ser sequenciados a fim de identificar alelos favoráveis específicos, relacionados a características de interesse ou a alelos não funcionais, com objetivo de preservá-los (EUSEBI; MARTINEZ; CORTES, 2020).

Marcadores microssatélites, especificamente, têm sido amplamente utilizados para estimar a magnitude da variabilidade genética em populações, bem como a estrutura genética populacional que busca a composição genética dentro de linhagens, fazendas e linhas de seleção, para que se facilite a identificação de reservatórios genéticos e a implementação do manejo genético (PEIXOTO et al., 2021; SEO et al., 2021). Para acompanhar os dados genealógicos, há uma necessidade de levantamentos genéticos moleculares em grande escala e pontuações quantificáveis para variação genética entre linhagens consanguíneas (BECK et al., 2000). As informações de linhagem têm sido utilizadas frequentemente para descrever populações de bovinos bem como para estabelecer uma base para esforços de conservação para raças e populações pequenas, com estudos relacionados a avaliação da taxa de endogamia, o tamanho efetivo da população, as contribuições relativas dos ancestrais e alguma quantificação da correspondência de desempenho (SIMMONS et al., 2021).

O conceito de linhagem pode ser entendido por indivíduos tanto descendentes diretos do genitor, como os que apresentam acentuada semelhança em determinados fenótipos e características genéticas advindas desse ancestral comum (FREUDENSTEIN et al., 2016; FUERST, 2017). O uso de linhagens em animais de produção tem como precursores estudos

dentro da avicultura e da suinocultura, onde um intenso processo de seleção buscando melhorias em características de interesse do mercado foram priorizadas e tiveram uma grande evolução devido ao isolamento de linhagens especializadas para machos e fêmeas e seus usos na criação de híbridos por meio de cruzamentos. A prioridade, dentro da criação de suínos hoje, é a criação de genótipos especializados com altos níveis de produtividade, qualidade de carne e carcaça e bom desempenho reprodutivo.

As linhagens especializadas são grupos genotipicamente diferentes de animais criados com base em uma ou várias raças e selecionados de acordo com uma ou duas características economicamente úteis, que podem aumentar significativamente o ganho genético dos animais. Quando essas linhagens são cruzadas entre si, uma progênie altamente produtiva é obtida com base no benefício da heterose (CHALOVA, 2021). No caso de aves se utiliza linhagens maternas e paternas detentoras de características de interesse para a finalidade que se destina na granja. No exemplo de poedeiras, são selecionados animais que apresentem boa qualidade dos ovos, tamanho de pintos e longevidade para que haja fixação dessas características (MUELLER et al., 2020).

A definição mais divulgada do termo dentro da raça Nelore considera linhagem como conjunto de animais, descendentes de um reprodutor ou de uma matriz que possuam entre si um maior grau de parentesco do que em relação ao restante da raça (OLIVEIRA et al., 2002). A exemplo de linhagem Nelore conservada e identificada, a linhagem Lemgruber tem destaque por compreender uma população que passa por seleção desde a época das importações, e que conseguiu manter seu rebanho fechado em busca da fixação de características que hoje são destaques, como tolerância a ectoparasitas e ganho de peso eficiente a pasto (LOPES et al., 2016; OLIVEIRA et al., 2011).

Quando há a possibilidade de identificação das variações de perfil genético dentro de outras possíveis linhagens, o direcionamento de acasalamentos se torna mais eficiente (TAKESHIMA et al., 2018), possibilita melhora no progresso genético do rebanho e reduz ou pelo menos melhora o controle de taxas de endogamia dentro da população. Visto que o grande uso de animais endogâmicos reduz a variabilidade genética e pode causar prejuízos do ponto de vista produtivo e reprodutivos (SANTANA et al., 2019).

A identificação de linhagens permite o aproveitamento da heterose por meio dos acasalamentos exógamicos e entre populações diferentes (AHMAD et al., 2017). Poder rastrear essas linhagens em uma população possibilita a melhora no uso de ferramentas de seleção e acasalamentos direcionados dentro do rebanho, fixando genes, definindo grupos genéticos, mas com o controle de endogamia dos animais.

2.6. Programas de Melhoria de Gado de Corte

Algumas linhagens foram desenvolvidas e mantidas principalmente em centros de pesquisa, como é o caso do desenvolvimento de linhagens Hereford, que começou no LARRL (Livestock and Range Research Laboratory) em 1934, onde houve o desenvolvimento de um total de 14 linhas que foram mantidas em rebanho fechado, e selecionadas intensivamente com base em desempenho, com o objetivo de se obter animais adaptados ao sistema de pastagem. Os critérios de seleção adotados pelo programa visavam a melhora de fertilidade, ganho rápido e econômico, conformação e altas porcentagens de cobertura, e apesar de apresentarem elevadas taxas de endogamia, não se observou perdas substanciais nas características de desempenho (SUMREDDEE et al., 2019). No entanto, seis dessas linhagens foram selecionadas para crescimento deficiente, onde os pesquisadores conseguiram observar que ao realizar o cruzamento entre estas linhas, suas progênies apresentavam um melhor desempenho. Existe hoje apenas a Linha 1, sendo considerado o experimento de seleção de gado de corte mais antigo do mundo segundo Newman (2016).

Além da Linha 1, houve o desenvolvimento de linhagens Hereford por criadores particulares, como a linhagem Trask, cuja origem vem de um mesmo criador do sudeste do Estados Unidos que foi intensamente selecionado para fertilidade com suplementação alimentar mínima, além de característica de tipo com maior profundidade de costelas do que os demais exemplares da raça (SIMMONS et al., 2021). Estudos conduzidos com bovinos espanhóis da raça Miura, demonstram a presença de pequenas linhagens com diferentes características comportamentais de temperamento em cada uma delas, e apesar de a subdivisão populacional ter sido conduzida de forma não intencional, se demonstrou um poderoso mecanismo para manter a variabilidade genética, pois diferentes características podem ter sido fixadas em diferentes linhagens e mantidos como fonte de variação (EUSEBI; MARTINEZ; CORTES, 2020).

O isolamento de linhagens dentro de raças também foi observado em animais da raça de bovinos Jutland na Dinamarca, onde quatro linhagens (Westergaard, Vesterbølle, Oregaard e Kortegaard), foram mantidas em isolamento de outros grupos semelhantes. Esses animais carregam variantes genéticas para a adaptação, importantes em condições que preveem muita alteração de ambiente devido às mudanças climáticas. Tais variantes podem ser críticas para sua preservação e contribuir para o aumento da tolerância ambiental em raças comerciais (STRONEN et al., 2019).

Populações da Raça Guzerá foram analisadas por meio de marcadores microssatélites, onde se verificou quais os principais animais responsáveis pela origem das três linhagens: corte,

leite e duplo propósito no estado de Minas Gerais (PEIXOTO et al., 2021). Esses estudos ajudam a entender o impacto dos cruzamentos planejados na diversidade genética e na distribuição dessa diversidade entre linhagens selecionadas para diferentes fins. Se observou que a variabilidade genética apresentava diferenças entre as propriedades estudadas, onde o desenvolvimento de linhagens próprias dentro de cada fazenda foi observada de acordo com o objetivo de produção, no entanto, o estudo em questão demonstrou que, apesar da estrutura populacional produzida por estratégias de seleção baseadas originalmente em linhagens relativamente fechadas, a diversidade genética conservada na raça Guzerá ainda é elevada (PEIXOTO et al., 2021).

Centros de pesquisa brasileiros tem mantido rebanhos experimentais, como é o caso do Instituto de Zootecnia do estado de São Paulo, que segundo Oliveira et al. (2002) é o Instituto mais antigo de experimentação na Raça Nelore, e possui linhagens formadas desde 1951, sendo elas: linhagem Nelore Seleção (NeS) e Nelore tradicional (NeT) selecionados para maior peso ao sobreano e Nelore Controle (NeC) para peso ao sobreano médio em relação à média dos contemporâneos de rebanho. O rebanho NeT é o único que recebeu reprodutores de rebanhos de terceiros, enquanto os outros permanecem fechados à introdução de animais de outras propriedades (INSTITUTO DE ZOOTECNIA, 2021). Esses três rebanhos com forma de seleção distinta permite a comparação e o desenvolvimento de inúmeras pesquisas ao longo do tempo, inclusive a formação das duas linhagens endogâmicas, NeS e NeC, com a fixação de características importantes para a pecuária de corte (FREITAS et al., 2021).

A possibilidade de identificação das variações de perfil genético dentro de possíveis linhagens em diferentes rebanhos, pode permitir o direcionamento de acasalamentos eficientes (TAKESHIMA et al., 2018), possibilitando melhora no progresso genético do rebanho e do controle de taxas de endogamia dentro da população, evitando prejuízos do ponto de vista produtivo e reprodutivos (SANTANA et al., 2019). A identificação de linhagens permite o aproveitamento da heterose por meio dos acasalamentos exógamicos e entre populações diferentes (AHMAD et al., 2017). Poder rastrear essas linhagens em uma população possibilita a melhora no uso de ferramentas de seleção e acasalamentos direcionados dentro do rebanho, fixando genes, definindo grupos genéticos, mas com o controle de endogamia dos animais.

Há evidências de diferentes padrões genéticos entre as linhagens bovinas para melhora de características de importância econômica. Dessa forma, os programas de melhoramento buscam estratégias para preservar as populações, caracterizar e monitorar o coeficiente de endogamia de todo o genoma tanto para manter a diversidade genética, como para que haja a conservação de longo prazo dos recursos genéticos (PERIPOLLI et al., 2018). Com a utilização

da seleção genômica, é possível eleger em linhagens previamente identificadas os traços complexos e genes candidatos que justificam a superioridade daqueles animais para determinadas características (RAMAYO-CALDAS et al., 2016).

A maioria das raças estudadas atualmente possuem diversidade genética marcadamente reduzida, o que permite a identificação de características importantes de forma mais acentuada como, alta fertilidade e longevidade em ambientes estressantes (PAPACHRISTOU et al., 2020). Identificar e agrupar permite também a criação de programas que visem a caracterização, conservação e valorização desses recursos genéticos únicos (PITT et al., 2019). A seleção combinada de características com boa diversidade genética, e a fixação delas na população em conjunto é possível quando conhecemos os animais que são destaques para transmitir-las às suas progênes.

2.7. Agrupamentos como forma de seleção

Os avanços nas áreas de pesquisa do melhoramento genético animal, têm permitido que grande número de características possua avaliação genética e se tornem passíveis de seleção. Verifica-se, portanto, que o melhoramento genético animal tem evoluído rapidamente para uma seleção mais objetiva, permitindo que as estimativas sejam mais acuradas e as características mais específicas para o objetivo de seleção. Em todo caso, é necessário às propriedades e às populações, ferramentas e parâmetros para determinar quais características devem ser consideradas, levando em consideração correlação entre elas e a variação populacional dentro de rebanhos de seleção.

Análises de dados multivariados são aliadas para entender a variação total da população e podem ser utilizadas como ferramentas de seleção, pois possibilitam uma visualização organizada em grupos similares dos animais por meio de seus valores genéticos, permitindo a observação de predominância de determinados animais com relação à fenótipos de interesse.

2.7.1. Análise de Componentes Principais

A análise de componentes principais (PCA) é uma técnica estatística multivariada, usada para descrever dados quantitativos. Permite reduzir a dimensionalidade do conjunto de dados, retendo a informação mais importante em novas variáveis, simplificando a descrição e possibilitando analisar a estrutura das variáveis ou dos indivíduos. Banco de dados com mais de seis mil animais e muitas variáveis como o utilizado nesse estudo, pode não ser analisado corretamente sem o uso dessas técnicas, porque muitas das informações podem não contribuir para a identificação de padrões ou relações dentro da população, sem uma prévia limpeza do

ruído contido em grandes volumes de informações. As informações inclusas no banco de dados correspondem à inércia total ou variabilidade dos dados que ele contém. O PCA consegue identificar as direções em que a variação de dados é máxima e criar novas variáveis que vão facilitar a leitura e visualização dos dados. Nos componentes, cada observação ou indivíduo é representado por um vetor em um espaço multidimensional (ABDI; WILLIAMS, 2010).

O primeiro componente principal é a combinação linear de todas as variáveis analisadas, que possuem variância máxima, ou inércia total, portanto, definimos coeficientes para o primeiro componente de forma que sua variância seja maximizada, sujeito à restrição de que a soma dos coeficientes quadrados seja igual a um. Essa restrição é necessária para que uma resposta única possa ser obtida (ABDI; WILLIAMS, 2010). Todos os componentes principais subsequentes têm a mesma propriedade. O segundo componente principal é a combinação linear de todas as variáveis que responde pelo máximo possível da variação restante, porém a correlação entre o primeiro e o segundo componente é nula. Os outros componentes são calculados também como combinações lineares que respondem pelo máximo possível da variação restante e não estão correlacionados com os outros componentes principais (LEVER; KRZYZWINSKI; ALTMAN, 2017).

O número de componentes principais escolhidos precisa ser suficiente para conter o máximo de variação dos dados, e pode ser justificado a partir de parâmetros característicos da análise como a contribuição de cada variável para o componente, cosseno ao quadrado e correlações entre variáveis e componentes. Esse último também conhecido como qualidade da representação, permite verificar a relevância de cada componente para as variáveis, ajudando a encontrar os componentes que são importantes para interpretar as observações, garantindo maior confiança na quantidade de componentes escolhidos, e que apenas a informação realmente relevante seja considerada. Essa medida pode ser descrita como o quadrado do cosseno do ângulo do triângulo retângulo formado pela origem, a observação e a projeção no componente. Quanto maior o valor de COS^2 em determinado componente, melhor a representação da variável no componente.

2.7.2. Clusterização

O processo de clusterização é uma técnica de mineração de dados multivariados, que permite o agrupamento automático a partir de medidas de similaridade ou dissimilaridade para agrupar os indivíduos de modo que os animais presentes no mesmo cluster sejam os mais semelhantes possíveis, e que a diferença entre clusters seja a maior possível. Dessa forma vários estudos de similaridade/dissimilaridade podem ser conduzidos para identificação de padrões ou

que auxiliem a descrever relações que não poderia ser visualizadas nos dados brutos. O agrupamento auxilia principalmente na redução das unidades dos dados, em grupos ou subgrupos permitindo maior exploração das informações. Tem aplicação nas mais diversas áreas, porque fornece uma solução para grande quantidade de dados coletados em todas as áreas de estudo.

A técnica exige uma medida de intervalo para quantificar as proximidades ou distâncias. A distância euclidiana é uma das mais frequentemente utilizadas, definida como a distância em linha direta entre dois objetos. Os métodos utilizados podem variar a depender do tipo de dados que estão sendo trabalhados, subdividindo os algoritmos em dois tipos de formação mais tradicionalmente usados: Particionais e Hierárquicos. O hierárquico organiza os dados em uma estrutura hierárquica de acordo com sua proximidade em relação às características especificadas (WARD, 1963), e o particional exige um número de clusters inicial previamente informado, para a subdivisão dos dados, onde a distância dos indivíduos/variáveis são calculados com base na distância do indivíduo e do centro dos clusters, que foi previamente decidido, a partir do número de clusters escolhidos.

Essa técnica permite o uso em uma variedade de dados, na produção animal, incluindo tipologia de propriedade rural, de consumo, diversidade genética, nível de tecnificação, etc. Amamou et al. (2018) utilizaram clusterização hierárquica com base no método de Ward para entender a percepção de produtores leiteiros sobre as mudanças climáticas na Tunísia. As propriedades foram agrupados em quatro clusters, permitindo melhor visualização das particularidades de cada. Verificou-se que em todos os clusters, independentemente do tamanho da propriedade ou do nível de tecnificação, eram observados pelos criadores impactos das ondas de calor na produção de leite e na fertilidade, além da disponibilidade e custo dos alimentos como as principais restrições, permitindo conhecer e ter uma estimativa de como os diferentes produtores reagiram e se adaptaram a esse problema.

O uso de valores genéticos e genômicos para determinar a similaridade genética de populações não é muito descrito, os mais utilizados são estudos com o próprio genoma bovino, buscando corridas de homozigose ou assinaturas de seleção (EUSEBI; MARTINEZ; CORTES, 2020). Mas com o uso de agrupamentos se observa que animais de raças compostas, podem ter suas raças formadoras individualizadas de modo a melhor entender o desempenho desses animais e os atributos que possuem. Bocchi (2006) utilizando análise de agrupamentos em animais da raça Montana, para definir tipos biológicos a partir de características genéticas, conseguiu observar quais raças apresentavam maior similaridade, de acordo com sua

proximidade de composição genética dentro dos clusters formados, permitindo melhor entendimento sobre a variabilidade do composto.

REFERÊNCIAS

- ABDI, H.; WILLIAMS, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, p. 433–459, 2010. <https://doi.org/10.1002/wics.101>.
- AHMAD, S. F. et al. Breed clustering analysis of six Indian cattle breeds using genome-wide SNP data. *Journal of Entomology and Zoology Studies JEZS*, vol. 5, no. 56, p. 1129–1134, 2017. Available at: <http://www.entomoljournal.com/archives/2017/vol5issue6/PartP/5-6-28-293.pdf>.
- ALVES, F. V. et al. Carne Carbono Neutro: um novo conceito para carne sustentável produzida nos trópicos. *Documentos* 210, , p. 29, 2015.
- AMAMOU, H. et al. Climate change-related risks and adaptation strategies as perceived in dairy cattle farming systems in Tunisia. *Climate Risk Management*, vol. 20, no. March 2017, p. 38–49, 2018. DOI 10.1016/j.crm.2018.03.004. Available at: <https://doi.org/10.1016/j.crm.2018.03.004>.
- ARAUJO NETO, F. R.; PEGOLO, N. T.; ASPILCUETA-BORQUIS, R. R.; PESSOA, M. C.; BONIFÁCIO, A.; LOBO, R. B.; DE OLIVEIRA, H. N. Study of the effect of genotype–environment interaction on age at first calving and production traits in Nellore cattle using multi-trait reaction norms and Bayesian inference. *Animal Science Journal*, vol. 89, no. 7, p. 939–945, 2018. <https://doi.org/10.1111/asj.12994>.
- ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS EXPORTADORAS DE CARNES [ABIEC]. Beef Report: Perfil da pecuária no Brasil. [S. l.: s. n.], 2021. Available at: <http://abiec.com.br/publicacoes/beef-report-2021/>.
- BECK, J. A. et al Genealogies of mouse inbred strains. *Nature Genetics*, vol. 24, no. 1, p. 23–25, 2000. <https://doi.org/10.1038/71641>.
- BOCCHI, A. L. Peso Pré-Desmama Em Bovinos De Uma População Peso Pré-Desmama Em Bovinos De Uma População. 2006.Tese, Doutorado,UNESP, Jaboticabal. 2006.
- CHALOVA, N. A. Obtaining Intermediate and Designed Animal Groups within Hybridization Systems. , p. 54–63, 2021. <https://doi.org/10.32743/kuz.mepa.2020.54-63>.
- DELVEAUX A. B. C. et al Association between residual feed intake, digestion, ingestive behavior, enteric methane emission and nitrogen metabolism in Nellore beef cattle. *Animal Science Journal*, vol. 91, no. 1, p. 1–9, 2020. <https://doi.org/10.1111/asj.13455>.
- ELER, J. P. Seleção. Pirassununga: Universidade de São Paulo, 2017a. <https://doi.org/10.11606/9788566404135>.
- ELER, J. P. Bases do Melhoramento Genético Animal. Pirassununga: Universidade de São Paulo, 2017b. <https://doi.org/10.11606/9788566404128>.
- EUSEBI, P. G.; MARTINEZ, A.; CORTES, O. Genomic tools for effective conservation of livestock breed diversity. *Diversity*, vol. 12, no. 1, 2020. <https://doi.org/10.3390/d12010008>.

FREITAS, A. P. et al. Different selection practices affect the environmental sensitivity of beef cattle. *Plos one*, vol. 16, no. 4 April, p. 1–18, 2021. <https://doi.org/10.1371/journal.pone.0248186>.

FREUDENSTEIN, J. V. et al. Biodiversity and the Species Concept—Lineages are not Enough. *Systematic Biology*, vol. 66, no. 4, p. syw098, 26 Oct. 2016. DOI 10.1093/sysbio/syw098. Available at: <http://academic.oup.com/sysbio/article-lookup/doi/10.1093/sysbio/syw098>.

FUERST, J. Lineage population: A concept needed by an observer of nature? *Mankind Quarterly*, vol. 57, no. 4, p. 590–631, 2017. <https://doi.org/10.46469/mq.2017.57.4.7>.

GINJA, C. et al. The genetic ancestry of American Creole cattle inferred from uniparental and autosomal genetic markers. *Scientific Reports*, vol. 9, no. 1, p. 1–16, 2019. <https://doi.org/10.1038/s41598-019-47636-0>.

HAGAN, B. A.; CUE, R. Generation intervals in Canadian dairy cattle herds. *Canadian Journal of Animal Science*, vol. 100, no. 1, p. 175–183, 2020. <https://doi.org/10.1139/cjas-2019-0053>.

INSTITUTO DE ZOOTECNIA. Plantel - Corte. 2021. Available at: <http://www.iz.sp.gov.br/pagina.php?id=23>. Accessed on: 30 Dec. 2021.

KUKUČKOVÁ, V. et al. Genetic diversity of local cattle. *Acta Biochimica Polonica*, vol. 65, no. 3, p. 421–424, 2018. https://doi.org/10.18388/abp.2017_2347.

KUMAR, M. et al. Heterosis—Genetic Basis and Utilization for Improvement of Dairy Cattle: A Review. *Research & Reviews: Journal of Dairy Science & Technology*, vol. 5, no. 2, p. 40–50, 2016. <https://doi.org/10.37591/rrjodst.v5i2.479>.

LEVER, J.; KRZYWINSKI, M.; ALTMAN, N. Points of Significance: Principal component analysis. *Nature Methods*, vol. 14, no. 7, p. 641–642, 2017. DOI 10.1038/nmeth.4346. Available at: <http://dx.doi.org/10.1038/nmeth.4346>.

LOPES, F. B. et al. Efeito da endogamia sobre pesos pré-desmame em bovinos da raça Nelore mocho criados a pasto no bioma Cerrado. *Archivos de Zootecnia*, vol. 65, no. 250, p. 177–182, 2016. <https://doi.org/10.21071/az.v65i250.485>.

MACNEIL, M. D. et al. E. Effects of inbreeding and heterosis in Hereford females on fertility, calf survival and preweaning growth. *Journal of animal science*, vol. 67, no. 4, p. 895–901, 1989. <https://doi.org/10.2527/jas1989.674895x>.

MALHADO, C. H. M, et al. Histórico genético e populacional do rebanho Nelore Puro de Origem no Sertão Nordeste. *Pesquisa Agropecuária Brasileira*, vol. 44, no. 7, p. 713–718, 2009. <https://doi.org/10.1590/S0100-204X2009000700010>.

MENEZES, A. C. B. et al. Feeding behavior, water intake, and energy and protein requirements of young Nellore bulls with different residual feed intakes. *Journal of animal science*, vol. 98, no. 9, 2020. <https://doi.org/10.1093/jas/skaa279>.

MKIZE, L. S.; ZISHIRI, O. T. Population genetic structure and maternal lineage of South African crossbred Nguni cattle using the cytochrome b gene in mtDNA. *Tropical Animal Health and Production*, vol. 52, no. 4, p. 2079–2089, 2020. <https://doi.org/10.1007/s11250-020-02231-8>.

MUELLER, S. et al. Growth, carcass, and meat quality of 2 dual-purpose chickens and a layer hybrid grown for 67 or 84 D compared with slow-growing broilers. *Journal of Applied Poultry Research*, vol. 29, no. 1, p. 185–196, 2020. DOI 10.1016/j.japr.2020.10.005. Available at: <https://doi.org/10.1016/j.japr.2019.10.005>.

NEWMAN, S. Genetic's History: USDA ARS. 2016. Available at: <https://www.ars.usda.gov/plains-area/miles-city-mt/larri/docs/genetics-history/>.

OLIVEIRA, J. H. F.; MAGNABOSCO, C. de U.; BORGES, A. M. de S. M. Nelore : Base Genética e Evolução Seletiva no Brasil. Embrapa Cerrados, , p. 50, 2002. Available at: <http://www.infoteca.cnptia.embrapa.br/bitstream/doc/566499/1/doc49.pdf>.

OLIVEIRA, P. S. et al. Estrutura populacional de rebanho fechado da raça Nelore da linhagem Lemgruber. *Pesquisa Agropecuária Brasileira*, vol. 46, no. 6, p. 639–647, 2011. <https://doi.org/10.1590/S0100-204X2011000600010>.

PAPACHRISTOU, D. et al. Genomic diversity and population structure of the indigenous Greek and Cypriot cattle populations. *Genetics, selection, evolution : GSE*, vol. 52, no. 1, p. 43, 2020. DOI 10.1186/s12711-020-00560-8. Available at: <https://doi.org/10.1186/s12711-020-00560-8>.

PEIXOTO, M. G. C. D. et al. Genetic diversity and population genetic structure of a guzerá (*Bos indicus*) meta-population. *Animals*, vol. 11, no. 4, p. 1–14, 2021. <https://doi.org/10.3390/ani11041125>.

PEREIRA, J. C. C. *Melhoramento Genético Aplicado à Produção Animal*. Belo Horizonte: FEPMVZ, 2008.

PERIPOLLI, E. et al Autozygosity islands and ROH patterns in Nellore lineages: Evidence of selection for functionally important traits 06 Biological Sciences 0604 Genetics. *BMC Genomics*, vol. 19, no. 1, p. 1–14, 2018. <https://doi.org/10.1186/s12864-018-5060-8>.

PITT, D. et al. Demography and rapid local adaptation shape Creole cattle genome diversity in the tropics. *Evolutionary Applications*, vol. 12, no. 1, p. 105–122, 2019. <https://doi.org/10.1111/eva.12641>.

RAMAYO-CALDAS, Y. et al. Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. *Genetics Selection Evolution*, vol. 48, no. 1, p. 1–9, 2016. <https://doi.org/10.1186/s12711-016-0216-y>.

ROSA, A. do N.; MARTINS, E. N.; MENEZES, G. de O.; SILVA, L. O. C. da. *Melhoramento genético aplicado em Gado de Corte*. 1st ed. Brasília: Embrapa, 2013.

SANTANA, M. H. A. et al. The genetic and genomic effects of Nellore lineages on feed efficiency, intake and performance. *Livestock Science*, vol. 228, no. March, p. 104–108, Oct.

2019. DOI 10.1016/j.livsci.2019.08.003. Available at:
<https://linkinghub.elsevier.com/retrieve/pii/S1871141318306176>.

SEO, D. et al. Identification of target chicken populations by machine learning models using the minimum number of snps. *Animals*, vol. 11, no. 1, p. 1–18, 2021.
<https://doi.org/10.3390/ani11010241>.

SIMČIČ, M. et al. Recovery of native genetic background in admixed populations using haplotypes, phenotypes, and pedigree information - Using cika cattle as a case breed. *Plos one*, vol. 10, no. 4, p. 1–20, 2015. <https://doi.org/10.1371/journal.pone.0123253>.

SIMMONS, M. A. et al. Distributional characterizations and testing for differences of relatedness and inbreeding of a subpopulation of American Hereford bulls. *Translational Animal Science*, vol. 5, no. 1, p. 1–10, 2021. <https://doi.org/10.1093/tas/txab008>.

STRONEN, A. V. et al. Genomic analyses suggest adaptive differentiation of northern European native cattle breeds. *Evolutionary Applications*, vol. 12, no. 6, p. 1096–1113, 2019. <https://doi.org/10.1111/eva.12783>.

SUMREDDEE, P. et al. Inbreeding depression in line 1 Hereford cattle population using pedigree and genomic information. *Journal of Animal Science*, vol. 97, no. 1, p. 1–18, 2019. <https://doi.org/10.1093/jas/sky385>.

SWALVE, H. H. Theoretical basis and computational methods for different test-day genetic evaluation methods. *Journal of Dairy Science*, vol. 83, no. 5, p. 1115–1124, 2000. DOI 10.3168/jds.S0022-0302(00)74977-0. Available at: [http://dx.doi.org/10.3168/jds.S0022-0302\(00\)74977-0](http://dx.doi.org/10.3168/jds.S0022-0302(00)74977-0).

TAKESHIMA, S. N. et al. Genetic diversity of BoLA-DRB3 in South American Zebu cattle populations. *BMC Genetics*, vol. 19, no. 1, p. 1–13, 2018. <https://doi.org/10.1186/s12863-018-0618-7>.

TWOMEY, A. J.; CROMIE, A. R.; MCHUGH, N.; BERRY, D. P. Validation of a beef cattle maternal breeding objective based on a cross-sectional analysis of a large national cattle database. *Journal of Animal Science*, vol. 98, no. 11, p. 1–15, 2020. <https://doi.org/10.1093/JAS/SKAA322>.

UPADHYAY, M. et al. Genomic relatedness and diversity of Swedish native cattle breeds. *Genetics Selection Evolution*, vol. 51, no. 1, p. 1–11, 2019. DOI 10.1186/s12711-019-0496-0. Available at: <https://doi.org/10.1186/s12711-019-0496-0>.

WARD, J. H. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, vol. 58, no. 301, p. 236–244, 1963. <https://doi.org/10.1080/01621459.1963.10500845>.

CAPÍTULO 2- SIMILARIDADE GENÉTICA EM POPULAÇÃO BOVINA DA RAÇA NELORE

Milena Aparecida Ferreira Campos¹, Adriana Luize Bocchi², Edgar Alain Collao Saenz³

¹Aluna do Programa de Pós Graduação Em Biociência Animal, Universidade Federal de Jataí, Jataí, GO, Brasil

²Professora Associada, Área de Melhoramento Genético Animal, Universidade Federal de Jataí, Jataí, GO

³Professor Associado, Área de Modelos de Simulação em Ruminantes, Universidade Federal de Jataí, GO

RESUMO

Explorar ao máximo a diversidade genética dentro das populações por meio de acasalamentos dirigidos garante longevidade a seleção e contínuo progresso genético. É conhecido que muitos dos touros formadores de linhagem deram espaço a novos reprodutores dentro de cada rebanho. A partir do banco de dados proveniente da empresa Genética Aditiva (n= 6.579 indivíduos com informações completas), objetivou-se identificar grupos geneticamente similares para características de interesse econômico. Todas as análises estatísticas foram realizadas por meio de linguagem R Core Team®. A análise de correlação de Pearson e de PCA foram utilizadas para verificar o comportamento das características e auxiliar na escolha das variáveis que seriam consideradas para o agrupamento de acordo com sua similaridade genética. Os quatro clusters formados mostraram diferenças de desempenho e superioridade predominante de alguns animais para as características analisadas. A verificação dos reprodutores responsáveis pelos animais superiores pode ser uma forma de entender melhor a estrutura populacional do rebanho. A taxa de endogamia encontrada dentro dos clusters foi baixa, demonstrando que há espaço para uma seleção intensa que vise a fixação dessas características. As análises indicam que agrupar animais de acordo com a similaridade genética possibilita a seleção de indivíduos para formação de linhagens especializadas, além de ser uma forma de direcionar o trabalho de seleção e verificar quais as deficiências dentro do rebanho e o potencial genético existente.

Palavras Chave: acasalamento, agrupamento hierárquico, cluster, seleção

ABSTRACT

GENETIC SIMILARITY IN NELORE CATTLE POPULATION

Making the most of genetic diversity within populations through targeted mating ensures longevity, selection and continuous genetic progress. It is known that many of the lineage-forming bulls gave rise to new sires within each herd. The database came from the company Genética Aditiva, (n= 6.579 individuals with complete information. The aim was to identify genetically similar groups for traits of economic interest. All statistical analyzes were performed using the R Core Team ® language. Pearson's correlation analysis and PCA's were used to verify the behavior of the traits and help in choosing the variables that would be considered for the grouping according to their genetic similarity. The four clusters formed showed differences in performance and predominant superiority of some animals for the analyzed characteristics. The verification of the reproducers responsible for the superior animals can be a way to better understand the population structure of the herd. The inbreeding rate within clusters was low, demonstrating that there is space to intense selection aimed at fixing these characteristics. The analyzes indicated that grouping animals according to genetic similarity allows the selection of individuals for the formation of specialized lines, in addition to being a way of directing the selection work and verifying the deficiencies within the herd and the existing genetic potential.

Keywords: mating, hierarchical grouping, cluster, selection

1 Introdução

Estima-se que o rebanho bovino corte e leite brasileiro possua cerca de 187,55 milhões de cabeças (ABIEC, 2021) e, desse total, segundo a Associação dos Criadores de Nelore do Brasil [ACNB] (2021), 80% são representados por animais de raça Nelore ou mestiços da raça. Para alcançar números tão expressivos o Nelore passou por processo de formação e permanece em seleção com o objetivo de obter animais cada vez mais eficientes para produção de carne de qualidade, boa reprodução, eficiência alimentar, adaptabilidade, resistência a doenças e bom temperamento.

As importações de *Bos taurus indicus* iniciaram na segunda metade do século XVI, e o gado Ongole, como era conhecido na Ásia, chamava atenção por sua rusticidade e adaptação ao clima tropical. Calcula-se que durante o período de 1868 a 1963 (PÉREZ-PARDAL et al., 2018) foram importados cerca de 7000 animais zebuínos, número bem menor quando comparado aos animais de raças taurinas trazidos durante a colonização, que somaram cerca de 800 mil cabeças (MUDADU et al., 2016).

As linhagens precursoras do rebanho Nelore vêm de touros como Karvadi, Taj Mahal, Golias, Godhavari, Rastã, Akasamu e Padhu (FERNANDES JÚNIOR et al., 2020), que foram touros que apresentavam fenótipos de interesse para a seleção inicial e por isso foram utilizados intensivamente, permitindo que houvesse a fixação de genes para características marcantes em seus descendentes. No entanto, essas linhagens formadas inicialmente utilizadas de modo combinado e deram espaço a novos reprodutores dentro de cada rebanho, tendo assim diferentes famílias responsáveis pelos touros e vacas superiores dos dias de hoje.

Com a evolução na pesquisa e o uso de novas tecnologias, a seleção ganhou novas ferramentas possibilitando que o processo de melhoramento genético animal de bovinos esteja em contínuo progresso. A seleção artificial e consequente maior intensidade de uso de determinados reprodutores melhoradores, reduz a frequência de genes na população permitindo que determinadas características sejam fixadas dentro do rebanho (CARDOSO et al., 2018). A endogamia é um sistema de acasalamento entre os indivíduos que é utilizada como estratégia para aumentar a prepotência nas linhagens, uma vez que indivíduos mais endogâmicos possuem menor variação de gametas tornando a progênie mais uniforme. Desta forma os animais, após um determinado número de gerações, dependendo da variabilidade genética inicial, apresentarão composição semelhante dentro de sua própria linhagem distanciando-se geneticamente de outras populações, mesmo que de animais de mesma raça.

Acasalamentos entre linhagens diferentes podem permitir a manifestação da heterose, expressa como o desempenho das progênies superior à média dos pais de linhagem pura,

causado pela restauração da perda de vigor adaptativo perdido pela endogamia (OLIVEIRA, et al., 2011). Estudar a estrutura de rebanho e a variabilidade genética é fundamental para implementação ou manutenção do sistema de seleção que vise longevidade do programa de melhoramento da raça, possibilita maior intensidade de seleção e consequente aumento no ganho genético, elevando a eficiência dos sistemas produtivos de gado de corte brasileiro (BARROS et al., 2018).

Nesse sentido, é importante conhecer famílias com maior representação dentro de rebanhos específicos de seleção e garantir maior objetividade e precisão para a seleção de indivíduos. No presente trabalho, objetivou-se identificar grupos geneticamente similares para características de interesse econômico em um rebanho de alto impacto na pecuária brasileira e países da América do Sul, que utilizam a genética da raça Nelore.

2 Material e Métodos

O banco de dados foi proveniente da empresa Genética Aditiva® por meio da Melhora Mais Consultoria Genética®, responsável técnica pela seleção e acasalamento dos animais na empresa. A Genética Aditiva participa do Programa de Melhoramento Genético Nelore Brasil da Associação Nacional de Criadores e Pesquisados (ANCP) e trabalha com seleção de animais a pasto. Os dados utilizados são referentes a avaliação genética do mês de fevereiro do ano de 2021. O banco original continha informações de 30.419 animais, incluindo pedigree, sexo e nascimento e diferenças esperadas na prole (DEPs) de avaliação genética e genômica para doze características. Dados com falta de informação de pedigree ou de avaliações genéticas para as características avaliadas foram excluídos da análise, totalizando 6.579 indivíduos com informações completas no arquivo final. A análise descritiva da população está descrita na Tabela 1.

Tabela 1. Análise descritiva das diferenças esperadas na progênie da população analisada

	Média	DP	Var	Mín	Máx
MP120	2,308	1,828	3,340	-5,930	9,150
MP210	3,037	2,488	6,190	-8,660	12,340
DP210	11,302	4,098	16,791	-1,630	27,940
DP365	20,513	6,041	36,494	-1,530	44,020
DP450	21,591	6,536	42,720	-4,230	47,290
DAOL	2,220	1,548	2,396	-3,120	8,870
DACAB	0,395	0,332	0,110	-0,870	1,950
D3P	76,271	7,712	59,479	38,950	95,770
DMAR	0,059	0,137	0,019	-0,480	0,500
DCAR	-0,013	0,079	0,006	-0,357	0,332
DIMS	0,036	0,162	0,026	-0,665	0,626
DIPM	-0,506	0,527	0,277	-2,524	1,771

DP: desvio padrão; Var: variância; Min: mínimo; Max: máximo; MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca; CV: coeficiente de variação;

As análises estatísticas foram realizadas por meio de linguagem R Core Team (2021). A análise de correlação de Pearson foi utilizada para verificar o comportamento e direções das características considerada de acordo com a metodologia descrita por Kendall (1938) e Becker et al. (1988).

Para reduzir a dimensionalidade dos dados e facilitar a interpretação das variáveis foi realizada análise de componentes principais (PCA), que retorna um número de componentes ou dimensões exatamente igual ao número de variáveis avaliadas e os classifica de acordo com a porcentagem de variação explicada por cada um deles. Apesar de não existir uma regra clara com relação a quantas dimensões devem ser consideradas, se recomenda que o número escolhido represente mais de 70% da variação total amostral (KASSAMBARA, 2017).

O pacote FactoMineR (HUSSON, et al., 2020) foi utilizado para a realização da PCA seguida da clusterização por meio da função HCPC- Hierarchical Clustering on Principal Components (HUSSON; JOSSE; JÉRÔME PAGÈS, 2010). Correlação entre as variáveis e os componentes e a contribuição das variáveis para os mesmos, são demonstrados como resultados da análise PCA permitindo melhor entendimento da composição dos componentes escolhidos. Os cluster foram formados com base nas distâncias euclidianas, por meio do método de Ward (WARD, 1963), agrupados de acordo com a similaridade de seus valores genéticos em quatro clusters.

A análise descritiva dos clusters levou em consideração as médias de cada agrupamento, número de animais, variância e desvio padrão, assim como valores mínimos e máximos. Os 10

touros com maior número de filhos dentro de cada cluster foram identificados, assim como sua representatividade com relação ao número de filhos. A matriz de parentesco e o coeficiente médio de endogamia foram obtidos pelo método de Meuwissen & Luo (1992), pacote GeneticsPed (GORJANC; HENDERSON, 2021) utilizando linguagem R. As informações de pedigree precisaram ser renomeadas e reordenadas de modo que os animais mais velhos aparecessem primeiro no arquivo, permitindo a formação da matriz de parentesco e cálculo do coeficiente médio de endogamia da população total e por cluster. Nesta análise foram considerados todos os pedigrees do arquivo inicial, independente dos dados faltantes de avaliações genéticas. A representatividade dos reprodutores como pais foi calculada por meio da proporção de filhos de cada touro pela razão do total de animais no cluster.

3 Resultados

3.1 Análises descritivas e correlações de Pearson da população analisada

A Figura 1 ilustra as correlações entre as características estudadas.

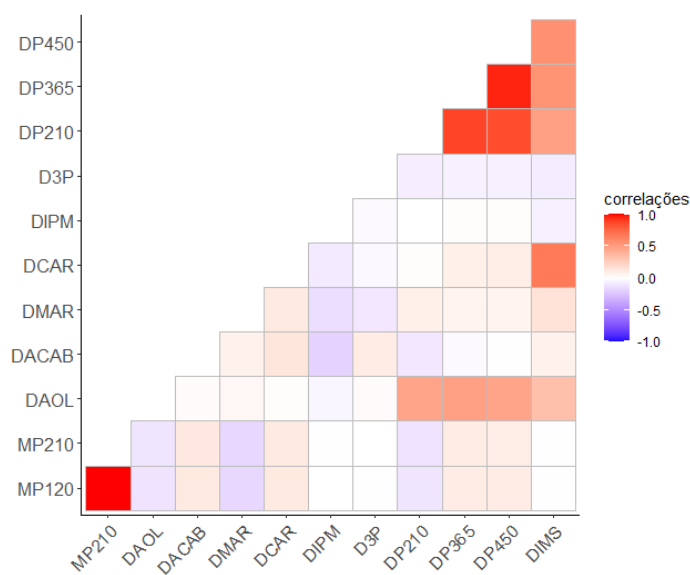


FIGURA 1. Correlações de Pearson entre características estudadas

Nota: MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

As características maternas foram altas e positivamente correlacionadas entre elas, assim como as características de peso em diferentes idades. As correlações entre pesos e AOL e, pesos e IMS foram medianas e positivas, igualmente entre AOL e IMS. A característica

DCAR possui correlação media a alta com IMS e baixa e positiva com DACAB. DACAB é correlacionada negativamente com DIPM, porém ainda considerada uma correlação baixa. D3P apresenta correlações baixas com todas as características, porem correlações negativas com MP120, DMAR, DCAR, DIPM, Pesos em diferentes idades e IMS, e correlação positiva com DAOL e ACAB.

A análise de correlações apresenta juntamente com a análise de PCA justificativas para que essas características tenham se comportado de forma similar ou antagônica, facilitando a interpretação quando da formação dos componentes principais.

3.2 Análise de componentes principais

As informações inclusas no banco de dados correspondem à inércia total ou variabilidade dos dados que ele contém, o número de componentes principais escolhidos precisa ser suficiente para conter o máximo de variação dos dados (ABDI; WILLIAMS, 2010). Os primeiros componentes são responsáveis por reter o máximo de variação possível, quando a perda de inércia é muito pequena significa que usar um maior número de componentes não é justificável. As FIGURAS 2 e 3 ilustram como a perda de inércia influencia na escolha da quantidade de componentes.

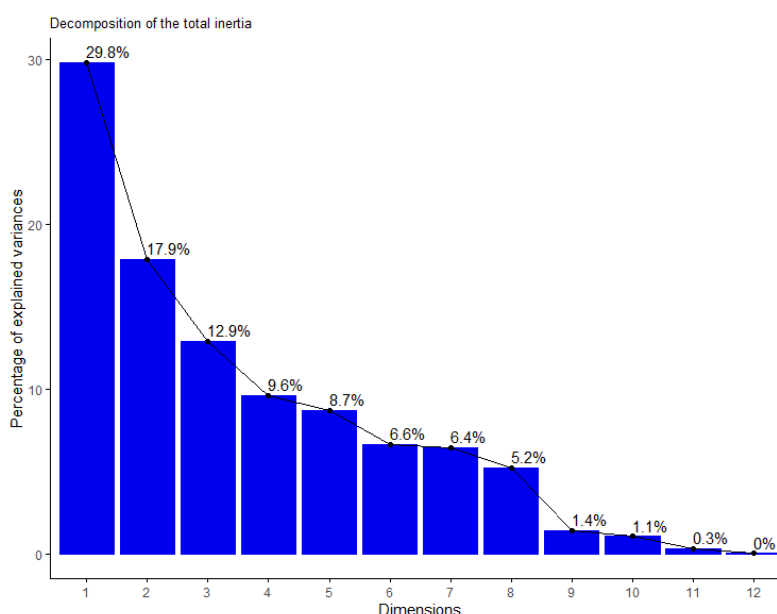


FIGURA 2. Porcentagem de variação explicada pelos componentes principais

Os 5 primeiros componentes representam 79% da porcentagem de variação.

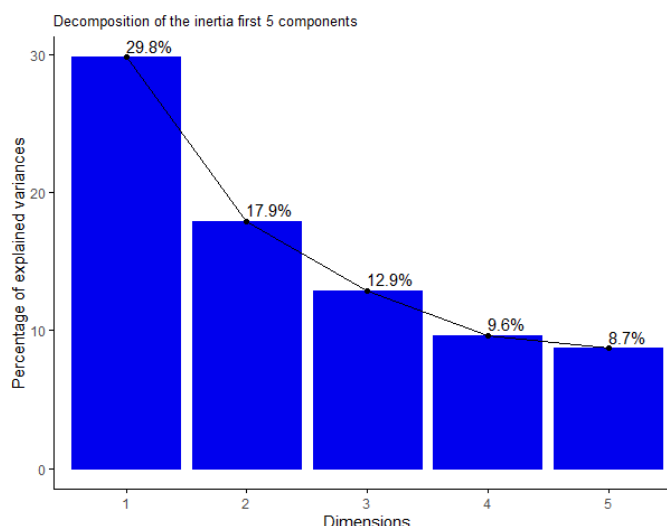


FIGURA 3. Porcentagem de variação explicada pelos 5 primeiros componentes principais

A importância de uma variável para o componente é demonstrada pela proporção de contribuição da variável para o componente, quanto maior a porcentagem, mais a variável contribui para o componente (FIGURA 4).

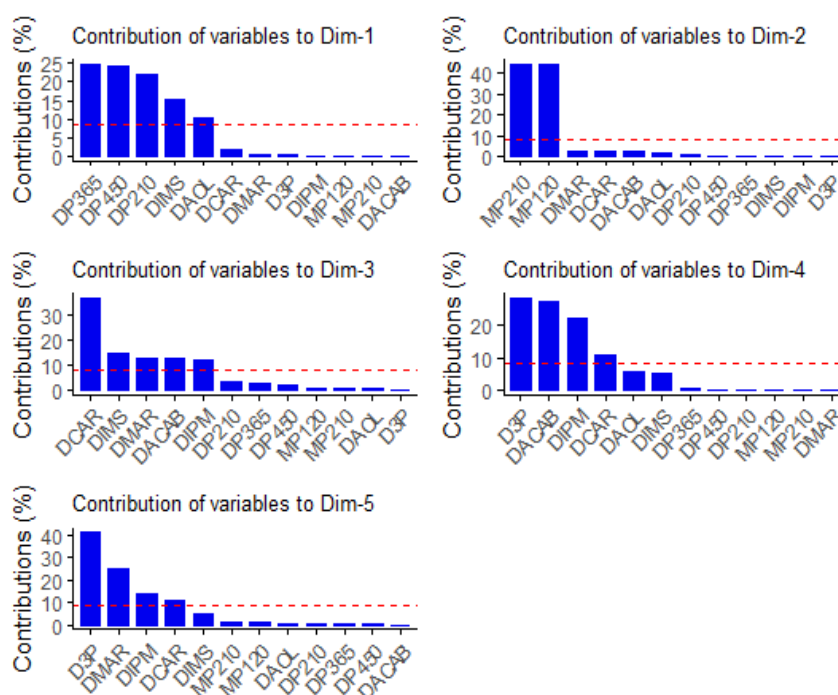


FIGURA 4. Contribuição das variáveis para as dimensões utilizadas

MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

O Componente Principal 1 é definido principalmente pelas características de peso em diferentes idades, DIMS e DAOL, enquanto o Componente 2 é caracterizado pelas duas

características maternas. O Componente Principal 3 pela característica de eficiência alimentar, DCAR. O Componente Principal 4 por DIPM e DACAB e o último componente considerado apresenta maior contribuição das características de D3P e DMAR.

As características possuem contribuições diferentes em cada componente principal, o que faz com que sejam melhor representados por determinadas características. As variáveis com maior contribuição nos primeiros componentes são responsáveis por explicar a maior variabilidade dos dados da população, não existe necessariamente uma correlação entre as características, elas apenas possuem variação similar dentro do conjunto de dados (HUSSON, François; JOSSE; JÉRÔME PAGÈS, 2010)

Na FIGURA 5 se observa a correlação entre características e componentes principais.

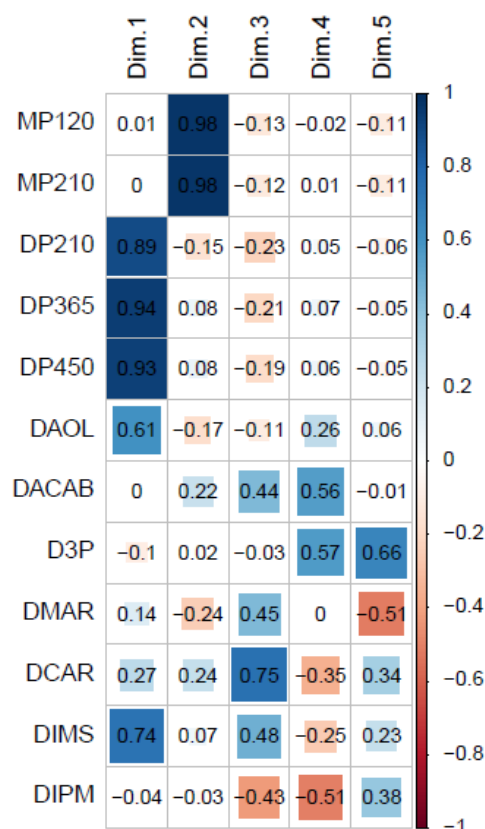


FIGURA 5. Correlação entre as variáveis e os componentes principais

MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

A análise de correlação e as contribuições das variáveis para os componentes (FIGURA 4) tem comportamento similar com relação às características que compõe cada componente principal. A descrição das variáveis é fornecida pela análise de PCA, e possibilita a observação de variância que pode ser explicado por cada variável nos componentes (Tabela 2).

Tabela 2. Proporção da variância explicada pelas variáveis em estudo (Eta²)

Característica	Eta ²	P-value
MP120	0,32	0,0E+00
MP210	0,32	0,0E+00
DP210	0,54	0,0E+00
DP365	0,56	0,0E+00
DP450	0,55	0,0E+00
DAOL	0,23	0,0E+00
DIMS	0,34	0,0E+00
DMAR	0,06	1,2E-88
DCAR	0,05	2,1E-71
D3P	0,05	2,0E-70
DACAB	0,04	1,4E-56
DIPM	0,03	6,4E-33

MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca; Eta²: proporção da variância associada ou contabilizada por cada um dos principais efeitos; p-value>0.05

3.3 Clusterização

Com base nas distâncias euclidianas, os animais foram clusterizados, a FIGURA 6 no canto direito inferior demonstra que a perda de inércia de cinco clusters para quatro é muito menor, quando comparada à perda de quatro para três clusters, determinando que o número de quatro clusters é o suficiente para agrupar os animais sem perda de informação. Os quatro grupos definidos no estudo permitiram observar variações com relação ao desempenho predito dos animais, em diferentes categorias, evidenciando conjuntos de animais mais especializados em determinadas variáveis.

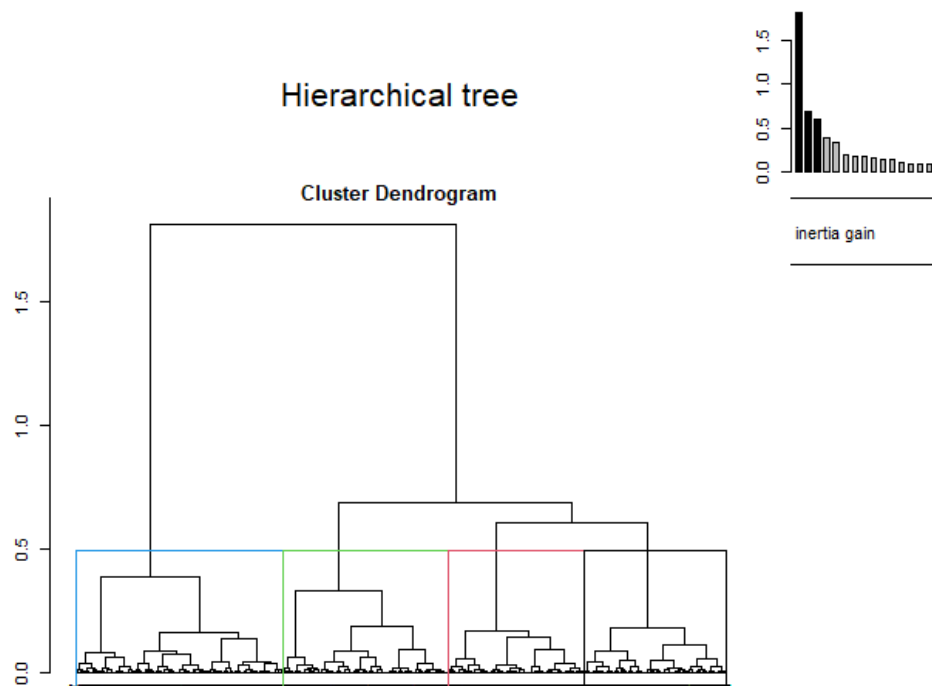


FIGURA 6. Dendrograma demonstrando clusterização e escolha de 4 clusters

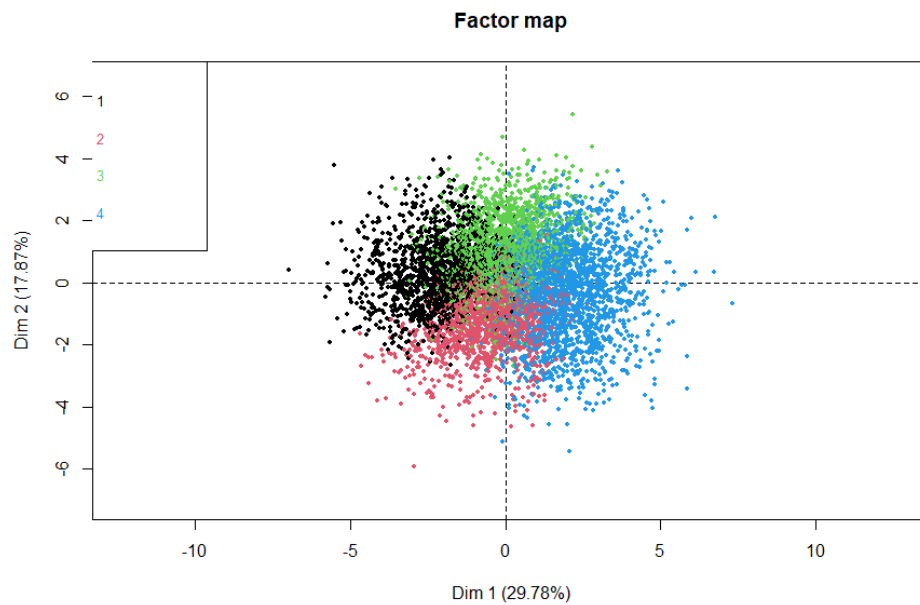


FIGURA 7. Indivíduos agrupados de acordo com o cluster a que pertencem, representando a primeira e segunda dimensão da análise de HCPC

A FIGURA 8 ilustra a distribuição dos animais nos clusters de acordo com o valores de similaridade genética para as características escolhidas.

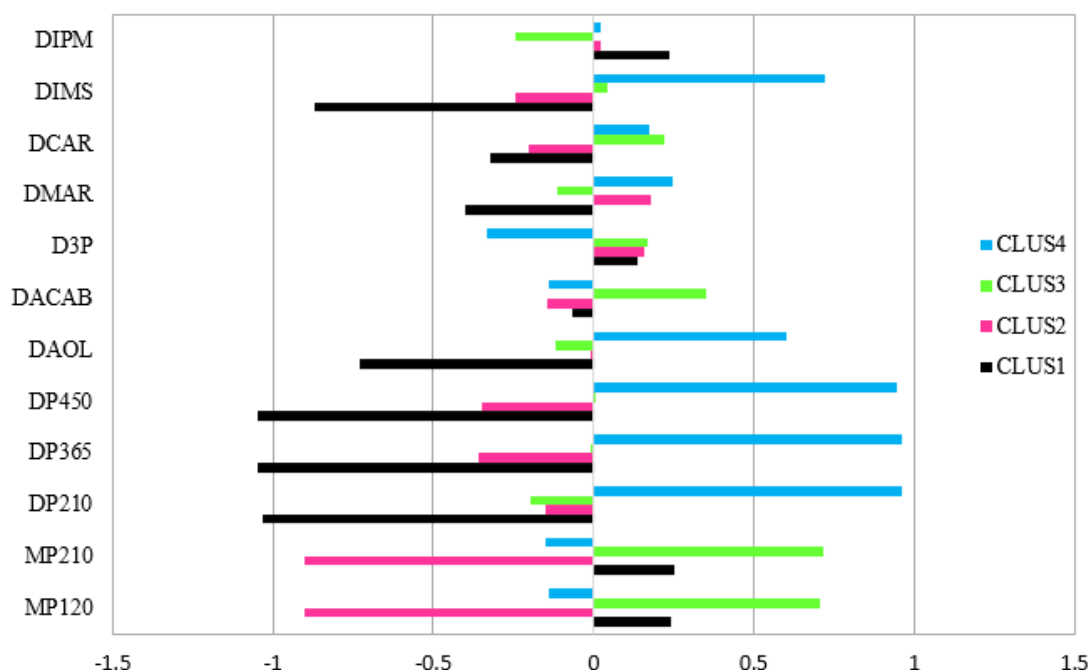


FIGURA 8. Representação do valor médio de DEPs dos animais que compõe os clusters.

MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

Os animais agrupados no CLUSTER 1 apresentam os menores valores para DIMS, DCAR, DMAR, DP210, DP365, DP450 e DAOL, e maior valor para DIPM. São, portanto, animais não tão precoces ($DIPM >$), com boa eficiência alimentar, porém quando comparados ao restante da população, apresentam valor genético inferior para pesos em diferentes idades e carcaça. O CLUSTER 2 apresenta de modo geral, valores medianos para todas as características e os menores valores para MP120, MP210 e DACAB. Não apresentam grande destaque para nenhuma das características, mas médias interessantes e equilibradas.

No CLUSTER 3 estão os animais com maiores médias para as características MP120, MP210, DACAB, D3P, DCAR e DIMS, e com as menores médias para DIPM. Foram agrupados aqui, animais com os maiores valores para as características maternas, superiores para precocidade sexual e de acabamento, porém com eficiência alimentar inferior. O CLUSTER 4 agrupou os animais com maiores médias para as três características de peso, DAOL e DMAR e as menores médias para D3P. Composto principalmente pelos animais destaques para valores de peso, com um bom indicativo para maior quantidade de carne na carcaça e com bons valores de gordura intramuscular

3.4 Representatividade dos reprodutores por grupo

Verificou-se a frequência dos progenitores dos animais em cada cluster, para determinar quais touros mais contribuem para a média de cada agrupamento. Os dez touros com maior número de filhos foram utilizados para melhor descrever cada cluster, como demonstrado na TABELA 3.

Tabela 3. Médias e porcentagens de representação dos dez touros com maior número de filhos em cada cluster, assim como o coeficiente médio de endogamia de cada cluster.

	CLUS 1	CLUS 2	CLUS 3	CLUS 4
Total animais no cluster	1428	1377	1682	2092
Total de filhos dos 10 touros	655	738	803	1174
% de representatividade desses touros como pais	46%	54%	48%	56%
Total de touros no cluster	137	111	123	116
% de representatividade desses touros no cluster	7%	9%	8%	9%
Coeficiente de endogamia médio	0,60%	0,61%	0,72%	0,71%

CLUS 1: CLUSTER 1; CLUS 2: CLUSTER 2; CLUS 3: CLUSTER 3; CLUS 4: CLUSTER 4.

Observa-se que os dez touros escolhidos, com o número de filhos somados representam 46% de todos os outros animais presentes no CLUSTER 1, 54% de todos no CLUSTER 2 e 48% e 56% para os CLUSTERS 3 e 4 respectivamente, demonstrando que esses touros apresentam grande representatividade dentro de cada agrupamento. Foi observado também que dentro desse grupo de 10 touros mais representativos, dois reprodutores aparecem em todos os clusters, outros dois animais aparecem em três dos quatro clusters, e outros seis aparecem em pelo menos dois clusters.

4 Discussão

O entendimento do comportamento das variáveis analisadas, possibilita compreender como as estratégias de seleção e as interpretações das análises podem ser influenciados por suas correlações. No caso de características maternas, as correlações positivas e altas indicam que a seleção para uma das características maternas (MP120 ou MP210) terá reflexo na outra de modo similar. Até o desmame há grande influência da vaca no desempenho da progênie, sendo um bom período para a obtenção de medidas que permitam antecipar o processo de seleção, além de possibilitar a predição da capacidade materna (BOLIGON et al., 2019).

De modo semelhante às características maternas, os pesos em diferentes idades se apresentam positivas e altamente correlacionadas, com valores de 0,88; 0,84 e 0,96 para DP210-DP365, DP210-DP450 e DP365-DP450, respectivamente. As altas correlações indicam que os

genes responsáveis pelo maior peso enquanto bezerro são praticamente os mesmos para maior peso adulto (BARROS et al., 2018), o que otimiza a seleção desde o início da vida desses animais.

Tramonte et al. (2019) também encontraram correlações positivas e altas para características de peso em bovinos da Raça Guzerá, 0,95; 0,91; 0,97 para DP210-DP365, DP210-DP450 e DP365-DP450 respectivamente. Valores semelhantes aos nossos, foram observados por Kluska et al., (2018) em uma população de bovinos Nelore com correlações entre DP210 e DP450 dias de 0,95 e correlações medianas à altas entre a característica de DAOL com DP210 e DP450, 0,48 e 0,47 respectivamente. A característica de peso ao ano é uma boa representação dos outros pesos ajustados considerados pela seleção, porque ajuda a analisar o desempenho do animal por conta própria após o desmame.

A seleção para DCAR mede a eficiência alimentar, e possui independência fenotípica do ganho diário (SEABURY et al., 2017). Não teve altas correlações com as características estudadas, com exceção de DIMS (0,67). Bonamy et al. (2019) observaram comportamento parecido entre DCAR e a característica reprodutiva D3P, com correlação de -0,04, enquanto no presente estudo foi de -0,03. É uma característica passível de seleção, visto que estimativas de herdabilidade por meio de marcadores moleculares em bovinos da Raça Angus, Hereford e SimAngus (Simental x Angus) sugerem valores moderados a altos para CAR, sendo 0,20; 0,49; 0,40, respectivamente para cada raça (SEABURY et al., 2017).

A característica D3P apresentou baixas correlações com todas as características, por ser uma característica reprodutiva sofre muita influência ambiental podendo justificar a correlação não tão expressiva. Kluska et al. (2018) encontraram correlações um pouco mais altas entre D3P e características de carcaça, D3P-DAOL (0,11) e D3P-DACAB (0,35), porém positivas, do mesmo modo das que foram encontradas aqui, D3P-DAOL (0,02), D3P-DACAB (0,10), e um pouco mais alta em D3P-DP210 (-0,03) quando comparada a nossa de D3P-DP210 (-0,08).

Silva Neto et al. (2020) encontraram correlações medianas e negativas entre D3P e DIPM (-0,24), assim como a encontrada no presente estudo que foi negativa, porém baixa (0,02). A correlação negativa entre essas duas características é favorável, porque demonstra que a seleção mesmo que em sexos diferentes favorece o progresso do rebanho em busca de animais mais precoces.

As características analisadas também podem ser descritas pelo Eta^2 que se refere a proporção de variância que cada variável representa. Desta forma, é possível observar o quanto cada característica é responsável pela variabilidade dos dados dentro da população analisada.

É estratégico conhecer a variabilidade genética das populações de bovinos para garantir o ganho genético, assim como o entendimento de quais características já estão com maior fixação de genes dentro do rebanho, e quais ainda precisam maior atenção.

As três características de peso às diferentes idades apresentaram maior valor de Eta^2 , indicando serem responsáveis por grande variação na população. Dentro do rebanho há animais com valores mínimos para DP365 de -1,53 no CLUSTER 1 e animais com valores máximos para a mesma característica de 44,02 no CLUSTER 4, demonstrando que apesar de serem características selecionadas há muito tempo, não são selecionadas para valores extremos, buscando um equilíbrio não só entre elas, mas também com os outros objetivos de seleção da propriedade.

As características maternas também apresentam alta variabilidade dentro da população total o que permite a identificação de animais superiores que podem agregar ao ganho genético da população, apesar de alguns animais já terem valores expressivos de DEPs para esses critérios. No entanto, as características reprodutivas e de carcaças aparecem com menores valores para Eta^2 . Apesar de não serem grandes responsáveis pela variância dos dados, se observa uma amplitude entre as médias desses critérios nos CLUSTERS. Um exemplo é a característica DACAB, que apresentou baixo valor de Eta^2 (0,04), no entanto quando se compara os valores para essa característica no cluster CLUSTER 3, temos valores de DEP mínimos de -0,87 e máximo de 1,95 indicando variabilidade e, portanto, possibilidade de seleção para melhora da qualidade da carcaça.

As características de carcaça frequentemente apresentam ganho genético com boa velocidade, visto que apresentam herdabilidades medianas e são mensuradas em vida tanto em machos como em fêmeas (KLUSKA et al., 2018). Mesmo critérios com menor variabilidade nem sempre sugerem que já houve a fixação dessas características no rebanho. A característica de DIPM, apresenta a menor variação (0,027), entretanto foi recentemente adicionada às avaliações e há poucos dados disponíveis em relação às demais características.

A PCA indicou que os cinco primeiros componentes são capazes de explicar porcentagem suficiente da variação dos dados, não sendo, portanto, necessária a utilização das 12 dimensões. Essa redução garante maior objetividade no momento da clusterização, permitindo que a análise agrupe de forma mais coerente os animais. De mesma forma, é necessária a escolha de um número ótimo de clusters que seja suficiente para descrição e interpretação dos dados, visto que uma variedade muito grande de grupos pode não permitir uma boa interpretação. Uma forma de eleger a quantidade de clusters a serem escolhidos, assim como nos componentes principais, é observar a perda de inércia ou variação total.

O uso dos animais agrupados no CLUSTER 1 apresentam propriedades focadas na melhora de eficiência alimentar, porque são os animais com maior destaque para características que visem bom desempenho com menor consumo de alimentos. A comparação entre os agrupamentos evidencia desempenhos superiores para características nos diferentes clusters, indicando fixação das mesmas nesses animais.

O CLUSTER 2 é um agrupamento de médias de DEPs equilibradas, onde apesar de nenhuma das médias apresentarem grandes destaques, os animais desse grupo possuem valores superiores à média nacional do programa de melhoramento genético da ANCP como demonstrado na TABELA 4 na maioria das características, e portanto, poderiam contribuir de modo uniforme em fazendas que estão no início da seleção ou buscando a padronização de seus rebanhos, usando reprodutores equilibrados para todas as características avaliadas.

Tabela 4. Comparação entre médias dos animais agrupados no CLUSTER 2 com a média nacional dos animais avaliados pelo programa de melhoramento genético Nelore Brasil da ANCP.

Característica	Média nacional*	Cluster 2
DACAB (mm)	-0,030	0,347
DAOL (cm ²)	-0,330	2,202
DMAR (%)	0,000	0,083
DCAR (kg de MS)	-	-0,029
MP120 (kg)	0,820	0,664
MP210 (kg)	1,050	0,790
DIPM (meses)	-	-0,495
D3P (%)	64,280	77,497
DP210(kg)	3,020	10,681
DP365(kg)	6,240	18,330
DP450 (kg)	3,370	19,298
DIMS	-	-0,003

Abreviações: DACAB: acabamento; DAOL: área de olho de lombo; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade de precocidade em machos; MP120: peso maternal aos 120 dias; MP210: peso maternal aos 210 dias; D3P: probabilidade de parto precoce; DP365: peso ao ano

*Valores retirados do sumário da associação nacional de criadores e pesquisadores referentes aos animais avaliados pelo programa de melhoramento genético deles em agosto de 2021.

O grupo de animais com melhores médias para as características maternas do CLUSTER 3, apresentam também valores desejáveis para precocidade sexual de fêmeas e machos. O uso de reprodutores ou matrizes desse grupo, pode resultar em uma escolha eficiente para boas mães no rebanho, iniciando vida reprodutiva mais cedo, quando comparadas aos outros agrupamentos, e com boa capacidade de criar o bezerro. Já o CLUSTER 4 destaca animais mais pesados nas diferentes idades, característica muito buscada na pecuária de corte, visto que a principal forma de remuneração do produtor é por peso do animal, assim como

animais com indicativo de maior proporção de carne na carcaça e de boa qualidade, por apresentar maiores médias para DAOL e DMAR.

A análise permite que grupos de animais com valores inferiores de DEPs sejam identificados, assim como a descendência deles. Permite utilizar os resultados como critério de descarte ou modo estratégico quando se deseja fazer acasalamentos corretivos. Quando pensamos em efeito aditivo de complementariedade, podemos afirmar que, a utilização de animais com alto desempenho em um cluster pode compensar baixos desempenhos em outro com o acasalamento entre clusters. Vê-se também a possibilidade de maior fixação das características, entre os animais de mesmo cluster.

Empresas que trabalham com seleção e acasalamento utilizam um grande número de informações combinadas para realizar os acasalamentos dirigidos, garantindo longevidade na seleção. Assim mesmo, evitam-se os prejuízos causados em processo de depressão endogâmica, o que justifica os valores baixos encontrados neste estudo para os coeficientes de endogamia. Os acasalamentos devem ser feitos de modo particular dentro de cada rebanho considerando o coeficiente de endogamia e a escolha de características que podem compor índices de seleção, onde o mérito genético de características do indivíduo é demonstrado em um único valor, facilitando o uso para os técnicos e melhoristas. Esses índices são particularmente vantajosos para obter ganho genético quando existem correlações genéticas antagônicas entre as características que limitam a capacidade de mover ambas em direções desejáveis (TWOMEY et al., 2020).

O coeficiente de endogamia fornece a probabilidade de que em um loci aleatório, dois alelos sejam idênticos por descendência (MALÉCOT, 1948). Espera-se que um indivíduo tido como endogâmico esteja conectado no pedigree por linhas de descendência de um ancestral ou mais ancestrais em comum (WRIGHT, 1922). Considera-se o limite de 6,25% para a taxa de endogamia em bovinos, valores acima podem comprometer desempenhos fenotípicos dos animais (DOEKES; BIJMA; WINDIG, 2021; GONZÁLEZ-RECIO; LÓPEZ; GUTIÉRREZ, 2007; LOZADA-SOTO et al., 2021) .

A taxa média de endogamia encontrada em todos os clusters apresentou valores abaixo do que é considerado preocupante, demonstrando que as técnicas de acasalamento dirigido utilizadas têm sido efetivas nesse controle. A metodologia de controle da taxa endogâmica utilizada nessa população em específico é baseada na referência obtida por meio dos programas de melhoramento e pela utilização de mais de um software de acasalamento, permitindo comparações e combinações que resultem em valores genéticos ótimos sem prejuízos endogâmicos.

Em meta-análise realizada com dados de 30 anos de pesquisas, Doekes et al., (2021) observaram que a cada 1% de aumento no coeficiente de endogamia há uma diminuição média de 0,13% da média de características de interesse para o setor pecuário. O coeficiente de endogamia médio dessa população demonstra que não houve grandes perdas de diversidade genética pelo uso intenso de alguns touros, o processo de seleção tem espaço para se desenvolver sem a necessidade de introdução de animais de outros rebanhos, garantindo que a formação de uma linhagem própria seja continuada sem prejuízos no progresso genético e uniformidade do rebanho.

Para o CLUSTER 1 por exemplo, verifica-se que 7% dos touros responsáveis pelos animais que constituem esse agrupamento são os pais de 46% de toda a progênie, com melhores valores para consumo alimentar residual da população. A maior taxa de endogamia nos clusters, foi de 0,72% no CLUSTER 3, não chegando nem a 1%, demonstrando que o processo de acasalamento dirigido utilizado pela propriedade tem sido eficiente no controle. A TABELA 3 demonstra que enquanto pais, eles são responsáveis por no mínimo 46% por cento dos animais agrupados (CLUSTER 1). Desempenhos superiores determinados por um número relativamente pequeno de touros, confirmam uma maior fixação dessas características nesses animais, que podem ainda ser utilizados de modo bastante intensivo, sem grandes incrementos de endogamia, desde que haja um controle nos acasalamentos.

5 Conclusões

As diferenças de desempenho dos descendentes de seis touros com maior número de filhos, apesar de passarem por processo de seleção semelhante, sugerem que caminham no mesmo sentido no aumento nos valores genéticos. Portanto, os animais filhos desses touros podem ser utilizados em acasalamentos entre clusters para corrigir desempenhos inferiores, e dentro do mesmo cluster, para maior fixação das características, visto que a endogamia é reduzida.

As análises permitiram identificar e caracterizar grupos de animais de acordo com a similaridade genética. Os grupos apesar de alta proporção de filhos de um número reduzido de touros, apresentaram coeficientes de endogamia baixos

A identificação dos grupos permite direcionar o trabalho de seleção e verificar quais as deficiências dentro do rebanho e o potencial genético existente.

6-Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, por meio de bolsa de estudos.

Fornecimento de dados pelas empresa Genética Aditiva® e Melhora Mais consultoria Genética ®.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS EXPORTADORAS DE CARNES [ABIEC]. Beef Report: Perfil da pecuária no Brasil. [S. l.: s. n.], 2021. Available at: <http://abiec.com.br/publicacoes/beef-report-2021/>.

ACNB. Associação dos Criadores de Nelore do Brasil, 2006. Disponível em: <<http://www.nelore.org.br/Institucional/ACNB>>. Acesso em: 20 de julho de 2021

BARROS, I. C. et al. Genetic evaluation of the growth of Polled Nelore via multi-trait models. *Revista Ceres*, vol. 65, no. 5, p. 402–406, 2018. <https://doi.org/10.1590/0034-737X201865050004>.

BECKER, R. A., Chambers, J. M. and Wilks, A. R. *The New S Language*. Wadsworth & Brooks/Cole. 1988.

BOLIGON, A. A., VICENTE, I. D. S., ROSO, V. M., SOUZA, F. R. P. Direct and maternal annual genetic changes for selected traits at weaning and yearling in beef cattle. *Acta Scientiarum - Animal Sciences*, 41(1), 1–9, 2019. <https://doi.org/10.4025/actascianimsci.v41i1.42572>

BONAMY, M. et al. Genetic association between different criteria to define sexual precocious heifers with growth, carcass, reproductive and feed efficiency indicator traits in Nelore cattle using genomic information. *Journal of Animal Breeding and Genetics*, vol. 136, no. 1, p. 15–22, 2019. <https://doi.org/10.1111/jbg.12366>.

CARDOSO, D. F. et al. Genome-wide scan reveals population stratification and footprints of recent selection in Nelore cattle. *Genetics Selection Evolution*, vol. 50, no. 1, p. 22, 2 Dec. 2018. DOI 10.1186/s12711-018-0381-2. Available at: <https://doi.org/10.1186/s12711-018-0381-2>.

DOEKES, H. P.; BIJMA, P.; WINDIG, J. J. How depressing is inbreeding? A meta-analysis of 30 years of research on the effects of inbreeding in livestock. *Genes*, vol. 12, no. 6, 2021. <https://doi.org/10.3390/genes12060926>.

FERNANDES JÚNIOR, G. A. et al. Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nelore (*Bos primigenius indicus*). *Scientific Reports*, vol. 10, no. 1, p. 1–7, 2020. <https://doi.org/10.1038/s41598-020-66272-7>.

GONZÁLEZ-RECIO, O.; LÓPEZ DE MATURANA, E.; GUTIÉRREZ, J. P. Inbreeding depression on female fertility and calving ease in spanish dairy cattle. *Journal of Dairy Science*, vol. 90, no. 12, p. 5744–5752, 2007. <https://doi.org/10.3168/jds.2007-0203>.

GORJANC, G.; HENDERSON, D. a. Calculation of genetic relatedness / relationship between individuals in the pedigree. vol. 32, no. 2002, p. 23–24, 2021. .

HUSSON, A. F.; JOSSE, J.; LE, S.; MAZET, J.; HUSSON, M. F. Package ‘FactoMineR.’ [S. l.: s. n.], 2020.

HUSSON, F.; JOSSE, J.; JÉRÔME PAGÈS. Principal component methods - hierarchical clustering - partitional clustering: why would we need to choose for visualizing data? *Formalized Mathematics*, vol. 18, no. 1, p. 17–26, 2010. <https://doi.org/10.2478/v10037-010-0003-0>.

KASSAMBARA, A. Practical guide to principal component methods in R. Montpellier: STHDA, 2017.

KENDALL, M. G. A new measure of rank correlation, *Biometrika*, 30, 81–93, 1938. doi: 10.1093/biomet/30.1-2.81.

KLUSKA, S. et al. Estimates of genetic parameters for growth, reproductive, and carcass traits in Nelore cattle using the single step genomic BLUP procedure. *Livestock Science*, vol. 216, p. 203–209, 2018. <https://doi.org/10.1016/j.livsci.2018.08.015>.

LOZADA-SOTO, E. A. et al. Trends in genetic diversity and the effect of inbreeding in American Angus cattle under genomic selection. *Genetics Selection Evolution*, vol. 53, no. 1, p. 1–15, 2021. DOI 10.1186/s12711-021-00644-z. Available at: <https://doi.org/10.1186/s12711-021-00644-z>.

MALÉCOT, G. Les Mathématiques de l'hérédité. Paris: Masson et Cie, 1948. <https://doi.org/10.1051/978-2-7598-2085-6.c061>.

MEUWISSEN, T. H. E.; LUO, Z. Computing inbreeding coefficients in large populations. *Genetics, Selection, Evolution*, vol. 24, no. 4, p. 305–313, 1992. <https://doi.org/10.1186/1297-9686-24-4-305>.

MUDADU, M. A. et al. Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *BMC Genomics*, vol. 17, no. 1, p. 1–16, 2016. DOI 10.1186/s12864-016-2535-3. Available at: <http://dx.doi.org/10.1186/s12864-016-2535-3>.

OLIVEIRA, P. S. et al. Estrutura populacional de rebanho fechado da raça Nelore da linhagem Lemgruber. *Pesquisa Agropecuaria Brasileira*, vol. 46, no. 6, p. 639–647, 2011. <https://doi.org/10.1590/S0100-204X2011000600010>.

PÉREZ-PARDAL, L. et al. Legacies of domestication, trade and herder mobility shape extant male zebu cattle diversity in South Asia and Africa. *Scientific Reports*, vol. 8, no. 1, p. 1–8, 2018. <https://doi.org/10.1038/s41598-018-36444-7>.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL <https://www.R-project.org/>.

SEABURY, C. M. et al. Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics*, vol. 18, no. 1, p. 1–25, 2017. <https://doi.org/10.1186/s12864-017-3754-y>.

SILVA NETO, J. B. et al. Genetic correlation estimates between age at puberty and growth, reproductive, and carcass traits in young Nelore bulls. *Livestock Science*, vol. 241, no.

September, p. 104266, 2020. DOI 10.1016/j.livsci.2020.104266. Available at: <https://doi.org/10.1016/j.livsci.2020.104266>.

TRAMONTE, N. C. et al. Genetic parameters, genetic trends, and principal component analysis for productive and reproductive traits of guzera beef cattle. *Revista Brasileira de Zootecnia*, vol. 48, 2019. <https://doi.org/10.1590/RBZ4820180034>.

TWOMEY, A. J.; CROMIE, A. R.; MCHUGH, N.; BERRY, D. P. Validation of a beef cattle maternal breeding objective based on a cross-sectional analysis of a large national cattle database. *Journal of Animal Science*, vol. 98, no. 11, p. 1–15, 2020. <https://doi.org/10.1093/JAS/SKAA322>.

WARD, J. H. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, vol. 58, no. 301, p. 236–244, 1963. <https://doi.org/10.1080/01621459.1963.10500845>.

WRIGHT, S. Coefficients of Inbreeding and Relationship Author (s): Sewall Wright
Source : *The American Naturalist* , Vol . 56 , No . 645 (Jul . - Aug . , 1922), pp . 330-338
Published by : The University of Chicago Press for The American Society of Naturalists Sta.
The American Naturalist, vol. 56, no. 645, p. 330–338, 1922.

APÊNDICE- Tabela de correlação de Pearson entre as características estudadas e análise descritiva dos agrupamentos

TABELA 1. Correlação de Pearson entre características estudadas

	MP120	MP210	DAOL	DACAB	DMAR	DCAR	DIPM	D3P	DP210	DP365	DP450	DIMS
MP120	1	1.000	-0.110	0.110	-0.170	0.110	0.000	-0.010	-0.110	0.110	0.100	-0.010
MP210		1	-0.110	0.120	-0.170	0.110	-0.010	0.000	-0.120	0.100	0.100	-0.010
DAOL			1	0.020	0.040	0.010	-0.030	0.020	0.470	0.490	0.470	0.330
DACAB				1	0.080	0.130	-0.190	0.100	-0.100	-0.020	-0.010	0.070
DMAR					1	0.110	-0.140	-0.100	0.080	0.060	0.060	0.150
DCAR						1	-0.090	-0.030	0.010	0.080	0.090	0.670
DIPM							1	-0.020	0.000	0.010	0.010	-0.060
D3P								1	-0.080	-0.060	-0.060	-0.080
DP210									1	0.880	0.840	0.490
DP365										1	0.960	0.550
DP450											1	0.570
DIMS												1

MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

Tabela 2. Análise descritiva dos animais que compõe o Cluster 1 para todas as características avaliadas

CLUSTER 1						
	Média	DP	Var	Min	Max	n
MP120	2,748	1,399	1,956	-0,87	7,490	1428
MP210	3,658	1,878	3,526	-1,280	9,590	1428
DP210	7,074	2,630	6,919	-1,630	14,910	1428
DP365	14,188	3,924	15,398	-1,530	24,070	1428
DP450	14,743	4,399	19,349	-4,230	26,420	1428
DAOL	1,091	1,334	1,780	-3,120	5,520	1428
DACAB	0,373	0,329	0,108	-0,66	1,430	1428
D3P	77,333	7,601	57,771	42,710	95,770	1428
DMAR	0,004	0,127	0,016	-0,48	0,37	1428
DCAR	-0,038	0,073	0,005	-0,302	0,177	1428
DIMS	-0,105	0,137	0,019	-0,665	0,311	1428
DIPM	-0,383	0,498	0,248	-2,099	1,349	1428

DP: desvio padrão; Var: variância; Min: mínimo; Max: máximo; n: quantidade de animais; MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

Tabela 3. Análise descritiva dos animais que compõe o cluster 2 para todas as características avaliadas

CLUSTER 2						
	Média	DP	Var	Min	Max	n
MP120	0,664	1,271	1,614	-5,930	4,580	1377
MP210	0,790	1,738	3,020	-8,660	6,270	1377
DP210	10,681	2,671	7,134	1,170	18,600	1377
DP365	18,330	4,036	16,290	3,850	29,250	1377
DP450	19,298	4,454	19,840	3,210	32,360	1377
DAOL	2,202	1,317	1,736	-3,000	6,840	1377
DACAB	0,347	0,305	0,093	-0,520	1,370	1377
D3P	77,497	6,536	42,716	49,630	94,480	1377
DMAR	0,083	0,141	0,020	-0,370	0,500	1377
DCAR	-0,029	0,070	0,005	-0,276	0,174	1377
DIMS	-0,003	0,128	0,016	-0,454	0,370	1377
DIPM	-0,495	0,519	0,270	-2,078	1,484	1377

DP: desvio padrão; Var: variância; Min: mínimo; Max: máximo; n: quantidade de animais; MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;

Tabela 4. Análise descritiva dos animais que compõe o cluster 3 para todas as características avaliadas

CLUSTER 3						
	Média	DP	Var	Min	Max	n
MP120	3,600	1,375	1,891	-1,340	9,150	1682
MP210	4,821	1,858	3,454	-2,050	12,340	1682
DP210	10,494	2,779	7,721	-0,140	19,420	1682
DP365	20,463	3,991	15,926	5,130	31,790	1682
DP450	21,607	4,306	18,538	4,900	34,540	1682
DAOL	2,039	1,372	1,883	-2,400	6,950	1682
DACAB	0,512	0,339	0,115	-0,870	1,950	1682
D3P	77,579	7,512	56,425	44,970	94,300	1682
DMAR	0,043	0,138	0,019	-0,370	0,500	1682
DCAR	0,005	0,087	0,008	-0,357	0,332	1682
DIMS	0,043	0,136	0,019	-0,392	0,502	1682
DIPM	-0,634	0,520	0,270	-2,524	1,077	1682

DP: desvio padrão; Var: variância; Min: mínimo; Max: máximo; n: quantidade de animais; MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca.

Tabela 5. Análise descritiva dos animais que compõe o Cluster 4 para todas as características avaliadas

CLUSTER 4						
	Média	DP	Var	Min	Max	n
MP120	2,049	1,808	3,269	-4,180	7,300	2092
MP210	2,657	2,452	6,010	-6,020	9,360	2092
DP210	15,247	2,939	8,637	6,400	27,940	2092
DP365	26,306	4,091	16,737	13,220	44,020	2092
DP450	27,762	4,432	19,643	12,040	47,290	2092
DAOL	3,147	1,382	1,911	-1,630	8,870	2092
DACAB	0,348	0,323	0,104	-0,810	1,410	2092
D3P	73,689	8,025	64,408	38,950	93,030	2092
DMAR	0,092	0,124	0,015	-0,310	0,500	2092
DCAR	0,001	0,076	0,006	-0,272	0,260	2092
DIMS	0,153	0,126	0,016	-0,362	0,626	2092
DIPM	-0,494	0,534	0,285	-2,426	1,771	2092

DP: desvio padrão; Var: variância; Min: mínimo; Max: máximo; n: quantidade de animais; MP120: materna 120 dias; MP210: materna 210 dias; DAOL: área de olho de lombo; DACAB: acabamento; DMAR: marmoreio; DCAR: consumo alimentar residual; DIPM: idade precocidade em machos; D3P: probabilidade de parto precoce; DP210: peso aos 210 dias; DP365: peso aos 365 dias; DP450: peso aos 450 dias; DIMS: ingestão de matéria seca;